

8
2022

The Data Sprint Approach to Research: Experiments, Protocols and Knowledge

Coordinated by: **Dr. Janna Joceli Omena**

University of Warwick, Centre for Interdisciplinary Methodologies, UK
ICNOVA, iNOVA Media Lab, Portugal
Public Data Lab

Dr. Beatrice Gobbo

Politecnico di Milano, Density Design Lab, Italy

Dr. Lorena Cano-Orón

Universitat de València, Spain

Dr. Ana Marta M. Flores

ICNOVA, Universidade NOVA de Lisboa, iNOVA Media Lab, Portugal



Dígitos. Revista de Comunicación Digital

Annual Journal (2015 -)
ISSN 2444-0132

Departamento Teoría de los Lenguajes y Ciencias de la Comunicación
Facultad de Filología, Traducción y Comunicación
Universitat de Valencia
Avenida Blasco Ibáñez, 32
46010 Valencia, España



This work is licensed under a Creative Commons Attribution 3.0 License.

EDITOR

Guillermo López García, Universitat de València, España

ASSISTANT DIRECTOR

Lorena Cano-Orón, Universitat de València, España

EDITORIAL BOARD

Dafne Calvo, Universidad de Valladolid, España
Vicente Fenoll Tome, Universitat de València, España
José Vicente Gámir Ríos, Universitat de València, España
Silvia Guillamón Carrasco, Universitat de València, España
Irene Liberia Vaya, Universidad de Sevilla, España
Germán Llorca Abad, Universitat de València
Janna Joceli Omena, Universidade Nova de Lisboa, Portugal
Laura Pérez Altable, Universitat Pompeu Fabra
Meritxell Roca Sales, Columbia University in the City of New York, EEUU
Lidia Valera Ordaz, Universitat de València, España

SCIENTIFIC COMMITTEE

José Ignacio Aguaded Gómez, Universidad de Huelva, Spain
María Bella Palomo Torres, Universidad de Málaga, Spain
Salomé Berrocal Gonzalo, Universidad de Valladolid, Spain
Rodrigo Browne Sartori, Universidad Austral, Chile
Daniel Cabrera Altieri, Universidad de Zaragoza, Spain
Eva Campos Domínguez, Universidad de Valladolid, Spain
María José Cantalapiedra González, Universidad del País Vasco, Spain
Andreu Casero Ripollés, Universitat Jaume I de Castelló, Spain
Jorge Catalá Carrasco, University of Newcastle, United Kingdom
Lluís Codina Bonilla, Universitat Pompeu Fabra, Spain
Carmina Crusafon, Universitat Autònoma de Barcelona, Spain
José Luis Dader García, Universidad Complutense de Madrid, Spain
Javier Díaz Noci, Universidad Pompeu Fabra, Spain
Concha Edo Bolós, Universidad Complutense de Madrid, Spain
José Alberto García Avilés, Universidad Miguel Hernández, Spain
Felip Gascón i Martín, Universidad de Playa Ancha, Chile
José Luis González Esteban, Universidad Miguel Hernández, Spain
Antonio Laguna Platero, Universidad de Castilla - La Mancha, Spain
Xosé López García, Universidad de Santiago de Compostela, Spain
Javier Marzal Felici, Universitat Jaume I de Castelló, Spain
Juan Julián Merelo Guervós, Universidad de Granada, Spain
Koldobika Meso Ayerdi, Universidad del País Vasco, Spain
Juan Carlos Miguel de Bustos, Universidad del País Vasco, Spain
Marcial Murciano Martínez, Universitat Autònoma de Barcelona, Spain
José Luis Orihuela Colliva, Universidad de Navarra, Spain
José Manuel de Pablos Coello, Universidad de La Laguna, Spain
José Manuel Palacio Arranz, Universidad Carlos III de Madrid, Spain
Montserrat Quesada Pérez, Universitat Pompeu Fabra, Spain
Javier del Rey Morató, Universidad Complutense de Madrid, Spain
Ramón Salaverría Aliaga, Universidad de Navarra, Spain
Víctor Sampedro Blanco, Universidad Rey Juan Carlos, Spain
José Manuel Sánchez Duarte, Universidad Rey Juan Carlos, Spain
Francisco Sierra Caballero, Universidad de Sevilla, Spain
Anna Tous Roviroso, Universitat Autònoma de Barcelona, Spain
Carlos del Valle Rojas, Universidad de La Frontera, Chile

EDITORIAL DESIGN

Lorena Cano-Orón, Universitat de València, Spain

DIGITAL COMMUNICATION

Marta Meneu Borja

GRAPHIC ELEMENTS

La portada ha sido diseñada usando imágenes de slidesgo / Freepik.com



Guillermo López García 5

Special Issue

The *Data Sprint Approach to Research: Experiments, Protocols and Knowledge*

What are data sprints for? 9

Janna Joceli Omena, Lorena Cano-Orón, Beatrice Gobbo, Ana Marta M. Flores

Data Sprint Learning. Exercising Proximity to Data in Teaching Situations 31

Laura Kocksch, Mace Ojala, Katharina Kinder-Kurlanda

“Hands On”: incorporación de *data sprints* y *digital methods* en la docencia de marketing digital 51

Òscar Coromina Rodríguez, Adrián Padilla Molina

El método *data sprint*: un ejercicio de reflexividad feminista sobre las prácticas de producción de conocimiento 67

María Concepción Castillo-González, Dorismilda Flores-Márquez, Gabriela Elisa Sued

Staging and storing data sprint-based research results: a communication design approach 87

Beatrice Gobbo, María de los Ángeles Briones Rojas, Elena Aversa, Tommaso Elli, Andrea Benedetti, Michele Mauri, Gabriele Colombo

Visualizing an image network without rendering files: A method to combine user hashtags with computer vision labels 109

Giulia Tucci

Exploring User Engagement with Portuguese Political Party Pages on Facebook: Data Sprint as Workflow 127

Jorge Martins Rosa, N. Gizem Bacaksızlar Turbic, Alda Magalhães Telles, Clara González Tosat, Cristian Jiménez Ruiz, Kalliopi Moraiti, Oğuz Özgür Karadeniz, Valentina Pallacci

Visual Communication strategy of populist leaders on Instagram in 2020 155

Ilya Lavrov, Chiara Miozzo, Marie Palaffre, Franziska Schranz

Open Zone

Game frames y populismo en las elecciones al Parlament de Catalunya de 2021. Actividad, temas y discursos de los candidatos en Twitter 179

Miguel Ibáñez-Cuquerella

Feminist discourses of Spanish digital influencers: An analysis of their activity on social media 201

Celina Navarro, Isabel Villegas-Simón

El tratamiento periodístico de la privacidad en las aplicaciones de rastreo de COVID-19 en España y Reino Unido 215

Sofiya Kamalova, Alfredo Moralejo

El consumo de alcohol, tabaco y drogas en los jóvenes: un estudio sobre las teen series españolas recientes (2015-2021) 231

Patricia Palomares-Sánchez, Tatiana Hidalgo-Marí, Jesús Segarra-Saavedra

La cobertura mediática de la ocupación de viviendas y el pánico moral. Análisis de los magacines matinales de TVE y Telecinco 251


Lourdes García Molina

The Story Behind the Story: Reporting Conditions in China During the Coronavirus Crisis as Told by the Correspondents of El País and ABC 273

Adrià Calatayud-Vaello



EDITORIAL

 **Guillermo López García**
Guillermo.Lopez@uv.es
Universitat de València

Comenzamos el proyecto de la revista *Dígitos* hace ahora ocho años. Desde entonces, y merced a un esfuerzo no menor, hemos conseguido que la revista llegara puntualmente a su cita anual con los lectores y la comunidad investigadora, albergando trabajos de diversa índole, encuadrados bajo el marco genérico de los estudios de comunicación, y con especial énfasis en el impacto de las tecnologías digitales sobre el ecosistema comunicativo. Buena parte de los trabajos, en cada número, han venido encuadrados como parte de un tema monográfico, lo que nos ha permitido trabajar con investigadores de diferentes instituciones, formación y líneas de investigación, encargados de coordinar cada monográfico, seleccionar las propuestas aprobadas por los revisores externos y darle forma de conjunto al monográfico.

Durante todo este tiempo, la revista *Dígitos* ha seguido creciendo. Hemos procurado siempre ser una ventana abierta a la comunidad investigadora, acogiendo las propuestas que revestían suficiente calidad, según el juicio de los editores y revisores de la revista, así como difundiéndonlas entre el público interesado. A lo largo de todos estos años, la revista *Dígitos* se ha mantenido como una publicación de acceso abierto y libre, en la que tanto los autores como los lectores puedan difundir y leer contenidos sin ningún tipo de trabas o requerimientos económicos. Comento esto, que desde mi punto de vista debería ser algo común en las publicaciones académicas, y que para nosotros constituye una apuesta estratégica y una defensa básica de la investigación y la actividad académica -tal y como la hemos entendido siempre-, porque desgraciadamente es una circunstancia cada vez menos habitual.

En este número de *Dígitos*, albergamos un monográfico editado por cuatro investigadoras: Janna Joceli Omena (University of Warwick), Beatrice Gobbo (Politecnico di Milano, Density Design Lab), Lorena Cano-Orón (Universitat de València) y Ana Marta M. Flores (Universidade NOVA de Lisboa). La temática, "El método Data Sprint para investigar: Experimentos, protocolos y conocimiento", da cabida a los siguientes artículos: "What are data sprints for?" por Janna Joceli Omena (University of Warwick), Lorena Cano-Orón (Universitat de València), Beatrice Gobbo (Politecnico di Milano, Density Design Lab) y Ana Marta M. Flores (Universidade NOVA de Lisboa); "Data Sprint Learning. Exercising Proximity to Data in Teaching Situations" por Laura Kocksch (Technoanthropology Lab, Aalborg University), Mace Ojala (IT University of Copenhagen) y Katharina Kinder-Kurlanda (University of Klagenfurt); "Hands On': incorporación de *data sprints* y *digital methods* en

la docencia de marketing digital” por Òscar Coromina Rodríguez (Malmö University) y Adrián Padilla Molina (Universitat Autònoma de Barcelona); “El método *data sprint*: un ejercicio de reflexividad feminista sobre las prácticas de producción de conocimiento” por María Concepción Castillo-González (Tecnológico de Monterrey), Dorismilda Flores-Márquez (Universidad De La Salle Bajío) y Gabriela Elisa Sued (Universidad Nacional Autónoma de México, Instituto De Investigaciones Sociales); “Staging and storing data sprint-based research results: a communication design approach” por Beatrice Gobbo, María de los Ángeles Briones Rojas, Elena Aversa, Tommaso Elli, Andrea Benedetti, Michele Mauri (Density Design Lab, Politecnico di Milano) y Gabriele Colombo (King’s College London); “Visualizing an image network without rendering files: A method to combine user hashtags with computer vision labels” por Giulia Tucci (Universidade Federal do Rio de Janeiro); “Exploring User Engagement with Portuguese Political Party Pages on Facebook: Data Sprint as Workflow” por Jorge Martins Rosa (Universidade Nova de Lisboa), N. Gizem Bacaksızlar Turbic (GESIS - Leibniz Institute for the Social Sciences), Alda Magalhães Telles (Universidade Nova de Lisboa), Clara González Tosat (Universidad de Navarra), Cristian Jiménez Ruiz (ISCTE-IUL & Lisbon University), Kalliopi Moraiti (University of Gothenburg), Oğuz Özgür Karadeniz (KU Leuven DTAI) y Valentina Pallacci (Politecnico di Milano); “Visual Communication strategy of populist leaders on Instagram in 2020” por Ilya Lavrov, Chiara Miozzo, Marie Palaffre y Franziska Schranz (Universidade NOVA de Lisboa).

Además, nuestra sección Zona Abierta de la revista alberga en este número 8 un total de seis artículos: “Game frames y populismo en las elecciones al Parlament de Catalunya de 2021. Actividad, temas y discursos de los candidatos en Twitter”, por Miguel Ibáñez-Cuquerella. Este artículo, además, fue el galardonado con el VI Premio Predoctoral de Investigación de la revista Dígitos. “Feminist discourses of Spanish digital influencers: An analysis of their activity on social media”, de Celina Navarro (Universitat Autònoma de Barcelona) e Isabel Villegas-Simón (Universitat Pompeu Fabra); “El tratamiento periodístico de la privacidad en las aplicaciones de rastreo de COVID-19 en España y Reino Unido”, por Sofiya Kamalova y Alfredo Moralejo (Universidad de Alicante); “El consumo de alcohol, tabaco y drogas en los jóvenes: un estudio sobre las teen series españolas recientes (2015-2021)”, de Patricia Palomares-Sánchez, Tatiana Hidalgo-Marí y Jesús Segarra-Saavedra (Universidad de Alicante); “La cobertura mediática de la ocupación de viviendas y el pánico moral. Análisis de los magazines matinales de TVE y Telecinco”, por Lourdes García-Molina (Universitat de València); y “The Story Behind the Story: Reporting Conditions in China During the Coronavirus Crisis as Told by the Correspondents of El País and ABC”, por Adrià Calatayud-Vaello (Universitat Pompeu Fabra).

Este octavo número es también, para mí, la ocasión oportuna para dejar la dirección de la revista Dígitos. Quisiera agradecer a mis compañeros del Consejo de Redacción de Dígitos, a los autores, revisores y coordinadores de monográficos, a los miembros del Comité Científico, y a la subdirectora de la revista, Lorena Cano, su apoyo todos estos años, así como al Departamento de Teoría de los Lenguajes y Ciencias de la Comunicación de la Universidad de Valencia, por su confianza en el proyecto y su patrocinio.

Ocho años es tiempo más que suficiente para desarrollar un proyecto, ver cómo evoluciona y, en su caso, si ha ido razonablemente bien, cómo continúa su camino y sigue creciendo. Máxime en este caso, porque la persona que a partir de ahora se encargará de liderar el proyecto como directora de la revista Dígitos, la profesora Lorena Cano-Orón, no sólo ha estado en su misma génesis y me ha acompañado en la gestión de la revista todos estos años, primero como secretaria y luego como subdirectora de la misma, sino que en la práctica se ha encargado en los últimos tiempos de la dirección de la revista.

En efecto, con gran generosidad, Lorena Cano ha ejercido *de facto* como directora de Dígitos en los últimos años, conforme a mí las vicisitudes de la vida (que podríamos resumir con dos palabras: “hijos” y “cargos”; no me hagan dilucidar aquí cuál de las dos supone más trabajo; más satisfacciones y felicidad, pueden ustedes figurarse que la primera) me llevaban por otros derroteros. Para mí es una gran alegría que ahora pase a ser también directora de Dígitos *de iure*, tras la aprobación unánime del Consejo Editorial de la revista. No es sólo que piense que a Dígitos le irá bien con ella al frente; es que tengo la certidumbre de que le irá mucho mejor que hasta ahora. Les deseo, a ella y a la revista, la mejor de las suertes en la nueva andadura que ahora se abre.

SPECIAL ISSUE

The Data Sprint Approach
to Research: Experiments,
Protocols and Knowledge



What are data sprints for?

¿Para qué sirven los data sprints?

ID Janna Joceli Omena

Janna.Joceli.Omena@warwick.ac.uk
Centre for Interdisciplinary Methodologies,
University of Warwick
ICNOVA, iNOVA Media Lab; Public Data Lab

ID Lorena Cano-Orón

Lorena.Cano@uv.es
Universitat de València

ID Beatrice Gobbo

beatrice.gobbo@polimi.it
Density Design Lab, Politecnico di Milano

ID Ana Marta M. Flores

amflores@fcsh.unl.pt
ICNOVA, Universidade NOVA de Lisboa

ABSTRACT

The data sprint approach enables various objectives for the digital methods community, from fostering interdisciplinary collaboration to providing teaching-learning spaces regarding creative methods. However, data sprints' purpose, advantages and concrete results are still little known across disciplines. Thus, this paper presents four facets *pertaining to* and *deriving from* data sprints to explain their prospects for different (non-) academic communities. First, we define the data sprint approach, providing a detailed description of what data sprints are and what they involve in practice and, in turn, propose guidance to facilitate the replicability of this work method. Second, we elucidate how data sprints are (1) *a means of teaching and learning digital methods research*, arguing that the data sprint environment is not only (2) *a space for methods and tools creation* but also (3) *a reflective tool* to understand the triad of data-, software- and platform-oriented research (from the standpoint of practice). Therefore, data sprints offer researchers situational ways to access and critique scientific knowledge production. Finally, we address a standard post-sprint procedure, (4) *the reutilisation of data sprint reports for producing scientific knowledge*, through academic and non-academic publications, as an established research practice. The four facets unpack the data sprint approach for a broader audience whilst indicating the possible takeaways during and after such events. We conclude with reflections from and on data sprints.

KEYWORDS

data sprint, digital methods, science production, data analysis, data visualisation, interdisciplinary research

How to cite this article:

Omena, J. J.; Cano-Orón, L.; Gobbo, B. & Flores, A. M. (2022). What are data sprints for?. *Dígitos. Revista de Comunicación Digital*, 8: 9-30. DOI: 10.7203/drdcd.v1i8.253

RESUMEN

El método *data sprint* permite varios objetivos para la comunidad de métodos digitales, desde el fomento de la colaboración interdisciplinaria hasta la provisión de espacios de enseñanza-aprendizaje en relación con los diseños metodológicos creativos. Sin embargo, el propósito, las ventajas y los resultados concretos de los *data sprints* son todavía poco conocidos en todas las disciplinas. Este artículo presenta cuatro facetas que pertenecen y se derivan de esta metodología de trabajo para explicar sus perspectivas para diferentes comunidades (no) académicas. En primer lugar, definimos el método de los *data sprints*, proporcionando una descripción detallada de lo que son y lo que implican en la práctica y, a su vez, proponemos una guía para facilitar la replicabilidad de este método de trabajo. En segundo lugar, exponemos cómo los *data sprints* son: (1) un medio de enseñanza y aprendizaje de la investigación de métodos digitales, argumentando que el entorno de los *data sprints* no es sólo (2) un espacio para la creación de métodos y herramientas, sino también (3) una herramienta de reflexión para entender la tríada de la investigación orientada a los datos, el software y la plataforma, desde el punto de vista de la práctica. Por lo tanto, los *data sprints* ofrecen a los investigadores formas delimitadas de acceder y criticar la producción de conocimiento científico. Finalmente, abordamos un procedimiento estándar posterior al *sprint*: (4) la reutilización de los informes del *data sprint* para producir conocimiento científico, a través de publicaciones académicas y no académicas, como una práctica de investigación establecida. Estas cuatro dimensiones desvelan el enfoque del *data sprint* para un público más amplio, a la vez que indican las posibles conclusiones durante y después de tales eventos. Concluimos con reflexiones procedentes de los *data sprints* y sobre ellos.

PALABRAS CLAVE

data sprint, métodos digitales, producción científica, análisis de datos, visualización de datos, investigación interdisciplinaria



What are data sprints for?

1. Introduction

Data sprints are a well-known, conventional approach within the digital methods community. They can be part of master's and Ph.D. programs whilst also being available to a broader audience or dedicated to specific working groups and research projects. As a burst and intensive week of work with online data and methods, a range of tools and creative visualisations, the data sprint approach has made a name for itself and achieved status over the past decade. Its interdisciplinary profile plays a crucial role in this. It is no wonder that participants and collaborators from different disciplines co-operate in data sprint projects and workflows, contributing to achieving a common end, from answering specific research questions to experimenting with new methods or building research tools.

The data sprint approach has also inspired scholars, researchers and students to conduct research *with* and *about* digital platforms, online data and software. A data sprint project, for instance, can mobilise disciplines such as communication, design, sociology, history and computer science for the simultaneous study of climate change issues on social media and algorithmic recommendation systems. However, the

purpose, advantages and concrete results of data sprints are still little known across fields of study. So, what purposes does the data sprint approach serve? What does it bring to the scientific community? In addition, how does the data sprint approach work in practice? What is expected from all participants involved? This paper focuses on unpacking what data sprints are for, being organised as follows: First, we define the data sprint approach, providing a detailed description of what data sprints are and what they involve in practice; in turn, we propose guidance to facilitate the replicability of this working method. Subsequently, we present the four facets that *pertain to* and *derive from* data sprints, namely an understanding of data sprints as: (1) a means of teaching and learning digital methods research, (2) a space for methods and tools creation, (3) a reflective tool and (4) a means for producing scientific knowledge through interdisciplinary collaborations. By so doing, we argue that data sprints are more than an interdisciplinary collaborative environment; they are also a pedagogic and scientific tool. This reflection arises from the experience of the coordinators of the Dígitos monograph section 'The data sprint approach for research: experiments, protocols and knowledge' concerning the participation in, organisation of and collaboration within data sprints; and also from the experience of the authors and their work included in this special issue.

2. Data sprint: definition, characteristics and format

A data sprint is an interdisciplinary research practice that has been gaining ground in different fields and growing over the years. Unlike hackathon proposals, which bring programmers and other stakeholders together to build software programs or codes, data sprints focus on social, cultural and media projects. Although there are multiple variations related to diverse interpretations and fields, some common elements can be identified. Data sprints are time-limited and project-oriented, and they encourage co-participation amongst participants of different expertise and career levels. Research projects are developed within a period typically spanning 5 days and carried out in person, online or in hybrid modes. These projects often rely on data from the web and its technological environment to answer specific research questions. Data sprints may have different goals, but they are usually appropriate for tackling so-called ill-defined problems, namely situations where it is challenging to define an initial design brief and that require multi-disciplinary effort and expert knowledge to address the issue (Mauri *et al.*, 2019).

Data sprints have been implemented to foster and promote collaborative, interdisciplinary and experimental workflows. In this environment, participants with different backgrounds engage on specific projects whilst 'work(ing) together on a set of data and research questions' (Venturini *et al.*, 2018: 1). By providing a space for intensive collaborative workflows, tool creation, coding and methodological workshops, exploratory data analyses and visualisations, data sprints reflect the learning-by-doing principle. They thus require practical work and some technical knowledge in method implementation and data visualisation. With many people participating in real-time data analysis, data sprint activities are far from a segmented and isolated mode of

work. The richness of this approach lies in its interdisciplinary nature, since it enables ‘social scientists, developers and data designers together with relevant domain experts to explore research questions and create prototype digital methods projects’ (Munk, Madsen, & Jacomy, 2019: 110).

However, not all is rosy in data sprints, as Venturini and Munk (2022) warn when they define data sprint research as “a quick and dirty” approach. This is because of the intensity and speed with which teams shape research and test methods and generate results. Configured as an open-ended project, research in data sprints is shaped collectively by team members as data become available and analyses are carried out. This is why Rogers and Lewthwaite (2019: 19) state that data sprint projects operate “on the fly”. For Mauri and Ciuccarrelli (2016), data sprints often take the form of “unfinished products”. Indeed, at the end of the data sprint week, many of them will be discarded, but thanks to them, it will have been possible to open up new hypotheses and make progress while also answering research questions.

3. Understanding the Data Sprint Approach

We now move on to what happens before, during and after a data sprint to describe the characteristics and formats of a data sprint, as well as to explain what it is. We unpack how the data sprint approach works in practice, focusing on its different phases of preparation, realisation and results.

3.1. Before | Data sprint preparation: a checklist

Table 1 presents a data sprint preparation checklist according to the tasks assigned to organisers, collaborators and participants. This helps indicate how different activities are expected from all those involved in a data sprint, yet drawing particular attention to the role of participants.

Data sprint organisers secure funding from both inside and outside the institution they are a part of, choose the conceptual and thematic proposal, invite keynote speakers and collaborators (including those responsible for project pitches) and plan the data sprint’s programme. Organisers should consider, or try to balance the impact of, the proposed programme on both the project’s development and the participants’ learning expectations. The event’s logistics and communication management are also the responsibility of the organisers, from booking rooms or creating web-conferencing links to communicating with collaborators and attendees. Organisers coordinate activities not only before but also during and after the data sprint.

Data sprint collaborators (or facilitators) are crucial players in making the sprint happen owing to their specific field or discipline expertise, and are generally researchers, media scholars, issue experts or organisations willing to pitch a project or join a project. Also, early-career or senior professionals, such as developers, designers (also design students in training), data analysts and digital methods researchers, enable hands-on workshops and join a project.

Data sprint	Organisers	Collaborators	Participants
Checklist	<ul style="list-style-type: none"> ✓ Secure funding ✓ Define theme and practical labs ✓ Invite keynote speakers and collaborators ✓ Plan data sprint schedule ✓ Manage communication and logistics ✓ Provide a list of (video) tutorials, worksheets or recipes, as well as suggested readings for participants ✓ Coordinate activities not only before but also during and after the data sprint 	<p>[regarding hands-on workshops]</p> <ul style="list-style-type: none"> ✓ Practical labs preparation and presentation (face-to-face or pre-recorded section) ✓ Prepare method recipes and shared data folder for practical exercises <p>[regarding project management]</p> <ul style="list-style-type: none"> ✓ Provide a project description ✓ Create a project folder and datasheets for datasets ✓ Collect data and treat the dataset (if required by the project) 	<ul style="list-style-type: none"> ✓ Download and install tools and software suggested by organisers and required for the practical labs sections ✓ Watch suggested tutorials (or at least those of personal interest) ✓ Spend some time following up on the recommended tutorials and tools or codes ✓ Read suggested readings ✓ Engage with project work during and after the data sprint ✓ Collect data if required by the project

Table 1. Data sprint preparation checklist (for organisers, collaborators and participants)

Collaborators giving practical labs should prepare hands-on material to help participants use software or web-based tools, such as the Memespector Graphical User Interface (GUI) (Chao, 2021). Tutorials assist participants in making sense of digital records (i.e., hashtags, URLs and image metadata) and methodological approaches while using worksheets or method recipes for research (i.e., Public Data Lab and SMART Data Sprint recipes¹). Additionally, collaborators are also responsible for teaching specific data collection, analysis and visualisation techniques and how to make sense of innovative textual and visual methodologies. *Collaborators managing data sprint projects* have an even longer list of responsibilities. They should provide a project description containing, for example, the project title, a brief introduction to contextualise the project, research questions, what the methods, datasets and tools are, preliminary findings (if any) and references. This is where ethical considerations may arise, remaining relevant until the dissemination of the results of the data sprint. In addition, it is often recommended that data collection be done before the data sprint, leaving more space for exploratory and analytical tasks during the sprint. At this point, a shared folder for the project is created and organised with subfolders such as datasets, resources, analysis and visualisations (Figure 1). The organisation of the project folder guides how and where participants should enter their collaboration whilst providing common access to the results of the team’s work.

1 <https://recipes.publicdatalab.org/>, <https://smart.inovamedialab.org/digital-methods-recipes/>

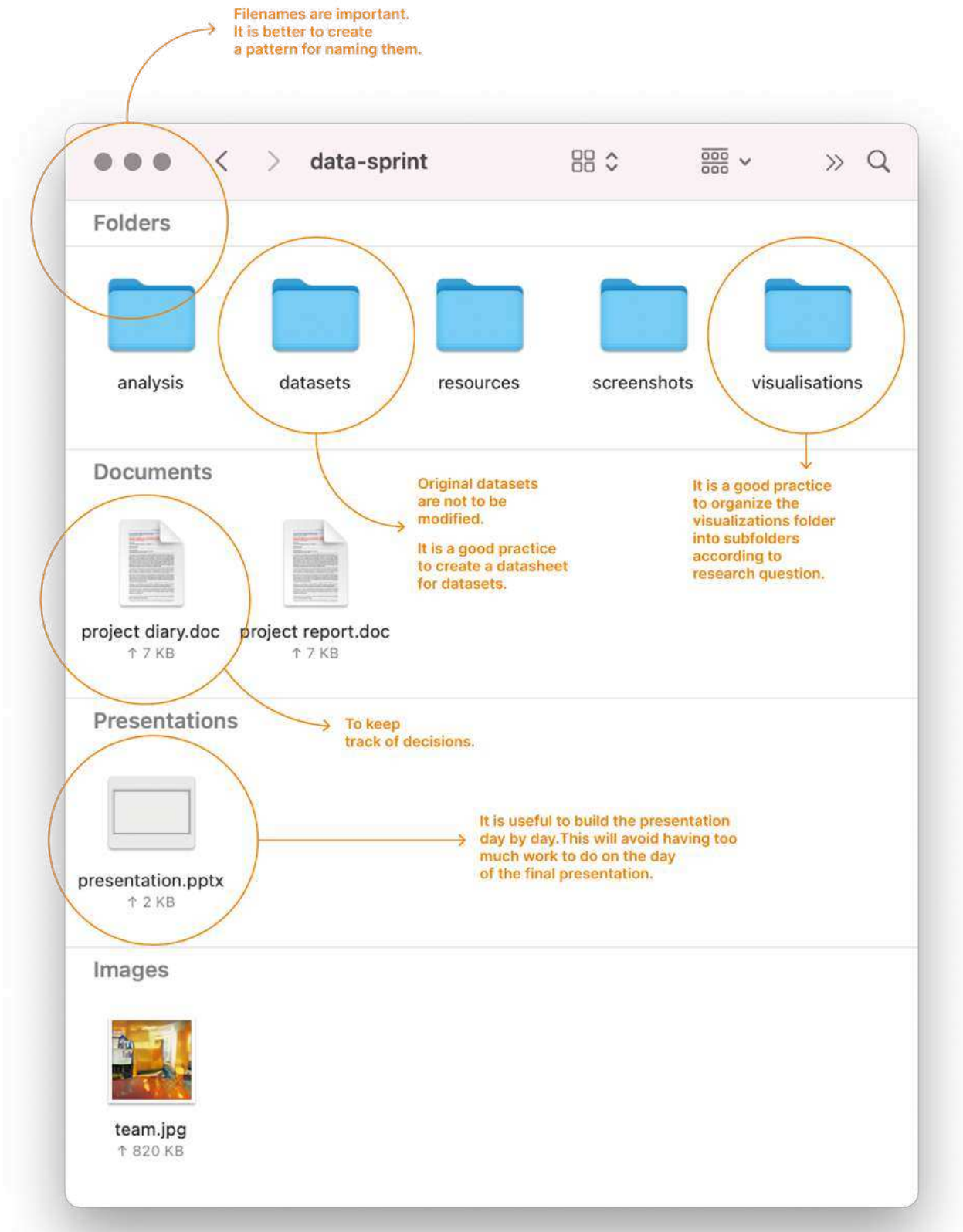


Figure 1. Sample folder for organising data sprint materials.

Project leaders are also expected to provide and look after good documentation for the dataset building (i.e., using dataset folders and datasheets for datasets) and project analysis and findings (i.e., using project diaries or shared files for analysis). Documenting practices before and during the sprint is crucial for facilitating data sprint workflows. For example, datasheets for datasets can document decisions concerning query design

Platform	Entry point(s)	Date— Data collection & extraction software	Parameters for data collection	Who collected the data?	Folder	Total number of images	Visualisation Folder	Computer Vision API & features in use	Memespector	Computer vision outputs
Facebook	"Covid Meme" Query	First week of January 2022 // Crowdtangle	(Not Meme Search) Pages and Groups Photos Only English Only January 1, 2020-January 3, 2022 Sorted by Total Interactions First 1000 Out of 8565		[link]	1005 images	[link]	Google Vision: web detection & web entities (text?)		[link]

Figure 2. Example of a *Dataset-building sheet* organised with participants one week before the 2022 Digital Methods Winter School. Project: What is a meme, technically speaking?, available at <https://wiki.digitalmethods.net/Dmi/WinterSchool2022WhatIsAMeme>

Platform/Case study	Folder-of-images	Total number of images	GCV API existing results from	Any computer vision output files from 2021?	CSV file with images annotated by GCV API	Subject of study	Descriptive info	Digital record	Entry point to data collection	Month/year of data collection & extraction software	Previous work
Tumblr - EleNÃO vs. EleSim	https://drive.google.com/drive/folders/1x-u5Yzrz8hpON5sB8tk0NHYY4fUw2fw?usp=sharing	567 images	October, 2018	No (also run web detection)	In this doc: annotated_eleao_elesim_20181003-2040	Women leading protests against Bolsonaro before 2018 Brazilian presidential elections	Images of people protesting, selfies, posters/banners, food	hashtags	#EleSim, #EleNÃO	October, 2018 (Tumblr Tool)	https://thesocialplatforms.wordpress.com/2018/10/22/elenao-vs-elesim/
Facebook - Portuguese Universities	https://drive.google.com/drive/folders/1pB0Ch11UJNDQOU3SYnmN7HC3YkwkEEy?usp=sharing	22,594 images	May, 2018	No	In this doc: annotated_universidades_20180608-1603 or here: https://drive.google.com/file/d/1gk8bGuixTcnw76ZoQck-cnGhVNE5jD_Q/view?usp=sharing	The imagery of Portuguese Universities on Facebook from 2009 to 2018. Timeline images of 15 official pages of Portuguese Higher education.	People attending academic events, sports, musical performance, buildings	image URLs from FB timeline publications	Page id	Between 2017 and 2018 (Netvizz)	https://icono14.net/ojs/index.php/icono14/article/view/1436
Instagram - Microcephaly	https://drive.google.com/drive/folders/1BbkcPCo5M0_LViWt5AhZSRcVgGqE7UJ7F?usp=sharing	10,797 images	February, 2018	Yes: June, 2021	https://drive.google.com/file/d/1iqz347c9a8zfnGjKEtjnUjaMob1gB/view?usp=sharing	The visualities associated with microcephaly in Brazil, how Brazilians families impacted by the virus shared their concerns and everyday life.	Images with infants with microcephaly accompanied by the mother or family, banners, events, spokespersons, food	hashtags	#microcefalia	Between June 2012 and October 2017 (Visual Tagnet Explorer)	Not published, but used as a reference in some presentations. Currently is part of our Disena paper.

Figure 3. Example of a *datasheet for datasets* organised weeks before the software-testing and cross-computer vision enquiry sprint at the Centre for Advanced Internet Studies in May 2021. Data sprint for testing the new version of Memespector Graphical User Interface (GUI) (Chao, 2021) and interrogating multiple computer vision APIs. Research software available at <https://github.com/jason-chao/memespector-gui>

strategies and data collection parameters, while serving as reminders of teamwork and decisions made before the subject of analysis. Figures 2 and 3 illustrate datasheets for datasets and the actions involving dataset building for studying meme collections (Figure 2), such as the choice of the platform, the query design strategy, entry points and parameters for data collection. Links to the original data folder (or files) and data processing go in this sheet. The different datasheets also show the documentation of datasets for software testing and cross-computer vision inquiry (Figure 3).

Data sprint participants extend across a wide range of academic and non-academic career levels, from master’s or Ph.D. students and early-career scholars to senior researchers and professors. Artists, journalists or data analysts are other examples of possible audience members. Data sprints demand a different level of self-sufficiency and anticipatory attitudes from their participants, unlike other academic events such

as conferences and symposiums. Therefore, the preparation of participants starts by committing to the instructions given by the data sprint organisers. Before the sprint, participants are often challenged to download and install research software and web-based tools, and spend some time following up on the recommended video tutorials and engaging with a list of suggested readings. Data sprint preparation for participants draws on the flipped learning approach (see Milman, 2012; Zuber, 2016). Students acquire knowledge before the class, using classroom time to practice and deepen understanding through interaction with peers. By following the preparatory instructions, participants can bypass the more overwhelming experience found in data sprint contexts, precisely because they have done some training in advance.

3.2. *During* | Data sprint realisation: hands-on tutorials, group work and final presentation

Keynote talks are often the opening activities of a data sprint, addressing the sprint theme and making room for reflections and discussions. However, the core activities are developed on the basis of the dynamic factors experienced by each participant. A data sprint implementation is composed of at least three phases: methods training and practising (hands-on tutorials), project development (group work) and presentation of the results (final presentation), as explained in detail below.

→ Phase 1 | Practical labs: training and practising methods

Firstly, there is the training phase. This is developed through attendance in the practical labs (or 'hands-on' tutorials) which explain the use of certain tools or the specific development of certain methods (Image 1). This is always useful, practical knowledge that is directly applicable to the participants' projects. Usually, the training is offered in parallel sessions so that a range of options can feed into different participant interests



Image 1. Practical lab on visualising data with RawGraphs during the 2019 SMART Data Sprint at NOVA University Lisbon, Portugal. Image source: https://smart.inovamedialab.org/wp-content/uploads/2019/03/IMG_0518.jpg

and project natures. It is here that data sprint collaborators share knowledge expertise, step-by-step protocols and other valuable teaching resources.

These meetings offer training that goes beyond formal education, utilising the methods of “learning by doing” and “learning from the community” methods. During data sprints, newbies and beginners acquire new skills and explore methods and tools that are not very familiar to them. This means stepping out of one’s comfort zone and opening one’s mind to new ways of approaching research in the digital environment.

➔ Phase 2 | Group work: project development

Secondly, there is the phase of project development on the basis of group work sessions, which is the longest part of the sprint, laying the foundation for the initiative. Data sprint projects are composed of the person or people responsible for the pitches, who have proposed and conceived the main objectives and methods of the project; the facilitators, who are the experts that collaborate and help participants to achieve the goals of the project (they can be communication designers, developers or methodological experts); and, of course, the participants, who are the working muscle of the project and who will shape the initial research over the course of the sprint. Participants are divided into projects according to their affinities for the topic or the methods they plan to use as input for each project. Once participants join a project, they all work together to achieve their goals.

These intense sessions usually last an entire working day. Participants work together on analyses and visualisations of research data. Here, participants and collaborators feed the project folder in a coordinated and organised way, keeping the original datasets intact whilst they work on copy versions. They discuss the methods used, try out different techniques and jointly interpret the results obtained, accounting for ethical considerations. During the group work sections, many types of output are produced: spreadsheets, tools, image collections, visualisations and (gif) animations, amongst others. This collaborative work contributes to the interdisciplinarity and richness of such study by allowing multiple points of view to contribute to the construction of the research.

The project members are encouraged to develop a draft report on a daily basis during the sprint, using for example a shared research diary or analysis files. This will make it much easier to return to the writing of the research later, when memory starts to fail regarding how and why certain decisions were made during the research process. In addition, as the research evolves over time, the report’s narrative changes; that is, new research questions and means of data analysis appear, adapting to the project’s evolution and the group work’s daily discussion. This is why writing up data sprint reports should be done with an open mind.



Image 2. Group working during the 2020 SMART Data Sprint at NOVA University Lisbon, Portugal. Image source: <https://smart.inovamedialab.org/editions/2020-digital-methods/photo-gallery/>

→ Phase 3 | Presentation of the results

Third, there is the presentation of the results to the entire forum of participants, facilitators, organisers and people responsible for pitches of that edition of the data sprint. This communication may be experienced with some anxiety. No matter how well organised the work is, there are always things to improve at the last minute. Nevertheless, this time constraint leads the projects' teams to produce clear and understandable conclusions for the audience, even if the results haven't met the initial goals. This exercise is fundamental after the hard work of the previous days so that the participants can give meaning to and create a narrative surrounding what they have developed.

Project groups usually present their research process, research questions and findings using slides, which are focused on highlighting the research path and describing the main findings and pitfalls, as well as future research opportunities. Here, researchers can insert hyperlinks to resources such as datasets and high-quality visualisations. At this stage, communication designers are key players as they present visualisations and support the organisation of the content. The presentation is a crucial moment when researchers have the opportunity to receive feedback and input for developing the project in the future.

At the end of the week, through formal, practical and collaborative learning, the acquisition of specific competences is achieved. In addition, engaging in data sprint projects not only offers ways of keeping up to date with the methods in use by the related research community but also stimulates future opportunities for collaboration and professional endeavours.

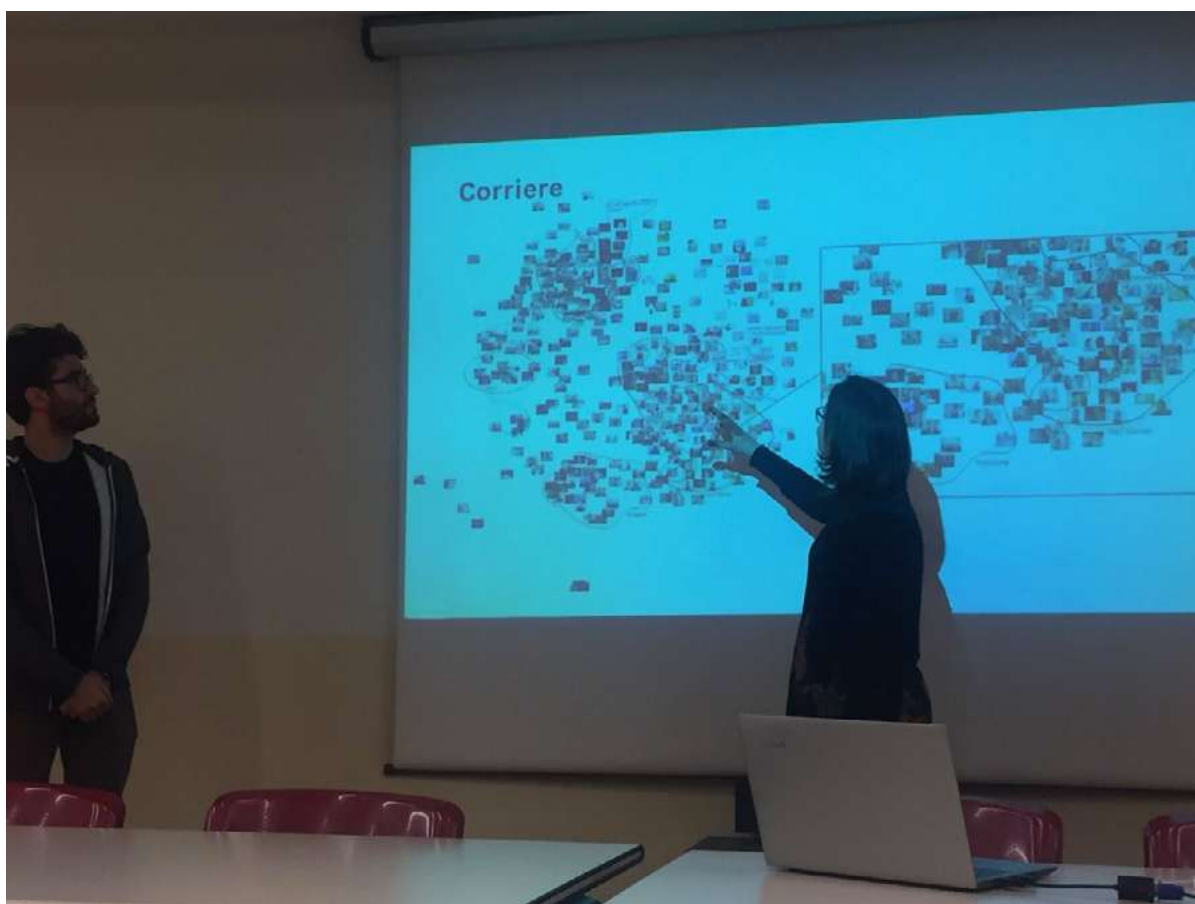


Image 3. Final presentation of project during the 2018 SMART Data Sprint at NOVA University Lisbon, Portugal. Image source: https://smart.inovamedialab.org/wp-content/uploads/2018/07/DVDCJY_XUAAtxJc.jpg

3.3. After | Data sprint results

Writing up data sprint reports

Once the results have been presented, the report that summarises the method and the results obtained to answer the research questions must be finalised. The data sprint is really only over when the project report is published on the venue provided by the organisers (i.e., the event's website). In this regard, it is good practice to draft the report throughout the sprint, and then refine its narrative, after the data sprint has ended.

Project reports derived from data sprints often have a particular structure that differentiates them from academic papers. Typically, their structure² is as follows:

1. First, **the summary of the main findings** in the research is presented. It is essential to write this part properly because it will spark interest in discovering the ins and outs of the research conducted (described throughout the report).

2. An example of a template for the Digital Methods Summer School 2022 report can be found here: <http://bit.ly/dmi22-ss-gdoc-template>; this short link is also available at <https://wiki.digitalmethods.net/Dmi/SummerSchool2022>

2. Second, **the contextualisation of the research** is displayed, detailing the subject matter and the research questions that motivated the work.
3. Third, **the methodology** followed is described: the data and methods used to answer the research questions. **Visual research protocols** (see Mauri et al., 2020) help to inform the platform(s) under investigation, dataset building, tools and analyses.
4. Fourth, **the results, discussions and conclusions** are addressed. These should be written in a concise but clear way, trying to use understandable language not only to those who know how the methods work but also to a wider audience. The conclusion section usually includes, if appropriate, possible applications of the methods and tools used in other types of research, as well as possible future research avenues to continue what was developed in the sprint.

Disseminating data sprint-based research

Research derived from data sprints, if well executed, is often rich and valuable to the academic community. This is because the research is conducted with a real multi-disciplinary perspective and an “open-mindedness” that is hardly found in more traditional work. This freshness allows work derived from data sprints to stimulate interest in the academic community, as it represents a “new” point of view on problems that have been worked on for a long time. Indeed, turning data sprint reports into academic articles has become the standard procedure for disseminating data sprint results. However, as Venturini and Munk (2022: 267) remind us, extensive refinement and documentation work is necessary to turn data sprints’ final presentations (slides or posters) and project reports into scientific research. As they state, without distilling the results, the “quick and dirty” data sprint approach would not meet the accuracy and robustness requirements of scientific research.

Disseminating the results of a data sprint may also require an interactive design process to make the content of the report browsable by a wider audience. Amongst them, web platforms and field guides are the most common communicative products in this context. Web platforms are web pages designed to collect the results of several data sprint experiments which are part of the same overarching research project (Venturini *et al.*, 2014). Unlike collections of slides and project reports, web platforms allow for the deep exploration of interactive digital content, such as networks, maps and huge collages of images. Whereas, field guides (Bounegru *et al.*, 2018) are aimed at stressing and narrating the methodological approach behind sprint-based research and packaging reusable practices to allow others to repeat them.

These are just some of the examples of how dissemination can occur as part of the current literature. However, the widespread use of the data sprint format and its various facets make it well suited as a source for results that are communicated to ever-varying audiences with diverse objectives, such as teaching resources and communication products addressed to the general public.

4. Four facets explain what data sprints are for

This section unpacks the four facets that *pertain to* and *derive from* data sprints, explaining their prospects for a broader audience (Figure 4). We first elucidate how data sprints are (1) a means of teaching and learning about digital methods research, arguing that the data sprint environment is not only (2) a space for methods and tools creation but also (3) a reflective tool to understand the triad of *data-*, *software-* and *platform-oriented research* (from the practice standpoint). Finally, we address a standard post-sprint procedure: (4) the reutilisation of data sprint reports for producing scientific knowledge, through academic and non-academic publications, as an established research practice.

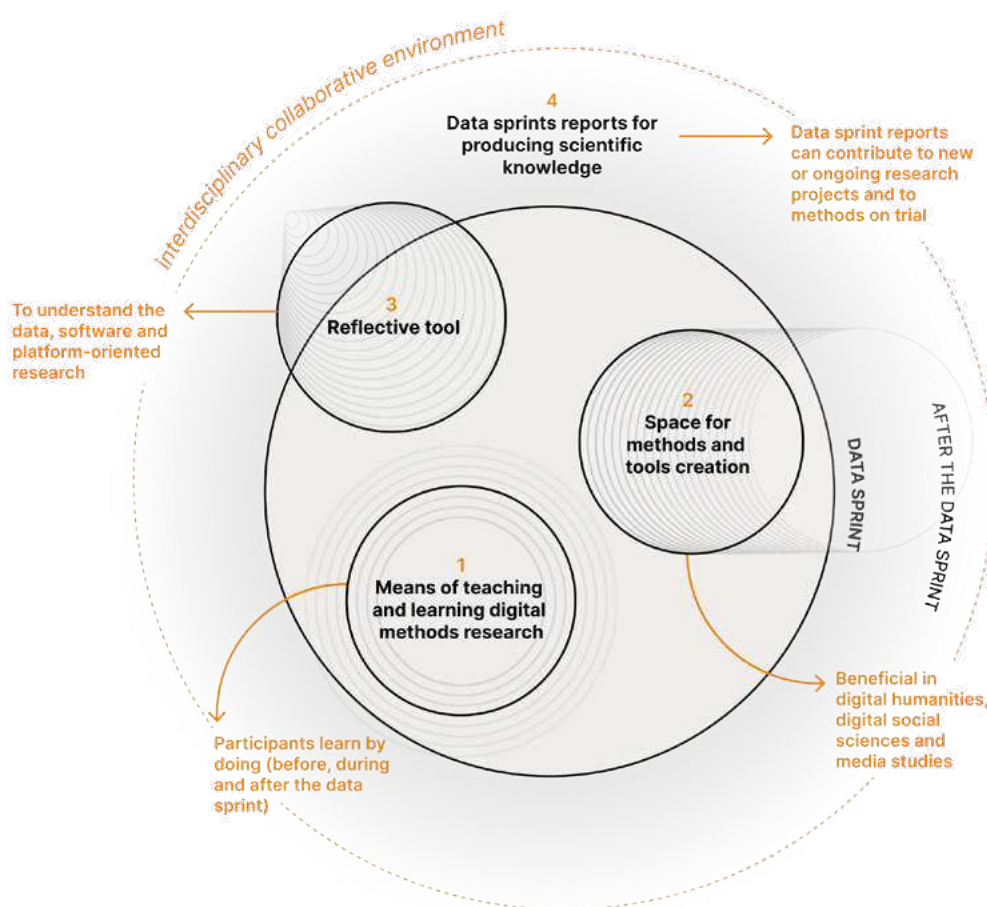


Figure 4. Four facets of data sprints

4.1. Data sprints as a means of teaching and learning digital methods research

After the a data sprint, competences in seeking out methods with different standpoints are achieved. This section elucidates data sprints as a means for teaching and learning digital methods (Omena, 2021; Rogers & Lewthwaite, 2019). Sprints offer a non-traditional sense of instructions, as Richard Rogers explains: the teaching method is tacit knowledge transfer, “where the students are gathering data, using tools, outputting analytical outputs, creating visualisations, and writing reports, all in a compressed timeframe”

(Rogers & Lewthwaite, 2019: 20). Data sprints can be considered research-led teaching devices as the support material (i.e., worksheets, methods' recipes, visual protocols and workshop slides) for the development of projects, and project themes themselves are often associated with broader lines of research and/or research projects followed by those who pitch project ideas/facilitators (Gray et al. 2022). Critical reflection on the role of platforms and software content whilst providing literacy in data analysis and visualisation is also a key teaching principle embedded in sprint activities.

Participants learn by doing by practising digital methods individually and with peers while engaging with a workgroup. By doing this, they understand in practice the rationale behind the digital methods approach (Omena, 2021). The learning process starts from both individual effort and the time pressure imposed on the collective actions and decisions during group work. If participants follow the data sprint preparation instructions (and not starting from scratch), they take significant advantage of the workshops (practical labs or tutorials). When attending these how-to sections, what is taught becomes less of a new element and more of an opportunity to clarify doubts and practice. Group work provides another opportunity for learning digital methods because projects offer "a good environment to understand not only how to make research questions but also learn in practical terms how to respond to these questions" (Omena, 2021: 42). Learning outcomes can thus be used in contexts other than the sprint.

In one way (teaching) or another (learning), being part of a data sprint also means understanding what makes sense when designing a research plan using digital methods and the reasoning behind this research practice. The relation between learning and teaching from data sprints is discussed by Mace Ojala, Laura Kocksch and Katharine Kinder-Kurlanda, who purposefully designed a sprint as a teaching situation. In their article 'Data sprint learning. Exercising proximity to data in teaching situations', they argue that data sprint learning is not exclusively for students; it can be for educators as well. They presented the data sprint preparation as part of a Ph.D. course in digital methods and data critique, and recognised that the teaching role is demanding and includes choosing a theme, collecting and cleaning data and informing students what they should do before the sprint.

Ojala, Kocksch and Kinder-Kurlanda brilliantly describe essential aspects of such learning processes without losing sight of their teaching role throughout the sprint. One situation informs when students should cope with feeling overwhelmed by "too many things to install, learn to operate and understand", from the intense and tight programming involved in data sprints to working with data. Unfortunately, as the authors explain, these tasks can also disappoint some students when they realise the extra effort demanded from them throughout the sprint. Another situation that refers to what the authors call "disconcertment over data" is when the sprint facilitators purposely cause students to feel unsettled, forcing them to, in turn, move from "moments of suspicion and scepticism and turn them into an analytic strategy". In this situation, the sprint facilitators should provide moments to engage with datasets as an opportunity to make sense of the collected data while also being aware of the data curation and collection

processes. These and other situations represent the first encounter of scholars at different carrier levels, not only that of the student, in a data sprint environment.

Òscar Coromina and Adrián Padilla, in their 'Hands On' article, incorporating data sprints and digital methods into digital marketing education (*'Hands On': incorporación de data sprints y digital methods en la docencia de marketing digital* [original language]), also reflect on their experience as teachers with master's degrees in digital marketing who have been organising data sprints with their students for 9 years. They see this practice as a way of presenting students with real challenges in which both students and working professionals collaborate. Coromina and Padilla emphasise that this type of workflow provides students with practical knowledge and digital and communicative competences that, beyond applied research, especially problem-solving practice, serves them in other environments outside the academy. In order for students to get the most out of the experience, the organisation already has workshops and specific teaching on data analysis to prepare them. Following common practice in sprint preparation, the authors include data collection in this phase prior to the sprint. In this way, they make sure that students participating in these *ad hoc* data sprints are at the same, or nearly the same, levels of digital literacy and make much better use of their time.

4.2. Data sprints as a space for methods and tools creation

This section demonstrates how data sprints are a space for methods and tools creation which has proven beneficial in digital humanities, digital social sciences and media studies. The making, testing and/or maintenance of research tools in sprint contexts has been substantial and has made a difference in several research fields, such as web archive research and (visual) network analysis (see Huuderman *et al.*, 2013 and Jacomy, 2022). Regarding methods made through data sprints, we start with Noortje Marres' Situational Analytics (2020), initially applied in a data sprint to analyse testing situations of Covid-19 on Twitter³. Her semi-automated method proposes an interpretative form of data analysis where researchers can actively curate the data and situation under study from large digital datasets. Another example is the three-layered (3L) perspective⁴ for addressing hashtag engagement by Omena, Rabello & Mintz (2020). The method was first proposed in 2017 and later refined and implemented in other projects and sprints. In addition to specific cases, data sprints play a role in a broader agenda for generating visual and vision-reproducible methods (see the work of D'Andrea & Mintz, 2019; Pearce & De Gaetano, 2021; Niederer & Colombo, 2019; and Rogers, 2021).

Giulia Tucci enters this space of developing new solutions and methodological outputs in the article "Visualizing an image network without rendering files: a method to combine user hashtags with computer vision labels". Her method proposal was developed during the 2021 Digital Methods Initiative Summer School. By avoiding the need for memory

3. Project report available at <https://wiki.digitalmethods.net/Dmi/SummerSchool2020COVIDTestingSituations>.

4. The first layer looks at potential differences in the use of hashtags by high-visible users and ordinary users. The second focusses on hashtag activity and the reutilisation of hashtags in social media databases. The final layer looks at the images and texts to which hashtags are related.

space on the computer when visualising networks with images, Tucci's work presents the potential of a creative technique for visualising networks of geolocated images without rendering the image files on the network. Specifically, the paper explores the Twitter case with geolocated images combining the #deepfake hashtags tweeted with the images and the Google Cloud Vision API best single expression to describe each image. She builds a network where nodes are *best-guess labels* and *country flags* while the edges are the images. In so doing, the author argues that it is possible to analyse large image datasets without necessarily seeing the images but with the help of computer vision.

The creative space of data sprints is not limited to methods or tools creation, as the work of Beatrice Gobbo and colleagues shows in "Staging and storing data sprint-based research results: communication design approach". The article's co-authors are a group of communication designer scholars, namely Maria de Los Ángeles Briones, Elena Aversa, Tommaso Elli, Andrea Benedetti, Michele Mauri and Gabriele Colombo. Faced with the challenges of better disseminating data sprint outcomes, the authors developed "a digital *place* that allows data sprints' results to be stored, staged, and accessed after the research activities". This *place* is a tailored platform to communicate the main results of projects developed in data sprints. As one feeds the platform, each step is an invitation to step back, reflect and define what and how to communicate the results. In other words, the authors operationalise ways of translating data sprint reports through the creation of such a platform.

As a source of creativity and spaces for searching for new solutions and building tools, data sprints allow the improvement of methodological processes.

4.3. Data sprints as a reflective tool

Being part of a data sprint working dynamic allows, at a later stage, reflection on how this has taken place, what decisions have been made and what the implications have been. Whether in an academic article or in a more informative way, some participants reflect on data sprints' modes of researching and relating to the community. This section discusses data sprints as a reflective tool to understand the triad of data-, software- and platform-oriented research. The post-sprint time is ideal for critical reflection on what was done, whether the results are helpful and if so how. The reflections come from the practice standpoint, when one should have at least a fair understanding of the rationale behind the adopted methods, proposed analysis and outcomes.

María Concepción Castillo-González, Dorismilda Flores-Márquez and Gabriela Elisa Sued address sprints as a reflective tool in "The data sprint method: An exercise in feminist reflectivity regarding the practices of knowledge production" (*'El método data sprint: un ejercicio de reflexividad feminista sobre las prácticas de producción de conocimiento'* [original language]). On the basis of their experience as well as a focus group with participants in a data sprint, they reflect on the phases of work, the difficulties encountered, the positive experiences of the participants and an evaluation from a gender perspective. Participating in these initiatives allows an "insider's view" of how the research takes

shape in a short period of time. From a critical perspective, they argue that these are also spaces with power differentials and social inequalities. In a sprint, participants with all kinds of backgrounds come together, and amongst them, there are also people who are practically strangers to these dynamics and working methods and have come precisely to learn. In a framework where speed is key to making decisions and getting results, this kind of digital skills gap creates feelings of frustration for some people on the team. The absence of skills is not only owing to lack of knowledge in the field of study, but also in terms of access to tools and training. As Castillo-González, Flores-Márquez and Sued state, tools of this kind are not intuitive and require a solid base for their full understanding, which cannot be achieved during a sprint. The authors also underline the positive effects of collaborative work, such as learning from peers, knowledge sharing and group support to achieve the goal. Regarding the gender perspective, they highlight that the participation of women in data sprints empowers them by promoting the appropriation of digital spaces in these activities.

Data sprints are therefore spaces where experience and observation allow participants to reflect on research processes and ways of working. In this way, thanks to this subsequent reflection, data sprints can also be drivers of change in and of themselves because they are also ideal scenarios for proposing new work dynamics on the basis of previous reflections.

4.4. Data sprint reports for producing scientific knowledge

This section reports another post-sprint activity: the project reports which often offer an updated version of the project's final presentation. Data sprint reports can contribute to new or ongoing research projects and to methods on trial. The reports also support researchers in grounding theory through empirical research and producing hands-on field guides (see Bounegru *et al.* 2018; Marres, 2020). Such reports do not necessarily have to remain just that; many academics appreciate their potential and continue to work on them afterwards, providing them with theoretical frameworks that support such work and that frame the research in a particular academic tradition. That is a scientific contribution. Nevertheless, the production of scientific knowledge generated in data sprints is not only about what is eventually published or even researched during the sprint. The concept is much broader, and also encompasses the participants' own appropriation of the methods they have learned to use and which they will then include in their subsequent research designs.

Jorge Martins Rosa, N. Gizem Bacaksızlar Turbic, Alda Magalhães Telles, Clara González Tosat, Cristian Jiménez Ruiz, Kalliopi Moraiti, Oğuz Özgür Karadeniz and Valentina Pallacci show how a data sprint can be a suitable framework for advancing scientific knowledge. As part of a much broader research project, they decided to use the collection of material to set themselves a challenge: to find the differences in how people engage with the content of the several Portuguese political party pages on Facebook during a period before and after two relevant election campaigns (one national and one European), and to detect what type of content received the most attention from Facebook users (understood as the number of reactions on that platform). This type of

research objective allowed the authors to apply methods and techniques developed in previous data sprints to explore content on Facebook.

Ilya Lavrov, Franziska Schranz, Chiara Miozzo and Marie Palaffre, in their contribution “Visual Communication strategy of populist leaders on Instagram in 2020”, present a study devoted to analysing and comparing the visual communication strategies of 12 populist leaders on Instagram. By tailoring the Tony Blair Institute methodology for identifying populist leaders, the research group analysed a pool of more than 3000 posts containing static images from various perspectives. After a comparative study considering the visual and textual content of the images using Vision API techniques, it turns out that populist leaders publish images that communicate national and personal power. Hence, they portray themselves as ordinary citizens when showing their strengths and reliability. Thanks to their work, it emerges that the use of Vision API and Digital Methods, combined with a qualitative categorisation and coding of emerging clusters, allows for the analysis of communication strategies on Instagram.

Whilst the data sprint workflow allows different methods to be explored collectively like a hive mind, in the post-sprint period, researchers can further analyse and refine the sprint outputs, approving or discarding what they are doing according to the needs of the study.

5. Reflections from and on data sprints

In this article, facets learned from the data sprint approach were identified. We have tried to elucidate important definitions for future applications in a proposal that compiles the anatomy of a data sprint, while taking into consideration its main purposes, uses and meanings that can be given to the practice. Considering the contemporaneity of teaching methods and learning tools in a highly digitalised and fragmented attention world, data sprints operate as a first-hand knowledge experience as few other tools can. The sense of experience and co-creation are perhaps the main differentiators of the practice, as the compact and intense format promotes the high performance and active involvement of the participants.

As a two-way street, the participation process of both those who develop or lead a project and those who are participants is a win-win situation. Collaboration in developing answers or solutions to issues focussed on a research question is key to a good outcome. More than solving technical or objective questions, a *Data Sprint serves as a means of teaching and learning digital methods research*. It is important to highlight that, when dealing with objects of an ever-changing nature, be it methodology, platform, software or empirical, the natural state of a researcher ready to participate is that of an eternal apprentice. This condition of learner that the sprints provide to all who are participating is especially reflected in the sense of being a space for the *creation of methods and tools*. In analogy to an algorithm which, in simplified terms, operates by means of a basic formula, one or more elements of this equation will always be variable. It is this variable that encourages a critical view and that allows us to fail and iteratively improve the research (Venturini & Munk, 2022). As in modern innovation processes, failure is itself

an engine for more challenging research with more interesting results.

Another important point to mention is an *analytical view of the sprinting processes* themselves as a meta-analysis exercise. Considering the short duration of sprints, it is not feasible to have a reflection space, because reflection requires time. In the post-sprint period, this exercise becomes possible. The critical evaluation of methodological and ethical decisions made can also be re-evaluated, generating new protocols and promoting a more solid base that can be replicated by peers and in future sprints. Besides specific technical issues, such as data access, information processing, writing and visualisation of results and so many other processes that are part of the work, it is important to consider lateral elements. Such elements include people management, sensitivity to deal with participants' varied cultural contexts, team development to collaborate and learn in a practical exercise of leadership and soft skills, processes that are rarely mentioned in scientific work in related areas.

Finally, the *variety of outputs* that can result from projects developed in data sprints are considered. Besides the dissemination patterns of scientific knowledge such as in articles and essays, case studies, conferences etc., the relationship of transdisciplinary visions suggests a valid and often unprecedented contribution to various areas of knowledge. In addition to the results themselves, the methodologies developed, the research frameworks and the guides for extraction, treatment and visualisation of data, as well as software and platforms that aid research in the area, are contributions that unite and enhance studies in digital methods on a global scale.

Acknowledgements

The authors want to acknowledge Thais Lobo for her comments and suggestions, and Nate Tkacz, Warren Pearce and Elaine Teixeira Rabello for their constructive feedback.

References

- Berry, D. M.; Borra, E.; Helmond, A.; Plantin, J. C.; Walker Rettberg, J. (2015). The data sprint approach: exploring the field of Digital Humanities through Amazon's application programming interface. *Digital Humanities Quarterly*, 9(3). Available at: <http://eprints.lse.ac.uk/65438/>
- Bounegru, L.; Gray, J.; Venturini, T.; Mauri, M. (2018). *A Field Guide to "Fake News" and Other Information Disorders: A Collection of Recipes for Those Who Love to Cook with Digital Methods*. Public Data Lab. Available at: <https://fakenews.publicdatalab.org/>
- Chao, T. H. J. (2021). Memespector GUI: Graphical User Interface Client for Computer Vision APIs (Version 0.2) [Software]. Available at: <https://github.com/jason-chao/memespector-gui>
- Colombo, G.; Niederer, S. (2021). Visual Methods for Online Images: Collection, Circulation, and Machine Co-Creation. *Diseña*, 19. <https://doi.org/10.7764/disena.19.Intro>

- D'Andrea, C.; Mintz, A. (2019). Studying the live cross-platform circulation of images with computer vision API: An experiment based on a sports media event. *International Journal of Communication*, 13: 1825–1845.
- Gray, J.; Bounegru, L.; Rogers, R.; Venturini, T.; Ricci, D.; Meunier, A.; Mauri, M.; Niederer, S.; Sánchez Querubín, N.; Tuters, M.; Kimbell, L.; Munk, K. (2022). Engaged research-led teaching: composing collective inquiry with digital methods and data. *Digital Culture & Education*, 14(3): 55–86 Available at <https://www.digitalcultureandeducation.com/volume-14-3>
- Huurdeman, H. C.; Ben-David, A.; Sammar, T. (2013). Sprint methods for web archive research. In *Proceedings of the 5th Annual ACM Web Science Conference* (pp. 182-190). ACM. Available at: <https://dl.acm.org/doi/10.1145/2464464.2464513>
- Jacomy, M. (2022, 18 June). Call for participants: Gephi code sustainability retreat 2022. *Gephi Blog*. Available at <https://gephi.wordpress.com/2022/06/18/call-for-participants-gephi-code-sustainability-retreat-2022/>
- Mauri, M.; Gobbo, B.; Colombo, G. (2019). O papel do designer no contexto do data sprint. In J.J. Omena (Ed.), *Métodos digitais: Teoria-Prática-Crítica* (pp. 161–180). ICNOVA.
- Mauri, M.; Briones, M. A.; Gobbo, B.; Colombo, G. (2020). Research Protocol Diagrams as Didactic Tools to Act Critically in Dataset Design Processes. In *INTED2020 Proceedings* (pp. 9034–9043). Available at: <https://library.iated.org/view/MAURI2020RES>
- Mauri, M.; Ciuccarelli, P. (2016). Designing diagrams for social issues. *Design+ Research+ Society-Future-Focused Thinking*, 3(5): 941-958.
- Marres, N. (2020). For a situational analytics: An interpretative methodology for the study of situations in computational settings. *Big Data & Society*, 7(2), 2053951720949571.
- Milman, N. B. (2012). The flipped classroom strategy: What is it and how can it best be used?. *Distance learning*, 9(3), 85.
- Munk, A. K.; Madsen, A. K.; Jacomy, M. (2019). Thinking Through The Databody: Sprints as Experimental Situations. In M. E. Å. Mäkitalo, T. Nicewonger (Ed.), *Designs for Experimentation and Inquiry: Approaching Learning and Knowing in Digital Transformation* (pp. 110–128). Routledge. <https://doi.org/10.4324/9780429489839>
- Niederer, S.; Colombo, G. (2019). Visual Methodologies for Networked Images: Designing Visualizations for Collaborative Research, Cross-platform Analysis, and Public Participation. *Diseña*, 14: 40–67. <https://doi.org/10.7764/disena.14.40-67>
- Omena, J.J. (2021). *Digital Methods and Technicity-of-the-Mediums. From Regimes of Functioning to Digital Research* [Doctoral Dissertation, Universidade Nova de Lisboa]. <http://hdl.handle.net/10362/127961>
- Omena, J. J.; Rabello, E. T.; Mintz, A. G. (2020). Digital Methods for Hashtag Engagement Research. *Social Media + Society*. <https://doi.org/10.1177/2056305120940697>
- Ozdamli, F.; Asiksoy, G. (2016). Flipped Classroom Approach. *World Journal on Educational*

Technology, 8(2), 98–105. <https://doi.org/10.18844/wjet.v8i2.640>

Pearce, W.; De Gaetano, C. (2021). Google Images, Climate Change, and the Disappearance of Humans. *Diseña*, 19, Article.3. <https://doi.org/10.7764/disena.19.Article.3>

Rogers, R. (2021). Visual media analysis for Instagram and other online platforms. *Big Data & Society*, 8(1), 20539517211022370.

Rogers, R.; Lewthwaite, S. (2019). Teaching Digital Methods: Interview with Richard Rogers. Interviewer: S. Lewthwaite. *Diseña*, 14, 12–37. <https://doi.org/10.7764/disena.14.12-37>

Venturini, T.; Meunier, A.; Munk, A.; Borra, E.; Rieder, B.; Mauri, M.; Azzi, M.; Ciminieri, D.; Ubaldi, G.; Kaltenbrunner, A.; Laniado, D. (2014). Climaps by EMAPS in 2 pages (a summary for policy makers and busy people). SSRN. Available at: <http://ssrn.com/abstract=2532946>

Venturini, T.; Munk, A. K. (2022). *Controversy Mapping: A Field Guide*. John Wiley & Sons.

Venturini, T.; Munk, A. K.; Jacomy, M. (2018). “Data-Sprinting: A Public Approach to Digital Research”. In E. Lury, C., Fensham, R., Clough, P., Heller-Nicholas, A., Lammes, S., Last, A., Michael, M., Uprichard (Ed.), *Handbook of Interdisciplinary Research Methods Routledge*. Routledge International Handbooks. <https://doi.org/10.4324/9781315714523>




Data Sprint Learning. Exercising Proximity to Data in Teaching Situations

Aprendizaje a partir de *data sprints*.
Ejercitar la proximidad a los datos en contextos docentes

 **Laura Kocksch**
laurak@ikl.aau.dk
Technoanthropology Lab, Aalborg University

 **Mace Ojala**
maco@itu.dk
IT University of Copenhagen

 **Katharina Kinder-Kurlanda**
Katharina.Kinder-Kurlanda@aau.at
University of Klagenfurt

Received: 23/12/2021

Accepted: 20/06/2022

ABSTRACT

This paper reports on a data sprint conducted as part of a PhD course on digital methods and data critique at the University of Klagenfurt. We reflect on how our data sprint contributed to this higher educational setting, and point to ways in which the data sprint method can be developed further based on our experience. The paper discusses how the sprint fabricated a moment of “critical proximity” for students that were mainly working with qualitative social science methods. The data sprint allowed them to put their critique on “big data” into practice by working with selected sets of data from Twitter and Scopus. We reflect on our collective experience and draw conclusions on the use of data sprints in teaching. Data sprints encourage us to engage with feelings of being underwhelmed and overwhelmed by data that provoke our social science way of critique. Our data sprint tangibly demonstrates that data work is in fact “messy”: transgressing ideals of good data management, biased, ambiguous and open-ended. But instead of turning away from this “wildness”, we urge to make use of it in teaching settings. This wildness allows to step out of conventional modes of critique, and into modes of action. We conclude with a protocol as a practical guide for everyone who wants to introduce data sprints in their teaching.

KEYWORDS

data sprint, data work, data analysis, pedagogy, critical proximity

How to cite this article:

Kocksch,L.; Ojala,M.; & Kinder-Kurlanda,K. (2022). Data Sprint Learning. Exercising Proximity to Data in Teaching Situations. *Dígitos. Revista de Comunicación Digital*, 8: 31-50. DOI: 10.7203/drdcd.v1i8.232

RESUMEN

Este artículo presenta los resultados de un 'data sprint' realizado como parte de un curso de doctorado sobre métodos digitales y crítica de datos en la Universidad de Klagenfurt. Con este artículo, reflexionamos sobre cómo nuestro *data sprint* contribuyó a este entorno educativo superior y señalamos las formas en las que el método de *data sprint* se puede desarrollar, aún más, a través de nuestras experiencias. El documento analiza cómo el *sprint* fabricó un momento de "proximidad crítica" en estudiantes que trabajaban principalmente con métodos cualitativos en el campo de las ciencias sociales. El *data sprint* permitió a los estudiantes poner en práctica críticas sobre la noción de *Big Data* al trabajar con conjuntos de datos seleccionados de Twitter y Scopus. A partir de una reflexión sobre nuestras experiencias colectivas, compartimos conclusiones sobre el uso de *data sprints* en la enseñanza. El método *data sprint* nos anima a interactuar con el sentimiento de estar decepcionados o abrumados por los datos, que provocan un modo de crítica perteneciente a las ciencias sociales. Nuestro *data sprint* demuestra, de manera tangible, que el trabajo de datos es, de hecho, "desordenado"; ya que transgrede los ideales de una buena gestión de datos, al ser sesgado, ambiguo y sin límites fijos. Pero en vez de alejarnos de este "estado salvaje", nosotros invitamos su uso en la enseñanza. Este "estado salvaje" permite exceder modos convencionales de crítica y entrar en modos de acción. De este modo concluimos ofreciendo un protocolo, a modo de guía práctica, para todo aquel que quiera introducir los *data sprint* en su docencia.

PALABRAS CLAVE

sprint de datos, trabajo con datos, análisis de datos, pedagogía, proximidad crítica



Data Sprint Learning. Exercising Proximity to Data in Teaching Situations

1. Introduction

This article reports on a one-day data sprint we conducted online as part of a PhD course in digital methods and data critique at the University of Klagenfurt. We address the educational potential of data sprints, particularly their ability to bring students into "critical proximity" with digital data (Latour, 2005: 253; Birkbak *et al.*, 2015) Our data sprint induced feelings of being overwhelmed in students, when they looked at a digital data set for the first time; encountering its sheer volume or being 'wooded' by colourful visualisations. On the other hand, students also expressed feeling underwhelmed by a lack of relevant questions the data could answer, feeling suspicious of the data and its reductionist nature. This prompted a discussion of the partiality of data with the students which was not tailored towards a distant critique, but turned into a set of analytic categories that imbued the subsequent analytic process. The data sprint facilitated a 'wildness' of working with data that was both disconcerting to students (being overwhelmed and underwhelmed at the same time), but also engendered playfulness and experimentation. Students reacted to the data sprint with both excitement and defiance; feeling at the same time intrigued to acquire technical skills, while also being forced to work with data that they found inherently lacking and biased. These reactions are possible only from a position of closeness to data, where its potential and risks are

out in the open. We describe this disconcerting during the sprint in relation to how it fostered a close-up and productive reflection of data - rather than abstract critique. We argue that it is worth facilitating under- and overwhelmedness, disconcerting and bewilderment in educational data sprints, and we show how we crafted them in the course of our exercise.

A data sprint is a limited-duration event where people are invited to come together to sort, interpret, and visualise a prepared dataset (Venturini *et al.* 2018b; Jensen *et al.*, 2021; Munk *et al.* 2019a). Compared, on the one hand, to workshops, data sprints are more open ended, non-committal and exploratory in nature. On the other hand, compared to hackathons, data sprints are focused on a data set or data sets selected and curated by the organisers prior to the event, fostering a more thorough analysis and conversation during the sprint. Data sprints are interdisciplinary, and can even be seen as non-disciplinary in the sense that what counts as success is left implicit, open for interpretation, and might not adhere to standards of any established field (Munk *et al.*, 2019b). Participants are invited to “mess around” with the data brought in by the facilitators, encouraging also those not proficient in data analysis, statistics or digital methods to get involved. The data thus serve as a boundary object (Star and Griesemer, 1989) during the sprint; while they are pre-configured they also allow flexibility. Issue experts, designers and software developers are invited to facilitate, provoke and guide the process. Controversy might arise as different parties attempt to gain authority in the interpretation of data or use it to support their specific claims about the world. In this regard, data sprints are not solely meetings where insights are extracted from data, but settings in which specific (differences in) epistemic politics become observable.

Data sprints generate extraordinary situations in which participants and organisers who often bring qualitative social science backgrounds can get into proximity with digital data. Sprints thus overcome common themes of critique towards data, and urge participants to “get their hands dirty” (Munk, 2021). This brings together the tradition of controversy mapping with that of participatory design (Venturini *et al.*, 2018b; Munk *et al.*, 2019; Jensen *et al.*, 2021; Venturini & Munk, 2021). In this approach, data sprints aim more at crafting a situation in which stakeholders come together and unfold their specific points of view based on the data at hand. In the process both participants and the organisers step on uncharted territory, eliciting often surprising and unanticipated insights in data (Jensen, 2020; Jensen *et al.*, 2021). Our sprint, however, drawing on this realisation, had pedagogic goals; the sprint familiarised students with a selection of digital methods while also giving them the chance to dip a toe into data work. This forged a contrast to the rest of the PhD course that was based on generic principles of data management and data ethics and scholarly critique of data-driven surveillance and big tech.

Although data sprints have been employed in a variety of domains (see for example Sanderhoff, 2014; Berry *et al.*, 2015) and including a range of stakeholders, or “issue experts” (Munk *et al.*, 2019a) (e.g., scientists, activists, public officials, educators, etc.), so far their potential as teaching situations has not been discussed explicitly. Most

commonly, data sprints are seen as ways to investigate a topic or “issue” (Marres, 2015) while also interfering in and reflecting upon the affordances of the platform and method in the process (Omena *et al.*, 2020; Pearce *et al.*, 2020). As data sprints have proven useful in prompting hands-on reflexivity, this paper aims at exploring their potential in educational settings. We argue that data sprints combine the useful role of “toy problems” and “real world” data, allowing students to try and test data before moving to their own research projects. In the case of PhD education, data sprints offer situations of extraordinary encounters with data, allowing “playing around” with data without committing to include results in a publication or thesis (which is usually the main concern of PhD students). Data sprints allow competent play with data instead of *testing* them for their utility in a thesis or publication. The data sprint’s format also allows reflection of the data analysis process as a whole in a very short time - which may be accomplished over years in the actual PhD project and thus much more difficult to oversee. To investigate the additive value of data sprints in educational settings, we ask: how would a data sprint trigger “critical proximity” in ways a theoretical course would not?

For our data sprint, we chose a topic that aligned with the topic of the digital methods and data critique focused PhD course: Open Data and Open Science. We anticipated students to have limited knowledge of Open Data and Open Science, but as digital humanities and social science researchers to be equipped with the conceptual tools to identify recurring themes or specific problematizations.

The choice of topic and desired outcome of the sprint was threefold. Firstly, we designed the sprint to act as a teaching situation where students would learn about Open Data as an ongoing controversy in science (Levin & Leonelli, 2016; Prainsack & Leonelli, 2018). Secondly, Open Data and sharing of research data was a topic that we, the organisers, were interested in and have published on (e.g., Kinder-Kurlanda *et al.*, 2017; Kinder-Kurlanda & Weller, 2020; Sørensen & Kocksch, 2021). Coming into the data sprint with our own research questions as scholars in science and technology studies, the data sprint allowed us to work on our specific questions towards Open Data controversies, e.g, what similarities and differences manifest in the published research literature and discourse on social media, how openness and concealment are discussed, what methods are (not) widely acknowledged etc. Lastly, it was our goal as educators to exercise the method of data sprints itself, familiarise ourselves with the nitty-gritty of putting together events like this and reflect upon their productiveness in educational settings. Whereas the three goals intertwine, this article specifically aims at a discussion of the last. In synthesis, data sprint learning provided learning for students and educators alike.

In section 2, we outline our data sprint procedure. The section stands in the place of a methodology chapter you might expect in a traditional article. However, it does not include a description of our participant observation data collection or close description of our analysis because we did not follow a strict methodological procedure as prescribed by e.g. ethnographic methods. We report instead as organisers, educators and participants of the data sprint, putting our past experiences in teaching and impressions

during and after the data sprint in conversation. We report how the situation was set up and how it encouraged specific forms of action. We thereby lean on description of teaching situations in science and technology studies (STS) that emphasise how learning is composed by specific socio-material assemblages in teaching situations (Sørensen, 2009). Rather than being interested in analysis or evaluation of the sprint, we aim at reflecting the data sprint method as a way to forge a specific learning setting.

Section 3 is more conventional to readers of ethnographic papers: we present three episodes from the sprint which we in the concluding section 5, emphasise as particularly relevant for understanding how the data sprint created an extraordinary teaching situation.

However leading up to the conclusion, we offer a protocol as a condensed description of our data sprint in the form of a set of instructions in section 4. The “protocol” fulfils a task different from that of the Digital Methods Initiative community where protocols are step-by-step descriptions of tools and techniques which can be reproduced to collect and analyse comparable data sets (Rogers, 2019; Venturini *et al.*, 2018a). Instead our notion of protocol points to what has been suggested by Ballesterro and Winthereik (2021:11) in the form of an “analytic protocol”, that is a “practical guide [...] to set up conditions to create analytic timespace” for ethnographic analysis. Protocols in that sense “invoke a sense of organised reflection” (Ballesterro & Winthereik, 2021:11) that can aid theorising and foster creativity in ethnographic practice. We find that our protocol engenders such systematic reflection of data sprints as moments of getting close to data while remaining reflexive. While Ballesterro and Winthereik argue that protocols should be used to prolong moments of analysis in ethnography and possibly invite others into the process, our protocol relies on a short-term method (quite literally “a sprint”), that fosters fast and messy decision making. As we point out, the wildness of the data sprint method has potential for ethnographic encounters with data. Specifically, we reflect how working with data close up can generate analytically productive *disconcertment* and *bewildering*; reactions that are cultivated by employing our protocol. It is intended to be re-used and developed in other teaching settings where data forms a matter of concern (Latour, 2004; Puig de la Bellacasa, 2017).

2. Method and structure of the data sprint

The data sprint was a 6-hour event conducted over a video-conferencing platform in June 2021. The data sprint was organised in two sessions. A morning session included introductions of research topics and interests by the participants and an introduction to the data by the organisers, and developing research questions. In an afternoon session students worked in groups of 2-4 that were interested in a specific question developed in the first.

Figure 1 shows the structure of the data sprint, including introductory sessions and open working sessions. Section 1 introduced the data sprint and gave an introduction into the data collection process for both data sets. Section 2 introduced students to network visualisations, specifically Gephi. Sections 3, 4 and 5 were designed as group working time with occasional prompts to “freeze” and report. The presentation deck

made use of the “sprint” metaphor (with both the background and the “freezes”). The exact time slots were adopted during the sprint.



Agenda	
10-10.40	Introduction to data sprint, Data collection 1 (research references) and 2 (social media)
10.40-10.50	Break
10.50-11.20	Intro to networks
11.20-11.30	Break
11.30-12.30	Activity 1+2 (inkl. Freezes): Question formulation
12.30-13.30	Lunch Break
13.30-14.30	Activity 3+4 (inkl. Freezes): Network analysis
15-15.20	Break
15.20-16	Activity 5+6: Finalizing prototypes and wrapup

Figure 1. Structure of the data sprint (as shown to participants, original slide deck)

The groups were supported by the organisers in handling data graphs and tweaking visualisations with Gephi (Bastian *et al.*, 2009). The second session closed with a discussion of interesting findings and further questions. The entire event was facilitated by and documented on Miro – a canvas tool for online collaborative work. Throughout the event we made various efforts to account for the fact that we as participants were all coming together online, while also sitting alone in very different dorm rooms, patios and offices. We tried to create elements of commonality (e.g. everyone was asked to bring a mouse in addition to the laptop to facilitate better navigation of the graphics) and different possibilities to engage with each other, by voice, in chats, on boards etc. to allow for different needs and wants regarding interaction.

Ten people gathered for the data sprint, including three organisers. The participating PhD students were already advanced in their projects. In the sessions leading up to the data sprint they had been provided with literature and presentations on a variety of issues around methodological, epistemological and ethical concerns of Big Data and their treatment in the more social science-oriented literature. For example, “platform politics” had been discussed as well as the issue of researchers becoming complicit in platform power and user data commodification. The sprint thus provoked the critical suspicion towards data that they had trained in the rest of the course. The event started with a round of introductions where participants stated their backgrounds and familiarity with data work, Digital Methods as advanced by Rogers (2013; 2019) and others, and the topic Open Data/Open Science. Half of the participants had previously worked with or considered working with digital methods, mainly with social media data and digital archives. The other half reported they were not familiar with digital methods and expressed doubts how valuable such methods could be for their research. Some

articulated that in order to be critical about digital methods, they would want to know more about their “inner” functioning.

We requested the participants to prepare by installing the network analysis tool Gephi (Bastian *et al.*, 2009) before the event and watch two video tutorials. We had prepared two independent data sets. The first was a set of academic references, sourced from Elsevier's bibliographic database Scopus. The second was a set of social media posts from Twitter, collected using 4CAT (2021) via the Twitter API v2 Academic Research Track, a data product the company made available in January 2021 (Parack, 2021). Prior to the sprint, both data sets were pruned and cleaned to allow easier handling on students' own computers. This entailed filtering out entries in the Scopus set which lacked semantic data, i.e., an abstract or summary, and narrowing the Twitter sample down to span no more than one week, and transforming from JSON to CSV. We intentionally left the data sets *semi-cleaned*, e.g., leaving in entries that were misaligned (information on authors in the wrong column in the Scopus set) or hard to discern (extensive Twitter links). We also left files in formats that we anticipated might turn out to be non-trivial to open with Excel. While we tested turning CSV files into the specialised network format using the Table 2 Net tool, we did not provide ready-to-go Gephi files. Data file formatting was not intended to be an essential part of the workshop, however, in hindsight we recognize it productive to leave data in a rough shape such that students experience some of the “underwhelming” chores of data work. The following section reflects in detail on how students interacted with the data sets.

3. Data Sprint Learning. Being under- and overwhelmed, disconcerted and bewildered

This section describes three aspects of what characterised the process of learning during the data sprint: manoeuvring feelings of being underwhelmed and overwhelmed disconcertment over biases, and a negotiation of both wildness and bewilderment.

3.1. Being both under- and overwhelmed

Discussion during the sprint and participant feedback collected afterwards expressed sensations of being both under- and overwhelmed by the data, the tools and the sprint format itself. Students reported that their first impression was that of too many things to install, learn to operate and understand. For most of them it was not only the first time working with social media data or a corpus of literature references and the first introduction to network analysis and Gephi, but also the first introduction to analysing and visualising digital data at all. In the foreground of much of the work during the data sprint was therefore familiarising students with how to encounter such data and tools in the first place: how to open a CSV file? What re-formatting is needed for it to open correctly (with the right separator)? How to spot that it is opened incorrectly?

These ‘alien encounters’ with data resulted in students being overwhelmed by the means and potential that lay before them. They were ‘wooded’ by some of the visualisations that they produced with our guidance, putting them in a state of amazement that initially

prohibited any questioning or critique.

But on the other hand students expressed being underwhelmed; they pointed out that the data was full of gaps, absences, biases and that much of data work seemed almost mundane. This became most apparent in the first session where we instructed students to open a CSV file with Excel or LibreOffice Calc. Most of them encountered a separator issue (between the “,” used in the generated data sets, and the locale setting of “;” in their program). Translating the data back and forth from Calc into Excel and Google Spreadsheet seemed random to students and a tedious process to overcome. The mundanity of data work stood in conflict with how they had imagined it: a thorough analytic process. Their underwhelming was added to when they had managed to open the CSV file as a table which still did not allow a lot of conclusions. Students lacked familiarity with *reading* a CSV file, leaving them with a table without any interesting insights.

For the students competent in qualitative methods these struggles led to being underwhelmed by the perceived reductionist and positivist nature of the methods, and the banality of the data (which was even deepened with the data changing seemingly randomly from a table into text and back). In other words, students were not impressed with the data and data work, even though they were amazed and also overwhelmed by the techniques in use and visualisations we showed them as a possible product of the sprint process.

3.2. Disconcertment over data bias

In order to provoke their social science interest more, we explored with them how the data had been collected and discussed what biases and data gaps were introduced in the process; or how absences “haunted” the data (Meldgaard Kjær *et al.*, 2021). Our intention was to make them see the data as carefully crafted rather than “raw” (Bowker, 2009; Gitelman & Jackson, 2013: 2), and therefore capture their interest in the data as the result of a specific selection process. We then asked them to identify traces of that process and additional gaps and biases in the data, hoping to train them in situating the data set.

In preparation for the event, we had compiled two lists of biases to seed discussion, one for each data set. For the first dataset from Scopus, we listed:

- Scopus contains mainly English language publications in their final stages
- My search prioritizes highest citation counts (last years may only be published next year due to scientists’ citation practices)
- Scientific Publications only

For the second dataset sourced from Twitter, we listed the following potential biases:

- Twitter contains intentional as well as casual conversation

- High volume limits the timespan for our query
- What does #opendata OR “open data” OR #openscience OR “open science” not match?
- How is the experience of using Twitter via the app different from using it via a spreadsheet program?
- Is all content human-made? Does this matter?

Stating that data is incomplete and potentially biased is not enough. It does not come as much of a surprise to any sociologist, critical data studies or science and technology studies scholar. It also risks being defeating. Rather, we took it as a starting point and stated more precisely how *the* particular data set at hand was crafted in a specific way and hence haunted by specific absences (Meldgaard Kjær *et al.*, 2021). During the sprint, data gaps familiarised the students with the process of collecting data and composing simplified data sets that would allow students easy access to methods and data. We did not bring the biases up to check them off our list - to make our encounter with data more “pure” as we have moved the necessary tedious data composition out of the way. Rather, we fostered a conversation about biases to reach “disconcertment” in students - a critical, yet, productive, reflexive analytic mode to discern between different epistemic systems (Verran, 2001).

We asked students to open the data sets on their computers and look at them in standard spreadsheet software. Students were prompted to analyse the data and make notes on: “What data is missing? Whose point of view is rendered absent? What would be nice to have?”

Students came back with various additions to our data, e.g., that they did not contain any information about gender, age and ethnicity of authors neither in Scopus nor on Twitter, which does not allow us to investigate further whether certain points of view are over-/underrepresented. Neither did the data contain any geographical information or institutions which could have helped to make transparent whether the conversation on open data was dominated by Western ideals of science, disregarding any precarities of opening indigenous knowledge, for example. Students also noted that Twitter data lacked real names and that, or that in comparison, Twitter data contained more controversy than the Scopus data set. See figure 2 for some of the notes taken during this session of the data sprint.

The exercise had a two-fold effect: it brought students closer to the data sets, letting them read through and think about what objects the data refer to. As students were not familiar with the dataset or the topic, this moment contributed to their feeling of overwhelmedness and underwhelmedness. Letting them go through the data with the sole goal to point out what was missing appeared arbitrary. It was left open to them what exactly to identify, leaving them space to employ any line of critique about digital data that they were familiar with. For example, one student suggested applying typical lines of critique from her social science background, e.g., biases based on gender, race, class or age as a strategy to identify biases in the data set.

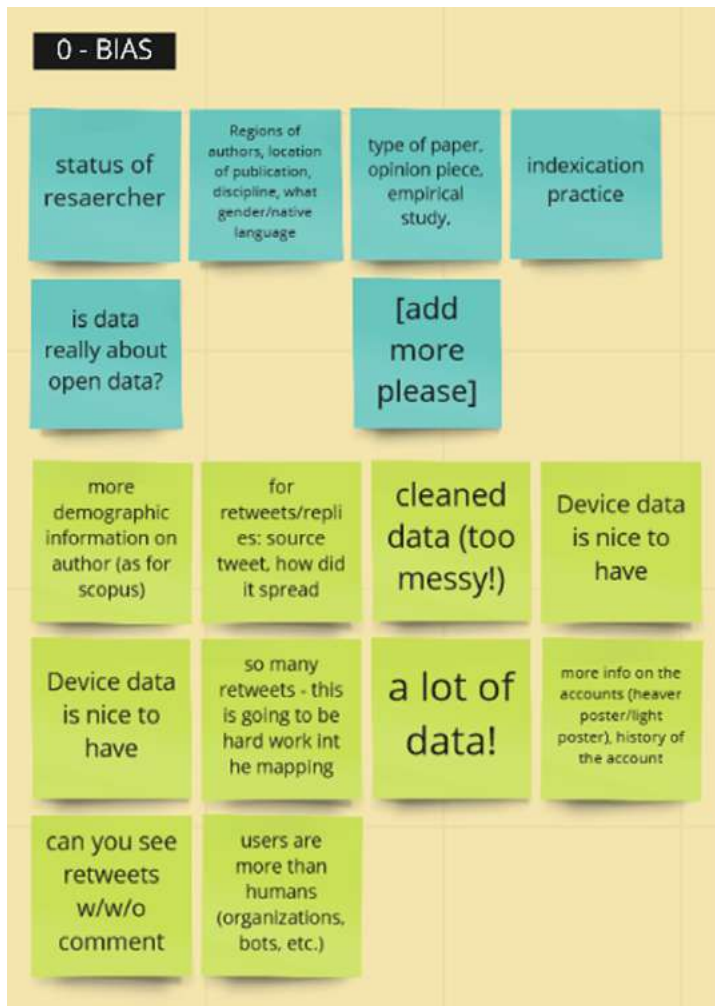


Figure 2. Collaborative note-taking work in progress

Secondly, letting students point to biases trained their critical gaze: asking what and how knowledge is imbued with a specific local and historical gaze is a key strategy in feminist STS (Haraway, 1988), something that we exercised here in a very pragmatic and concrete form. Students were asked to localise their “disconcertment” with data in the data set, turning it into an analytic strategy. Disconcertment allows us to notice moments of suspicion and scepticism and turn them into an analytic strategy. Localising “biases” was a task of critiquing up close.

Thirdly, we could have concluded the work here, and continued a conversation about dramatic effects of biases and gaps in all data (which would have certainly been worthwhile, see Criado Pérez, 2019). But we did not. We did not want either ourselves or the students to lean back and ruminate about biases and their consequences. Rather

the sprint format prompted us all to move on and stay with what we had, however disconcerting.

In the next step, we asked students how to deal with the biases and gaps they identified. We started from four responses to biased data: “(1) Ignore them, (2) Don’t use the data, (3) Reflect upon them in a later analysis or (4) Contextualise the data employing the biases.” See figure 3 top.

Subsequently we asked students to come up with more ways of addressing biases in the specific data set. Together we compiled a new list, with more perspectives. The extended list in figure 3 included (5) throwing statistics at the biases (quantifying whether they are statistically significant) (6) consulting with interdisciplinary colleagues, non-academics, minorities or domain experts (qualifying whether they are relevant or not and what harm/benefit the data set could do nevertheless) (7) searching for existing societal biases like race, gender, class, etc. (linking the data to larger societal problems making it an exemplary result of power structures) or (8) throwing more data at it (adding more data to present a more “complete” picture).

The feeling of disconcertment was later re-examined when students were analysing



Figure 3. Suggestions for how to deal with biases in data.

the data sets by themselves: some decided to take out “big” nodes in their Gephi visualisations to see how that would change the relations between other nodes – accounting for making hierarchies in data appear in a different light. Others suggested it needed more data to make their initial findings more sound, or one suggested that it needed domain experts to interpret the data they had visualised.

So, how did the students experience data gaps and biases differently through the proximity with data during the sprint? Firstly, they were prompted to localise biases and data gaps in the concrete data sets at hand. This shifted generalised and abstract critique of data bias to very concrete questions of authorship, platform politics, collection methods and who was present.

Secondly, broad strokes critique of data bias was not a point of closure, but a disconcertment turned into a starting point for future queries. Besides quantifying the biases or de-biasing the data with more complete data, others suggested to rather make biases a strength and employ them analytically, e.g. by inviting affected publics in to comment and utilise the data. When visualising the data with Gephi, some proposed to delete some of the data to see other relations more clearly. These wild ideas demonstrate that the data sprint urged students to exercise agency and materially engage with the partiality and composition of data sets pragmatically and productively. The sprint format afforded balancing material work with a feeling of disconcertment. We stayed with the biased data, making biases part of the pragmatic data work. Biases became nothing to evade, circumvent or be defeated by, but something that is part of the analytic process. Rather than imagining a clean – complete, unbiased, absolute – data set, students made use of their disconcertments surrounding data’s reductionism and partiality. Due to - not despite - their partiality, data became a relevant object of inquiry to students (Jacomy, 2021; Haraway, 1988).

3.3 Wildness and bewilderment

Rather than being well-disciplined, accountable and tidy, in other words tame and domesticated, data sprints are a rowdy mess. This is their virtue. Sprints focus on whichever doings seem most plausible for short-term goals. As the name suggests, sprints are relatively short, intense and single-minded events relaxed from the concerns of what will follow afterwards. Isolated from other data work, a sprint is thus a safe place to experiment. A data sprint typically avoids having pre-defined and well delineated success criteria. As a research situation the research questions, if formulated at all, are typically exploratory. A data sprint does not aim at, and cannot reach robust findings, consideration of nuance and edge cases, or final solution to all the biases. Hence, data sprints are especially appropriate when interfacing with participants (Munk *et al.*, 2019a; Jensen *et al.*, 2021), or for generating ideas at the early stages of a research.

Rather than, for example, teaching and disciplining the participants to the more well-defined issues of traditional data management the impetus is to find new questions and to allow for surprise and *bewilderment* to happen without becoming incapable of acting. A simple but illustrative example of how a certain bewildered wildness manifested in our data sprint, observe some of the data filenames used during the sprint:

- 4CAT Dataset/convert-csv-excel-70212a65460708be9088c35e73f72b36.csv
- 4CAT Dataset/convert-csv-excel-70212a65460708be9088c35e73f72b36.actually-c-separated.csv
- 4CAT Dataset/opendata-or-open-data-a884822c87a999a5f95d8a521846c7c5.ndjson
- Scopus Dataset/Scopus-2/scopus-2.csv
- Scopus Dataset/Scopus-2/scopus2-nurAutJahCit.csv
- Folders Scopus Dataset/new, Scopus Dataset/medium and Scopus Dataset/waste

Some of the above shown filenames index the data source or the device of collection (*Scopus**, *4CAT**), some document the procedures taken (*convert-csv-excel**; multilingual **-nurAutJahCit**, **.actually-c-separated**), some document versioning (**-2**; because this was improvised, there is no **-1**), some recycle unique identifiers such as *70212a65460708be9088c35e73f72b36* from the data sets.

The filenames are idiosyncratic, improvised and do not follow an established naming scheme as good data management would expect – they are more wild than that. However, they are neither random nor without any management; for instance, we observed the application of practices of no-space, CamelCase and generally conveying meaning in the filenames, but the practices are unsystematically applied. This wildness of filenames generalises to other decisions done while conducting loosely governed data work without the constraints of documentation and specification requirements, provenance, metadata management and security requirements of a more serious

data management regime of a more sustainable research situation, especially in collaborations.

The data sprint did not encounter wildness as it may be common in technology design - where research "in the wild" refers to research or design conducted outside of experimental settings (Hutchins, 1995; Rogers, 2011). The wildness composed during our sprint was that of an experiment; a setting of contrived and liberating unruliness. Instead of getting wildness into the design process, the sprint aimed at introducing wildness as a moment of becoming familiar with and uninhibited by data work while reflecting about it.

4. A checklist to data sprint learning

In this article we argue that facilitating moments of underwhelmedness, overwhelmedness, and bewilderment is productive for learning situations. The data sprint forged encounters with data that affected students; both exciting them to work with data in the future and confronting them with the nitty-gritty of data work. We were able to exercise in the data sprint these moments because our group (both teachers and students) was composed of scholars who were used to taking a critically reflective stance as a default. In participants from very different academic backgrounds the data sprint may have had a different effect, possibly engendering less, rather than more reflection and critique. The elusive state that we wanted to capture in this exercise was the experience of familiar notions and habits being challenged and possibly being turned upside down. Engendering similar moments of bewilderment in, say, most engineers would probably have required a completely different sprint setting, for example a reading and writing of philosophy texts or the generation of a theory. Our data sprint reached such an encounter with data by following some principles we will summarise in the following.

The article leaves you with lessons learnt from our data sprint: five dos and two don'ts. We ask you to take it as a device to slow down judgement during your own data sprints in teaching contexts; a device to prompt revision, critique, stumbling; a physical object you can toss around, cut up, rework, leave coffee stains on and discard if necessary.

Five dos

- *Choose a very particular case*, grounded in research. This allows to mitigate overwhelming and underwhelming, as students are capable of responding to and articulating data in regard to a concrete issue at hand.
- *Curate a small data set* that is comprehensive to students when looking at it without any visualisation techniques (in order to be able to point to gaps and biases).
- *Clean up data beforehand* - there will still be plenty that is new and bewildering to students.
- *Assign preparatory homework*, in moderation. Get participants to install and try out tools beforehand. When conducting a course on data management, think

about how to present the “ideal” case and how the data sprint can productively counternarrate formal data management rules.

- *Prepare students* that the data sprint will be exploratory and they are not being *tested*. We observed some defiance and self-deprecating tendencies that could be mitigated by making clear that the data sprint is about bringing forth an individual’s expertise rather than testing their competencies in a fixed set of technical skills. We had some trouble enacting that; while the organisers had more technical expertise, the goal was to have students select and design visualisations that were useful to them (enacting their own expertise in their field of study).

Two don’ts (some of which we failed at)

- *Don’t prioritise explanation time* over workshop time; be aware that both overlap but the sprint should give enough time to explore data independently. Let students know the sprint is a playground and they should not mistake it as a test of their abilities (forming an utterly different academic encounter than they are used to). We see this point related to the last point stated above: With extensive explanations of the data collection method, students saw organisers as experts on the topic - which is not the case. We missed the goal to have students become experts of the data themselves by using it for their individual research projects.
- *Don’t hesitate to set up dynamic, multimodal prompts* such as “chat storms” (asking for written answers in the chat), background music, and different camera perspectives. In an online sprint, processes of group formation and a sense of commonality can be aided by these methods. They help to loosen the atmosphere and enhance the feeling of a playground.

5. Conclusion and implications

Data work in contemporary technoscience requires dealing with ambiguity, vagueness, inconclusiveness, effects of being overwhelmed and underwhelmed by data, and its untameable biases. We have argued that our data sprint did not only teach students about such messiness in a matter-of-fact or distanced critical mode, but let them experience it first-hand. Furthermore, it prompted them to put some aspects of the received critique on platforms, methods or quantitative data to work. This was reached through the nature of “wildness” that the temporary, extraordinary data sprints events embody: students don’t have to be inhibited by what they read – concepts, best practices, principles, methods or critique – nor do they need to cite literature, maintain metadata or be precise about their terminology. They are allowed, enabled and indeed expected to leave inhibitions aside to explore, do and *tinker* (Mol *et al.*, 2010; Law, 2011), and generate preliminary, affective, contradictory or undertheorized insights.

Pedagogically, we argue, *wildness* of a data sprint is valuable. Firstly ignoring the time-consuming 80% of wrangling of the so-called “80/20 rule”, and jumping straight into the 20% of analysis allows getting further along the data processing “pipeline”. But the data sprint also showed that the practices of cleaning and analysis can not be neatly

separated and that, in fact, leaving some roughness in the data urged students to ponder about the composition of data. This point has been made in relation to local traces, such as the separator (Loukissas, 2019). Students found themselves confronted with the tedious side of data work (for example translating formats, making local copies, etc) which fostered analytic overwhelming and underwhelming as well as disconcertment over the banality and politics of data at the same time.

Secondly, by doing so, wild data management enables first-hand experience of engaging with the struggles of data work, rather than disables that experience like *a-priori* critical distance tends to do. Proximity invites the material to present some of its own, inherent troubles. Facing the temptations, passions and excitements is part of engaging the Zuckerbergian “move fast and break things” experience in a critical proximity (Latour, 2005: 253; Birkbak *et al.*, 2015; Herbrechter, 2017). Learning to manoeuvre between proximity and distance is necessary for reflection, and zooming in and zooming out is productive for critique and understanding not only data as such, but also its conditions. Establishing a habit for moving between proximity and distance can, in longer term, move an individual researcher or a research project toward *critical technical practice* (Agre, 1997) – and given that most contemporary academic work is deeply involved with technology (Latour, 1987) – what we might call a *critical technoscientific practice*.

Thirdly, wildness levels the playing field by decentering data work expertise and opens the process to wider participation and people with a diversity of skill sets. Students were able to apply their critique without staying back, turning into analytic strategies *with* the data.

If the bewildering outputs of a data sprint – provisional insights, sketches, visualisations, provocations, hypotheses, personal contacts, data versions and scripts – are to be integrated into a more structured research project with the usual research accountability needs this mess needs to be tamed. If left undomesticated, the sprinting will not be able to contribute to moving the research forward in productive and sustainable ways. As such, wild data management is a teachable moment of taking “technical debt” to be paid later, for the benefit of gaining momentum to achieve short-term goals (Cunningham, 1992). In the educational context of the PhD course, balancing this tradeoff was safely contained in the isolated experimental setting, with the data sprint as a “lab”, rather than directly imposed on the students’ own PhD research projects. The reflection and insights gained in the data sprint could be taken up by students in different ways and settings later on - and the ‘debt’ be paid there - rather than, e.g. be addressed in a longer data sprint.

As authors, we recognize that encouraging wild data management is itself a risky proposition; it suggests that throwing out good practices under pressure, “in a hurry”, is fine. This is a hazardous lesson since all research work is under pressure. Not only is research intrinsically hard epistemic work, but also the conditions of research are a source of pressure. The sum of the funding schemes (including enrollment in a PhD programme), the toxic ‘publish or perish’ doctrine and other hybrid entanglements within which these PhD students will conduct their work now and in the foreseeable

future seem to suggest that there are always reasons to cut corners. Perpetual sprinting and hurrying seems not only a rational idea, but under these conditions an imperative! This is a reasonable, but unfortunate conclusion. We by no means suggest that data work in research should replace its ideals for stringent standards with wildness typical of data sprints, or valorize accumulating “technical debt” without strategies to pay it back. We argue however that gaining the experience first-hand in critical proximity (Latour, 2005: 253; Herbrechter, 2017; Birkbak *et al.*, 2015) connects a researcher to the temptations, excitements and doubts of data work, and helps scholars develop a critical technical practice instead of, or in addition to, categorical rejection of data work. Facilitating a stance of critical technical practice in the different phases of ‘sprinting’ was surely the biggest challenge for us as teachers and brought its own moments of over- and underwhelmedness with the intended, perceived and/or achieved results.

The data sprint urged us all to continue working with data despite its biases, and thus stay with the trouble for a while in critical proximity. The result of surfacing specific biases during the sprint was not that data was less biased afterwards, nor solely that we can “tick the box” by saying we reflected on biases and thus legitimise using the data nonetheless. Rather than surrendering to biases and sensations of overwhelmedness and underwhelmedness, sprint participants stayed with the vagueness and slipperiness of data through modes of tinkering with the composition of data, trying out different “views” and swapping data out for each other. The data sprint’s material orientation as well as playful wildness facilitated students not simply to think about or expose biases in the data as an empirical finding, but turned biases into pragmatic work.

The mode of critique elicited through sprinting was that of engagement and productiveness, of *thinking-with* (Puig de la Bellacasa, 2017: 71), proposing an analytic tool to ask fruitful questions rather than establish closure. The sprint presented data not as an explanation for anything but much more as something that requires suspicion, that fosters disconcertment, halting and questioning. Student’s sensations of being underwhelmed and not inclined at first sight was immensely valuable; they did not fall into the sheer promises of data work, but pointed to gaps and biases.

We emphasise that data spring learning was not limited to the students learning digital methods, but also allowed us, as organisers and authors to learn crafting situations in which data operates as a device to foster critical questioning and slows down judgement.

Acknowledgements

The data sprint was funded by the Center for Advanced Internet Research (CAIS NRW). We owe our gratitude to the group of PhD fellows and colleagues from D!ARC (Digital Age Research Center) and RUSTlab (Ruhr University Science and Technology Studies Lab) who participated in the data sprint; for without their engagement, expressions of bewildering and provoking reflections this paper would not exist. We are also indebted to our anonymous reviewers and the editors who significantly improved the structure and argument of this manuscript. Further, we are grateful to Barbara Nino Carreras who generously translated our abstract into Spanish.

References

- 4CAT Capture and Analysis Toolkit [Computer software]. (2021). OILab and Digital Methods Initiative at the University of Amsterdam. Retrieved from <https://github.com/digitalmethodsinitiative/4cat>
- Agre, P., Toward a Critical Technical Practice: Lessons Learned in Trying to Reform AI. In Bowker, Geoffrey, ed. (1997). *Social Science, Technical Systems and Cooperative Work: Beyond The Great Divide*. USA: L. Erlbaum Associates Inc. ISBN 978-0-8058-2403-2.
- Ballestero, A., Winthereik B. (2021). *Experimenting with Ethnography. A Companion to Analysis*. Duke University Press.
- Bastian M., Heymann S., Jacomy M. (2009). Gephi: an open source software for exploring and manipulating networks. *International AAAI Conference on Weblogs and Social Media*.
- Berry, D. M., Borra, E., Helmond, A., Plantin, J. C., & Rettberg, J. W. (2015). The data sprint approach: exploring the field of Digital Humanities through Amazon's application programming interface. *Digital Humanities Quarterly*, 9(4).
- Birbak, A., Petersen, M. K., & Elgaard Jensen, T. (2015). Critical proximity as a methodological move in techno-anthropology. *Techné: Research in Philosophy and Technology*, 19(2), 266-290.
- Borgman, C. L., Scharnhorst, A., & Golshan, M. S. (2019). Digital data archives as knowledge infrastructures: Mediating data sharing and reuse. *Journal of the Association for Information Science and Technology*, 70(8), 888-904.
- Bowker, G. C. (2009). *Memory Practices in the Sciences*. MIT Press
- Corti, L., Van den Eynden, V., Bishop, L., & Woollard, M. (2019). *Managing and sharing research data: a guide to good practice*. Sage.
- Criado Perez, C. (2019) *Invisible Women. Exposing Data Bias in World Designed for Men*. Vintage Books.
- Cunningham, W. (1992). The WyCash Portfolio Management System. OOPSLA '92. <http://c2.com/doc/oopsla92.html>. Accessed 1 November 2021.
- Gitelman, L. and Jackson, V. Introduction. In Gitelman, L. (ed.) (2013). *"Raw Data" is an Oxymoron*. MIT Press.
- Godfrey-Smith, P. (2003). *Theory and Reality. An Introduction to the Philosophy of Science*. Chicago: University of Chicago Press.
- Haraway, D. (1988). Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective. *Feminist Studies*, 14(3), 575-599. <https://doi.org/10.2307/3178066>
- Haraway, D. (2016). *Staying with the trouble: Making Kin in the Chthulucene*. Duke University Press.
- Herbrechter, S. (2017) Critical proximity, *Journal for Cultural Research*, 21:4, 323-336, DOI: [10.1080/14797585.2017.1370485](https://doi.org/10.1080/14797585.2017.1370485)

- Hutchins, E. (1995). *Cognition in the Wild*. MIT Press.
- Jacomy, M. (2021). *Situating Visual Network Analysis*. Aalborg Universitetsforlag, Aalborg Universitet. Det Humanistiske Fakultet. Ph.D.-Serien
- Jensen, T. E. (2020). Exploring the Trading Zones of Digital STS. STS Encounters - DASTS working paper series, 11(1), 89-116. https://www.dasts.dk/wp-content/uploads/4_Trading_FV_1.pdf
- Jensen, T. E., Birckbak, A., Madsen, A. K., & Munk, A. K. (2021). Participatory Data Design: Acting in a digital world. In *Making and Doing STS*. MIT Press.
- Kinder-Kurlanda, K., Weller, K., Zenk-Möltgen, W., Pfeffer, J., & Morstatter, F. (2017). Archiving information from geotagged tweets to promote reproducibility and comparability in social media research. *Big Data & Society*. <https://doi.org/10.1177/2053951717736336>
- Kinder-Kurlanda, K. E., & Weller, K. (2020). Research Ethics Practices in a Changing Social Media Landscape. *AoIR Selected Papers of Internet Research, 2020*. <https://doi.org/10.5210/spir.v2020i0.11251>
- Latour, B. (1987). *Science in Action: How to Follow Scientists and Engineers Through Society*. Harvard University Press. ISBN 0-674-79291-2
- Latour, B. (2004). Why has critique run out of steam? From matters of fact to matters of concern. *Critical Inquiry*, 30(2), 225-248.
- Latour, B. (2005). *Reassembling the Social. An Introduction to Actor-Network Theory*. Oxford University Press.
- Law, J. (2011). Heterogeneous Engineering and Tinkering. <http://www.heterogeneities.net/publications/Law2011HeterogeneousEngineeringAndTinkering.pdf>. Accessed 1 November 2021.
- Levin, N., & Leonelli, S. (2017). How Does One “Open” Science? Questions of Value in Biological Research. *Science, Technology & Human Values*, 42(2), 280–305. <https://doi.org/10.1177/0162243916672071>
- Loukissas, Y. A. (2019). *All data are local: Thinking critically in a data-driven society*. MIT Press.
- Marres N. (2015). Why Map Issues? On Controversy Analysis as a Digital Method. *Science, Technology, & Human Values*, 40(5):655-686. doi:10.1177/0162243915574602
- Meldgaard Kjær, K., Ojala, M., Henriksen, L. (2021). Absent Data: Engagements with Absence in a Twitter Collection Process. *Catalyst*, 7(2). <https://doi.org/10.28968/cftt.v7i2.34563>
- Mol, A., Moser, I. and Pols, J. (eds) (2010). *Care in Practice: on Tinkering in Clinics, Homes and Farms*. Transcript verlag, Bielefeld.
- Munk, A. (2001). The digital minced meat. *EASST Review*. 40(1). <https://easst.net/article/the-digital-minced-meat/>
- Munk, A. K., Tommaso, V., & Meunier, A. (2019a). Data Sprints: A Collaborative Format in Digital Controversy Mapping. In J. Vertesi, & D. Ribes (red.), *Digital STS: A Field*

Guide for Science & Technology Studies (s. 472-496). Princeton University Press.
<https://doi.org/10.2307/j.ctvc77mp9.34>

Munk, A. K., Madsen, A. K., & Jacomy, M. (2019b). Thinking Through The Databody: Sprints as Experimental Situations. In Å. Mäkitalo, T. Nicewonger, & M. Elam (Eds.), *Designs for Experimentation and Inquiry: Approaching Learning and Knowing in Digital Transformation* (1 ed., pp. 110-128). Routledge. <https://doi.org/10.4324/9780429489839>

Omena, J. J., Rabello, E. T., & Mintz, A. G. (2020). Digital Methods for Hashtag Engagement Research. *Social Media and Society*, 6(3), 1-18. <https://doi.org/10.1177/2056305120940697>

Parack, S. (2021). Introducing the new Academic Research product track. *Twitter Developer forum*. Retrieved from <https://twittercommunity.com/t/introducing-the-new-academic-research-product-track/148632>.

Pearce, W., Özkula, S. M., Greene, A. K., Teeling, L., Bansard, J. S., Omena, J. J., & Rabello, E. T. (2020). Visual cross-platform analysis: digital methods to research social media images. *Information Communication and Society*, 23(2), 161-180. <https://doi.org/10.1080/1369118X.2018.1486871>

Prainsack, B., & Leonelli, S. (2018). "Responsibility". In *Science and the politics of openness*. Manchester, England: Manchester University Press. Retrieved Dec 22, 2021, from <https://www.manchesteropenhive.com/view/9781526106476/9781526106476.00013.xml>

Puig de la Bellacasa, M. (2017). *Matters of care: Speculative ethics in more than human worlds*. University of Minnesota Press.

Rogers, R. (2013). *Digital Methods*. MIT Press.

Rogers, R. (2019). *Doing Digital Methods*. Sage.

Rogers, Y. (2011). Interaction design gone wild: striving for wild theory. *Interactions*, 18(4) <https://doi.org/10.1145/1978822.1978834>

Sanderhoff, M. (ed.). (2014). *Sharing is Caring. Openness and Sharing in the Cultural Sector*. Statens Museum for Kunst. Copenhagen, Denmark. <https://www.smk.dk/en/article/the-sharing-is-caring-anthology/>

Star, S., Griesemer, J. (1989). Institutional Ecology, 'Translations' and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. *Social Studies of Science*, 19 (3): 387-420. DOI: <https://doi.org/10.1177/030631289019003001>

Sørensen, E. (2009). *The Materiality of Learning: Technology and Knowledge in Educational Practice* (Learning in Doing: Social, Cognitive and Computational Perspectives). Cambridge: Cambridge University Press. DOI:<http://doi.org/10.1017/CBO9780511576362>

Sørensen, E. and Kocksch, L. (2021). Data Durabilities: Towards Conceptualizations of Scientific Long-Term Data Storage. *Engaging Science, Technology, and Society*, 7(1): 12-32. DOI: <https://doi.org/10.17351/ests2021.777>


- Venturini, T., Bounegru, L., Gray, J., & Rogers, R. (2018a). A reality check(list) for digital methods. *New Media & Society*. 20(11): 4195-4217. DOI: <http://doi.org/10.1177/1461444818769236>
- Venturini, T., Munk, A., & Meunier, A. (2018b). Data-Sprinting: a Public Approach to Digital Research. In: Celia Lury; Rachel Fensham; Alexandra Heller-Nicholas. *Routledge handbook of interdisciplinary research methods*, Routledge, Routledge international handbooks, 978-1-138-88687-2. ff10.4324/9781315714523-24ff. Ffhal-01672288f
- Venturini, T., & Munk, A. K. (2021). *Controversy Mapping: A Field Guide*. John Wiley & Sons.
- Verran, H. (2001). *Science and an African logic*. University of Chicago Press.



“Hands On”: incorporación de *data sprints* y *digital methods* en la docencia de marketing digital

“Hands On”: incorporating data sprints and digital methods into digital marketing education

 **Òscar Coromina Rodríguez**
oscar.coromina@mau.se
Malmö University

 **Adrián Padilla Molina**
adrian.padilla@uab.cat
Universitat Autònoma de Barcelona

RESUMEN Este artículo recopila la experiencia acumulada durante 9 años en el diseño y ejecución de *data sprints* en el marco del Máster en Comunicación y Marketing Digital de la Universitat Autònoma de Barcelona. Con más de 90 proyectos distintos a sus espaldas, estas jornadas han sido un punto de encuentro para profesionales, académicos y estudiantes, posibilitando el intercambio de conocimiento y la transferencia de técnicas entre métodos de investigación del ámbito académico con la investigación aplicada al marketing digital. Al mismo tiempo, el formato *data sprint* facilita el conocimiento práctico y la adquisición de competencias comunicativas, organizativas y el proceso de resolución de problemas. El artículo presenta tres ejemplos de proyectos realizados para ilustrar la actividad llevada a cabo y aborda su impacto a partir de un focus group realizado con varios participantes, para señalar los beneficios de la experiencia, así como su dificultad de aplicar el formato en el entorno profesional.

PALABRAS CLAVE

Métodos digitales, data sprint, marketing digital, visualización de datos, análisis de datos

ABSTRACT This article compiles the experience accumulated over 9 years in the design and execution of data sprints within the framework of the Master’s Degree in Digital Communication and Marketing at the Universitat Autònoma de Barcelona. With more than 90 different projects behind them, these conferences have been a meeting point for professionals, academics and students, enabling the exchange of knowledge and the transfer of techniques between academic research methods and research applied to digital marketing. At the same time, the data sprint format facilitates practical knowledge and the acquisition of communication, organisational and problem-solving skills. The article presents three examples of projects to illustrate the activity carried out and discusses its impact based on a focus group with several participants, to point out the benefits of the experience as well as the difficulty of applying the format in the professional environment.

KEYWORDS

Digital Methods, data sprint, digital marketing, data visualization, data analysis

Recibido: 23/01/2022 | Aceptado: 25/05/2022

Cómo citar este artículo:

Coromina Rodríguez, Ò. y Padilla Molina, A. (2022). “Hands On”: incorporación de data sprints y digital methods en la docencia de marketing digital. *Dígitos. Revista de Comunicación Digital*, 8: 51-66. DOI: 10.7203/drdcd.v1i8.243



"Hands On": incorporación de data sprints y digital methods en la docencia de marketing digital

1. Introducción

Durante los últimos 9 años, en el marco del Máster en Comunicación y Marketing Digital (MCMD) de la Universitat Autònoma de Barcelona, se han venido celebrando anualmente una serie de jornadas llamadas "Hands On", cuya filosofía es la transferencia e intercambio de conocimiento científico-técnico adaptando el formato organizativo de los *hackathons* y *data sprints*. El objetivo principal de estas jornadas es formativo: que los estudiantes adquieran competencias en el ámbito del análisis y visualización de datos al mismo tiempo que se familiarizan con los conceptos teóricos y las técnicas propias de los métodos digitales. Adicionalmente, las jornadas "Hands On" han sido una plataforma de experimentación y exploración en dos direcciones: incorporar distintas aproximaciones de los métodos digitales a las técnicas de investigación de los mercados digitales, y explorar nuevas vías para que las herramientas, conceptos y técnicas que los profesionales del marketing digital emplean en su día a día puedan ser apropiados para la investigación académica en el ámbito de las ciencias sociales y las humanidades. Este segundo objetivo conecta con la tradición de los métodos digitales de seguir y utilizar métodos y datos del medio para la investigación de fenómenos sociales y culturales (Rogers, 2019), y reclama el valor de apropiarse y aprovechar prácticas y herramientas del marketing digital para la investigación académica, además de aportar un nuevo punto de observación, una interfaz metodológica (Marres y Gerlitz, 2016), que adopta las técnicas de investigación de los mercados digitales, la especificidad de las plataformas, los datos y las prácticas de los usuarios y su contexto de aplicación.

En el presente artículo recopilamos la experiencia acumulada en el diseño y ejecución de *data sprints* en el ámbito académico, una metodología de trabajo basada en el "*learning by doing*" que se caracteriza por su intensidad y componente experimental del trabajo que se lleva a cabo. Para ello, realizamos una contextualización teórica de los *data sprints* y su conexión con los Métodos Digitales. Seguidamente, describimos la adaptación realizada de este formato a un contexto de máster en comunicación y marketing digital, cuya vocación es la capacitación de profesionales en este ámbito. Para comprender mejor la aplicabilidad de los *data sprints* en este contexto formativo, presentamos tres ejemplos de proyectos realizados, que combinan herramientas, datos y conceptos propios de la investigación de mercados digitales con los Métodos Digitales. Seguidamente, presentamos los resultados de un Focus Group en el que participantes en las sesiones "Hands On" reflexionan sobre los resultados formativos y las posibilidades que surgen de aplicar métodos digitales en el ámbito del marketing digital. En el apartado de conclusiones, incidimos en los aprendizajes, aportaciones e impacto que esta experiencia ha supuesto a los participantes y también sus contribuciones a los métodos digitales.

2. *Data Sprints y Digital Methods*

Los *data sprints* tienen su origen en los *hackathons* y *barcamps* que a finales del siglo pasado se introducen en el ámbito del desarrollo del software y más concretamente el entorno del software libre (Venturini, Munk y Meunier, 2018). Como su propio nombre indica los *hackathons* se inspiran en la comunidad hacker o, más concretamente, en los modelos organizativos y colaborativos alternativos que textos como “la catedral y el bazar” (Raymond, 1998) y “la ética hacker y el espíritu de la sociedad de la información” (Himanen, 2001) reivindican por fomentar la creatividad y la innovación. Ambos escritos presentan una visión de los hackers que dista de la imagen de criminal informático o activista de Anonymous que el término evoca en la actualidad y toman como referencia la cultura de las primeras comunidades hackers surgidas en los campus universitarios de Estados Unidos en la década de los 60. Estos primeros hackers eran estudiantes apasionados de la informática que establecieron una relación simbiótica con los ordenadores al mismo tiempo que desarrollaban un código moral y de conducta no escrito que promovía el acceso libre a los ordenadores y a la información, el conocimiento práctico (*hands on*), la descentralización, la organización del trabajo fuera de los horarios laborales, el imperativo de la productividad y las jerarquías. Su aproximación a los ordenadores y la programación se caracterizaba por ser eminentemente exploratoria y experimental y fue fundamental en el desarrollo de la informática más allá del terreno científico extendiendo su uso y funcionalidades al ámbito lúdico, artístico y profesional (Levy, 1984). Los *hackathons* toman de la ética hacker su carácter experimental, lúdico y exploratorio, así como un formato que elude los corsés de las rutinas empresariales tradicionales y las jerarquías con el objetivo de implementar soluciones innovadoras y creativas en el ámbito del desarrollo del software (Briscoe y Mulligan, 2014).

Los *data sprints* son la adaptación del formato *hackathon* al ámbito académico y coinciden con la eclosión del fenómeno web 2.0, que facilita el acceso y análisis de datos a niveles hasta el momento sin precedentes, y distintos llamamientos a aprovechar esta oportunidad para el desarrollo de una ciencia social computacional (Lazer *et al.*, 2009) que analice las trazas digitales para interrogar sobre las fuerzas que dan forma a fenómenos sociales y culturales (Latour, 2007a). En este contexto es pertinente señalar el papel pionero de la *Digital Methods Initiative* de la Universidad de Ámsterdam, y su propuesta metodológica que busca ‘seguir los métodos en evolución del medio’ (Rogers, 2013) así como a los académicos que trabajan en esta tradición, que han ilustrado el valor de reutilizar datos y métodos de plataformas como Facebook (Poell *et al.* 2016) o Twitter (Bruns y Burgess, 2011) teniendo en cuenta tanto los datos como la infraestructura a través de la cual se producen. Los *Digital Methods* beben de las teorías desarrolladas en el ámbito de los estudios de Ciencia, Tecnología y Sociedad (Jensen *et al.*, 2021) y la teoría actor-red (Latour, 2007b). No es casual, por tanto, que los *data sprints* surjan en paralelo en el seno de la *Digital Methods Initiative* en Amsterdam y en el *Média Lab* de Sciences Po Paris, fundado por Bruno Latour. Ambos proyectos se caracterizan por la incorporación de las trazas digitales al repertorio de objetos de análisis y el uso de métodos computacionales para extraer, analizar y visualizar los datos. Al igual que los *hackathons*, los *data sprints* se sirven de la unidad en el tiempo y espacio para contrarrestar la heterogeneidad de los participantes, y la priorización de soluciones

“quick and dirty” por encima de la exhaustividad (Venturini, Munk y Meunier, 2018). Pese a que la conexión de los *hackathons* y los *data sprints* es clara y reconocida, estos últimos se caracterizan por una serie de rasgos distintivos, tales como la participación de “stakeholders” (personas con conocimiento, experiencia práctica en las materias investigadas), que aportan contexto y cercanía en las distintas fases del sprint y ayudan a enfocar la investigación en los resultados más relevantes (Madsen y Munk, 2019). En definitiva, los *data sprints* consisten en pasar de la abstracción a la concreción y a dotar a los actores implicados en los ámbitos de estudio de voz propia y capacidad de incidir en el desarrollo de investigación (Birkbak, Petersen y Elgaard Jensen, 2015).

3. Las Jornadas “Hands On” del MCMD

Al igual que los *data sprints*, las jornadas “Hands On” se caracterizan por ser un punto de encuentro entre profesionales en activo en el ámbito de la comunicación y el marketing digital, académicos y estudiantes, de perfiles formativos muy variados, que se integran en equipos multidisciplinares, con diferentes competencias y también con sensibilidades y objetivos diversos. Los talleres “Hands On” tienen una duración menor a la que suele ser habitual en un *data sprint*. Esto es consecuencia de la adaptación del formato a los horarios y calendario del MCMD, se extienden a lo largo de 3 días y comprenden 12 horas de reunión presencial. En este tiempo los diferentes grupos desarrollan un proyecto que culmina con la presentación de los resultados. Posteriormente los principales resultados se resumen en un breve artículo que se publica en el blog del MCMD. Estos artículos hacen las veces de informe técnico, y permiten a los nuevos equipos comprender las aproximaciones realizadas en el pasado, retomar metodologías, contrastar resultados o profundizar en casos de estudios que todavía merecen nuestra atención.

Puesto que el formato de las “Hands On” es especialmente intensivo y acotado en el tiempo, exige una formación propedéutica previa dentro del currículum del máster en el ámbito del análisis y visualización de datos así como una introducción teórica a los métodos digitales y las técnicas y herramientas que los distinguen. En estas sesiones de preparación los estudiantes se familiarizan con conceptos, métodos y software para gestionar, analizar y visualizar grandes volúmenes de información. Al mismo tiempo, durante la fase preparatoria se realiza la selección de temas y la recolección de datos, de modo que los participantes disponen desde el primer minuto de los datos listos para su análisis. Inicialmente la recolección de datos formaba parte de los proyectos y se realizaba durante las jornadas, pero al disponer de menos tiempo del habitual y para agilizar el trabajo, se tomó la decisión de empezar las jornadas con las colecciones de datos preparadas, focalizando los *sprints* en la exploración y análisis.

En los últimos 9 años se han llevado a cabo cerca de 90 proyectos durante las “Hands On”. En cada edición se celebran dos *data sprints*, con 4 o 5 proyectos distintos en cada uno de ellos. Cada equipo de trabajo está compuesto por entre 5 y 7 estudiantes por lo que cerca de 250 alumnos han participado en estos talleres. Las jornadas “Hands On” han acogido gran variedad de proyectos. Desde trabajos de investigación de mercados digitales, hasta investigaciones de corte académico, pero todos se han caracterizado por integrar una gran variedad de metodologías y herramientas centradas en la recolección,

tratamiento y análisis de grandes volúmenes de datos, que habitualmente no forman parte del repertorio de herramientas de los profesionales del marketing digital. Algunas de estas herramientas de recolección de datos, han sido desarrolladas por la *Digital Methods Initiative*, como es el caso de "TCAT" (Borra y Rieder, 2014) para la captura de datos de Twitter, "Netvizz" (Rieder, 2013) para la extracción de datos de Facebook y "Youtube Data Tools" (Rieder, 2015) para la extracción de datos de YouTube. Otras, han sido desarrolladas *ad hoc* para las "Hands on", como "Twitter for Academic Research" (Padilla, 2020a), para la extracción sistematizada y masiva de datos históricos de Twitter o "Instaloader Scripts" (Padilla, 2020b) para el procesamiento de datos obtenidos de la plataforma Instagram. El desarrollo de software propio requiere de la existencia de cierta infraestructura técnica: hardware capaz de albergar los programas de captura y las bases de datos. Pero también es necesario contar con la colaboración de personal técnico capaz de escribir este software, por lo que es necesario planificar un *data sprint* con mucha antelación. Las "Hands On" también se han servido de herramientas específicas del marketing digital. Software como Ahrefs, Xenu o ScreamingFrog, muy utilizados en la optimización para motores de búsqueda (SEO), y Google Analytics o Google Search Console, propios de la analítica digital, aportan datos sobre el comportamiento de los usuarios. Los datos e información obtenida a través de estas herramientas complementa los datasets de análisis utilizados durante los talleres.

4. Proyectos

Las plataformas digitales y las redes sociales han sido el espacio de investigación de la mayoría de estos proyectos, que han sido testigo y se han adaptado a los cambios en las políticas de acceso a los datos de cada plataforma. El caso más paradigmático es Facebook, que inicialmente permitía el acceso a los datos y metadatos de páginas y perfiles a través de su API. Pero tras el estallido del caso Cambridge Analítica, en 2018, Facebook retiró todas las vías de acceso y sometió a un fuerte control cualquier iniciativa de estudio académico de su entorno, imposibilitando (a día de hoy) cualquier forma de acceso a la información de la plataforma. En el otro lado de la balanza tenemos a Twitter, una plataforma que tradicionalmente ha limitado el acceso a sus datos, pero que en el año 2020 abrió su archivo histórico con fines de investigación, otorgando un acceso muy amplio a todo su contenido de forma prácticamente ilimitada. En cualquier caso, es un hecho que el acceso a los datos mediante APIs es cada vez más restringido para los académicos (Bruns, 2019), dificultando el escrutinio de las plataformas y situando a la investigación ante dilemas éticos (Perriam, Birckbak y Freeman, 2020). No obstante, las APIs siguen siendo esenciales en el modelo de monetización de las plataformas digitales, y el acceso a datos por parte de profesionales del marketing y la publicidad no se ha visto afectado de la misma manera, por lo que la apropiación de estos entornos de transferencia de datos representa una oportunidad para los métodos digitales.

A continuación, presentamos brevemente tres casos de estudio, que ejemplifican el trabajo realizado durante las "Hands On" y también muestran cómo se han integrado distintas herramientas y metodologías profesionales con los métodos digitales.

4.1. Análisis multiplataforma de la estrategia de comunicación de Ada Colau

Coincidiendo con la irrupción de Ada Colau en la política municipal y su proyección también en el ámbito estatal se llevaron a cabo varios proyectos que se servían de datos de distintas plataformas como materia prima del análisis.

Objetivos del proyecto: comprender la estrategia comunicativa de Ada Colau así como los resultados cosechados por la misma, basándose en el engagement generado por los contenidos publicados a través de las distintas cuentas vinculadas a Ada Colau, alcaldesa de Barcelona, en Twitter y Facebook. Al mismo tiempo se buscaba contrastar esta estrategia con los contenidos que los usuarios publicaban mencionando a Ada Colau y sus iniciativas.

Metodología:

1. Extracción de todas las publicaciones del perfil de Ada Colau en Twitter.
2. Captura de Tweets mencionando al usuario de Ada Colau en Twitter así como las palabras clave "colau", "ada colau" o "adacolau", entre octubre del 2015 y marzo del 2016.
3. Se realiza un análisis longitudinal de la presencia de hashtags en las publicaciones para observar cómo evolucionan los distintos temas en los que está involucrada la alcaldesa.
4. Recolección de las páginas de Facebook relacionadas con el perfil de Ada Colau en dicha plataforma mediante Netvizz. Visualización de la like-network de dicha página.
5. Extracción, también mediante Netvizz, de los posts y datos asociados publicados en las distintas páginas de Facebook utilizadas habitualmente por Ada Colau con fines comunicativos.
6. Exploración cuantitativa de los datos mediante técnicas de visualización con el fin de observar tendencias y patrones recurrentes.

Herramientas: Además de Netvizz y TCAT para la recolección de datos se utilizó Gephi para el análisis visual de redes, RankFlow (Rieder, 2017) para observar la evolución y continuidad de hashtags a lo largo del tiempo y Tableau para la exploración y visualización de datos cuantitativos.

Resultados: El proyecto permitió observar un uso muy intensivo de las redes dentro de la estrategia comunicativa de Ada Colau que se caracterizaba por la capacidad de generar polémicas de forma recurrente que conseguían involucrar tanto a detractores como partidarios de la alcaldesa de Barcelona. Asimismo, se pudieron detectar y cuantificar los eventos y temáticas que generaban picos de actividad e interacción mediante el conteo de "comentarios", "retweets" o "likes". El análisis de la *like-network* reveló toda la constelación de personajes públicos, políticos, instituciones y organizaciones afines a la alcaldesa y los vínculos entre ellos. A destacar el papel primordial que desempeñó en esta primera etapa la Plataforma de Afectados por la Hipoteca (PAH). Recurrir a la polémica proporcionaba réditos importantes en materia de interacción con los usuarios en las plataformas digitales sugiriendo que la estrategia comunicativa de la primera

etapa de Ada Colau se basaba en la generación constante de mensajes controvertidos para marcar la agenda mediática.

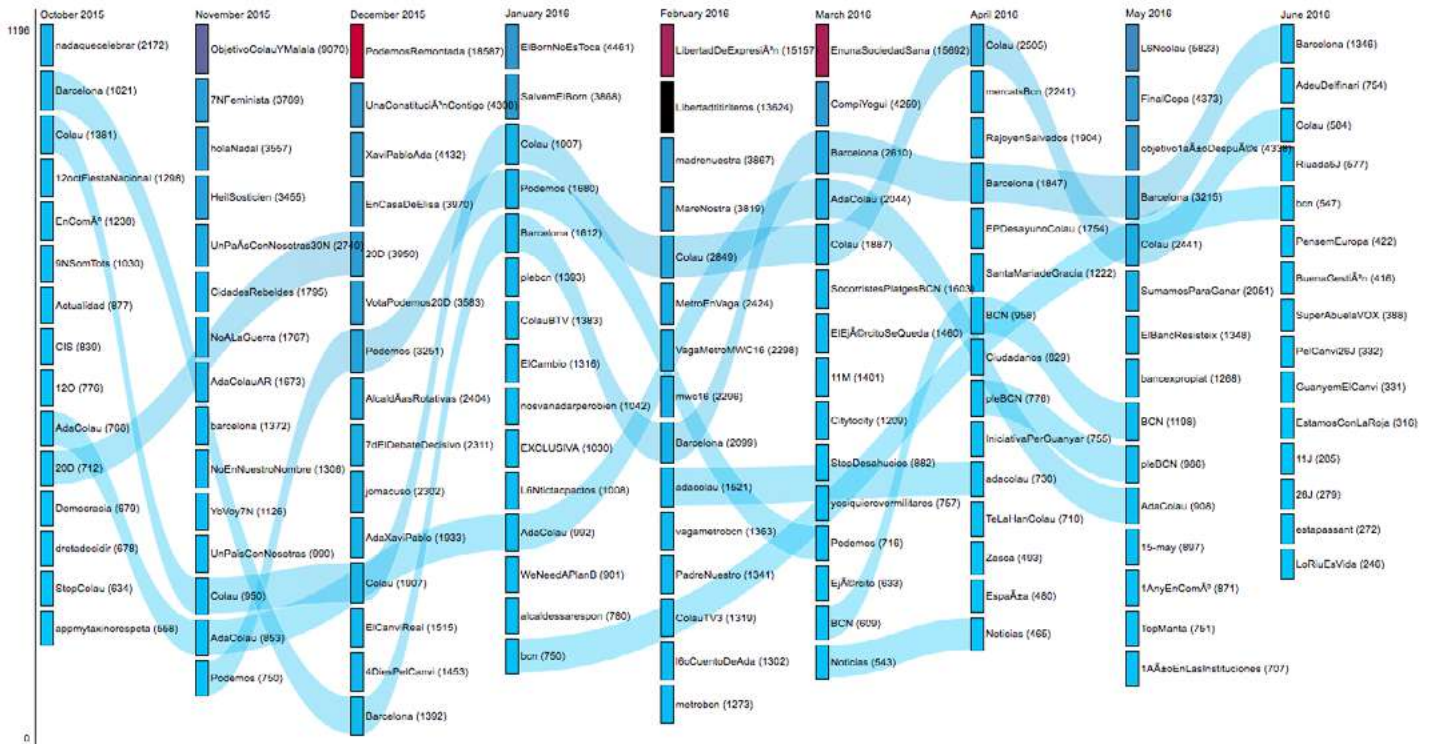


Figura 1: Rankflow de los diferentes hashtags a lo largo de 9 meses de análisis

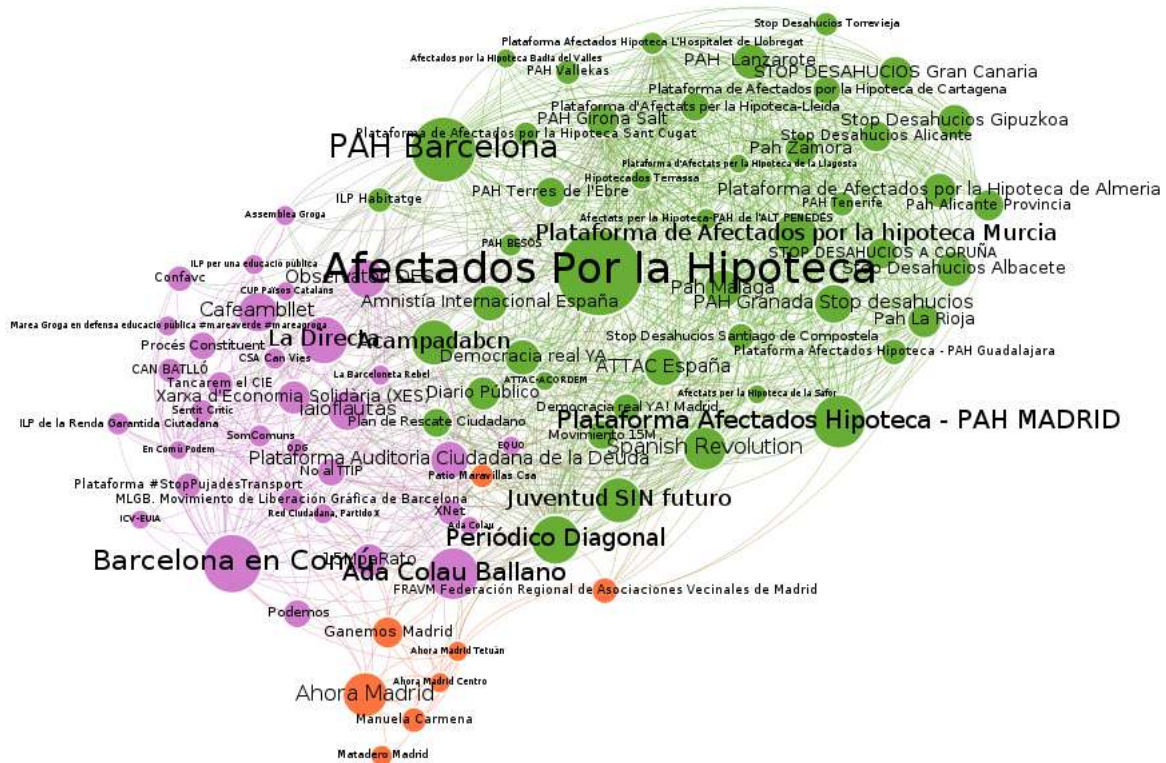


Figura 2: Grafo de comunidades de usuarios vinculados a Ada Colau

Aportaciones: El proyecto permitió incorporar nuevas fuentes de datos y métodos de análisis para investigaciones relacionadas con el marketing político y, más concretamente, con la gestión de la imagen y la reputación digital en redes sociales y plataformas digitales. Los hallazgos realizados durante las "Hands On" tuvieron continuidad con la realización de un Trabajo de Fin de Máster por parte de uno de los alumnos participantes en el *data sprint*, que dio a pie a una publicación académica (Coromina, Casacuberta y Renom, 2018) que analizó la polémica generada por la ubicación de unas estatuas franquistas en el Centro Cultural del Born con el motivo de una exposición que invitaba a reflexionar sobre la huella del franquismo en la ciudad de Barcelona.

4.2. Search Engine Optimization y Digital Methods

Dentro del ámbito del marketing y la comunicación digital existen disciplinas que tradicionalmente han requerido de procesos exhaustivos de análisis e investigación aplicada. Uno de los ejemplos más claros es el SEO (*Search Engine Optimization*) que engloba diferentes técnicas para obtener mayor visibilidad en los motores de búsqueda como Google. Se trata de un área especialmente estratégica para obtener visitas en los sitios web. El SEO comparte con los métodos digitales la utilización de datos y objetos nativos digitales (hipervínculos, páginas web, dominios y meta-etiquetas). En distintas "Hands On" se han realizado proyectos de auditoría SEO donde se ha explorado diferentes vías para acercar la investigación académica (los métodos digitales) con la investigación aplicada (SEO). Una de las vías más fructíferas ha sido incorporar técnicas y herramientas de análisis visual de redes para auditar la salud y la arquitectura interna de sitios web.

Objetivos del proyecto: Auditar los contenidos del sitio web de la asociación animalista People for the Ethical Treatment of Animals (PETA) para evaluar y detectar oportunidades para mejorar el posicionamiento en buscadores del sitio web www.peta.org.

Metodología:

1. Rastreo y extracción de información de las distintas páginas que componen el sitio web y de todos los hipervínculos del sitio web.
2. Construcción de la red interna de enlaces.
3. Análisis visual del entramado de enlaces internos.

Herramientas: Para la extracción de información de las páginas del sitio web se utilizaron herramientas propias de la investigación en el ámbito del SEO: Xenu, Screaming Frog y Ahrefs. Para el análisis visual de la red de enlaces, Gephi.

Resultados: El proyecto desveló una estructura de enlaces centralizada en el que la página principal de las distintas versiones idiomáticas acumulaba la práctica totalidad de hipervínculos en detrimento del resto de páginas. A efectos de posicionamiento en buscadores esto se traduce en una baja visibilidad de los proyectos que impulsa esta

organización. La investigación culminó con la creación de un documento de trabajo en formato “informe” que permitía identificar y priorizar con rapidez los puntos de mejora y los problemas críticos de la arquitectura del sitio web.

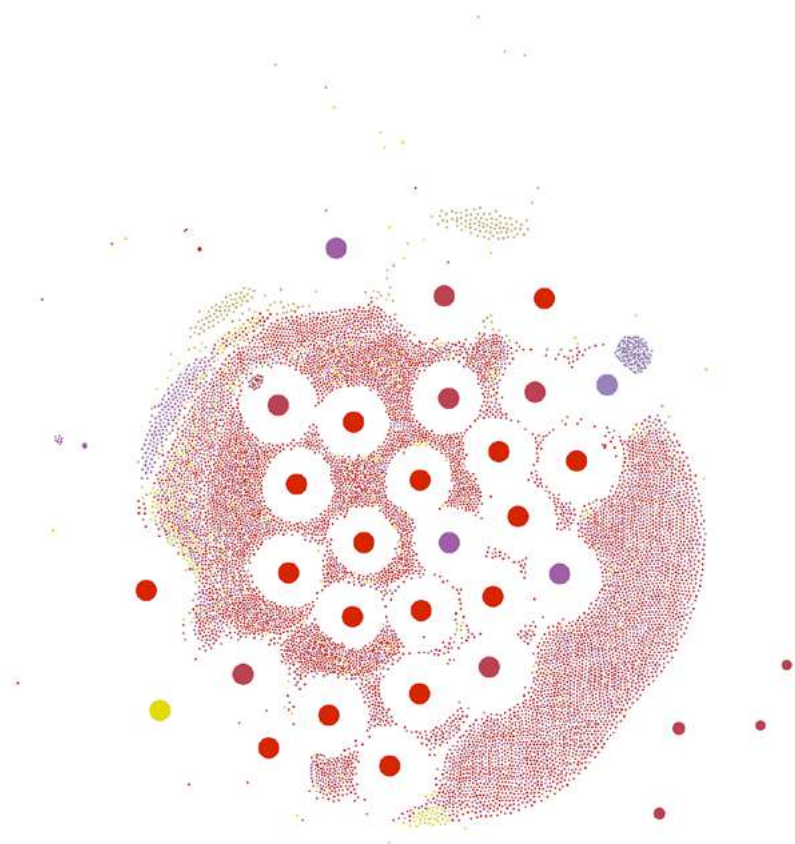


Figura 3: Grafo de hipervínculos y páginas web en el sitio web peta.org

Aportaciones: La principal contribución de este proyecto debe situarse dentro de la investigación aplicada en el campo del SEO, ya que los métodos utilizados y el análisis visual de redes permiten la exploración eficaz del perfil de los enlaces internos, señalando oportunidades para mejorar la visibilidad de los contenidos en buscadores, así como la indexabilidad, la usabilidad, la arquitectura de la información o la accesibilidad web. Al mismo tiempo, el uso de *crawlers* como Xenu y Screaming Frog puede incorporarse al repertorio de herramientas de los métodos digitales para el análisis de hipervínculos.

4.3. Optimización de contenidos en YouTube

YouTube es la única de las grandes plataformas digitales que ofrece a los creadores de contenido la posibilidad de obtener ingresos mediante el programa de monetización *YouTube Partners Program*. Parámetros como la duración de los videos, la temática o la frecuencia de publicación influyen decisivamente en la visibilidad y, por consiguiente, en la monetización de los videos. Aunque YouTube no hace público el peso de los distintos factores, los creadores profesionales adaptan sus contenidos para ganarse el favor de los algoritmos de recomendación (Coromina, Matamoros-Fernández y Rieder, 2020). Este proyecto parte del análisis de los YouTubers con más suscriptores y visualizaciones

en sus videos para determinar qué factores son percibidos como más determinantes en el proceso de optimización.

Objetivo del proyecto: Analizar la evolución de los distintos parámetros de los contenidos en YouTube y su relación con los ingresos generados partiendo de la base que estos están estrechamente relacionados con las visualizaciones obtenidas.

Metodología: El proyecto se benefició del acceso a la base de datos utilizada para el artículo “Mapping YouTube, a Quantitative Exploration of a Platformed Media System” (Rieder, Coromina y Matamoros-Fernández, 2020), que reúne más de 5 millones de videos publicados entre los años 2007 y 2019 por los canales localizados en España con más de 100.000. suscriptores. El análisis se centró en los contenidos publicados bajo la categoría “entretenimiento”, “gaming”, “how to and style” y “people and blogs”, y aquellos elementos sobre los que un creador de contenidos puede incidir de forma directa: la duración del contenido, la extensión del título, las descripciones y tags, así como el momento de publicación. Además se utilizaron las métricas “likes”, “dislikes”, “comentarios” y “visualizaciones” como indicadores para comprender cuál era la combinación óptima de elementos.

Herramientas: Python para la creación de un script para el tratamiento estadístico de los datos, y Tableau para la exploración y creación de visualizaciones que permitieran expresar gráficamente las tendencias.

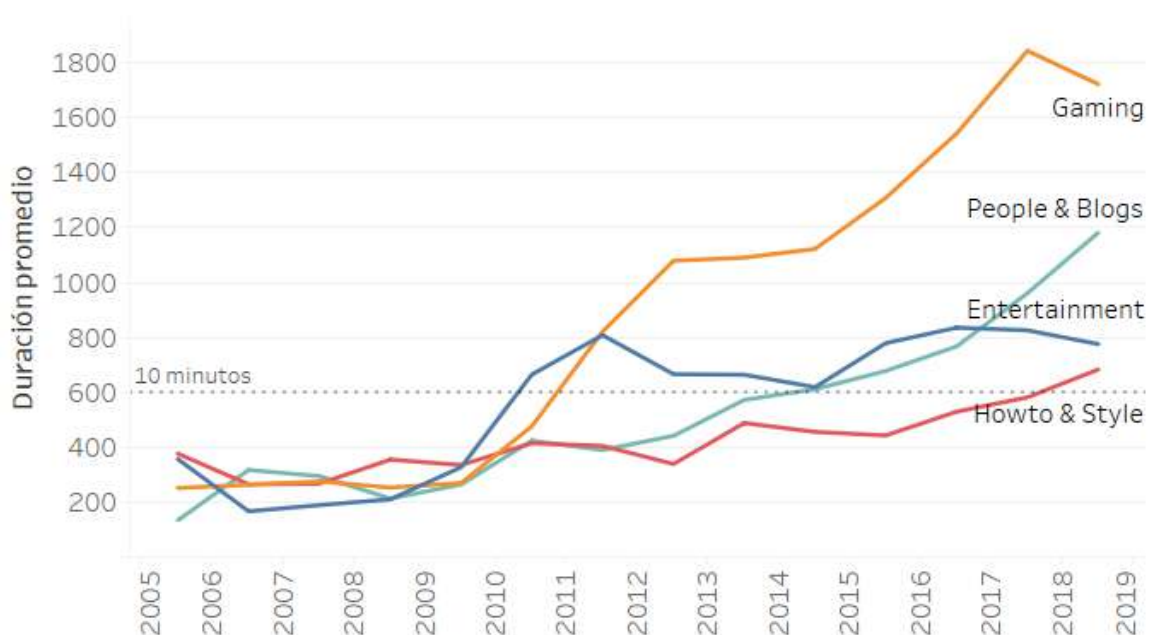


Figura 4: Evolución de la duración promedio de los contenidos de YouTube para 4 categorías.

Resultados: El proyecto permitió determinar no solamente qué estrategia de contenidos y publicación es óptima para los creadores de contenido, sino también qué evolución ha seguido este “punto óptimo” a lo largo del tiempo, desvelando el modo en cómo la plataforma ha modificado sus criterios de visibilidad y ha intervenido de forma

indirecta en las rutinas de producción de los creadores, premiando aquellos quienes cumplen con ciertas métricas de duración, horarios, frecuencia de publicación y nivel de interacción.

Aportaciones: Este proyecto muestra cómo YouTube ha modificado a lo largo del tiempo los distintos criterios que deben cumplir los contenidos para lograr una alta tasa de visibilidad, y al mismo tiempo cómo los creadores se han adaptado a estos criterios. De este modo, aporta información a los creadores de contenidos sobre las estrategias óptimas en la creación de vídeos, y cuál es la tendencia en España dentro de las categorías de contenidos analizadas.

5. Percepción de los participantes

Para comprender el impacto de las “Hands On” se ha realizado un Focus Group con un total de 8 participantes (4 hombres y 4 mujeres), el grupo incluía 6 alumnos de anteriores ediciones del Máster en Comunicación y Marketing Digital y 2 facilitadores. El encuentro se dividió en tres bloques: la utilidad del modelo organizativo, el uso del data-sprint como método de aprendizaje y, por último, la transferencia de conocimiento entre el ámbito académico y el profesional.

5.1. Modelo organizativo

Los participantes apuntan que la metodología del *data sprint* de las “Hands On” ayuda a desarrollar habilidades sociales y comunicativas, que son competencias vitales en un entorno laboral colaborativo. Así mismo el modelo organizativo horizontal obliga a todos los miembros del equipo a asumir un cierto grado de responsabilidad e implicarse de forma directa con el proyecto, lo que los alumnos consideran fundamental para fomentar competencias relacionadas con la comunicación, organización y negociación que permiten superar las tensiones que genera el proyecto.

“El hecho de no solo realizar el proyecto, sino también defenderlo frente a los compañeros, encontrar los puntos débiles y responder a preguntas, saber comunicarlo. Esto es útil en el mundo laboral” (Martí, 28 años).

“Desde pequeños se hacen muchos trabajos en grupo, pero el aprendizaje de las “Hands On” es una experiencia de trabajo real, donde debes organizar personas, roles y optimizar tiempo. Esto es algo que se ajusta mucho a la realidad del trabajo” (Martí, 28 años).

Los participantes también apuntan al factor “tiempo” como elemento de presión decisivo para tomar decisiones y guiar el proyecto que se ejecuta durante las “Hands On”.

“El formato sprint añade una presión interesante que obliga a los alumnos a sacar adelante un trabajo en un tiempo muy limitado. El tiempo se aprovecha mucho, al ser muy limitado. Se aprecia mucho tener a los profesores de cerca y el trabajo en pequeños grupos” (Antoine, 44 años).

5.2. El *data sprint* como método de aprendizaje

Todos coinciden en que las "Hands On" son una oportunidad para aprender de forma activa, al ser el facilitador un mero acompañante y no ejercer autoridad sobre el proyecto, lo que obliga a los alumnos a tomar decisiones consensuadas.

"En las clases se aprende de forma pasiva. Depende mucho de la personalidad de los alumnos el participar más o menos. En cambio, en las "Hands On" debes asumir responsabilidades, y esto obliga a involucrarse directamente y tener una posición activa en el aprendizaje" (David, 33 años).

Todos los participantes señalan la utilidad de las "Hands On" y, más concretamente, los *data sprints*, como un método para aproximarse al análisis y visualización de datos desde una perspectiva práctica y haciendo uso de ejemplos reales.

"Las "Hands On" aportan un aprendizaje en formas de pensar y organizar el trabajo en torno a un conjunto de datos" (Carmen, 25 años).

Las temáticas elegidas son de interés de los alumnos, aunque estas no tengan una relación directa con el marketing digital. Los temas elegidos y el despliegue metodológico se perciben cómo "una excusa para hacer uso de ciertas herramientas", que pueden ser utilizadas en otros casos de estudio más próximos a su interés.

5.3. Transferencia de conocimiento científico-técnico al ámbito laboral

En los que concierne a los métodos digitales y el apartado académico de los talleres, todos los alumnos coinciden en la complejidad de trasladar o adaptar el rigor científico de las "Hands On" al entorno laboral. Los participantes señalan que las empresas no entienden o no consideran útil procedimentar los análisis y describir el método utilizado. También se apunta a la falta de tiempo en el entorno laboral para profundizar en la metodología y el desconocimiento de los métodos digitales por parte de los cargos intermedios y los clientes de las agencias digitales.

"Los Métodos Digitales son útiles para aprender sobre temas de repercusión, pero es muy complicado llevar este tipo de métodos al ámbito profesional real, ya que los superiores no comprenden lo que se está haciendo. No es un conocimiento tangible" (David, 33 años).

"A veces no es posible aplicar los Métodos Digitales en el trabajo debido al tiempo que tienes para realizar este tipo de proyectos en el ámbito profesional. Tampoco se suelen enseñar al cliente y eso hace que no sean útiles" (Antoine, 44 años).

Aunque los alumnos reconocen el valor de los métodos digitales, a efectos prácticos no les aporta ningún valor adicional o ventaja competitiva como trabajadores, al ser este un campo totalmente desconocido para los reclutadores o responsables de las empresas. Sin embargo, coinciden que la formulación de preguntas de investigación y el diseño de métodos para responderlas son aplicables en el entorno laboral, especialmente en el análisis de datos.

"A lo largo del máster, en general, se adquieren frameworks para trabajar y mejorar en cualquier tema. Los Métodos Digitales y las "Hands On", en concreto, inculcan una forma de pensar" (Carmen, 25 años).

Los participantes se muestran optimistas en cuanto al potencial de los sistemas de organización horizontales para resolver en equipo tareas complejas en un breve espacio de tiempo, aunque reconocen que a menudo las empresas son contrarias a este tipo de modelos organizativos debido a las jerarquías existentes.

“Las “Hands On” son un buen ejemplo de cómo llevar a cabo un proyecto en tiempo récord, pero es muy complejo que las instituciones acepten estos métodos, porque no los entienden” (Jordi, 28 años).

6. Conclusiones

A lo largo de los 9 años que han tenido lugar las “Hands On” se ha demostrado la utilidad del *data sprint* en múltiples contextos y modalidades. Su aproximación práctica al análisis de datos para la realización de trabajos de investigación ha permitido a los estudiantes adquirir competencias y conocimientos con el objetivo de aplicarlas a su actividad profesional y más concretamente a la investigación de mercados digitales. Todo ello, en un formato organizativo que permite a los asistentes implicarse y aproximarse a distintos fenómenos de una manera creativa e incluso con un importante componente lúdico. La articulación de equipos de trabajo donde participan distintos perfiles: estudiantes, académicos y profesionales, ha contribuido de forma determinante a la adquisición de competencias transversales.

La realización de este tipo de actividades en el contexto de un máster de vocación profesional ha requerido una fuerte adaptación que se ha llevado a cabo sin renunciar a las señas de identidad de los *data sprints* (Venturini, Munk y Meunier, 2018). En el caso que nos ocupa es preciso situar este *data sprint* dentro de una propuesta formativa y didáctica con un marco de actuación muy específico, la investigación de mercados digitales. La introducción de sesiones propedéuticas para familiarizar a los alumnos en el uso de ciertas herramientas, así como la inclusión de una introducción a los principios que definen los métodos digitales (Rogers, 2019) ha sido indispensable para aplicar técnicas y procesos para el análisis y visualización de datos.

Al mismo tiempo, las “Hands On” han sido un espacio de experimentación y de adopción de métodos innovadores con el potencial de enriquecer las prácticas de los profesionales del marketing digital, un marco para la transferencia de conocimientos del ámbito académico al profesional. Los tres proyectos expuestos muestran distintas oportunidades para enriquecer el repertorio metodológico de la investigación aplicada de los mercados digitales. Poniendo el foco en la obtención de datos de las plataformas digitales mediante sus distintas API y en el análisis y visualización de datos como instrumento exploratorio. A menudo, invitando a recorrer caminos poco transitados en el marketing digital, como el análisis visual de redes, que se ha revelado especialmente útil en el ámbito del SEO. No obstante, según indican los participantes el ámbito profesional no es especialmente receptivo a la inclusión de estas nuevas metodologías debido a su desconocimiento.

Estos *data sprints* también han permitido investigar oportunidades para continuar con la apropiación de métodos de los medios para la investigación académica en el ámbito

de las ciencias sociales y humanidades. La incorporación de nuevas fuentes de datos cobra especial relevancia en la coyuntura actual, de restricción de acceso a datos para las investigaciones académicas (Bruns, 2019; Perriam, Birkbak y Freeman, 2020) por parte de las plataformas. En este sentido se ha experimentado con distintos servicios y herramientas como Xenu, Screaming Frog o Ahrefs, que aportan una gran cantidad de información sobre los sitios web. Además de los datos que pueden proporcionar dichos recursos, el marketing digital proporciona marcos y conceptos que los métodos digitales pueden también adoptar. En este aspecto el formato *data sprint* con la participación de *stakeholders*, expertos en las distintas disciplinas que integran el marketing digital, propicia un diálogo entre la investigación aplicada y la investigación académica que se seguirá explorando en futuras ediciones del máster.

A estas contribuciones hay que sumar todo el software creado para la extracción y análisis de datos que se ha desarrollado para los *data sprint* y que, siendo fiel al espíritu hacker que da pie a los *hackathons*, se ha compartido con la comunidad académica y profesional en formato de código abierto para que otros investigadores y desarrolladores puedan observar los caminos ya recorridos y continuar desarrollando las distintas herramientas.

Referencias

- Birkbak, A., Petersen, M.K. y Elgaard Jensen, T. (2015). Critical Proximity as a Methodological Move in Techno-Anthropology. *Techné: Research in Philosophy and Technology*, 19(2): 266–290. doi:10.5840/techne201591138.
- Borra, E. y Rieder, B. (2014). Programmed method: Developing a toolset for capturing and analyzing tweets. *Aslib journal of information management*, 66(3):262-278. <https://doi.org/10.1108/AJIM-09-2013-0094>
- Briscoe, G. y Mulligan, C. (2014). *Digital innovation: The hackathon phenomenon*. Available at: <http://www.creativeworkslondon.org.uk/wp-content/uploads/2013/11/Digital-Innovation-The-Hackathon-Phenomenon1.pdf>
- Bruns, A., y Burgess, J. (2011). The use of Twitter hashtags in the formation of ad hoc publics. En *Proceedings of the 6th European consortium for political research (ECPR) general conference 2011* (pp. 1-9). The European Consortium for Political Research (ECPR).
- Bruns, A. (2019). After the 'APIcalypse': social media platforms and their fight against critical scholarly research. *Information, Communication & Society*, 22(11): 1544–1566.
- Coromina, Ò., Casacuberta, D. y Renom, I.G. (2018). La batalla del Born. Relat polític i estratègies de posició ideològica a Twitter. *Comunicació: revista de recerca i d'anàlisi*, 35(2): 25–44.
- Coromina, Ò., Matamoros-Fernández, A. y Rieder, B. (2020). "Follow the Money: A Large-Scale Investigation of Monetization and Optimization on Youtube". Comunicación presentada en *AoIR 2020: The 21th Annual Conference of the Association of Internet Researchers*. Virtual Event: AoIR <https://journals.uic.edu/ojs/index.php/spir/article/view/11194/9803>
- Himanen, P. (2001) *The Hacker Ethic*. London: Viking.

- Jensen, T. E., Birkbak, A., Madsen, A. K., y Munk, A. K. (2021). Participatory Data Design: Acting in a digital world. En G. Downey y T. Zuiderent-Jerak (Eds.), *Making & Doing: Activating STS through Knowledge Expression and Travel* (pp. 117-136). MIT Press.
- Latour, B. (2007a). Beware, your imagination leaves digital traces. *Times Higher Literary Supplement*, 6(4): 129-131.
- Latour, B. (2007b). *Reassembling the Social: An Introduction to Actor-Network-Theory*. OUP Oxford.
- Lazer, D. et al. (2009). Computational Social Science. *Science*, 323(5915): 721–723. doi:10.1126/science.1167742.
- Levy, S. (1984). *Hackers: Heroes of the computer revolution (Vol. 14)*. Garden City, NY: Anchor Press/Doubleday.
- Madsen, A.K. y Munk, A.K. (2019). Experiments with a data-public: Moving digital methods into critical proximity with political practice. *Big Data & Society*, 6(1): 2053951718825357. doi:10.1177/2053951718825357.
- Marres, N. y Gerlitz, C. (2016). Interface methods: renegotiating relations between digital social research, STS and sociology. *The Sociological Review*, 64(1): 21–46. doi:10.1111/1467-954X.12314.
- Padilla, A. (2020a). *Twitter for Academic Research*. Disponible en: <https://github.com/AdriaPadilla/Twitter-API-V2-full-archive-Search-academics>
- Padilla, A. (2020b). *Instaloder Scripts*. Disponible en: <https://github.com/AdriaPadilla/InstaloderScripts>
- Perriam, J., Birkbak, A. y Freeman, A. (2020). Digital methods in a post-API environment. *International Journal of Social Research Methodology*, 23(3): 277–290.
- Poell, T., Abdulla, R., Rieder, B., Woltering, R., y Zack, L. (2016). Protest leadership in the age of social media. *Information, Communication & Society*, 19(7): 994-1014.
- Raymond, E.S. (1998). *The cathedral and the bazaar*. *FirstMonday*. <https://firstmonday.org/ojs/index.php/fm/article/download/1472/1387?inline=1>
- Rieder, B. (2013). "Studying Facebook via data extraction: the Netvizz application". En *Proceedings of the 5th annual ACM web science conference* (pp. 346–355).
- Rieder, B. (2015). *YouTube Data Tools*. Disponible en: <https://tools.digitalmethods.net/netvizz/youtube/>.
- Rieder, B. (2017). *RanFlow*. Disponible en: <https://labs.polsys.net/tools/rankflow/>.
- Rieder, B., Coromina, Ò. y Matamoros-Fernández, A. (2020). Mapping YouTube: A quantitative exploration of a platformed media system. *First Monday*, 25(8). <https://doi.org/10.5210/fm.v25i8.10667>
- Rogers, R. (2019) *Doing Digital Methods*. Sage Publishing.
- Venturini, T., Munk, A. y Meunier, A. (2018). "Data-sprinting: A public approach to digital research". En C. Lury, R. Fensham y A. Heller-Nichola (Eds.) *Routledge Handbook of Interdisciplinary Research Methods*. Routledge.



El método *data sprint*: un ejercicio de reflexividad feminista sobre las prácticas de producción de conocimiento

The data sprint approach: an feminist reflexivity exercise on the practices of knowledge production

id **María Concepción Castillo-González**
ccastill@tec.mx
Tecnológico de Monterrey

Recibido: 13/01/2022
Aceptado: 24/05/2022

id **Dorismilda Flores-Márquez**
dfloresm@delasalle.edu.mx
Universidad De La Salle Bajío

id **Gabriela Elisa Sued**
gabriela.sued@gmail.com
Universidad Nacional Autónoma de México, Instituto De Investigaciones Sociales

RESUMEN

Recientemente se ha popularizado el método *data sprint*: talleres intensivos donde convergen participantes de diversos ámbitos para trabajar colaborativamente en el análisis de datos con propósitos que oscilan entre la lógica abierta vinculada con la ética *hacker* y la lógica de mercado orientada a la productividad. Este artículo reflexiona sobre este método de trabajo a partir de experiencias de investigadoras que participaron activamente en un *data sprint* de una semana de duración, con el propósito de potenciar la reflexión crítica sobre la metodología en la producción de conocimiento. Esta exploración se fundamenta teóricamente en dos elementos clave: el *data sprint* como método y la reflexividad en la práctica científica que incluye la perspectiva feminista de la tecnología. Para ello se recuperaron las experiencias de siete participantes mediante un cuestionario con preguntas abiertas. Con los datos se realizó un análisis narrativo temático. Los resultados obtenidos se estructuraron en cuatro categorías: las fases de trabajo, las dificultades, las experiencias positivas y la perspectiva de género. Se destacan aspectos positivos, negativos y a mejorar. Entre los primeros se encuentran los aprendizajes metodológicos y técnicos, la posibilidad de generar producción académica y especialmente, aspectos relacionados con los vínculos y el sentido de colectividad entre las participantes. Entre los segundos, las dificultades generadas por la velocidad y los cambios de horario ocasionados por el formato híbrido. Entre los aspectos a mejorar se encuentra la consideración de las desigualdades presentes en la práctica científica para el planteamiento de diseños alternativos que promuevan apropiaciones tecnológicas equitativas.

PALABRAS CLAVE

data sprint, metodología, experiencias, reflexividad, feminismo

Cómo citar este artículo:

Castillo-González, M. C.; Flores-Márquez, D. y Sued, G. E. (2022). El método *data sprint*: un ejercicio de reflexividad feminista sobre las prácticas de producción de conocimiento. *Dígitos. Revista de Comunicación Digital*, 8: 67-86. DOI: 10.7203/drdcd.v1i8.235

ABSTRACT

The data sprint method, defined as workshops where participants converge to work together on data analysis for different purposes, has recently become popular. Data sprint work ranges from open-source logic linked to hacker ethics, to market logic oriented towards productivity. This article reflects critically on the data sprint as a method through the experiences of researchers who participated in the Smart Data Sprint of iNova Media Lab in 2021. This exploration is theoretically based on two key elements: the data sprint as a method and reflexivity in scientific practice, which includes a feminist perspective on technology. For this purpose, the experiences of seven participants were recovered through a questionnaire with open-ended questions. Data were examined through thematic narrative analysis. The findings were structured in four categories: work phases, difficulties, positive experiences, and gender perspective. Positive, negative, and improvement aspects are highlighted. The first includes methodological and technical learning, the possibility of generating academic production, and aspects related to the participants' links and sense of community. The latter involves the difficulties generated by speed and the schedule changes caused by the hybrid format. Among the aspects to improve, it is necessary to consider the inequalities present in scientific practice to configure alternative designs that promote equitable technological appropriations.

KEYWORDS

data sprint, methodology, experiences, reflexivity, feminism



El método *data sprint*: un ejercicio de reflexividad feminista sobre las prácticas de producción de conocimiento

Una práctica científica que omite cuestionarse a sí misma no sabe, en realidad, lo que está haciendo.

Pierre Bourdieu

1. Introducción

El modelo de investigación de los *data sprints*, definido como talleres donde participantes de diversos ámbitos del conocimiento trabajan colaborativamente a partir de conjuntos de datos y preguntas de investigación (Venturini *et al.*, 2018), ha adquirido relevancia en el campo de los *digital methods*, los cuales se contraponen a los métodos virtuales que “digitalizan los métodos socio-científicos existentes y los trasladan a la web, adaptándolos al medio” (Rogers y Lewthwaite, 2019:14). Los métodos digitales no buscan estudiar la cultura de los medios, sino que se plantean realizar investigaciones sociales con la Web y para ello desarrollan técnicas especializadas. Procesamiento informático de set de datos y metadatos, diseños de claves de búsqueda de información orientados a la investigación social, tratamiento de tendencias y reacciones como “likes” y “compartidos”, son ejemplos de estas técnicas.

Estas aproximaciones son parte del giro computacional de los estudios de Internet (Rogers, 2019). Los métodos digitales consideran a la Web como un archivo listo para ser explorado y a sus datos como sofisticadas fuentes de información sociocultural.

Este cambio de paradigma estimula la imaginación sociológica y da lugar a preguntas de investigación que pueden responderse con técnicas nativas que escudriñan cantidades robustas, pero manejables, de información multimedia y de metadatos organizadas en capas y en redes complejas (Omena *et al.*, 2021).

El *data sprint* como método tiene su origen en el año 2007 en el Digital Methods Initiative (DMI) de la Universidad de Ámsterdam (Laursen, 2017; Digital Methods Summer School Program, 2022). Desde sus inicios, la propuesta se planteó como una alternativa a los estudios de *Big Data*, a partir de un modelo de trabajo experimental *mezzo-level* (Berry *et al.*, 2015). Una de las características emblemáticas del modelo es la conformación de equipos con base en competencias en lugar de credenciales académicas. Este rasgo revela la filosofía pragmática y la visión sistémica de la producción del conocimiento que plantean.

Además de las ventajas que se obtienen en términos de experimentación, originalidad, invención y colaboración (Laursen, 2017) el modelo de trabajo acelerado es afín al paradigma ideológico del capitalismo tardío, que prioriza la productividad y la eficiencia. El modelo *sprint* no inicia en el campo de la producción científica, sino de la innovación tecnológica, motivada por el interés industrial de reducir los tiempos dedicados al proceso de diseño de nuevos productos (Suss *et al.*, 2011), así como por la inquietud de grupos de programadores por desarrollar software de forma colaborativa.

Respecto al *data sprint* como método de investigación académica, la literatura hace referencia a sus vínculos con los *booksprints*, *barcamps* (Berry, *et al.*, 2015) y *hackatones* (Venturini *et al.*, 2018), modelos colaborativos de trabajo acelerado que culminan en el diseño de prototipos para convertirse en productos y/o servicios innovadores sin fines de lucro (Rogers y Lewthwaite, 2019).

Los *hackatones* provienen de la cultura de software libre afín a la defensa de los comunes de Internet, de ahí que han sido utilizados por movimientos sociales como el feminismo (D'Ignazio *et al.*, 2020), aunque actualmente se han abierto a muchos temas y el sector privado los utiliza como método para fomentar la innovación acelerada con fines de lucro, con lo cual se va perdiendo su sentido original.

Esta tensión entre el paradigma del mercado y el de la cultura de software libre no exime a las prácticas investigativas. El poder y la opresión también las cruzan (D'Ignazio *et al.*, 2020). Los *data sprint* pueden ser permeados por lógicas competitivas, extractivistas y excluyentes, de ahí que convenga problematizar estas innovadoras metodologías. Existen trabajos que han examinado los *hackatones* con perspectiva feminista señalando asimetrías y alternativas (Hardin, 2021; D'Ignazio *et al.*, 2020), pero los *data sprint* no han sido analizados con esta perspectiva.

En esa línea y desde el conocimiento situado que propone el feminismo (Haraway, 1988), nuestras preguntas son: ¿Cómo es la experiencia de las mujeres participantes en la lógica del *data sprint*?, ¿cómo se compatibiliza y confronta la perspectiva de género con los procesos de producción de los *data sprint*? Más allá de la producción de datos y conocimiento que se da en estos encuentros en torno a las temáticas específicas,

se pueden revelar otros tipos de aprendizaje, además de vivenciar las tensiones que existen entre los paradigmas que enmarcan el desarrollo de estas metodologías.

El objetivo de este artículo es reflexionar críticamente y desde una perspectiva feminista y situada en el Sur Global (Milán & Treré, 2019) acerca del método *data sprint* en la producción de conocimiento, desde la experiencia de quienes participan. Para ello, se recuperaron las reflexiones de siete investigadoras que participaron en el *Smart Data Sprint* de iNova Media Lab en 2021 con el proyecto *#vivasylibresnosqueremos: Feminist digital practices and public debates in Latin America*¹, a través de un cuestionario con preguntas abiertas que les permitieron articular cavilaciones en torno a cinco ejes: los sentidos de participar, las fases de trabajo, las experiencias, el balance de ellas y las implicaciones de género.

2. Genealogía de los *data sprint*

En la literatura académica hay pocas referencias al *data sprint* como modelo acelerado de indagación digital que se ejecuta de forma colaborativa a partir de preguntas de investigación (Laursen, 2017; SMART Social Media Research Techniques, 2018). Esta poca incidencia de los estudios invita a aportar estudios originales al campo. Los artículos publicados se agrupan en dos líneas: aquellos que presentan investigaciones generadas en un *data sprint* y aquellos que sistematizan y fundamentan teóricamente el modelo de trabajo en el campo de los *digital methods* (Rogers, 2019).

Los primeros describen innovaciones metodológicas que incorporan técnicas nativas para la colección y el análisis de datos, desarrollados en uno o más *data sprint*. Así, Berry *et al.* (2015), documentan los resultados de un *data sprint* realizado en 2013, para mapear los campos de las Humanidades Digitales y Literatura Electrónica a partir del análisis de las recomendaciones de tiendas Amazon de diferentes países. Omena *et al.* (2020) desarrollaron, en dos *data sprint* (2018 y 2019), un método original para explorar *hashtags* en su dimensión sociotécnica², trascendiendo el tradicional criterio de selección de corpus (Omena *et al.*, 2020). Finalmente, Huurdeman *et al.* (2013) muestran el potencial del método para explorar y recuperar información relevante archivada en la Web, al analizar un sitio de noticias holandés. En suma, estos trabajos tienen en común el uso de herramientas digitales nativas y el diseño de originales protocolos de investigación que se construyen a partir de la exploración de datos y del análisis de visualizaciones.

Por otro lado, existen trabajos que fundamentan teóricamente los *data sprint* en el campo de los métodos digitales. Entre ellos se encuentran el de Venturini *et al.* (2018),

1. Las siete investigadoras fueron parte del equipo original que diseñó el proyecto *#vivasylibresnosqueremos: Feminist digital practices and public debates in Latin America* en respuesta a la convocatoria lanzada por iNova Media Lab 2021.

2. Las autoras consideran a los *hashtags* no sólo como marcadores temáticos a partir de los cuales es posible recolectar un corpus de datos, sino como objetos socioculturales. *"We approach hashtags as sociotechnical formations that serve social media research not only as criteria in corpus selection but also displaying the complexity of the online engagement and its entanglement with the technicity of web platforms"* (Omena *et al.*, 2020:1)

quienes describen el origen, las características y las fases de esta metodología; los textos y las entrevistas realizadas a Rogers, que permiten comprender la genealogía y características del modelo (Rogers, 2019; Rogers y Lewthwaite, 2019) y el trabajo de Laursen (2017) que sistematiza el método a partir de entrevistas realizadas a cuatro directores de Laboratorios de Datos de Copenhague.

Se identifica un vacío de conocimiento en torno a la experiencia de quienes participan en un *data sprint*, con la intención de observar si se cumplen los propósitos planteados y de conocer cuáles son las principales aportaciones, retos y posibilidades del modelo de trabajo. Además identificamos la oportunidad de revisar dicha metodología desde una perspectiva feminista, como ocurre en el caso del hackatón (D'Ignazio *et al*, 2020).

Cabe señalar que la literatura antes expuesta se basa en laboratorios en su mayoría europeos ya que en Latinoamérica aún no existen experiencias consolidadas en torno a los *data sprint*. Por consiguiente, tampoco existe suficiente literatura de referencia. Sin embargo, en los últimos años, participantes latinoamericanos han ganado presencia en los *data sprint* europeos, por lo que identificar sus experiencias resulta relevante.

Este caso de estudio contempla la experiencia de un equipo de investigadoras latinoamericanas que participaron en *data sprint* en una universidad europea. Por motivos de la pandemia de la COVID-19, la participación del equipo se llevó a cabo vía Zoom.

3. Marco teórico

En este apartado se relacionan dos elementos clave: el *data sprint* como método que implica distintas fases y actividades y la reflexividad en la práctica científica en tanto proceso que permite recuperar las experiencias para producir conocimiento sobre la producción de conocimiento. Se incorpora la perspectiva feminista en la reflexividad, ya que permite situar el conocimiento e identificar los sistemas de opresión presentes en la producción científica.

3.1 El *data sprint* como método

Venturini *et al.* (2018) afirman que el *data sprint* es una adaptación del modelo del *hackatón* en la investigación académica. Se retoma la idea de un equipo diverso de expertos reunido físicamente para trabajar exhaustivamente en un periodo breve en torno a un reto coordinado por uno o más investigadores que establecen un tema de trabajo y sus preguntas de investigación. El propósito de los hackatones es principalmente producir software o hardware. El resultado es un prototipo, no un proyecto terminado. Los autores señalan que la diferencia estriba en que el *data sprint* tiene un proceso más largo y estructurado; además, no se trata de un evento abierto, ya que los equipos deben tener las competencias necesarias para llegar a la meta propuesta.

Los autores identifican seis fases dentro del modelo *data sprint*. Previo al evento: 1) se diseñan y 2) operacionalizan las preguntas de investigación luego, 3) se reúnen y depuran los datos digitales. Durante el *sprint*, el equipo 4) adapta las herramientas y el

código de programación existente para dar respuesta a las preguntas establecidas³ y 5) diseñan visualizaciones e interfaces que ayudan a realizar interpretaciones. Al final, 6) se genera el compromiso con la co-producción del conocimiento, pues el equipo deberá seguir trabajando para afinar y complementar la información y así fundamentar científicamente el trabajo.

Laursen (2017) plantea que, a diferencia del *hackatón*, orientado a concluir con un producto específico, el *data sprint* requiere de mayor atención en el proceso porque se desarrolla en un ambiente académico. Janna J. Omena, directora del iNova Media Lab, coincide con esta apreciación: "*para mí, el proceso es lo más importante, el intercambio de ideas, el intercambio de conocimientos y el aprendizaje mutuo*" (Omena, 2018). Por su parte, Richard Rogers, fundador del Digital Methods Initiative, destaca el potencial pedagógico que tiene este método en su práctica docente: "...trabajamos en *sprints* de datos, en los que los alumnos recolectan datos, utilizan herramientas, obtienen resultados, crean visualizaciones y escriben informes, todo dentro de un plazo muy ajustado" (Rogers y Lewthwaite, 2019:20).

En este sentido, el reto planteado por el *data sprint* funciona como una motivación para el trabajo colaborativo, el aprendizaje y la práctica experimental conjunta (Laursen, 2017). El resultado no es un producto finalizado en términos científicos, sino una serie de bocetos y protocolos de investigación.

La estructura flexible de estos métodos, su potencial pedagógico y sus orígenes vinculados con la cultura "*open source, copy left and open publishing licenses*" (Venturini *et al.*, 2018:5) pretenden favorecer la inclusión social. Sin embargo, Caroline Hardin (2021) revela, a partir de un estudio con perspectiva de género sobre la experiencia de participantes de hackatones, que estos espacios siguen siendo excluyentes para las mujeres. Hardin confronta esta marginación con el resto de sus hallazgos que exponen a los hackatones como lugares que abren oportunidades para el empleo, generan redes de colaboración y desarrollan competencias. En ese sentido D'Ignazio, *et al.* (2020) cuestionan la estructura y los supuestos epistemológicos de estas metodologías y destacan el trabajo del *Feminist HCI* en el rediseño de hackatones inclusivos y orientados a la justicia social. La perspectiva feminista de la interacción humano-computadora (HCI) reconoce explícitamente que este tipo de exclusiones responde a un histórico vínculo entre la ciencia y la tecnología con el patriarcado del norte global que es el epicentro del capitalismo (D'Ignazio and Klein, 2020).

En el caso específico del *data sprint* no existen cavilaciones de corte feminista en la literatura, pero Laursen (2017) enfatiza la necesidad de reflexionar sobre estos métodos. Considera que su emergencia podría vincularse con el menosprecio que existe hacia las humanidades y se pregunta sobre el tipo de ciencia que se performa con estos modelos de trabajo acelerado, que rompen fronteras disciplinares, que requieren de

3. De acuerdo con los autores, los proyectos que se llevan a cabo en un *data sprint* tienen objetivos tan específicos que no suelen plantear el desarrollo de software genérico. Sin embargo, las raíces comunitarias del *data sprint*, deben asegurar que los datos, contenidos y códigos de programación generados queden abiertos para que puedan ser reutilizados y adaptados por la comunidad académica.

competencias específicas y que invitan a pensar a partir de datos. Para Berry *et al.* (2015), el *data sprint* requiere de una determinación agresiva para producir un proyecto de investigación creativo en un periodo de tiempo limitado. Estas características deberían estudiarse desde la perspectiva situada de las personas participantes para identificar si estos espacios favorecen la inclusión social y los principios pedagógicos y comunitarios planteados. En esta misma lógica, es necesario detonar procesos de reflexividad feminista acerca de la teoría metodológica y epistemológica que sustenta al *data sprint*.

3.2 La reflexividad feminista en las prácticas científicas

En este artículo abordamos la experiencia de participación en los *data sprints* desde la reflexividad en la práctica científica. Margaret Archer (2007) define la reflexividad como la capacidad y el ejercicio de las personas de mantener conversaciones internas cruciales para los modos en que hacemos nuestro camino en el mundo. Estos procesos se dan hacia adentro, pero implican también la posibilidad de entenderse en relación con los contextos sociales, de modo que se relacionan con la posición que asumimos en la sociedad.

Entendemos la investigación como una serie de prácticas, la reflexividad es inherente a ella: “mediante la reflexividad se pone en duda, se evalúa, se reelaboran las herramientas que garantizarán la confiabilidad y validez en los datos que se producen” (Sordini, 2021:4). Si bien la reflexividad se ha abordado principalmente en torno a los enfoques cualitativos (Corlett & Mavin, 2018; Soedirgo & Glas, 2020; Sordini, 2021), todas las prácticas de investigación implican motivaciones, posiciones y modos de hacer que debieran ponerse en discusión.

Sandra Corlett y Sharon Mavin (2018) entienden la reflexividad en la práctica científica como un auto-monitoreo y auto-respuesta a nuestros pensamientos, sentimientos y acciones, cuando nos involucramos en proyectos de investigación. Esto requiere reconocer nuestro posicionamiento en relación con las preguntas que nos hacemos.

La reflexividad metodológica reconoce que las y los investigadores tomamos decisiones y que los métodos que empleamos no son neutros. La autorreflexividad permite plantear preguntas sobre las motivaciones, los roles, la identidad y las relaciones de poder que se dan entre investigadores (Corlett & Mavin, 2018).

Jessica Soedirgo y Aarie Glas (2020) proponen hablar de reflexividad activa, que implica: reconocer nuestro posicionamiento, el modo en que es leído por otros y los supuestos sobre nuestras conclusiones. Las autoras sostienen que el posicionamiento comprende varios elementos relacionados que llaman características demográficas —tales como raza, edad, género y clase—, las experiencias personales y profesionales, las posturas políticas e ideológicas, así como otros elementos de la biografía social y los mundos de vida.

Incorporar la perspectiva feminista en la reflexividad implica tomar en cuenta que estas características están cruzadas por sistemas de opresión que dan lugar a desigualdades

sociales que afectan la producción, uso y reapropiación de la producción de conocimiento científico y tecnológico. Estas desigualdades han sido estudiadas desde la sociología de la ciencia y la tecnología. Wacjman (2007) destaca que la tecnología se encuentra marcada por el género masculino. Maffia (2007) señala que las mujeres han sido históricamente expulsadas de los ámbitos de producción de conocimiento para evitar su emancipación. Por su parte, D'Ignazio y Klein (2020) reflexionan desde una perspectiva interseccional acerca de los privilegios y opresiones en la producción de la ciencia. Esto revela las múltiples fuerzas que operan estructural y culturalmente para excluir a ciertas subjetividades y grupos sociales de las prácticas científicas y tecnológicas, las cuales se han desarrollado por una élite de "*straight, white, able-bodied, cisgender men from the Global North*" (p.8).

La reflexión sobre las aportaciones del feminismo al desarrollo de la ciencia y la tecnología abarca el reconocimiento de teorías, principios y formas de interacción, que son considerados en el diseño de métodos que intentan desarticular las exclusiones. Bardzell (2010) sostiene la necesidad de incorporar la epistemología feminista en el diseño de la ciencia y la tecnología por su compromiso con la agencia, la igualdad y la justicia social. En esa línea, la "*feminist standpoint theory*" reconoce que el conocimiento está socialmente situado en una matriz de poder (D'Ignazio & Klein, 2020) arraigada en sociedades patriarcales capitalistas y coloniales. De acuerdo con esta teoría, es importante revalorar las experiencias de las personas que son frecuentemente excluidas, pues su propio punto de vista revela y, a la vez, desarticula los sistemas que mantienen la desigualdad.

A partir de estas reflexiones se ha destacado el potencial de las tecnologías para mantener unidos a quienes son sistemáticamente excluidos, para que puedan apropiarse de los espacios que tienden a expulsarles (Plant, 1998; Haraway, 1988 y ; D'Ignazio & Klein, 2020). Una reapropiación tecnológica equitativa implica el reconocimiento de las ambivalencias y la construcción de otras formas de relación. Investigadoras como Bradzell (2010), D'Ignazio y Klein (2020) plantean principios que recuperan la interacción feminista en el diseño metodológico, tales como: la examinación y el desafío del poder, la reflexividad, la contextualización, el conocimiento situado y encarnado, entre otros. Por su parte, Leurs (2017), Linabary y Corple (2019) señalan que la reflexividad feminista profundiza la dimensión ética de las investigaciones, pues desafía los sistemas de opresión a partir de un enfoque interseccional. Además asume el compromiso de responder a la desigualdad desde una ética del cuidado que considera que los vínculos son valores que prevalecen sobre ideas abstractas.

D'Ignazio y Klein (2020) afirman que la toma de conciencia feminista tiene mucho que ofrecer al diseño de los *sprints* pues se pueden considerar sus principios dentro de su diseño para priorizar el proceso desde una ética del cuidado que considere el contexto, refuerce las relaciones, rompa jerarquías y cuestione el reparto inequitativo de tareas.

4. Metodología

Para responder a las preguntas y objetivo de la investigación, llevamos a cabo un estudio de caso en torno a la experiencia de 7 investigadoras que participaron en el *Smart Data Sprint* de iNova Media Lab en 2021. La exploración se enfocó en la reflexividad activa y con perspectiva de género, al considerar el punto de vista de las investigadoras para dar cuenta de los procesos desiguales de producción de conocimiento, en este caso, en la dinámica del *data sprint*.

Para ello, elaboramos un cuestionario con preguntas abiertas. Si bien los cuestionarios suelen tener connotaciones cuantitativas, en este trabajo lo usamos como una guía para que las participantes tuvieran oportunidad de narrar sus experiencias de forma anónima, a partir de ejes clave que orientaron la reflexión: 1) los sentidos de participar en un *data sprint*, 2) las fases de trabajo experimentadas, 3) las experiencias, 4) el balance de ellas, y 5) las implicaciones de género en los *data labs*.

Se envió el cuestionario a las 14 mujeres participantes del proyecto. Se recibieron siete respuestas. Con los datos hicimos un análisis narrativo temático (Riessman, 2008) que, por un lado, nos permitió recuperar elementos clave de los datos para crear categorías a partir de las experiencias narradas y, por otro lado, abrió la posibilidad para mantener la "historia" contada por las participantes, en términos de fases. En esta exploración se reconoce el carácter colectivo de los datos, puesto que las participantes aportaron diferentes miradas sobre una experiencia en común.

Las categorías y ejes discursivos fueron identificados en *Taguette* (Rampin y Rampin, 2021), una herramienta de software libre para el análisis cualitativo de datos.

5. Resultados

El análisis del cuestionario abierto permitió la creación de categorías a partir de las narraciones aportadas por las participantes. En esta sección recuperamos las cuatro principales que permiten reconstruir y articular las experiencias: 1) Las fases de trabajo en las cuales se articulan aprendizajes y posibilidades, 2) las dificultades y tensiones, 3) las experiencias positivas y 4) la reflexión con perspectiva de género.

5.1 Las fases de trabajo en el *data sprint*: aprendizajes y posibilidades

En coincidencia con Venturini *et al.* (2018), las participantes identifican tres fases de trabajo: una previa al evento, otra que se realiza durante el mismo y una más que se da después de este (figura 1). En la narración de estas fases se articulan los aprendizajes y las posibilidades identificadas.

En la fase previa, además de los elementos señalados por Venturini *et al.* (2018), las participantes enfatizan dos bloques de acciones: la integración del equipo de trabajo, que resulta fundamental para el diseño y ejecución del proyecto, a partir de la reunión con colegas con diferentes formaciones y áreas de *expertise*, y la preparación para los talleres, de carácter individual, que consistió en actividades como descargar y/o actualizar *software*, revisar tutoriales y experimentar con los recursos.



Figura 1: Fases de trabajo en el *data sprint*. | Fuente: Elaboración propia, con datos de Venturini et al. (2018) y datos del proyecto.

La fase del *lab* o *data sprint* es la más reconocida. Esta consiste en los días de trabajo intensivo y colectivo. Si bien lo central de esta fase se encuentra en los proyectos de los equipos, como plantean Venturini et al. (2018), también hay otras actividades, como conferencias y talleres que son valorados entre las integrantes.

Los primeros días fueron muy intensos, dado que las charlas y workshops se presentaban uno detrás del otro [...] disfruté mucho de esa etapa porque me permitió reconectar con las metodologías que luego tuve que poner en práctica en la etapa de investigación (participante 1) ⁴.

La mayoría de los aprendizajes narrados por las participantes (figura 2, color azul) se localizan en esta fase. Algunas enfatizaron los aprendizajes técnicos, en torno al uso y utilidad de herramientas digitales, otras los metodológicos, como la posibilidad de hacer análisis *cross-platform* (Rogers, 2017; Pearce et al., 2018), la importancia de una metodología sólida, las decisiones metodológicas, la experiencia de diseño y análisis colectivo.

El segundo [aprendizaje] es más técnico y refiere a herramientas que circularon en el lab y que desconocía. Todas han sido útiles y he de destacar su funcionamiento y limpieza. Todas pudieron ser reutilizadas cuando el lab terminó para mi trabajo académico (participante 3).

La experiencia significó una ventana a un mundo de metodologías basadas en datos, así como al ambiente que viven lxs investigadorxs que se dedican a ello. No es que no lo supiera antes, pero como experiencia inmersiva me permitió sentir el ambiente, las temporalidades, las formas de ver el mundo, de significarlo, de aproximarse al estudio de los fenómenos sociales (participante 2).

Dentro de la misma fase de *lab*, las participantes reconocen dos partes en el desarrollo

4. Con el fin de mantener el anonimato en el equipo, nuestro cuestionario no incluyó preguntas para caracterizar a las participantes. Por esta razón, en los testimonios empleamos "participante 1", "participante 2" y así sucesivamente.

del proyecto: en primer lugar, el trabajo con los datos y las visualizaciones; en segundo lugar, la preparación de la presentación final que se expuso el último día del *data sprint*.

La última fase *-post lab-* es una extensión del proyecto sin límite de tiempo. De acuerdo con Venturini *et al.* (2018), consiste en el compromiso con la co-producción del conocimiento. Curiosamente, es la fase menos reconocida explícitamente por las participantes, pero las actividades realizadas y/o por realizar fueron mencionadas al hablar de posibilidades (figura 2, color verde). En esta fase se ubica la elaboración del reporte del proyecto, que se entregó mes y medio después del *data sprint*; además, se considera la preparación de publicaciones que se desprenden del trabajo ejecutado en esos días, sin limitarse a él.

Finalmente todas las participantes coincidieron en que el *data sprint* les abre múltiples oportunidades ya sea en su práctica docente, en sus publicaciones científicas o en la exploración y manejo de nuevas herramientas. Estas posibilidades encuadran a los *data sprint* en la aplicación del conocimiento adquirido en el ámbito académico.

5.2 Experiencias negativas y dificultades

Las dificultades que las participantes enunciaron (figura 2, color gris) se enmarcaron como experiencias negativas. Pueden agruparse en dos categorías: las vinculadas con la participación virtual provocada por la pandemia Covid 19 y las que son las inherentes al modelo *data sprint*.

La experiencia de participar vía Zoom con un equipo de trabajo conformado por miembros de diversos países que no compartían el mismo huso horario, complicó la asistencia y la participación activa en los talleres ofrecidos por los organizadores, además dificultó la logística del trabajo del equipo.

Mi mayor dificultad era estar en línea todo el tiempo, porque tenía problemas con mis conexiones a Internet (participante 7).

La experiencia fue un poco frustrante. Algunas cosas naturales por el horario tan distinto entre México y Europa (participante 6).

La mayor parte de las experiencias negativas narradas se vinculan con las características mismas del modelo. Las competencias que se deben tener y desplegar a gran velocidad, la fragmentación de las fases y los diferentes niveles de conocimiento técnico existentes entre los miembros del equipo. Estas características pueden fomentar la percepción de que el modelo acelerado de trabajo no reconoce las relaciones de poder y las desigualdades sociales existentes.

Una de las dificultades era tratar de guiar a otras personas que parecían estar perdidas o no manejaban las herramientas (participante 2).

Lamentablemente, todavía existe una brecha grande en el nivel de producción en data-driven-research entre América Latina y Europa que cuesta mucho reducir. Si pudiera cambiar algo, sería tener un poco más de tiempo para analizar datos (participante 3).

No me gustó quedarme con una actividad solitaria. Entiendo esto por la velocidad del trabajo, pero se puede mejorar, sobre todo si apostamos realmente por la colaboración y la sororidad (participante 4).

Llama la atención que uno de los comentarios más repetido hizo alusión a la frustración, al sentimiento de insuficiencia y a la falta de agencia para participar activamente. Esto revela que un modelo basado en competencias, que apuesta por la velocidad en torno a un reto, puede resultar violento para personas que no tienen determinadas habilidades. Lo cual coincide con lo dicho por Berry *et al.* (2015), respecto a la inherente agresividad que requiere el despliegue de un *Lab*.

La experiencia fue interesante, también estresante. No solo por el horario, los lenguajes (inglés y computacional) y la rapidez. En ocasiones me sentí profundamente insuficiente para este mundo de análisis de datos (participante 5).

Aunque se dice que es un taller para principiantes o para todo mundo, no lo es. Creo que se requiere que llegues al menos con experiencia en scraping en diferentes plataformas (participante 6).

Las respuestas de las participantes materializan la tensión entre aceleración y aprendizajes propios del capitalismo informacional (Berry *et al.*, 2015) así como al diseño poco accesible que suelen tener estas metodologías (D'Ignazio *et al.*, 2020). En este sentido, puede decirse que estos valores se encuentran anidados en el método.

5.3 Experiencias positivas

La valoración positiva fue una de las categorías más relevantes en las narraciones de las participantes (color naranja, figura 2) y, en coincidencia con Laursen (2017) y Omena (2018), estuvo vinculada con la experiencia de trabajo colectivo.

Algunas participantes enfatizaron que la colaboración les permitió tanto aprender de las otras como enseñarles algo en función de las propias áreas de *expertise* y destacaron como experiencia positiva, las prácticas de apoyo y cuidado mutuo que emergieron en el equipo y que ayudaron a sortear las dificultades.

La mayor satisfacción fue volver a ponerme en contacto con investigadoras y profundizar con ellas mis conocimientos de métodos digitales (participante 1).

Los resolví [problemas de conexión] gracias a la solidaridad de mi equipo (participante 7).

Trabajar juntas en torno al feminismo latinoamericano impactó en la motivación y fortaleció los vínculos afectivos entre las participantes y con las las organizadoras.

Muy interesante [...] también por la posibilidad de participar en un proyecto significativo con un grupo de mujeres que estimo (participante 3).

Especialmente las participantes expresaron interés por las nuevas experiencias que pueden tener como grupo en el futuro y por el significado positivo que tiene para ellas el vínculo creado a partir de esta vivencia.

En otros laboratorios no lo he visto de igual manera [perspectiva de género]. Había pocas mujeres pitchers, y era más difícil que las mujeres trabajáramos en equipos con hombres para producir conocimiento. Los oradores tendían a ser hombres también (participante 3).

Definitely, I can notice that [gender perspective] in particular with all the support my team had from the organizers and also of the selected topics and the facilities to work on them (participante 7).

Sin embargo, se destacó cierta división del trabajo en la presentación de los talleres: mientras las mujeres enseñaron el manejo de herramientas creadas por terceros, los hombres presentaron desarrollos de software para el análisis de datos. Estas lecturas pueden orientar en un futuro acciones afirmativas orientadas a la equidad.

...en la constitución de equipos, en políticas de inscripción que lo favorezcan, en la demanda o recomendación de introducir perspectiva de género en todas las investigaciones, en talleres sobre perspectiva de género en métodos digitales dentro de la experiencia (participante 6).

Las participantes mostraron interés en proyectos de temática feminista, que se ven potenciados, por un lado, por el feminismo de datos propuesto por D'Ignazio y Klein (2020), y por el otro, por las importantes acciones de resistencia y visibilidad que el feminismo realiza habitualmente en las plataformas sociales (Sued et al., 2021).

Por último, se destacó también que la velocidad del proceso, la competitividad entre equipos, el manejo del idioma inglés y el requerimiento de resultados rápidos presenta una tensión con proyectos de índole inclusiva donde las mujeres del sur global tienen muchas brechas por superar para conseguir una apropiación tecnológica.

6. Discusión y conclusiones

Este artículo abordó el método *data sprint* a partir de las experiencias de un grupo de investigadoras que participaron en el *Smart Data Sprint* de iNova Media Lab en 2021. La finalidad de este trabajo fue potenciar la reflexión crítica, desde la perspectiva situada e interseccional que aporta el feminismo, sobre la metodología en la producción de conocimiento, de modo que sea también objeto de indagación y, en esa línea, de producción de conocimiento. Para ello se abordó teóricamente el *data sprint* como método y la reflexividad activa y con perspectiva de género en la práctica científica. Esto se tradujo metodológicamente en la recuperación de experiencias, a través de un cuestionario con preguntas abiertas, con cuyas respuestas se hizo un análisis narrativo temático. Los resultados se presentaron en cuatro categorías —las fases de trabajo donde se articularon aprendizajes y posibilidades, las dificultades, las experiencias positivas y la reflexión con perspectiva de género—, lo anterior permite problematizar y comprender los desafíos y las oportunidades de estas metodologías.

Los resultados de esta investigación no pueden generalizarse porque parten de un caso. Su principal aportación es la de recuperar la experiencia y la reflexividad de las participantes, situadas en el género femenino y en el Sur Global. Esta aproximación no

había sido abordada en análisis previos de *data sprints* y abre posibilidades para nuevas investigaciones, a la vez que ofrece un mapa de ruta para el diseño alternativo de experiencias más inclusivas y más sensibles al poder (Leurs, 2017) pues la producción científica requiere ser enriquecida por las aportaciones de quienes han permanecido en los márgenes.

En relación con la experiencia de las participantes se identifican tres fases en coincidencia con Venturini *et al.* (2018). Aunque hay diferencias menores en las actividades que se consideran dentro de cada fase, hay mucha claridad en la relevancia de cada una. El método *data sprint* se concentra en el evento, con gran cantidad de actividades, tales como conferencias y talleres, así como el desarrollo colectivo del proyecto.

Las participantes coinciden en que la experiencia de estas fases de trabajo, les deja aprendizajes metodológicos y herramientas para el procesamiento de datos, para nuevas publicaciones y para la inclusión de técnicas digitales en la docencia, lo anterior encuadra las posibilidades de los *data sprint* en el ámbito académico.

Algunas de las dificultades experimentadas se relacionan con la coyuntura del Covid-19. Si bien, se trata de un caso excepcional que no forma parte del modelo original que es presencial (Venturini *et al.*, 2018) es importante considerar el diseño de experiencias no-presenciales más flexibles que respondan, desde la ética del cuidado, al contexto de las participantes.

Se considera que la velocidad, lenguajes y competencias técnicas requeridas en un *data sprint* limitan la experiencia y comprometen los resultados. Vale la pena discutir si la lógica de trabajo, basada en la rapidez y la productividad, contribuye a desarrollar procesos reflexivos de creación de conocimiento encuadrados en una ética del cuidado que considere las inequidades sociales, la protección de la privacidad en el uso de datos, y la construcción de vínculos equitativos (Linabary & Corple, 2019).

En este sentido, los resultados problematizan la relevancia que la literatura señala acerca de la dimensión pedagógica del proceso (Laurson, 2017). Para varias participantes, el modelo acelerado se convirtió en un obstáculo para su aprendizaje y en un factor que desmotivó su interés en apropiarse de los métodos digitales. Se materializan en esta experiencia las desigualdades sociales y las tensiones entre la aceleración de los tiempos de producción y los procesos de aprendizaje que requieren de mayor elaboración. Estos resultados coinciden con los hallazgos que obtuvo Carolin Hardin (2021) en el estudio de la participación en hackatones, en el cual, únicamente las mujeres reportaron sentirse intimidadas y revelaron preocupaciones vinculadas con el cuidado de sí y de los otros.

La perspectiva feminista y autoreflexiva adoptada en este estudio permitió observar que las dimensiones de género construidas espontáneamente en la experiencia del trabajo y materializada desde la elección de una temática feminista, una organización mayoritariamente femenina y vinculación social y cognitiva establecida desde las fases iniciales del *lab*, contrarestó la lógica agresiva y poco accesible de este tipo de modelos de trabajo. En este sentido, habría que cuestionar si es el reto lo que genera

la motivación para el trabajo colectivo (Laursen, 2017), o son los vínculos, la experiencia de colectividad y el sentido que tiene el tema de investigación lo enriquece el proceso y fortalece la motivación para cumplir el reto.

Este hallazgo permite vislumbrar la posibilidad de diseñar experiencias *data sprint* a partir de principios feministas (D'Ignazio, et al, 2020) desde los cuales se gesten experiencias inclusivas, reflexivas y afines al pensamiento crítico que deben seguir aportando las humanidades en el capitalismo tardío. Esto podría responder a la inquietud de Laursen (2017) acerca del rol de las humanidades en el mundo contemporáneo y de Venturini, et al. (2018) sobre el deber de mantener la filosofía de los comunes que es parte de las raíces de esta metodología. Se podría considerar que, junto al giro computacional que ha impactado en los métodos de investigación (Rogers, 2019), es necesario plantear un giro crítico sobre el desarrollo tecnológico asumido como equitativo para todos los grupos sociales. En este punto el feminismo tiene mucho por ofrecer, pues las inequidades en la producción, uso y reapropiación de tecnologías han sido planteadas por el tecnofeminismo desde hace tiempo, (Wajcman, 2007) y, en términos de interseccionalidad, han sido registradas por el feminismo de datos recientemente (D'Ignacio y Klein, 2020).

La perspectiva de feminista fomenta la reflexión activa (Soedirgo y Glas, 2020) sobre el posicionamiento de las participantes en el *data sprint* y sobre el tipo de práctica social que promueve y performa dicha experiencia, lo que abre rutas metodológicas para futuras indagaciones. Este trabajo puede considerarse, en términos de Leurs (2017), una hoja de ruta para futuros diseños de *data sprints* más inclusivos y sensibles a los sistemas de opresión que están presentes en las prácticas investigativas.

En suma, el método de *data sprint* es relevante y permite abrir oportunidades para la investigación sociocultural de la Web. Estas oportunidades se fortalecen cuando se dan por fuera de las lógicas del mercado y desde una perspectiva feminista que es sensible a las desigualdades, y que, desde la responsabilidad y el cuidado (Leurs, 2017), establece otras formas de relación y reflexión sobre el tipo de ciencia que se performa con la ecuación velocidad más productividad propia de los sistemas capitalistas que promueven estructuras competitivas y desiguales.

Agradecimientos

Las autoras agradecen el apoyo para la realización de este artículo del Inova Smart Data Sprint (Universidad Nova de Lisboa) y de sus participantes. Gabriela Sued, becaria postdoctoral, agradece el apoyo para la realización de este artículo del Programa de Becas Postdoctorales de la UNAM, Coordinación de Humanidades, del IIS-UNAM y de su asesora académica, la Dra. Judith Zubieta García.

Referencias

Archer, M. (2007). *Making our way through the world. Human reflexivity and social mobility*. Cambridge: Cambridge University Press.

- Bardzell, S. (2010). Feminist HCI: taking stock and outlining an agenda for design. En *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'10)* (pp. 1310-1310). Association for Computing Machinery. <https://doi.org/10.1145/1753326.1753521>
- Berry, D., Borra, E., Heldmond, A., Plantin, J. C. y Walker, J. (2015). The data sprint approach: exploring the field of Digital Humanities through Amazon's application programming interface. *Digital Humanities Quarterly*, 9(4): 1-17. Disponible en: <http://www.digitalhumanities.org/dhq/vol/9/3/000222/000222.html>
- Corlett, S. y Mavin, S. (2018). "Reflexivity and researcher positionality". En C. Cassell; A. Cunliffe y G. Grandy (Eds.). *The Sage Handbook of Qualitative Business and Management Research Methods* (pp.377-389). SAGE.
- Digital Methods Summer School Program. (2022). Digital Methods Initiative. <https://wiki.digitalmethods.net/Dmi/DmiSummerSchool> [consultado el 8 de enero de 2022]
- D'Ignazio, C., y Klein, L. F. (2020). *Data feminism*. Cambridge, Massachusetts:MIT Press.
- D'Ignazio, C., Michelson, R., Hope, A., Hoy, J., Roberts, J., y Krontiris, K. (2020). «The Personal is Political»: Hackathons as Feminist Consciousness Raising. *Proceedings of the ACM on Human-Computer Interaction*, 4(150): 1-23. <https://doi.org/10.1145/3415221>
- Haraway, D. (1988). Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective. *Feminist Studies*, 14(3): 575-599. <https://doi.org/10.2307/3178066>
- Hardin, C. D. (2021). Gender Differences in Hackathons as a Non-traditional Educational Experience. *ACM Transactions on Computing Education*, 21(2): 13-30. <https://doi.org/10.1145/3433168>
- Huurdeman, H., Ben-David, A. y Sammar, T. (2013). Sprint methods for web archive research. En *Proceedings of the 5th Annual ACM Web Science Conference (WebSci '13)* (pp. 182-190). Association for Computing Machinery. <https://doi.org/10.1145/2464464.2464513>
- Laursen, C. (2017, febrero 15). "What is a data sprint? An inquiry into data sprints in practice in Copenhagen". *The ETHOS Lab Blog*. Disponible en: <https://ethos.itu.dk/caecilie-laursen/> [Consulta: 18 de noviembre de 2021]
- Leurs, K. (2017). Feminist Data Studies: Using Digital Methods for Ethical, Reflexive and Situated Socio-Cultural Research. *Feminist Review*, 115(1): 130-154. <https://doi.org/10.1057/s41305-017-0043-1>
- Linabary, J. R., y Corple, D. J. (2019). Privacy for whom?: A feminist intervention in online research practice. *Information, Communication & Society*, 22(10): 1447-1463. <https://doi.org/10.1080/1369118X.2018.1438492>
- Maffía, D. (2007). Epistemología feminista: la subversión semiótica de las mujeres en la ciencia. *Revista venezolana de estudios de la mujer*, 12(28): 63-98. Disponible en: http://saber.ucv.ve/ojs/index.php/rev_vem/article/view/2181
- Milan, S., y Treré, E. (2019). Big Data from the South(s): Beyond Data Universalism. *Television & New Media*, 20(4): 319-335. <https://doi.org/10.1177/1527476419837739>

- Omena, J.J. (2018). "#SMARTDataSprint" En: SMART Social Media Research Techniques [video] YouTube. Disponible en: <https://www.youtube.com/watch?v=bveMpEtAvug&t=1s> [Consulta: 18 de noviembre de 2021]
- Omena, J.J., Rabello, E.T. y Mintz, A.G. (2020). Digital Methods for Hashtag Engagement Research. *Social Media + Society*, 6(3):1-18. <https://doi.org/10.1177/2056305120940697>
- Omena, J. J., Pilipets, E., Gobbo, B., y Chao, J. (2021). The Potentials of Google Vision API-based Networks to Study Natively Digital Images. *Diseña*, 19(1). <https://doi.org/10.7764/disena.19.Article.1>
- Pearce, W., Özkula, S. M., Greene, A. K., Teeling, L., Bansard, J. S., Omena, J. J., y Rabello, E. T. (2018). Visual cross-platform analysis: Digital methods to research social media images. *Information, Communication & Society*, 23(2): 161-180. <https://doi.org/10.1080/1369118X.2018.1486871>
- Plant, S. (1998). *Zeros and Ones: Digital Women + the New Technoculture*. London: Fourth Estate.
- Rampin R. y Rampin, V. (2021). Taguette: open-source qualitative data analysis. *Journal of Open Source Software*, 6(68): 3522. <https://doi.org/10.21105/joss.03522>
- Riessman, C.K. (2008). *Narrative methods for the human sciences*. SAGE.
- Rogers, R. (2017). "Digital Media and Cross Platform Analysis". En J. Burges, A. Marwick, y T. Poell (Eds.). *The SAGE Handbook of Social Media* (pp. 91-110). SAGE.
- Rogers, R. (2019). *Doing Digital Methods*. SAGE.
- Rogers, R. y Lewthwaite, S. (2019). Enseñando métodos digitales: Entrevista a Richard Rogers. Entrevistadora: S. Lewthwaite. *Diseña*, 14: 12-37. <https://doi.org/10.7764/disena.14.12-37>
- Soedirgo, J. y Glas, A. (2020). Toward active reflexivity: Positionality and practice in the production of knowledge. *PS: Political Science & Politics*, 53(3): 527-531. <https://doi.org/10.1017/S1049096519002233>
- Sordini, M.V. (2021). Los retos de la reflexividad en la práctica. Revisiones sobre las estrategias de indagación y análisis. *Revista Latinoamericana de Metodología de la Investigación Social, ReLMIS*, 11(22): 4-7. Disponible en: <http://relmis.com.ar/ojs/index.php/relmis/article/view/presentacion>
- Soria-Guzmán, I. (2021). Mujeres hacker, saber-hacer y código abierto: tejiendo el sueño hackfeminista. *LiminaR, Estudios Sociales y Humanísticos*, XIX(1): 57-74. <https://doi.org/10.29043/liminar.v19i1.806>
- Sued, G. E., Castillo-González, M. C., Pedraza, C., Flores-Márquez, D., Álamo, S., Ortiz, M., Lugo, N., y Arroyo, R. E. (2021). Vernacular Visibility and Algorithmic Resistance in the Public Expression of Latin American Feminism. *Media International Australia*. <https://doi.org/10.1177/1329878X2111067571>
- Suss, S., Hisarciklilar, O. y Thomson, V. (2011). Management Methods for Reducing Span Time during New Product Development. *Proceedings of the 2011 Industrial Engineering Research Conference*. Georgia, USA.

- Venturini, T., Munk, A., y Meunier, A. (2018). Data-Sprint: a Public Approach to Digital Research. En C. Lury; R. Fensham; A. Heller-Nicholas; S. Lammes; A. Last; M. Michael y E. Uprichard (Eds.). *Routledge Handbook of Interdisciplinary Research Methods* (pp. 158-163). London / New York: Routledge.
- Wajcman, J (2007). From women and technology to gendered technoscience. *Information, Communication & Society*, 10(3): 287-298. <https://doi.org/10.1080/13691180701409770>
- Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. New York :Springer-Verlag.



Staging and storing data sprint-based research results: a communication design approach

Organización y almacenamiento de los resultados de la investigación basada en *data sprints*: un enfoque desde el diseño de la comunicación

ID Beatrice Gobbo
beatrice.gobbo@polimi.it
Density Design Lab, Politecnico di Milano

ID Elena Aversa
elena.aversa@polimi.it
Density Design Lab, Politecnico di Milano

ID Andrea Benedetti
andrea.benedetti@polimi.it
Density Design Lab, Politecnico di Milano

ID Gabriele Colombo
gcolombo@iuav.it
King's College London

ID María de los Ángeles Briones Rojas
mariade.briones@polimi.it
Density Design Lab, Politecnico di Milano

ID Tommaso Elli
tommaso.elli@polimi.it
Density Design Lab, Politecnico di Milano

ID Michele Mauri
michele.mauri@polimi.it
Density Design Lab, Politecnico di Milano

Received: 14/01/2022

Accepted: 30/05/2022

ABSTRACT

Data sprints are multidisciplinary, time-limited, practice-oriented, and group activities that explore complex issues through data gathering, visualization, and analysis. Data sprints usually conclude with final presentations where participants showcase the research process they followed and the achieved results. Research centers conducting data sprint activities store research results repurposing existing formats (i.e., Wikis) not directly tailored to organizing, archiving, and disseminating on-line data sprints' output. Starting from the experience of a data sprint on Digital Methods, the paper describes the design process of a digital format which condenses and reinterprets good practices already in use in current formats for disseminating data sprints results. The design process of the format is structured in two steps: (i) designing and (ii) staging the format. The former aims to dissect and recompose data sprint results in hierarchical block structures filled with variable contents; the latter describes outcomes that emerged during a preliminary evaluation phase with experts and a stress-test activity. The format is a container that participants can fill at the end of a data sprint to organize, archive, and present research results. Evaluation activities confirmed the efficiency of the format, which is distinguished by its nested structure and flexibility with different content.

KEYWORDS

digital methods, communication design, scientific dissemination, data sprint

How to cite this article:

Gobbo, B.; Briones Rojas, M. A.; Aversa, E.; Elli, T.; Benedetti, A.; Mauri, M. & Colombo, G. (2022). Staging and storing data sprint-based research results: a communication design approach. *Dígitos. Revista de Comunicación Digital*, 8: 87-108. DOI: 10.7203/drdcd.v1i8.241



RESUMEN

Los *data sprints* son actividades multidisciplinares, de duración limitada, orientadas a la práctica y en grupo, que exploran cuestiones complejas mediante la recopilación, visualización y análisis de datos. Los *data sprints* suelen concluir con presentaciones finales en las que los participantes muestran el proceso de investigación que han seguido y los resultados obtenidos. Los centros de investigación que llevan a cabo actividades de data sprint almacenan los resultados de la investigación reutilizando formatos existentes (por ejemplo, Wikis) que no están directamente adaptados para organizar, archivar y difundir los resultados de los *data sprints* en línea. Partiendo de la experiencia de un sprint de datos sobre *Digital Methods*, el artículo describe el proceso de diseño de un formato digital que condensa y reinterpreta las buenas prácticas ya utilizadas en los formatos actuales para difundir los resultados de los *data sprints*. El proceso de diseño del formato se estructura en dos pasos: (i) el diseño y (ii) la puesta en escena del formato. La primera tiene como objetivo diseccionar y recomponer los resultados del *data sprint* en estructuras jerárquicas de bloques llenas de contenidos variables; la segunda describe los resultados que surgieron durante una fase de evaluación preliminar con expertos y una actividad de prueba de esfuerzo. El formato es un contenedor que los participantes pueden llenar al final de un sprint de datos para organizar, archivar y presentar los resultados de la investigación. Las actividades de evaluación confirmaron la eficacia del formato, que se distingue por su estructura anidada y su flexibilidad con diferentes contenidos.

PALABRAS CLAVE

métodos digitales, diseño de comunicación, difusión científica, *data sprint*



Staging and storing data sprint-based research results: a communication design approach

1. Introduction: disseminating data sprint-based research

Although the format of the sprint was born in the computer science context, over the years sprints have been applied in different areas, from service design to the more recent academic context for data gathering and analysis (Mauri et al., 2019, p. 162). Within the academic context, data sprints are multidisciplinary events where teams of researchers, doctoral and master students from various disciplines, such as Science and Technology Studies, Future Studies, Sociology, Media Studies and Digital Humanities (Bounegru et al., 2018; Ciuccarelli and Elli, 2019) come together to best achieve a given goal through practice-based research. As stated by Mauri in 2019, data sprints usually address wicked problems (Mauri et al., 2019, p. 163), namely situations where it is not easy to define an initial design brief and where different professionals must agree on the type of result they want to realize (Rittel and Webber, 1973). In this regard, communication design and data visualization skills are considered crucial when dealing with controversy mapping and prototyping activities (Mauri and Ciuccarelli, 2016). Moreover, data sprints are time-limited, practice-oriented, and driven by case study activities where communication design researchers play a pivotal role in *informing* and *tracking* the research process while guiding and tutoring visualization activities within the teamwork (Mauri et al., 2019, pp. 161–180). Data sprints usually conclude with a

final presentation, in which the teams show the results achieved during the days of intensive work.

However, results produced during data sprints are fragmented and scattered across disparate media (slides, notebooks, repositories, shared folders), thus they are not always easily accessible and consistently organized, and the risk that some parts might be lost is frequent. Existing dissemination formats employed during data sprints (Wikis, web reports, slides) require users to define a structure for presenting outputs and often have to make compromises to include different specific features (i.e., slides do not allow insert interactive visualizations, Wikis do not generate a visual index). Hence, it seems appropriate to design a platform that inherits from existing formats the flexibility of content publication, while integrating their specificities (i.e., hypertexts from Wikis, visual content from presentations) into a single system. In this regard, the paper delineates the role of communication designers in the context of data sprints in the ideation of a dissemination format, namely a digital *place* that allows data sprints' results to be stored, staged, and accessed after the research activities.

1.1. The *Algocount* data sprint

The starting point of the paper is the work carried out during the *Algocount* data sprint held online from January 18th to 22nd, 2021.¹ Emulating the successful approaches of the Digital Methods Initiative Schools in Amsterdam and the Smart Data Sprint in Lisbon, all activities in the data sprint were conducted using Digital Methods (DM), a collection of web-native techniques for research into cultural change and societal conditions (Rogers, 2013). The presented data sprint is one of the main activities conducted within a broader multidisciplinary project that seeks to expand the current understanding of the role of algorithms in society and raise awareness about them, with a specific focus on public opinion in Italy. The data sprint was framed within the *infodemic* theme^{1,2}, and it exploited the discussions that emerged on social networks during the pandemic to study the influence of algorithms in the formation of Italian public opinion on social media platforms (Facebook, Twitter, Instagram³, TikTok, Reddit, and YouTube). The participants were divided into five sub-groups and, as a result, the data sprint produced a total of seven “experiments”, namely single investigations addressing a research question. Indeed, each group carried on at least one experiment during the data sprint. Materials of the seven experiments have been used as resources for designing the dissemination format presented in the paper which is digital, web-based, and ought to preserve features of data sprint results (i.e., datasets, static and interactive visualization, video animations, GIFs, and texts). Indeed, although the seven experiments had common features, research processes were not always generalizable or inscribable into definite categories, thus it was important to design a format that on the one hand was flexible, but on the other left room for users to be thorough.

In Section 2 authors present a review of formats for dissemination of data sprint-based research, reflecting on opportunities for innovating existing solutions; Section 3 describes in detail the design process from the collection of data sprint results to the

preliminary evaluation of the format. Section 4 shows the strengths and weaknesses of the format, using one of the experiments carried on during the data sprint as a case of application. Finally, the authors envision future research opportunities and usage scenarios.

2. Existing dissemination formats for data sprint-based research

When it comes to disseminating the outcomes of data sprints, every research team is interested in disclosing the technical details of the conducted experiments, and in conveying discovered findings. Authors identified four types of formats to disseminate sprints-based research: final presentations (1), written reports (2), web platforms (3), and field guides (4). These four dissemination formats describe current practices in disseminating research results in data sprints, and they can inform how the design of a new format could integrate their specificities to build on consolidated habits of documentation.

(1) **Final presentations** organize process and findings in slides that mostly include visual contents, to encourage verbal exposition and discussion in front of other teams of researchers⁴. Presentations can be recorded and shared on streaming platforms to preserve the discussions⁵. Researchers can insert hyperlinks to resources like datasets and high-quality pictures in slides, but the format is not suitable for storing them. Researchers are used to simultaneously preparing slides using cloud-based software. Research runs fast and heuristically, therefore they hardly have the time to organize the produced materials as the work proceeds, involving the risk, after the sprint ends, of losing resources as they are moved or deleted from shared folders. Consequently, the body of work needs to be recollected to be preserved.

(2) A part of the research team is usually assigned the task of writing a **comprehensive report** after presentations and research activities are concluded. Reports include written and detailed descriptions that are not included in presentations and may link to external resources. However, they usually do not include interactive features, nor store materials⁶.

(3) Another example of dissemination format is a custom **web platform**⁷: a website designed to host the outcomes of multiple data sprint experiments, which share the same context or topic because part of the same overarching research project (Venturini, 2014). A web platform resonates with other attempts aimed at renovating dissemination formats (Lupo et al., 2021). It provides comprehensive descriptions and gives space to visual and interactive contents. Additionally, it stores and gives access to materials like datasets, source files (such as networks in Gephi format, or vector-based

4. Example of a data sprint final presentation <https://bit.ly/3JbB43e> (accessed: 21/12/2021)

5. Example of a record of data sprint presentation <https://bit.ly/3mk68UD> (accessed: 21/12/2021)

6. Examples of written reports can be accessed from <https://bit.ly/3yOPxgl> or <https://bit.ly/3J1kjYy> (accessed: 21/12/2021)

7. Climaps is an exemplary case of a platform that collects sprint-based research. <http://climaps.eu/> (accessed: 21/12/2021)

projects files), high-quality images, and code snippets. Web platforms proved capable of appropriately displaying contents and surviving the passing of time, however, they require considerable efforts to be designed, implemented, and maintained.

(4) **Field guides** are a further dissemination format that is aimed at reflecting on the research process conducted during sprints (Bounegru et al., 2018). Guides stress the methodological approach behind sprint-based research and aim at packaging reusable procedures to allow the adoption by other practitioners. As for web platforms, field guides require significant efforts in that researchers must consistently analyze and reproduce previous research.

While presentations and reports are produced by individual research teams during or immediately after sprints, web platforms and field guides are the results of a post-sprint re-elaboration. In particular, the design and implementation of platforms requires significant resources, and we can't assume that all data sprints result in a dedicated website.

3. Design Process

Building on the previous considerations, authors identified the need for a format designed to simplify dissemination and preservation of sprint-based research. The proposed format requires researchers to give structure to their works, it allows for collaboration between multiple users, and, thanks to a dedicated technical infrastructure, it produces a functional website that can be hosted on any server. The design process of the presented dissemination platform includes two main phases. The first phase (Figure 1), finalized to design the dissemination format interface (3.1), involves: the design of a data sprint report structure, the reorganization of all the reports into a matrix, and the iterative design of a format mock-up. The second phase (3.2) is focused on staging the format and involves focus groups and stress-test activities.

3.1. Shaping the format: restructuring data sprints results into a dissemination format

As mentioned above, results of data sprints can have different dissemination formats, often like compositions of semi-finished visuals and text in the form of slides or reports with variable structures. Hence, all research carried out during the data sprint this paper refers to have in common that they are based (1) on data gathered from social networks, (2) deal with addressing ill-defined problems, (3) they are carried out with DM, (4) are driven by the main research question and eventual sub-research questions, (5) and they include a significant number of visual outputs (i.e., visualizations, videos, GIFs).

Designing the report structure — The 20 participants of the presented data sprint — a team of researchers from the social sciences and a communication design team — worked in five groups, each studying one or two platforms and experimenting with DM auditing and reverse engineering techniques (Ziewitz, 2016) combined with data visualization. Right after the sprint, each group stored slides containing the data sprint results, and afterwards the communication design team drafted a report structure to be filled by each group.

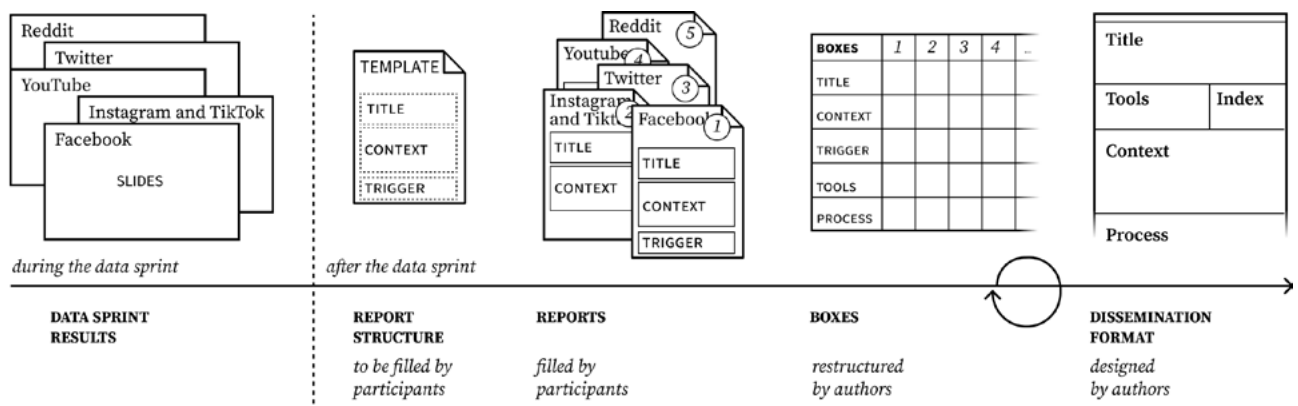


Figure 1. Diagram depicting the whole design process started from the AlgoCount data sprint. On the left, the starting point of the process, namely the collections of slides designed by participants. Post-data sprint activities included: the design of a report template to be filled by participants, the dissection of the filled reports, the re-organization of the gathered material into boxes and the design of the dissemination format.

The report structure was intended to collect processed information from all the groups consistently to have an arrangement that would make the experiments comparable to each other. Inspired by existing DM reports, the format required researchers to complete five main sections:

- **Introduction:** section dedicated to describing the research object (i.e., the Facebook Newsfeed) and the research problem triggered by an inexplicable algorithmic situation.
- **Experiment(s):** section meant to collect the experiment(s) carried on during the data sprint. While some working groups produced just one experiment, others were able to carry out more.

Each experiment includes:

- the *research question*, specifying the research object and the addressed wicked problem.
 - the *experiment setting*, describing the digital environment in which the data collection took place, with which devices, over what time range.
 - the *outline* of the process, depicting the step-by-step procedure for carrying out the data-driven experiment.
 - the *dataset(s)* output, containing the list of gathered data.
 - the discursive *description* of the experiment.
 - the main *findings* of the experiment, emerging from visualizations.
- **General findings:** if the research is comprised of multiple research questions, this space includes a general finding that summarizes the data sprint-based research outcomes.
 - **New research questions,** such as future research opportunities arising from the outcomes and outputs of the data sprint.

- **Envisioning for a dissemination format:** here, participants are invited to imagine a format for sharing their work to gather ideas that will be useful in the next phase, dedicated to the dismantling and reassembling of the works.

At the end of this process, once that all the participants have filled the report structure, authors collected five reports (one for each group), respectively including at least one experiment.

Packing into boxes: from reports to content blocks — Once the reports were completed, authors reviewed and dissected them having an overview of each and all of them at the same time. From five reports authors identified seven experiments. Given the structure of the reports previously described, it was easy to identify similar aspects in the work processes of each group. However, the diversity with which each experiment used DM also emerged. Through an activity that the authors called “*packing into boxes*”, they recomposed excerpts from the reports into a matrix (Figure 2. screenshot of the “packing into boxes” output: a matrix that reconfigured the information of the experiments.). Rows of the matrix represented boxes (i.e., platform, title, tools, teaser, contextualization), while each column was dedicated to a report (Facebook, TikTok, YouTube, Twitter, Reddit, Instagram). The goal was to get an overview of the five reports and the diverse experiments within them: *what are the similarities and differences in content, and what are the existing information hierarchies?*

Thus, the matrix enabled authors to organize results according to shared components and highlight the peculiarities of each experiment. For example, the authors identified tactics and tools that have been used in more than one experiment, such as *RAWGraphs* (Mauri et al., 2017) — a data visualization tool, and Fake Name Generator — a website

	1	2	3	4
	https://docs.google.com/document/d/1kkIq9D1Y-eXdfwa6W9SLipJe9YT3fJJZSMaA8ssdfw/edit?usp=sharing	https://drive.google.com/file/d/10enxJMZjk7NuaJVEIhoUm6qRp1_BEYz0/view?usp=sharing	https://drive.google.com/file/d/10z4cDIBRBSV-HsaeUGLpUOY_Qw5_uReb/view?usp=sharing	https://drive.google.com/file/d/1HNpGly1f54uEYHHUwTPIVbe/view?usp=sharing
Experiment type	AUDIT			
Platform	Facebook	TikTok	Youtube	Twitter
Title [trigger + platform + studied digital object]	Why I am seeing this infodemic on my Facebook's newsfeed?	The spread of polarization through opposite infodemic # on my TikTok's For you page	Recommendation system on YouTube platform and its specific functioning.	Twitter: hashtags and Who to follow
Tool tag	<ul style="list-style-type: none"> - OBS - Apify - RAWGraphs - Illustrator-Inkscape - Premiere - Fake name generator it.fakenamegenerator.com 	<ul style="list-style-type: none"> - RAWGraphs - Illustrator-Inkscape - Fake name generator it.fakenamegenerator.com 	<ul style="list-style-type: none"> - YouTube Data Tools suite - Google Spreadsheet/Excel - Gephi - RAWGraphs - Illustrator/Inkscape 	<ul style="list-style-type: none"> - Python (Spyder environment) - Google Spreadsheet/Excel - R (RStudio environment) - Gephi - RAWGraphs - Illustrator/Inkscape
Teaser	The project focuses on: - exploring the day by day development of a newborn facebook News Feed interested in no mask related issues in Italy; - collecting data about news feed and use activity; - visualize and tell the radicalization process following platform affordance.	The project focuses on: - Exploring how the For You page becomes polarised for a user in favour of the Covid-19 vaccine and for a user against it; - Understanding how the TikTok algorithm can contribute to the circulation of Covid-19 related content;	In this project we will put the algorithmic radicalisation logic through a "stress test", using the topic of the Covid-19 in Italy.	The project focuses on: - How infodemics structure themselves o - How researchers can map and visualize infodemic in order to be read and analyse - How Twitter's affordances influence the developments of infodemics
Contextualization [trigger + platform + studied digital object] > provide all the information for understanding the relevance of the studied digital object and the trigger.	Newsfeed description > "each of this buttons" from https://www.dummies.com/social-media/facebook/facebook-news-feed/ News Feed is the centerpiece of your Home page. When you log in to Facebook, you see the familiar blue bar on top and the left sidebar, but mostly you see News Feed. It's a constantly updating list of stories by and about your friends. Stories here refers to actions that your friends have taken on Facebook — things like writing a status update, sharing a photo or link, or becoming friends. You may also see stories from Pages that you follow. Facebook sometimes refers to stories as posts. More accurately, News Feed is an algorithm. It doesn't show you everything from your friends; instead, it tries to show you things it thinks you will find interesting.	[TikTok description based on https://www.bbcnews.com/better/lifestyle/what-tiktok-guide-clueless-parents-ncna1066466] TikTok is a social media app and platform you can use to create 15 second videos with hashtags, sound, filters, usually by lip-synching or dancing along with top songs. You can share the videos you create with friends, or with a wider circle of TikTok users. It was created by a chinese company in 2018. You can also follow other creators on the app without posting your own content. People can become virally famous on TikTok — many minor TikTok celebrities have 100,000-plus followers, and the biggest stars break 10 million and more. It is widely used by the "Gen Z". TikTok works through 3 main components	YouTube is an American online video sharing and social media platform launched in 2005. Around the world, its users watch more than one billion hours of videos each day. YouTube content creators, popularly referred to as YouTubers, upload over 100 hours of content per minute.	Twitter introduces itself as "where people out what's happening in the world right now" invites new users to "come to Twitter to see people are talking about and join the conversation. Through the use of hashtags, Twitter use immediately access the debate on a certain and discuss about it with others, in many strangers. Through the retweet function, decide to take a public stand and act as a of a certain information or position. Twitter has launched a program for academic researchers that allows — after the submission approval of a research project by the platform to access through the API v2 to the entire database of Twitter from its birth onward. passages could take some time, as they manual examination and approval by the

Figure 2. screenshot of the “packing into boxes” output: a matrix that reconfigured the information of the experiments.

which randomly generates fake biometric data. Hence, the authors compiled the matrix grouping tactics and tools inside a *box*. The reoccurrence of concepts laid the foundation for the design of a *glossary* of tools and tactics that would serve as a transversal means for all the experiments. Moreover, while *packing into boxes*, authors made ontological arguments about the structure of the experiments. Indeed, authors identified three nested levels of information:

- **Main-blocks** represent the higher level of information that structures type of contents transversally such as the *experiment presentation*, namely the overall description of each experiment provided by the groups. Authors identified *experiment presentation*, *context*, *research question*, *experiment setup*, *notes*, *step by step process*, *finding* and *references* as main blocks.
- **Sub-blocks**, namely the nested structure of each main block. Sub-blocks characteristics depend on the parent main block. Thus, some main blocks can lack sub-block. For example, while the *experiment presentation* can be composed by title, a brief description, a table of contents, and tools and techniques employed, the *references* block does not require an additional nesting level because it is a simple content list.
- Finally, the authors identified the **content blocks** that represent the most peculiar information, namely the text, images, videos, and multimedia contents that participants employed for describing their experiments.

Format layout design — The defined blocks were the starting point for translating the information into a graphic layout. Hence, by comparing the blocks across the experiments, authors identified common practices. For example: participants used

INFORMATION BLOCKS

① Main blocks that structure the information.

② Second level of information blocks that are specific to each experiment.

Type of content:
 VISUAL-DYNAMIC
 VISUAL-STATIC
 TEXTUAL

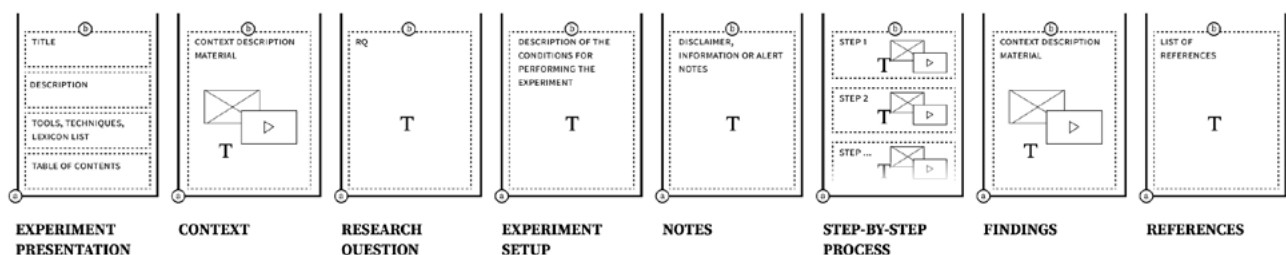


Figure 3. Diagram of information blocks that structure the experiments' content.

multiple media with various graphic styles and colors (static and animated images and visualizations, videos, links, among others) to narrate the process of the experiments. Moreover, authors noticed that experiments' blocks could be nested with different degrees of complexity and that it is necessary that the format allows the research to be read considering both the hierarchy and the sequence of the actions.

Following these observations, authors defined the design Requirements (**R**) of the format:

- **R1:** The graphic style of the format must not compete with the visual material and while emphasizing the structure of the information.
- **R2:** The format must give relevance to the use of visual and multimedia material present in the reports of each experiment.
- **R3:** The format must guarantee a fluid fruition of the contents from a desktop device.

The blocks identified above were essential to build the layout. Indeed, the authors translated them to fit into a clean visual grid that would allow to combine and nest different types of content (**R1**). Thus, while comparing blocks across the experiments, the authors decided to design a layout that could accommodate content in both one column and more than one column. This allowed authors to create diverse compositions including text, bullet points, and multimedia content. Generally, each *main-block* corresponds to a layout however, given the variety of *sub-blocks* and *content-blocks*, authors designed other layout options to meet any requirements.

Hence, as shown in Figure 4. format layout for composing the information blocks.

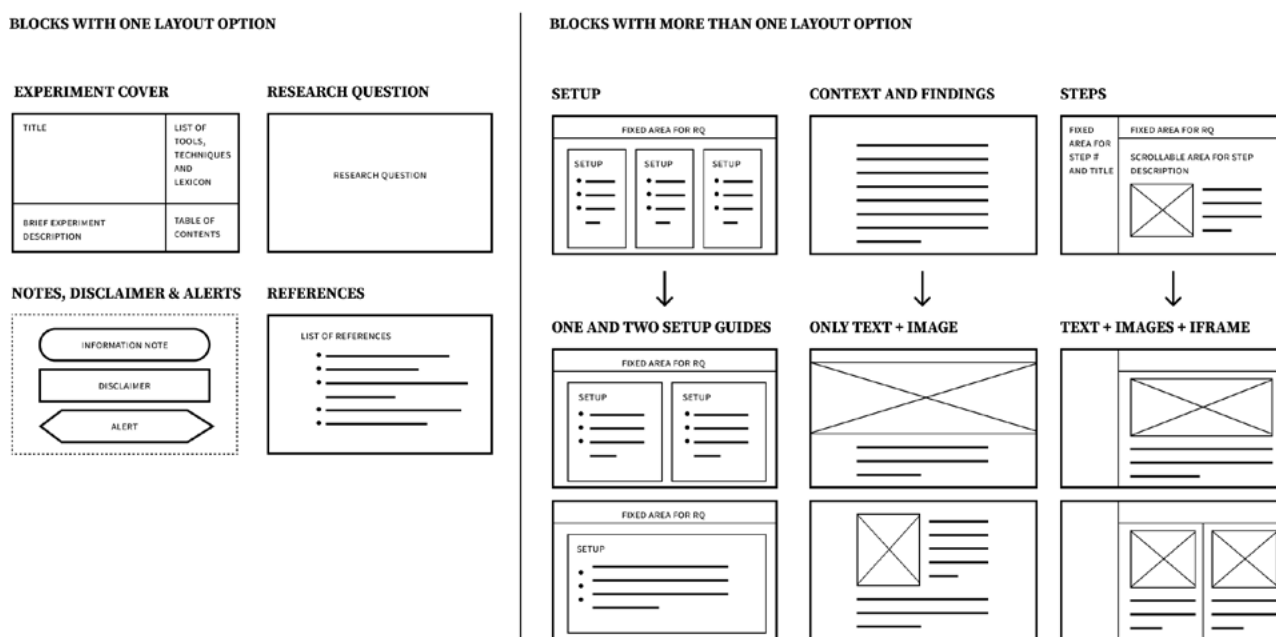


Figure 4. format layout for composing the information blocks. Some blocks have more than one layout depending on the specificity of the content. The diagram provides a sample of blocks that have been used in one experiment. Possible connections between experiments are depicted later (Figure 5. Overall structure of the dissemination format. Each project is a collection of experiments sharing a glossary which can be organized into categories, such as tools and techniques.).

Some blocks have more than one layout depending on the specificity of the content. The diagram provides a sample of blocks that have been used in one experiment. Possible connections between experiments are depicted later (Figure 5); the *experiment presentation*, the *research question*, the *notes*, and the *references* blocks have a fixed

layout. Differently, the *experiment setup*, the *context*, the *findings*, and the *step-by-step* blocks require multicolumn layouts or multimedia compositions to properly address design requirements (R2).

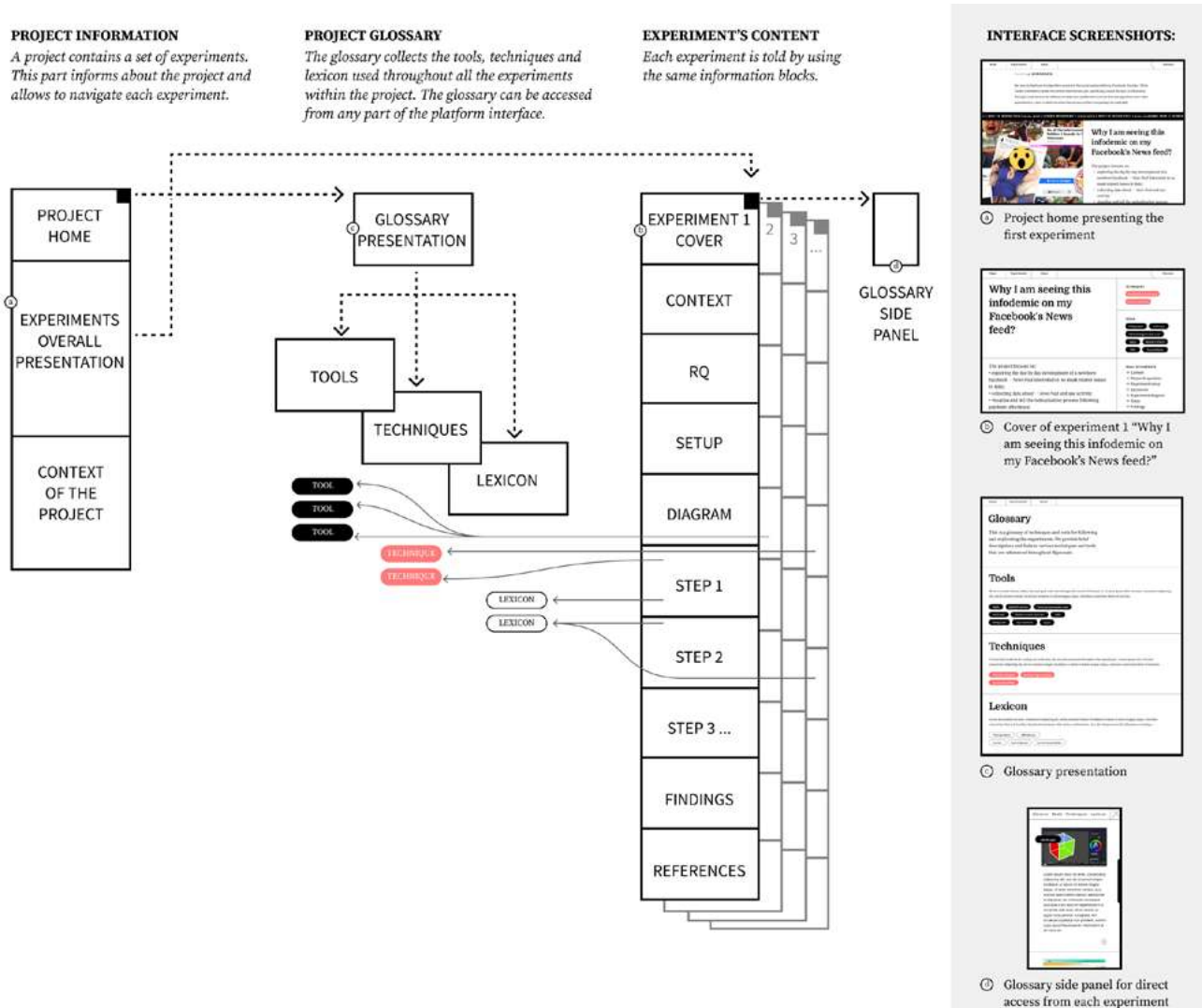


Figure 5. Overall structure of the dissemination format. Each project is a collection of experiments sharing a glossary which can be organized into categories, such as tools and techniques.

To guarantee fluid fruition of the contents, authors exploited interactivity giving priority to scroll, designing sticky and fixed elements, and including tools as magnifying lenses and zoom for multimedia contents. For example, the layout envisioned for the step-by-step process (Figure 4. format layout for composing the information blocks. Some blocks have more than one layout depending on the specificity of the content. The diagram provides a sample of blocks that have been used in one experiment. Possible connections between experiments are depicted later (Figure 5), last column) includes sticky elements (R3) which could support and improve the readability of the contents embedded into the scrollable area.

Finally, to make connections between experiments, the authors designed a layout

dedicated to the *Glossary* where tools and techniques could be gathered and described. The process of standardization of the possible compositions of the information is fundamental to achieve graphic consistency between all experiments, as well as to guarantee to those who will read the experiments a consistent way to find the needed information (Figure 5. Overall structure of the dissemination format. Each project is a collection of experiments sharing a glossary which can be organized into categories, such as tools and techniques.).

2.2. Staging the format: assessment through discussion and simulation of use

Once the layout of the format was defined, the authors addressed the focus group and stress-test activities to assess the format layout design.

Focus group — A wide range of disciplines (Silverman, 2016) has proposed *focus group methodologies* as “an informal discussion among selected individuals” (Becket al. 1986: 73) to gather valuable information about specific subjects. In such groups, participants are free to articulate their opinions in a spontaneous manner, which may facilitate the collection of more in-depth insights. On such intentions, the authors conducted a focus group (see Figure 6. Focus group participants from Political Sciences and Design fields.) of twelve participants (5 facilitators and 7 participants; almost all of them participated in the data sprint) belonging to both social sciences and communication design research fields, to test the structure, layout, usability, and flexibility of the dissemination template. The general aim was to collect feedback on those aspects which were not anticipated as critical during the design development. The focus group lasted approximately three hours and was structured in two main activities: a guided exploration of the format mock-up and a participatory discussion.



Figure 6. Focus group participants from Political Sciences and Design fields.

The first activity was carried out by a facilitator. During this phase, participants were invited to critically reflect on the different features showed. The guided exploration focused on different aspects concerning the layout design. For example, the facilitator asked participants to reflect on the flexibility of the column layout, on the design choices

of the Glossary, and on the arrangement of the visual elements in the cover. Participants were likewise provided with clear-cut questions to direct and nourish the discussion on the aforementioned aspects. For instance, when asked “What is the most suitable layout for your research?” most participants confirmed they would benefit from having the possibility to choose between a single-column or a two-column layout. At this stage the comments were rather positive and focused on the visual choices, the effectiveness of the layout and the usability of the platform.

Throughout the second part of active feedback exchange, the discussion evolved organically and was moderated by the facilitators. This activity was conceived to assess the completeness and flexibility of the template by asking the participants to imagine the application of other case studies. More specifically, they were asked to identify three features they found ineffective or complex along with the three they found useful. The discussion started with the latter, which proved to be a fruitful starting point for thinking about how to further improve specific details. For example, while the Glossary was found to be immediate and essential, the participants also pinpointed the following fine tunings: a) adding a customizable section and b) standardizing the layout for each section.

This phase was critical to confirm the effectiveness of the structure and the layout of the format to narrate the research process. Moreover, it emerged that the structure and the layout were reasonable flexible to accommodate different types of research. Finally, it allowed to identify the elements of the template to be improved.

Stress-test — Following the collection, analysis, and implementation of the feedback that emerged from the focus group, the succeeding activity involved a *stress-test*. The stress-test activity focused on evaluating the overall performance of the format. To assess it, authors populated a prototype version of the web-based format⁸ with other contents. Although in the stress testing methodology, systems are intensively tested above normal operational capacity through automated techniques (Beizer, 2003; Nelson, 2009), authors’ testing method consisted in creating a high-fidelity mockup of the format with two sample projects carried out during the DMI Summer and Winter schools and stored as reports on the website of the school. Such a method allowed the authors to examine the efficiency of the format when dealing with experiments with variable structures and fields of application.

As a result, the focus group and the stress-test assessed the overall usability of the format, both when using experiments coming from the *infodemic* data sprint and from other similar contexts. Criticalities and suggestions emerged from the feedback session have been progressively solved and integrated within the format. For example, the authors refined the reference section and enriched the glossary creating cross-references between techniques and tools.

8. The prototype version of the web-based format is an early-stage Content Management System enabling authors to compose each experiment according to the designed blocks and layouts.

3. Using the format: a case study from the *Algocount* data sprint

As emerged from focus groups and stress tests, the format helps in disseminating data sprint research's processes and results, besides answering storage and durability issues of the information from the experiments over time, since it fills the gap described in Section 2 arisen both from literature and authors' experiences in the field. To better illustrate its usage, in the following section the authors will describe one of the seven experiments developed in the context of the data sprint, shedding light on the aspects of the format that make it a promising tool during usage.

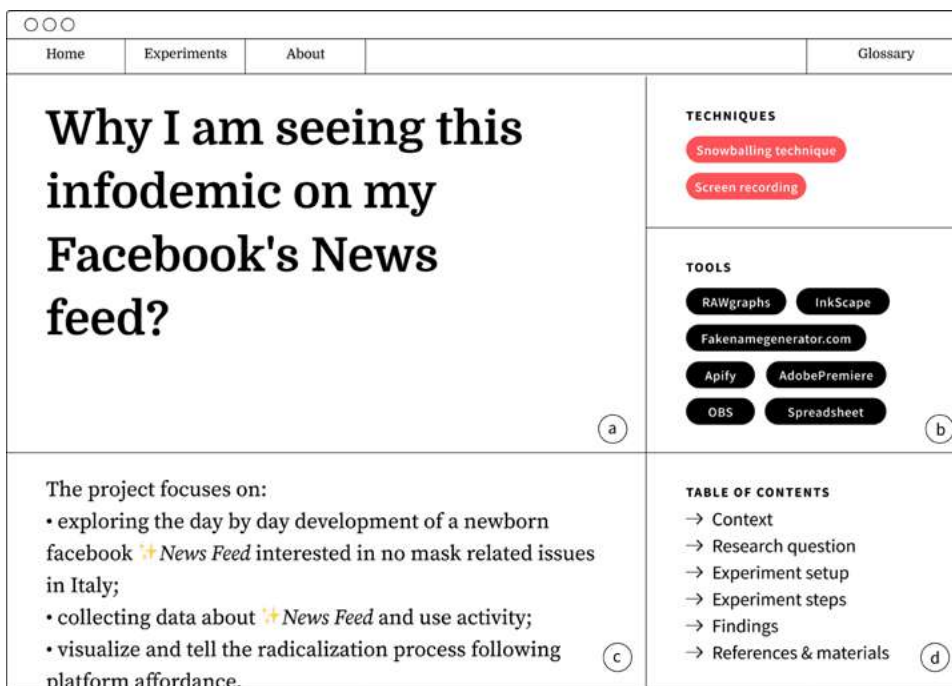


Figure 7. General presentation of the experiment "Why am I seeing this infodemic on my Facebook's News Feed" proposed as case study for this paper.

3.1. Case study: "Why am I Seeing this Infodemic on my Facebook News Feed"

The "Why am I seeing this *infodemic* in my Facebook's News feed?" experiment belongs to a collection of experiments developed during the data sprint. The experiment tries to answer the question "How does the Facebook algorithm facilitate the radicalization process of a newborn user through suggestions started from the query "no-mask Italia". As explained in the previous section, the format autogenerates this page, building it on top of the contents. "Why am I seeing this *infodemic* on my Facebook's News Feed?" focuses on (1) exploring the day-by-day development of a newborn Facebook account interested in no-mask related issues in Italy, (2) collecting data about newsfeed and user's activity manually, (3) visualizes and tells the radicalization process of the user.

Figure 7. General presentation of the experiment "Why am I seeing this infodemic on my Facebook's News Feed" proposed as case study for this paper. shows the general presentation of the experiment: the title (Figure 7. General presentation of the experiment "Why am I seeing this infodemic on my Facebook's News Feed" proposed as case study for this paper.a), the techniques and tools used to refer to the shared glossary (Figure 7. General presentation of the experiment "Why am I seeing this infodemic on my Facebook's News Feed" proposed as case study for this paper.b), the

brief description (Figure 7. General presentation of the experiment “Why am I seeing this infodemic on my Facebook’s News Feed” proposed as case study for this paper.c), and the table of contents (Figure 7. General presentation of the experiment “Why am I seeing this infodemic on my Facebook’s News Feed” proposed as case study for this paper.d), enabling a quicker navigation through the experiment.

Scrolling through the general presentation, as stated in the table of contents, the experiment contains the following blocks: the **context**, the **research question**, the **setup**, the **disclaimer**, the **visual index**, the research **steps**, and the **findings**. Among these, some are more relevant than others and deserve an in-depth description of their characteristics through the application of the case study under consideration. As emerged from the authors’ experiences and focus groups’ results, the main strengths of the format are the following:

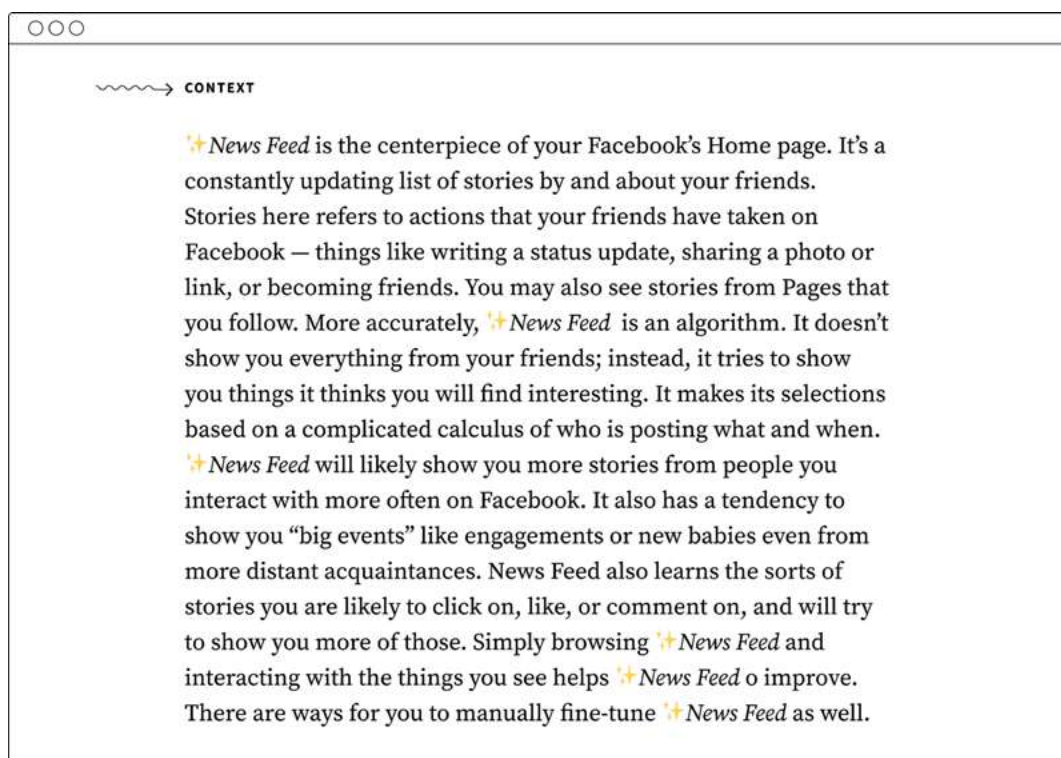


Figure 8. Block describing the context, namely the research object and its features.

- It invites users to *contextualize the research object*, situating it in space and time.
- It helps in *structuring the experiment setup* suggesting the items to be considered.
- Its *visual index* summarizes the research steps in a diagram.
- It allows for combining different layouts, promoting the use of *parallel columns* for comparing data, processes, visualization, or findings.
- It allows to create *a glossary*, shared by all the experiments of the project.

Contextualizing the research object — The News Feed is the research object around which the experiment revolves. It is the centerpiece of Facebook’s home page and is constantly updated with a list of stories by friends, advertisements, and pages

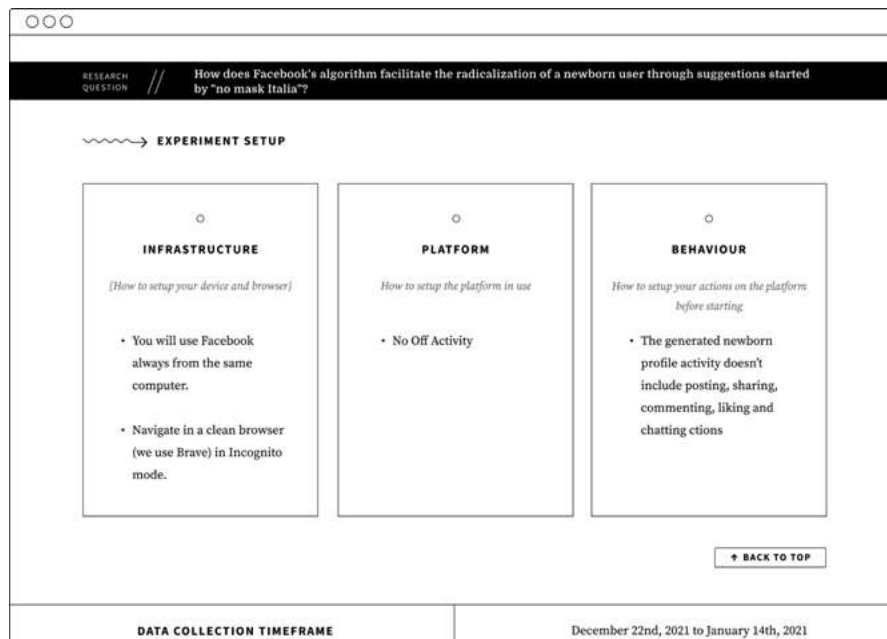


Figure 9. Experiment set-up building box in use. From left to right: “Infrastructure” is about describing how to set up the device and the browser. The “Platform” box is dedicated to settings related to the platform being analyzed. “Behaviour” defines which actions to perform within the platform.

updates. The variety of contents appearing on the News Feed is managed by a complex algorithm, which relies on many factors including user profiling, then users’ choices and advertisements. At the time of the experiment (January 2021), given the General Data Protection Regulation (Regulation (EU) 2016/679 of the European Parliament and of the Council), users can also discover the reason why some contents such as suggested ADVs appear on their News Feed by clicking on the “Why Am I Seeing This” button. If in January 2021, the button was active just for mobile devices, one year later it is also active for desktops. Indeed, research objects are frequently evolving: they are revised and modified in terms of interface or privacy policy; hence it is important to keep track of their state of the art at the time of research.

Thus, the **context** block available in the format, which is usually plain text, guarantees an in-depth description of the research object, where users are invited to report its specificities at the time of the data sprint (Figure 8. Block describing the context, namely the research object and its features.).

Setting up the experiment — As emerged from previous discourse it is crucial to report the set-up of the experiment, namely what type of infrastructure and which devices have been used, on which digital platform the research has been done and what type of behavior was held by the researcher(s) carrying on the experiment. Indeed, defining the experiment setup is the first stage for making the procedure transparent and repeatable. In this case, regarding the infrastructure, participants decided to access Facebook always from the same device — a 15" MacBook Pro, and to navigate in a clean browser with the incognito mode on (Figure 9. Experiment set-up building box in use. From left to right: “Infrastructure” is about describing how to set up the device and the browser. The “Platform” box is dedicated to settings related to the platform being analyzed. “Behaviour” defines which actions to perform within the platform.).

Telling the process —Tracking and presenting the research process carried on during a data sprint is task-driven usually by designers (Mauri et al., 2019). According to Mauri et al., designing visual protocols is a common procedure for summarizing processes, making them accessible, transparent and repeatable, both during data sprints (Mauri et al., 2019) and as a didactic activity (Mauri et al., 2020). Drawing inspiration from

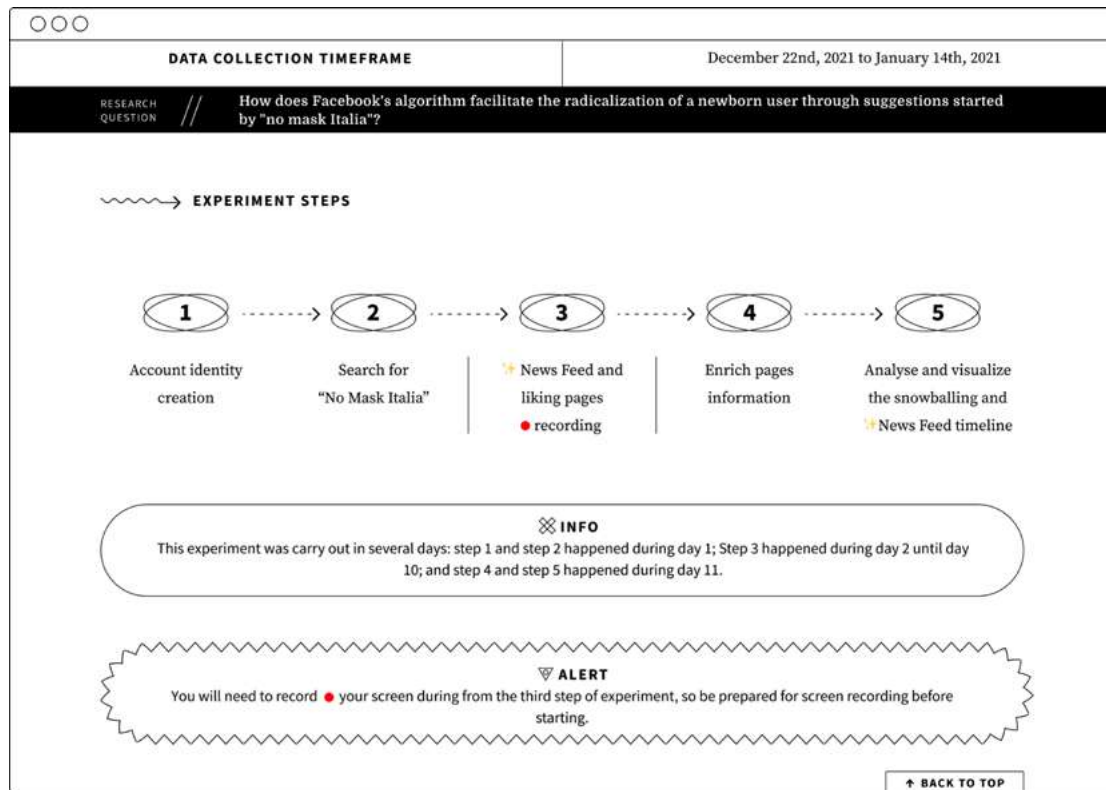


Figure 10. Experiment's visual index summarized in steps.

this tradition, the format allows users to self-generate a diagram working as a visual summary and index to quickly access the diverse steps of the research process. This feature not only makes the fruition simpler, but also helps to organize the experiment in a modular way, making it repeatable and transparent. Moreover, the format allows also to include supplementary information and disclaimers related to the process (Figure 10). As a first action, before starting to gather data, participants created a fake name and a new account on Facebook as depicted in Figure 11. Account creation for the fake Facebook user. The first fixed column on the left summarizes the step, while the double column layout allows to compare two actions, namely the fake name generation and the actual account creation.. During the two weeks before the data sprint, News Feed and user activity data were collected every two days employing techniques such as *screen recording* and *snowballing*. Through *screen recording* researchers stored a visual log of the actions done by the newborn user when collecting data appearing on the News Feed. *Snowballing*, namely the procedure whereby the researchers cascade likes or follows to what is suggested to them by the algorithm, was used to simulate a user activity based on the suggestions of the platform. Hence, given the long data collection process, a large part of the experiment was dedicated to constructing and enriching the

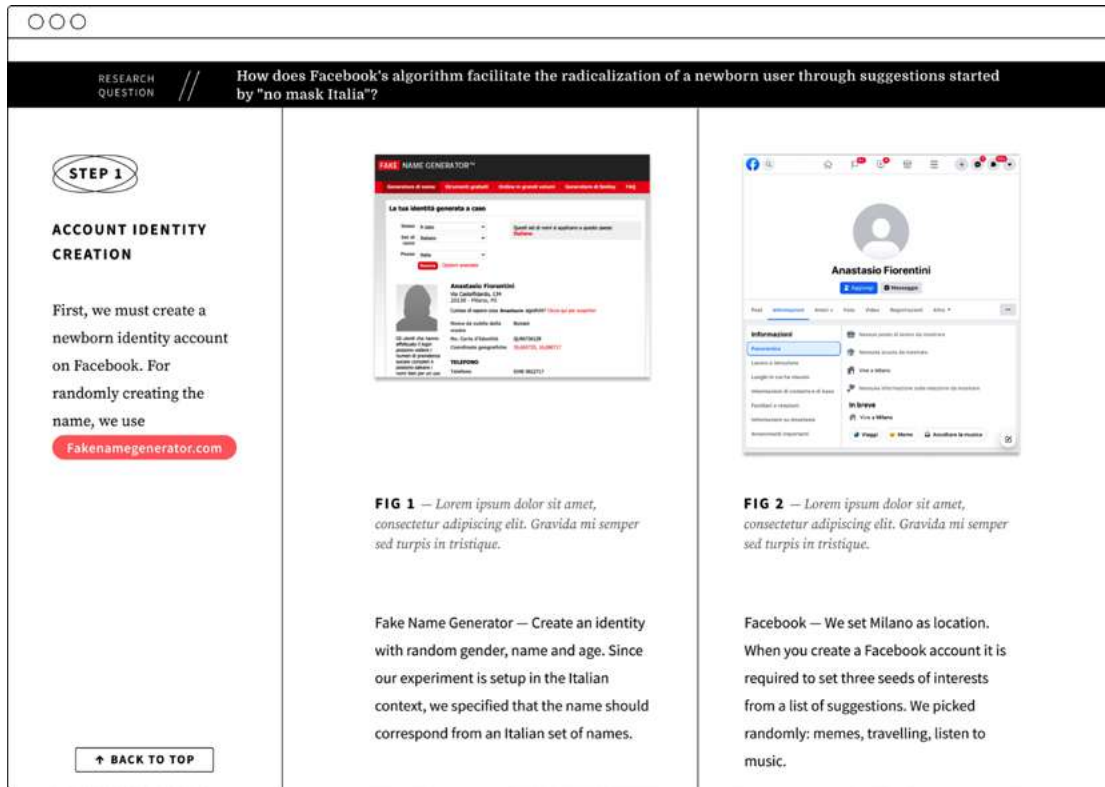


Figure 11. Account creation for the fake Facebook user. The first fixed column on the left summarizes the step, while the double column layout allows to compare two actions, namely the fake name generation and the actual account creation.

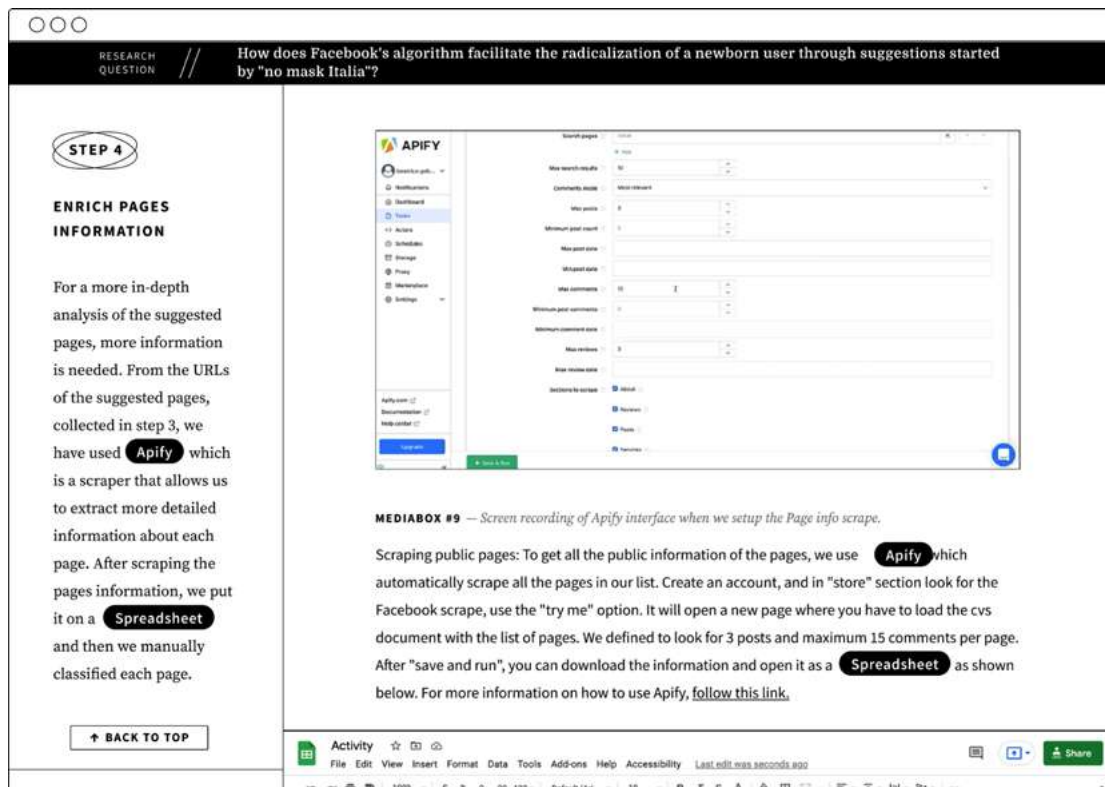


Figure 12. The column on the left summarizes the step, while the central column comprises two columns (top), or a single column with an animated GIFs and a caption.

main dataset. Indeed, to present how a new user interested in no-mask-related issues in Italy radicalizes on Facebook, researchers first collected the list of items appearing on the screen by scrolling through the News Feed. At the same time, they liked the suggested pages and, finally, they enriched the list of liked pages by crossing information coming from pages' descriptions. During the two weeks before the data sprint, News Feed and user activity data was collected every two days employing techniques such as *screen recording* and *snowballing*. Through *screen recording* researchers stored a visual log of the actions done by the newborn user when collecting data appearing on the News Feed. *Snowballing*, namely the procedure whereby the researchers cascade likes or follows to what is suggested to them by the algorithm, was used to simulate a user activity based on the suggestions of the platform.

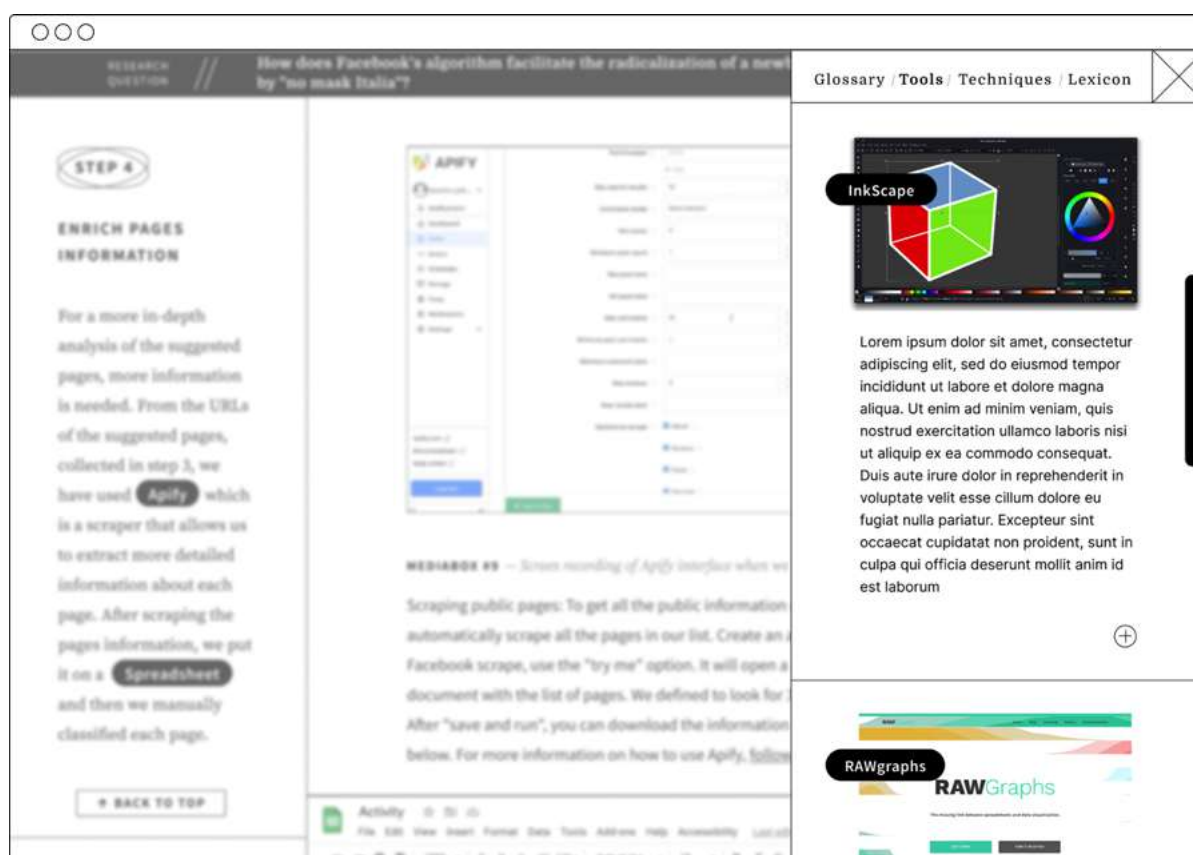


Figure 13. Screen of the Glossary which can appear both as a sliding panel or as an independent page, as depicted previously in Figure 5.

Comparing actions — When dealing with Digital Methods in data sprint-based research it is common to find experiments where some actions are compared in pairs: actions related to the same experiment step, the behavior of different social network accounts or the results emerging from two correlated hashtags —such as “program” and “anti-program” (Rogers, 2017, p. 82). To meet this need, the format allows users to populate a two-column layout where contents can be simultaneously read and compared. (Figure 11). Moreover, the format includes also a single-column layout, that users picked in this context for showing animated GIFs of a visualization, for embedding external video resources and for writing texts (Figure 12. The column on the left summarizes the step, while the central column comprises two columns (top), or a single column with an animated GIFs and a caption.).

Hence, given the long data collection process, a large part of the experiment was dedicated to constructing and enriching the main dataset. Indeed, to present how a new user interested in no-mask-related issues in Italy radicalizes on Facebook, researchers first collected the list of items appearing on the screen by scrolling the News Feed. At the same time, they liked the suggested pages and, finally, they enriched the list of liked pages by crossing information coming from pages' descriptions.

Generating a glossary — As emerged from authors' experience and from past data sprints documentation, it happens that experiments share tools, techniques, and lexicon. For instance, the previously described *Snowballing* technique has been used in other experiments as well as almost all the participants relied on RAWGraphs for designing visualizations.

Thus, the format integrates a feature for building a glossary and referring to its items throughout the whole experiment (Figure 5 and Figure 13). In this case, the shared glossary allowed users to self-create entries, take advantage of those integrated by participants who have worked on other experiments, and access items to be informed and inspired.

4. Conclusions: A flexible and thorough format for disseminating data sprint results

The paper describes the peculiarities of a new dissemination format by design practitioners for Digital Methods data sprints results. As described in Section 1, the necessity to design the format emerges from the authors' experience in the field. Indeed, the format condenses and reinterprets good practices already in use in current formats for disseminating data sprints. Authors acknowledge that the format presented in the paper requires users to try outside their usual way of collecting, compiling, and presenting the process and results. As stated before, contextualizing the *research object*, and structuring the *experiment setup* requires additional research and readings. Describing the *steps* of the research process forces users to include succinct text and accompanying images. Moreover, it is necessary for users to become familiar with the concept of building blocks for filling it. Working with different *layouts* and ordering blocks requires a clear perception of the results that are being presented. However, as emerged from the focus group, the stress test and the application case, the format gives users room to be thorough, and it provides complete information while still maintaining flexibility. The suggestion to contextualize the research object situates the experiments in space and time. Providing information about the characteristics of the research object at the time the experiment was done becomes significant when imagining researchers reading the experiment in a couple of years. Thanks to the possibility of adding datasets, tables, and animated images, the research process has more space in the format. Explaining the steps using screenshots simulating the actual use of tools not only gives researchers the freedom to describe the research as thoroughly as possible but helps future researchers repeat the process. Creating a glossary characterizes the project, enriching it with specific terms and techniques which can gain the attention of even readers who never used some of the mentioned techniques.

To be remembered, brought back to life and re-interpreted, knowledge needs to be dismantled, re-structured and stored in a *place* (Anceschi, 1992, p. 99). Similarly, the format presented in the paper is a digital *place* that allows data sprints' results to be stored, staged, and accessed after the research activities flexibly and thoroughly. Indeed, the format is designed by versed practitioners who condensed all their knowledge and experience to make future dissemination of data sprints more accessible. The format, available as an open-source Content Management System (CMS), will generate web pages to publish the research results online⁹. Even if the case study presented is framed within the context of a data sprint inquiring the public perception of algorithms on social networks, the authors foresee that the usage scenario of the dissemination format could be extended to others DM data sprints and advanced research carried on with DM. Thus, the next steps will be dedicated to testing with (1) participants of another data sprint and (2) students carrying on DM research activities, (3) DM researchers outside data sprints and (4) researchers not familiar with DM.

Acknowledgements

The paper is one of the results of the *Algocount* project, a collaboration between the Department of Political Sciences (Università degli Studi di Milano) and the Density Design Lab (Politecnico di Milano, Design Department). The *Algocount* project has been founded by Fondazione Cariplo.

Authors would like to thank Alessandro Gandini, the scientific coordinator of the project and the researchers from the Political Sciences department involved: Alessandro Gerosa, Luca Giuffrè, Diletta Huyskes, Silvia Keeling and Urbano Reviglio.

Besides, authors would like to thank participants of the described data sprint: Antonella Autuori, Irene Avossa, Laura Bruschi, Ludovica Corponi, Alessandra Facchin, Emma Fortunati, Giulia Giorgi, Michele Grilli, Giovanni Lombardi, Vincenzo Luise, Andrea Elena Febres Medina, Riccardo Pronzato, Camilla Volpe.

References

- Anceschi, G. (1992). *L'oggetto della raffigurazione*. Etaslibri.
- Bounegru, L., Gray, J., Venturini, T., Mauri, M. (2018). *A Field Guide to "Fake News" and Other Information Disorders*. Zenodo.
- Ciuccarelli, P., Elli, T. (2019). Beyond visualization. In Hotson, H., Wallnig, T. (Eds.), *Reassembling the Republic of Letters in the Digital Age: Standards, Systems, Scholarship* (pp. 299–314). Göttingen University Press, Göttingen.
- Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:02016R0679-20160504&from=EN>
- Lupo, E., Gobbo, B., Lonardo, E. (2021). Towards a new design culture of scientific production – Innovating the formats of scientific publication of design. In *Design*

9. <https://github.com/densitydesign/placplac>

Culture(s). Cumulus Conference Proceedings Roma (pp. 1082–1097). Volume #2. .

- Mauri, M., Briones, Á., Gobbo, B., Colombo, G. (2020). Research Protocol Diagrams as Didactic Tools to Act Critically in Dataset Design Processes. In *INTED2020. Proceedings of the 14th International Technology, Education and Development Conference* (pp. 9034–9043). <https://re.public.polimi.it/retrieve/handle/11311/1132634/498643/2470.pdf>
- Mauri, M., Ciuccarelli, P. (2016). Designing diagrams for social Issues. Presented at the *Design Research Society Conference 2016*. <https://www.drs2016.org/185>
- Mauri, M., Elli, T., Caviglia, G., Uboldi, G., Azzi, M. (2017). RAWGraphs: A Visualisation Platform to Create Open Outputs. In *Proceedings of the 12th Biannual Conference on Italian SIGCHI Chapter*. ACM, Cagliari. <https://dl.acm.org/doi/10.1145/3125571.3125585>
- Mauri, M., Gobbo, B., Colombo, G. (2019). O Papel do Designer no Contexto do Data Sprint. In Omena, J. J. (Ed.). *Métodos Digitais. Teoria-prática-crítica*. Livros ICNOVA.
- Rittel, H. W., & Webber, M. M. (1973). Dilemmas in a general theory of planning. *Policy sciences*, 4(2), 155-169.
- Rogers, R. (2013). *Digital methods*. The MIT Press, Cambridge, Massachusetts.
- Rogers, R. (2017). Foundations of Digital Methods: Query Design. In M. T. Schäfer, & K. van Es (Eds.). *The datafied society: Studying culture through data* (pp. 75–94). Amsterdam: Amsterdam University Press.
- Venturini, T. (2014). Climaps by Emaps in 2 Pages. *SSRN Journal*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2532946
- Ziewitz, M. (2016). Governing Algorithms: Myth, Mess, and Methods. *Science, Technology, & Human Values*, 41(1), 3–16.



Visualizing an image network without rendering files: A method to combine user hashtags with computer vision labels

Visualización de una red de imágenes sin renderizar archivos: Un método para combinar hashtags de usuarios con etiquetas de visión computacional

 **Giulia Tucci**
giuliatucci@ufrj.br
Universidade Federal do Rio de Janeiro

Received: 13/01/2022

Accepted: 10/05/2022

ABSTRACT This article presents a method for visualizing networks of geolocated images without rendering the image files on the network. The path I followed to develop this method is the result of an intensive “data sprint” which took place during the University of Amsterdam Digital Methods Initiative Summer School 2021. During the data sprint, I developed a methodological framework to generate a network of Twitter geolocated images combining the hashtags tweeted with the images and the Google Cloud Vision API best single expression to describe each image (*best guess label*). Considering the limitations of working with a massive amount of image data and the computational memory required to generate network visualizations, the possibility of using description tags to create image networks is promising. The images analyzed during this study were extracted from Twitter filtering for the #deepfakes and #deepfake and tagged with country code location. Thus, the hashtags included in the tweets by Twitter users provide the context and the user description of the image. This information was combined in a bipartite network with a computer vision entity, the computer vision description of the image, to generate a networked description of the whole image set. I point that this method can be considered in exploratory research when working with large sets of images.

KEYWORDS Digital Methods, Cloud Vision API, Network Analysis, Twitter hashtags, Deepfakes

RESUMEN Este artículo presenta una metodología de visualización de redes de imágenes geolocalizadas sin renderizar los archivos de imagen en la red. El desarrollo del método, aquí descrito, resulta del “sprint de datos” intensivo logrado durante la *Digital Methods Initiative Summer School 2021* de la University of Amsterdam. Durante el *data sprint*, fue desarrollado un marco metodológico para generar una red de imágenes geolocalizadas de Twitter que combina los hashtags twitteados con las imágenes y la mejor expresión unitaria de Google Cloud Vision API a describir cada imagen (*best guess label*). Considerando las limitaciones al trabajar cantidades masivas de datos de imágenes y la memoria computacional requerida para generar visualizaciones de redes, el método promete la posibilidad de usar etiquetas de descripción para la creación de redes de imágenes. Las imágenes analizadas durante este estudio se extrajeron del filtrado de Twitter para #deepfakes y #deepfake, etiquetados a la ubicación del código de país. Así, hashtags incluidos en los tuits de usuarios de Twitter aportan el contexto y la descripción de la imagen por parte del usuario. Esta información se combinó en una red bipartita con una entidad de Computer Vision y la descripción de Computer Vision de la imagen, al fin de generar una descripción en red de todo el conjunto de imágenes. Señaló que el método es considerable a la investigación exploratoria en casos de grandes conjuntos de imágenes.

PALABRAS CLAVE Métodos digitales, API Cloud Vision, análisis de redes, hashtags, Twitter, Deepfakes

How to cite this article:

Tucci, G. (2022). Visualizing an image network without rendering files: A method to combine user hashtags with computer vision labels. *Dígitos. Revista de Comunicación Digital*, 8: 109-126. DOI: 10.7203/drdcd.v1i8.237





Visualizing an image network without rendering files: a method to combine user hashtags with computer vision labels

1. Introduction

In this article, I propose an exploratory method for the visualization of a network of geolocated images without the necessity to render the image files in the graph. To interpret the images without plotting the files, I built upon a recipe proposed by Chao and Omena (2021) to generate and analyze image networks without rendering the images within the network. My strategy was to combine the Computer Vision API output label that defines each image, as indicated in the recipe, with the hashtags posted by Twitter users to describe and contextualize the images. Therefore, I created a bipartite network where the edges that connect the hashtags nodes and the computer vision labels nodes represent the images. In situations where there is a small set of images to be processed and rendered as nodes within a network, there is no requirement for massive dedicated computer memory, and the execution of the task is relatively easy. However, as the volume of images comprising a study dataset increases, the process of showing images on a network becomes more difficult and computer memory-demanding. Although I developed this method during a study conducted with a small set of images, it is recommended to test it, validate it, and apply it in network analyses of a large set of image files.

The crescent profusion of Internet data produced by digital platforms users and of information generated by artificial intelligence tools associated with these platforms, as well as their application in multidisciplinary academic research, gave rise to an extensive demand for the development of research strategies, tools, and methods to extract, process, and visualize information of (sometimes) massive sets of data. That demand intensified a not recent debate concerning the choice for qualitative, quantitative, or mixed-methods approaches. In addition, the increasing “societal relevance of applications” raised the argument that computer science “needs to define itself as a socio-technical discipline that contributes to the solution of social problems in context” (Stevens et al., 2018: 23).

Digital Methods, a path of study for doing Internet research advanced by Richard Rogers (2009), repurposes traditional social sciences and humanities research methods and emphasizes data collection and analysis methods developed within the Internet medium and for Internet data research. In other words, the path pioneered by Rogers suggests that, instead of adapting traditional research methods that focus on collecting information from users (interview, survey, etc.), we study and learn from the features, affordances, and methods developed by the leading platforms (Google, Facebook, Twitter, Telegram, etc.). Digital Methods “seeks to move Internet research beyond the study of online culture and beyond the study of the users of ICTs [Information and

Communication Technologies] only” (Rogers, 2013: 17) and surpasses the qualitative or quantitative research dilemma as a research practice that reworks digital platforms data, web services mechanisms, and platforms’ structures to do Social Sciences research (Omena, 2019).

Every year, the University of Amsterdam’s Digital Methods Initiative (DMI) holds winter and summer data sprints. Participating in a data sprint is part of an intense and fast-paced learning experience. After being trained to apply research tools and techniques related to a project, the goal is to spend a couple of days using the data and the resources to identify relevant findings and help answer the proposed research questions.

Working with a small amount of time generates a need to be immersed in the data, testing different ways to explore the dataset and create visualizations. This process outcome is motivating and, in my case, was the inspiration for combining geolocation data, hashtags, and computer vision outputs as nodes within a network. Presenting the results and participating in group discussions oriented by experienced mentors were fundamental steps in perceiving that I could systematize the analyses I was conducting into an image analysis technique.

The 2021 Digital Methods Initiative Summer School happened virtually during the coronavirus pandemic from 5 to 16 July 2021. The DMI Summer School theme was Fake everything: Social media’s struggle with inauthentic activities, and the project I worked on during the first half of the program was Mapping deepfakes with digital methods and visual analytics, facilitated by Richard Rogers, Lucia Bainotti, Sarah Burkhardt, Gabriele Colombo, Janna Joceli Omena, and Jason Chao. We aimed to interpret the sociological and cultural understandings of deepfakes by using visual analyses and Digital Methods.

Designated as the “Photoshop of videos” by Bimber and Gil de Zúñiga (2020), deepfake technology was the subject of an alerting announcement created and published by BuzzFeed (Mack 2018) in which a fabricated Barack Obama makes a pronouncement voiced by an actor. The BuzzFeed piece that warned about the new technology danger circulated widely on social media (Vaccari and Chadwick 2020), generating concern in the general public regarding new ways of manufacturing content and contributing to an environment of uncertainty.

Initially employed by pornographic content creators to overlap celebrities’ faces in sexual content videos (Cole, 2017), deepfake is a technique to algorithmically create artificial videos by swapping faces starting from two people video footages (Westerlund, 2019). When this new technology emerged, it created a fuss concerning the falsification of discourses, sabotage activities, and national security matters (Chesney, Citron, 2018; Westerlund, 2019) since deepfake is a tool capable of distorting the boundaries of what seems real for disseminating disinformation.

The rise of deepfakes disquieted scholars resulting in works focused on deepfakes effects and ways of mitigating them. Kietzmann et al. (2020) created a framework for managing deepfakes risk and diminishing the technology’s dangerous effects. Vaccari and Chadwick (2020) conducted an experiment to check if the BuzzFeed Obama deepfake video had the potential to delude individuals and alter their perception of its

authenticity. The researchers did not find evidence suggesting that the manipulated video misled the participants. However, they advised that unchecked deepfake content can increase the levels of distrust in the online informational environment.

My group project goal was to understand whether the conversation on the Internet about deepfakes is related to threats on the informational environment or, despite the deceptive and damaging potential of this new technology for society, it focuses on describing the techniques used and on debunking the hype surrounding deepfakes. Thus we designed a research project to identify the related topics and the context of the conversation about deepfakes from 2017 to 2021.

The project was divided into three sub-projects since there were three main paths to cover: Discourse mapping, Image vernaculars and trends, and Computer vision. The Computer Vision team was divided again into three groups to examine the discourse related to deepfakes using Twitter and Google Cloud Vision data. The three sub-groups project titles were: Google Vision web entities for Google images over time, Sites of image circulation: Google images and Google Vision's fully matching images, and Geographical mapping: Twitter hashtags and Google Vision web entities.

During the first data sprint week, my working team was responsible for the Geographical mapping with Twitter hashtags and Google Vision entities study and formulated two research questions in order to geographically map the images used on conversations about deepfakes on Twitter: (i) Where do #deepfake or #deepfakes images circulate geographically on Twitter over time? and (ii) Are there country-specific discourses around deepfake? If yes, what are the themes found in these country-specific discourses?

The first research question was methodologically addressed by Dr. Martin Roth, with the development of a choropleth world map indicating deepfakes circulation across location and time (see Tucci, Roth, Saxler, 2021). I was responsible for exploring the data and conducting network analyses to answer the second research question. While discussing with my subgroup data sprint mentor, Janna Joceli Omena, about the partial results of the network analysis graphs I created, she enlightened me about the sophisticated way I had just encountered to represent image networks without the need to dedicate all computer memory to Gephi process image files. The solution I unfolded to define the images was to merge Twitter hashtags chosen by who tweeted to identify the image and a Computer Vision API outputs.

Gephi, a traditional software for network analysis, demands high dedicated computer memory to run. On Gephi's GitHub page, there are popular issues (Chezsick 2017; Ghost 2016; KallyopeBio 2018) related to memory crashes after users follow the software installation guide's specific instruction to increase memory default settings. Accordingly, the larger the number of elements contained in a network (nodes and edges), the more complex the tasks of generating, configuring, and saving the visualizations produced in Gephi. When plotting images in a network, depending on the number of files to be rendered as nodes, the work can be difficult to execute (or even be undoable). Doing visual research with digital platforms data requires specific methodologies, considering the importance of "approaching these collections of images as data" and not as content,

as proposed by Niederer and Colombo (2019) in an article that describes the process of creating visual tools for digital research.

To address these issues, there is a demand for establishing methodological frameworks and strategies that treat image content as data and require less execution time and less dedicated computer memory. Thus, this paper aims to outline the methodological procedure that I have developed to generate a bipartite network that describes a set of images without the necessity to render the image files in the graph. In the following section, the dataset creation and the methods applied to run the analyses are presented, followed by the results and discussion introduced in the third section. In the last section, final considerations and a research agenda are discussed.

2. Materials and Methods: geographical mapping of Twitter data & computer vision labels

This section first describes how to curate a dataset of Twitter hashtags. Then, it explains how to identify geolocated tweets containing images. Subsequently, the section demonstrates how digital methods recipes can serve as inspirational tools for new methodological experiments.

2.1. Dataset description

Before the beginning of the data sprint, the original dataset was collected via Twitter API by filtering tweets containing image files and the hashtags #deepfake or #deepfakes, from January 1st, 2017 to June 1st, 2021, and excluding retweets and replies. The acquired dataset comprises 98.831 tweets that originated 19.713 unique image URLs.

Hashtags are an established resource created on Twitter and used to tag and group content related to specific topics, events, entities, etc. Posting a hashtag on Twitter generates a hyperlink that points to an aggregate of tweets containing the tag, grouping content on that topic. Computer Vision APIs are image recognition technologies that allow researchers to process images using machine learning algorithms and OCR technology to extract information regarding the content of the analyzed image. In this work, I use Google Cloud Vision API as the source of a computer vision description to the image files in my dataset. The feature I chose to work with was the best guess label, which selects a single expression to describe an image, "the service's best guess as to the topic of the request image. Inferred from similar images on the open web", as Google (2020) defined it.

Considering the aim to study how conversations about deepfakes travel around the globe, the working dataset was created by combining the 459 geolocated tweets, the 459 associated image files, and the best guess label results for 416 images since the vision API did not process 43 video thumbnail files. The Google Cloud Vision API feature was obtained via Memespector GUI (Chao, 2021).

To explore the Twitter dataset, I summarized the hashtags related to the conversation about deepfakes by using the R package rtweet (Kearney, 2017). Since the data was

collected by filtering for the #deepfakes and #deepfake, they are the most frequent hashtags (234 and 232 occurrences, respectively), followed by #ai (63), #fakenews (36), and #protectmyimage (28). The ranking of most frequent hashtags indicate that the majority of images about deepfakes posted with geolocation on Twitter comprises conversations on fake news and artificial intelligence as well as discussions on the protection of social media profile pictures and other shared photos protection against manipulations and stealing (Graham et al., 2014).

2.2. Network analysis with computer vision

Based on a DMI recipe (Chao, Omena, 2021) on how to work with the results obtained after processing the images with a computer vision tool - the MemeEspector-GUI outputs - I followed this procedure to create the networks: I imported raw data to Table 2 Net tool, selected the type of nodes and respective attributions, and imported the network files generated into Gephi. Then, I processed the data to create each network visualization presented in this article, according to the respective objective of the analysis. The specific parameters used to generate each graph are described below.

3. Results and Discussion

3.1. Analysing a network of Twitter hashtags and images

In order to interpret and visualize the dataset of images, the first step was to plot a bipartite network graph of Twitter hashtags and images (Figure 1). To understand how users combine images and hashtags in a single tweet and how this information correlates to specific countries' discourses, I applied a color code to identify countries in the network. Since all 459 images are unique and tweeted from specific countries, I colored the image nodes and their output edges by the tweet country code. Bearing in mind the Twitter hashtags can repeat throughout the dataset and that every tweet contains the #deepfake and/or #deepfakes, hashtags nodes were colored white. In addition, the .jpg files corresponding to the images were rendered as nodes by using the ImagePreview Gephi Plugin (Xue, 2012).

According to Figure 1 there are hyperconnected hashtags shared by Twitter users from diverse countries (i.e. #fakenews, #ai, #deeplearning, and #journalism) and these tags tend to indicate a conversation about tools, methods and subjects used for the creation, characterization and/or debunking of deepfakes.

By zooming into the US (Figure 2) as it is the country with the most locational data (168 images) and the country where dominant discourses appear, whereas there are some images that refer to conferences and academic presentations, the more dominant trend related to the discourse of deepfakes are pop culture and celebrities and commercial brands.

Aiming to explore images' visual elements and to add another visual component to the study of the US discourse surrounding deepfakes, I conducted a visual analysis sorting the 168 US images by color (Figure 3) using PicArrange software (Jung, 2021). The analysis of the US tweeted images associated with #deepfakes or #deepfake provides a sense of image trends and vernaculars (Rogers, 2021). Although it is not possible to identify

3.2. Analysing a network of Twitter geolocated images and best guess labels

To test the existence of country-specific Google vision vernaculars related to the deepfake debate on Twitter, the strategy I have chosen was to link tweeted images' geolocation data with Google Cloud Vision's automated description of the images. Hence the second step of the research process was to generate a bipartite network to relate the Cloud Vision API feature of each image to the country code returned by Twitter API (Figure 4).

In Figure 4, nodes that represent best guess labels were colored 'light blue', and country flags images were rendered as the respective country node. The edges represent the images. All nodes were sized by occurrence number, and the labels were scaled proportionally. In addition, I have colored the nodes and the edges of the three most frequent countries in our data: United States (168 images, color = royal blue), Great Britain (45 images, color = green), and France (32 images, color = red). The graph was generated with country flags files obtained in a GitHub flags repository (Borgos, 2021).

The network connecting the computer vision best guess label and the image geolocation shows specific clusters of Google Cloud Vision labels for different countries, demonstrating that the images tweeted from diverse locations were tagged by Cloud Vision API with specific groups of visual entities. In other words, the majority of Google visual vernaculars for #deepkafes and #deepkafes are country-specific. Taking into account that platforms vernaculars are shaped by platforms mediated practices and users habits of communication (Gibbs et al., 2015) and that visual vernaculars are patterns of images used to articulate and communicate an issue (Pearce, Colombo, 2019), Figure 4 indicates that Twitter users from different countries have particular ways of communicating about the deepfakes issue. However, some best guess labels appear in the content tweeted in more than one country, being the most frequent Cloud Vision labels in this data: presentation, communication, display device, and seminar.

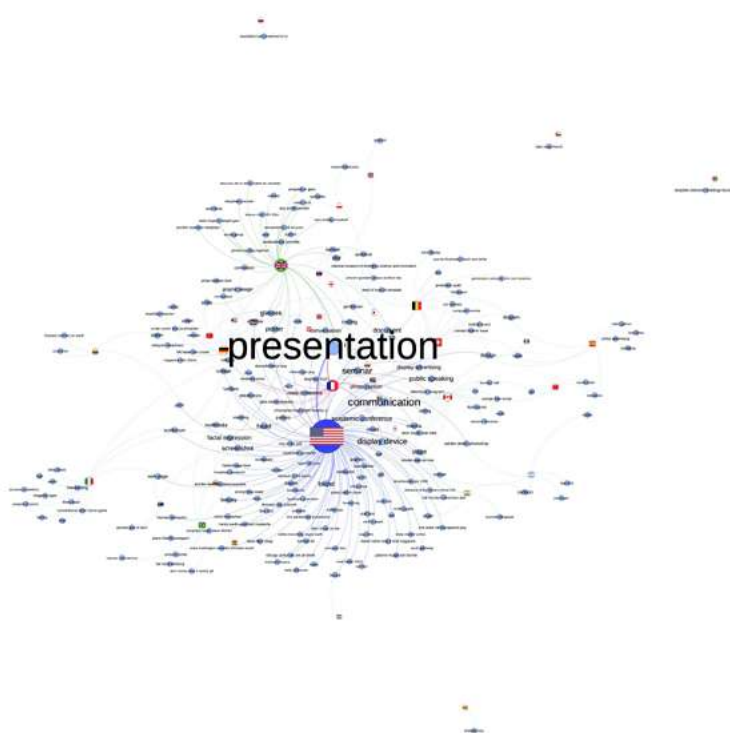


Figure 4: Bipartite network graph of Cloud Vision best guess label of 416 images and 41 countries where the images were tweeted, represented as 246 nodes and 328 directed edges. Network visualization was created using Gephi's Force Atlas 2 layout, nodes were sized by degree; Cloud Vision labels were sized by frequency; country nodes and edges were colored (US = royal blue, France = red, Great Britain = green) and the respective county flag .jpg file was rendered with Gephi ImagePreview Plugin. Best guess label nodes were colored light blue.

An important consideration about the relevance of the context provided by the Twitter hashtags, is that without this information and based only on the Cloud Vision tags, this data would not be understood as a group of images shared under the topic #deepfakes. Even with the removal of the #deepfake and #deepfakes, the other hashtags tweeted in combination provide context locating the images under the 'deepfake' topic. For example, #protectmyimage and #fakecelebs indicate that the theme relates to fake content about celebrities and a campaign on the protection of social media profile pictures and other shared photos. In Figure 7 there are other hashtags that give context, like #fakebrands, #fakealgorithms, and #facialrecognition, among others. Although there are specific platform vernaculars for different platforms, the Twitter hashtags can be considered an example of a form of expression established by convention through the community (Gibbs et al., 2015) and they are an appropriate fit to give context to images.

Analyzing the bipartite network graphs (Figures 1-3, and 5-7), it is possible to point out that the discourse constructed surrounding deepfakes has a similar structure in computer vision API images' description and in users' choice of hashtags. In this way and based on the existence of specific vernaculars for images shared on Twitter in specific countries, I have noticed that it is possible to adopt this methodological strategy when trying to understand a (very large) set of images that are geographically located. In other words, these findings indicate that computer vision APIs can be used to describe a set and extract relational structures and common topics of these images without having to render and visualize them. Google Cloud Vision best guess labels give meaning to images while hashtags provide context to the pictures (Cloud Vision API did not return any web entity result relating the images to deepfake).

3.4. Using Twitter's geolocated images and Google Vision best guess label as networks: a method protocol

I have proposed a methodological strategy (shown in Figure 8) to use network analysis in the description of geolocated images utilizing user description data (Twitter hashtag) and computer vision API data (best guess label Cloud Vision API classification).

The interpretation of the images-hashtags geolocated network should be carried out by mixing social and computational elements. As Marres (2020) indicates, in the study of situations in computational settings, the researcher should consider that social processes happening in digital environments are influenced by the technical specificities of the medium. Therefore, I propose that the methodological strategy schematized in Figure 8 applies to interpreting and situating geolocated Twitter images. Combining image artificial intelligence descriptions with Twitter hashtags contemplates "both the visual elements of an image and the contextual elements encoded through the hashtag practices of networked publics" (Geboers and Van De Wiele 2020). From a Digital methods perspective, researchers developing visual methodologies to extract information from digital platforms data shall create tools that facilitate the interpretation of this data and produce "visualizations that make complex information legible and ready for further analysis" (Niederer and Colombo 2019).

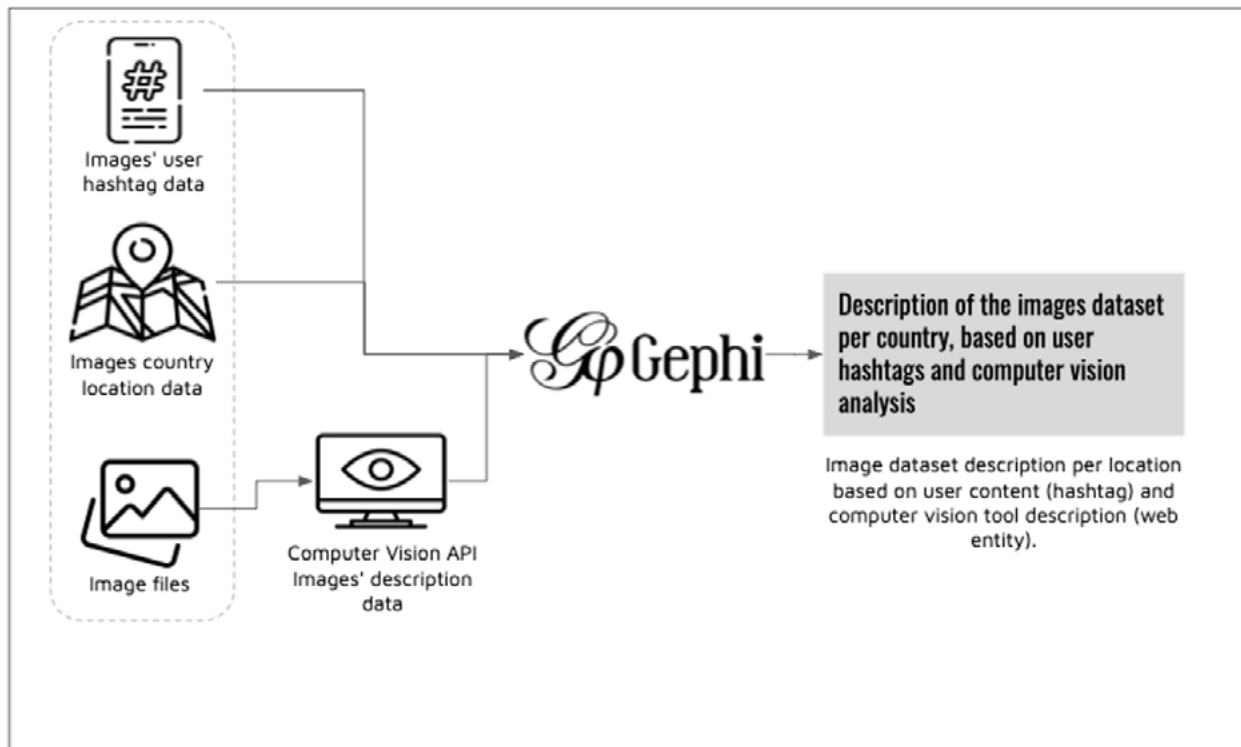


Figure 8: Schematic methodological strategy developed to use network analysis in the description of geolocated images utilizing user description data (Twitter hashtag) and computer vision API data (best guess label Cloud Vision API classification).

4. Limitations

During the development of this work, I have considered several limitations. I generated the network analyses presented in this article starting from a dataset that contained scarce geolocated data related to the debate about deepfakes on Twitter. Considering a total of 19.713 unique images extracted from the platform, only 459 images were tagged with their respective country code metadata. Although the small amount of data available was the motivation for the development of the method presented in this article, a small set of images could be analyzed qualitatively for a more robust interpretation and to give more precise answers to the research questions.

In addition, since our sample was narrowed to unique country-coded tweeted images, the working dataset comprises a small number of images shared during four years (2017 - 2021). This lack of robust data can bias the results of what type of discourses were happening about deepfakes in every location and over these years. The absence of geographical information on the majority of tweets, nevertheless, is an important disadvantage since “the attributes of language and location are crucial for understanding the geographies of online flows of information” (Graham et al., 2014). For the dataset used in this study, only 2,3 % of tweets contained country information and it is unlikely that they form a representative sample of the broader universe of content.

Another reason for interpreting the findings related to specific country-vernaculars with attention is that the dataset is biased towards particular language spaces, potentially related to the search terms #deepfakes and #deepfake. When analyzing Twitter

hashtags it is important to consider that the data can contain slang and abbreviations (Kim et al., 2013) and/or hashtags not related to the topic (i.e. call-for-action hashtags) since these are elements that are often present on user-generated content.

Furthermore, despite the visual consistency provided by computer vision (Pearce, Gaetano, 2019), an aspect of adopting artificial vision techniques is the loss of specificity in the description of the image versus the qualitative approach, which (maybe) should be more appropriate to analyze a small dataset. By admitting that reading images via different visual APIs can generate diverse outcomes, Mintz and Silva (2019) tested Google, Microsoft, and IBM computer vision services. They verified that Google Vision API tends to be more specific when describing details. However, Google Cloud Vision failed to annotate relevant elements in a war context, for example labeling a picture of a human corpse as a 'zombie' (Geboers and Van De Wiele 2020). The existence of bias in image interpretation by Google Cloud Vision or every other Computer Vision service must be taken into account. These automated annotation tools can (and probably do) reproduce cultural stereotypes and generalizations, as shown by Silva et al. (2020).

Technical limitations to be considered when using Google's automated image reading are the inconsistency of Cloud Vision API results when analyzing noisy images (Hosseini, Xiao, and Poovendran 2017) and the susceptibility of the service to black-box attacks (Brunner et al. 2019; Ilyas et al. 2018), ways of altering neural network-based classifiers by insistently querying the system with a manipulated input information.

Addressing the socio-cultural bias tendency and the technical related issues is desirable to achieve more consistent automated image reading results. A possibility is to use a combination of computer vision services to interpret the results. Finally, there is the need to use larger datasets to apply and validate this method of reading images in bipartite networks without rendering the files.

5. Final considerations and future agenda

The environment of intensive hands-on research maintained in data sprints provides the opportunity for the development of new research strategies and for adapting and studying new forms of applying data extraction and analysis tools and techniques. The ideas exchanges, the interdisciplinary discussions, and collaborative efforts stimulate discoveries and the creation of knowledge.

The research developed by the sub-group I participated in during the DMI Summer School 2021 data sprint aimed to combine Twitter and computer vision data to observe how the debate about deepfakes on Twitter traveled the world over time. So, after following the methodological steps described in this paper and analyzing the obtained results, we were able to infer that deepfakes related content has spread around the world and is concentrated in the United States. The approach we adopted was able to expose location-based - and potentially cultural and/or language-based - differences and similarities in the discourses surrounding deepfakes. Even considering the loss of specificity and nuances as well as the vagueness of semantic interpretation when relying on computer vision techniques, the combination of Cloud Vision labels with

user-created data and Twitter metadata (geolocation) added a semantic layer to our analysis.

We have found that the images representing the conversation about the deepfake topic on Twitter link specific themes and actors, considering a geographic location. When analyzing the US results, the ensemble formed by pictures of celebrities and hashtags like #fakecelebs, #fakenewsmedia, and #stalkarazzi emerge, which is consistent with other research. Westerlund (2019) argues that celebrities are easy targets for fabricated content since they have an abundance of free video, image, and audio content circulating over the internet and that content availability facilitates deepfakes creation. Dasilva et al. (2021) analyzed the debate around deepfakes on a Twitter network. They have shown that journalists and media are bridging information nodes and that celebrities figures are among the most referenced and virialized users in the networks.

The Cloud Vision labels and countries' bipartite network analysis indicates users tweet images related to deepfakes that were described by the computer vision tool as specific vernaculars for different geographical locations. Coming out of this conclusion, I developed a methodological framework for reading networks of images without rendering the files in the graph. In other words, I propose that combining the image description and context given by a digital platform user (Twitter hashtags) and the image definition given by a computer vision API (Google Cloud Vision best guess label) is an efficient way to characterize geolocated images. Applying this method can serve as a solution for computational memory issues while creating networks of massive sets of images since the network analysis software (in my case, Gephi) can experience computer memory-related difficulties managing large volumes of data. This methodological path can be thought out for exploratory research since it could generate insights and ideas serving as a starting point to additional analyses.

As I mentioned before, further testing and validation of this methodological path with larger datasets are required. Another future agenda raised by this study is to look beyond Twitter and apply this method with data extracted from other platforms since "the affordances and performances that constitute a vernacular are not necessarily specific to a platform" (Gibbs et al., 2015). For example, an application for this method could be the analysis of a large set of pictures shared on Instagram in a country-specific context by combining hashtags users post as a description of the images with computer vision retrieved image information. Finally, since some techniques allow combining computer vision services interpretations in network analysis (Silva et al. 2020), future experiments merging images annotations from multiple sources are encouraging to achieve more robust results and reduce bias.

Acknowledgments

I am grateful to all those who have contributed to the development of this work. I thank Professor Richard Rogers for all the excellent work developed by the Digital Methods Initiative, for conducting this research experience memorably, and for sparking my interest in this topic; Lucia Bainotti, Sarah Burkhardt, Gabriele Colombo, Janna Joceli

Omena, and Jason Chao for facilitating the Mapping deepfakes with digital methods and visual analytics project, for their assistance during every step of the work, and for stimulating rich discussions and debates during the DMI Summer School; Martin Roth and Flavia Saxler for accepting the challenge of participating in the Geographical mapping with Twitter Hashtags and Google Vision entities sub-group, for all the knowledge exchanged during the process, and for all the hard work we have done together. In addition, I would like to thank the two anonymous reviewers as well as this special edition editors for their valuable suggestions on the manuscript.

References

- Bimber, B.; Gil de Zúñiga, H. (2020). The Unedited Public Sphere. *New Media & Society*, 22(4):700–715. doi: 10.1177/1461444819893980.
- Borgos, J. B. (2021). country-flags Available at: <https://github.com/hampusborgos/country-flags> (Accessed: 08 Jul. 2021).
- Brunner, T.; Diehl, F.; Truong Le, M.; Knoll, A. (2019). “Guessing Smart: Biased Sampling for Efficient Black-Box Adversarial Attacks”. Paper presented at the 2019 *IEEE/CVF International Conference on Computer Vision (ICCV)*. Seoul, South Korea. doi: 10.1109/ICCV.2019.00506.
- Chao, T. H. J. (2021). *Memespector GUI: Graphical User Interface Client for Computer Vision APIs (Version 0.2)* [Software]. Available from <https://github.com/jason-chao/memespector-gui>. (Accessed: 20 May 2022).
- Chao, J.; Omena, J.J. (2021). *Networks of Image Description*. Available at: <https://docs.google.com/document/d/1ccBH691tW-6jVTrjQtMjd2aDTalZollbRBtMqlGISlg/> (Accessed: 05 Dec. 2021).
- Chesney, R.; Citron, D.K. (2018). Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security. *SSRN Social Science Research Network*. Rochester, NY. doi: /10.2139/ssrn.3213954.
- Chezsick (2017). “Gephi 0.9.1 Can Never Increase the Memory · Issue #1705 · Gephi/Gephi.” *GitHub*. Retrieved April 8, 2022 (<https://github.com/gephi/gephi/issues/1705>).
- Cole, S. (2017). AI-Assisted Fake Porn Is Here and We’re All Fucked. *Vice*. Available at: <https://www.vice.com/en/article/gdydym/gal-gadot-fake-ai-porn> (Accessed: 05 Dec. 2019).
- Dasilva, J. P.; Ayerdi, K. M.; Galdospin, T. M. (2021). Deepfakes on Twitter: Which Actors Control Their Spread?. *Media and Communication*, 9(1): 301–312.
- Geboers, M. A.; Van De Wiele, C. D. (2020). Machine Vision and Social Media Images: Why Hashtags Matter. *Social Media + Society*, 6(2). doi: 10.1177/2056305120928485.
- Gibbs, M.; Meese, J.; Arnold, M.; Nansen, B.; Carter, M. (2015). #Funeral and Instagram: death, social media, and platform vernacular”. *Information, Communication & Society*, 18(3), 255–268. doi: 10.1080/1369118X.2014.987152.

- Ghost (2016). "Gephi Memory Issue · Issue #1609 · Gephi/Gephi." *GitHub*. Retrieved April 8, 2022 (<https://github.com/gephi/gephi/issues/1609>).
- Google (2020). *Google.Apis.Vision.v1.Data.WebDetection Class Reference*. Available at: https://developers.google.com/resources/api-libraries/documentation/vision/v1/csharp/latest/classGoogle_1_1Apis_1_1Vision_1_1v1_1_1Data_1_1WebDetection.html#afddf64d68931805b23d9c863dd47b9c7 (Accessed: 05 Dec. 2021).
- Graham M.; Hale S. A.; Gaffney D. (2014). Where in the World Are You? Geolocation and Language Identification in Twitter. *The Professional Geographer*, 66(4): 568-578, doi: 10.1080/00330124.2014.907699.
- Ilyas, A.; Engstrom, L.; Athalye, A.; Lin, J. (2018). "Black-Box Adversarial Attacks with Limited Queries and Information." In *Proceedings of the 35th International Conference on Machine Learning* (pp. 2137–46). PMLR. <https://proceedings.mlr.press/v80/ilyas18a.html>.
- Jung, K. (2021). PicArrange -- Visually Sort, Search, and Explore Private Images on a Mac. arXiv:2111.13363
- KallyopeBio (2018). "Gephi UI Nearly Unresponsive with Large Network (100% CPU Utilization on One Core) · Issue #1947 · Gephi/Gephi." *GitHub*. Retrieved April 8, 2022 (<https://github.com/gephi/gephi/issues/1947>).
- Kearney, M.W. (2017). rtweet. Available at: <https://github.com/mkearney/rtweet.download> (Accessed: 23 Mar. 2018)
- Kietzmann, J.; Lee, L.W.; McCarthy, I. P.; Kietzmann, T. C. (2020). Deepfakes: Trick or Treat?. *Business Horizons*, 63(2):135–46. doi: 10.1016/j.bushor.2019.11.006.
- Kim, A.E.; Hansen, H.M.; Murphy, J.; Richards, A.K., Duke, J.; Allen, J.A. (2013). Methodological Considerations in Analyzing Twitter Data. *JNCI Monographs*, 47: 140–146. doi: 10.1093/jncimonographs/lgt026.
- Mack, D. (2018, April 17). "This PSA About Fake News From Barack Obama Is Not What It Appears." *BuzzFeed News*. Retrieved April 8, 2022 (<https://www.buzzfeednews.com/article/davidmack/obama-fake-news-jordan-peelee-psa-video-buzzfeed>).
- Marres, N. (2020). For a Situational Analytics: An Interpretative Methodology for the Study of Situations in Computational Settings. *Big Data & Society*, 7(2). doi: 10.1177/2053951720949571.
- Mintz, A.; Silva, T. (2019). *Interrogating Vision APIs*. Lisboa: Universidade Nova de Lisboa | NOVA FCSH | iNOVA Media Lab.
- Niederer, S.; Colombo, G. (2019). Visual Methodologies for Networked Images: Designing Visualizations for Collaborative Research, Cross-Platform Analysis, and Public Participation. *Diseña*, 14:40–67. doi: 10.7764/disena.14.40-67.
- Omena, J.J. (2019). Introdução: O que são Métodos Digitais? In Janna Joceli Omena (Ed.), *Métodos Digitais: Teoria-Prática-Crítica*. Lisboa: ICNOVA.
- Pearce, W.; Colombo, G.; De Gaetano, C. (2019). Using computer vision to see Google's

visual vernacular of climate change (2008-19). <https://research.hva.nl/en/publications/using-computer-vision-to-see-googles-visual-vernacular-of-climate>.

Rogers, R. (2009). *The end of the virtual: Digital Methods*. Inaugural lecture, University of Amsterdam. doi: 10.5117/9789056295936.

Rogers, R. (2021). Visual media analysis for Instagram and other online platforms. *Big Data & Society*. doi: 10.1177/20539517211022370.

Silva, T.; Mintz, A.; Omena, J.J.; Gobbo, B.; Oliveira, T.; Takamitsu, H. T.; Pilipets, E.; Azhar, H. (2020). Investigando mediações algorítmicas a partir de estudo de bancos de imagens. *Logos*, 27(52):30. doi: <https://doi.org/10.12957/logos.2020.51523>.

Stevens, G.; Rohde, M.; Korn, M.; Wulf, V. (2018). Grounded Design: A Research Paradigm in Practice-Based Computing. In V. Wulf; V. Pipek; D. Randall, M. Rohde; K. Schmidt; G. Stevens (Eds): *Socio Informatics –A Practice-based Perspective on the Design and Use of IT Artefacts*. Oxford University Press, Oxford. doi: 10.1093/oso/9780198733249.003.0002.

Tucci, G.; Roth, M.; Saxler, F. (2021). Geographical mapping: Twitter hashtags and Google Vision web entities. *Digital Methods Initiative*. Available at: https://wiki.digitalmethods.net/Dmi/WinterSchool2021Deepfakes#A_5.3.3._Geographical_mapping:_Twitter_hashtags_and_Google_Vision_web_entities (Accessed: 05 Dec. 2021).

Vaccari, C.; Chadwick, A. (2020). Deepfakes and Disinformation: Exploring the Impact of Synthetic Political Video on Deception, Uncertainty, and Trust in News. *Social Media + Society*, 6(1):205630512090340. doi: 10.1177/2056305120903408.

Westerlund, M. (2019). The Emergence of Deepfake Technology: A Review. *Technology Innovation Management Review*, 9(11): 40–53. doi: 10.22215/timreview/1282.

Xue, C. (2012). *gephi-plugin-image-preview*. Available at: <https://github.com/chrisxue815/gephi-plugin-image-preview> (Accessed: 08 Jul. 2021).



Exploring User Engagement with Portuguese Political Party Pages on Facebook: Data Sprint as Workflow

Explorando el *engagement* de los usuarios con las páginas de los partidos políticos portugueses en Facebook: El *data sprint* como flujo de trabajo

- | | |
|---|---|
| <p>Jorge Martins Rosa
dedalus.jmmr@gmail.com
ICNOVA, Universidade Nova de Lisboa</p> | <p>N. Gizem Bacaksizlar Turbic
gizem.bacaksizlarturbic@gegis.org
GESIS - Leibniz Institute for the Social Sciences</p> |
| <p>Alda Magalhães Telles
aldatelles@fcsh.unl.pt
ICNOVA, Universidade Nova de Lisboa</p> | <p>Clara González Tosat
cgonzalez@unav.es
Universidad de Navarra</p> |
| <p>Cristian Jiménez Ruiz
cristian_camilo_ruiz@iscte-iul.pt
ISCTE-IUL & Lisbon University</p> | <p>Kalliopi Moraiti
kalliopimor22@gmail.com
University of Gothenburg</p> |
| <p>Oğuz Özgür Karadeniz
oguzozgur.karadeniz@kuleuven.be
KU Leuven DTAI</p> | <p>Valentina Pallacci
vpallacci@gmail.com
Politecnico di Milano</p> |

Received: 12/01/2022 | Accepted: 15/05/2022

ABSTRACT In the broader context of a research project about political participation on Facebook during the Portuguese electoral year of 2019 - for the European Parliament in May, and for the National Parliament in October -, a dataset of more than 9,000 posts and corresponding reactions and comments had been previously retrieved using Facepager. Participating in a data sprint came up as an opportunity to explore some specific questions in an intensive and time-bound context, using a subset of the data tailored to the purpose of discovering: 1) how differently official parties tried to engage with Facebook users before, during, and after the campaigns for the major elections periods (“supply side”), and 2) what type of content received the most attention and engagement (“demand side”). Our results show that regardless of the party’s age and position in the political spectrum, the persistent main categories are “National Politics”, “Fundamental Rights”, and “Self-Promotion” for both elections. Also, we found that posts with images attract more attention, and apparently this may be leveraged if the textual content of the post is in topics of “National Politics”. However, this finding requires further investigation. Along with details about the research during the data sprint and the main findings, this paper is also a testimony about the singularities and learnings of a process built upon the constraints of taking data sprints situation as a workflow.

KEYWORDS Facebook, political engagement, elections, political parties, data science

How to cite this article:

Martins Rosa, J.; Bacaksizlar Turbic, N. G.; Magalhães Telles, A.; González Tosat, C.; Jiménez Ruiz, C.; Moraiti, K.; Özgür Karadeniz, O. & Pallacci, V. (2022). Exploring User Engagement with Portuguese Political Party Pages on Facebook: Data Sprint as workflow. *Dígitos. Revista de Comunicación Digital*, 8: 127-154. DOI: 10.7203/drdcd.v1i8.233

RESUMEN En un contexto amplio de un proyecto de investigación sobre participación política en Facebook durante el año electoral portugués de 2019 - para el Parlamento Europeo en mayo, y para el parlamento Nacional en octubre -, un *dataset* de más de 9.000 posts y sus correspondientes reacciones y comentarios fue previamente recolectado usando Facepager. A través de la participación en un *data sprint*, se tuvo la oportunidad de explorar cuestiones específicas en un contexto intensivo de tiempo limitado. Mediante el uso de un *subset* de los datos mencionados, se plantearon las siguientes preguntas de investigación: 1) ¿Cómo fueron las estrategias de los partidos políticos para interactuar e involucrar a los usuarios antes, durante y después de las campañas? (“oferta”); y 2) ¿Qué tipo de contenidos recibieron más *engagement*? (“demanda”). Nuestros resultados muestran que, independientemente de la antigüedad de los partidos y sus posturas en el espectro político, las categorías más populares respecto a la temática de las publicaciones son “National Politics”, “Fundamental Rights”, and “Self-Promotion” para ambas elecciones. Además, se aprecia en un análisis preliminar que las publicaciones con imágenes atraen más la atención, y aparentemente, esto puede ser aprovechado al máximo si el contenido en texto de la publicación es, sobre todo, acerca de temas de la categoría “National Politics”. Este primer contacto con el conjunto de datos, junto con otros detalles acerca de la investigación durante el *data sprint* y sus principales hallazgos presentan un testimonio sobre las singularidades y aprendizajes de un proceso construido sobre la base de utilizar el *data sprint* como flujo de trabajo.

PALABRAS CLAVE

Facebook, *engagement* político, elecciones, partidos políticos, data sprint, data science



Exploring User Engagement with Portuguese Political Party Pages on Facebook: Data Sprint as Workflow

1. Introduction

With the continuous rise in the use of online communication by the general public, political parties became increasingly present on social media platforms (Macnamara, 2011; Enli and Skogerbø, 2013; Nahon, 2015; Kalsnes, 2016; Vogels *et al.*, 2021). The online conversations and interactions on social media provide an exhaustive source of data on political engagement, opinion dynamics, and online community structures compared to the traditional sources, such as interviews and surveys, allowing new research approaches fueled by tools and methods native to the digital universe (for a comprehensive and still relevant systematic review, cf. Kapoor *et al.*, 2018). In the field of political communication, a very significant part of the existing literature is focused on how political actors such as parties and candidates use social media to reach out to the electorate (Effing *et al.*, 2011; Bennett, 2012; Larsson and Kalsnes, 2014; Lilleker *et al.*, 2015). Campaign periods tend to be considered particularly relevant time frames for analysis (Enli and Skogerbø, 2013; Magin *et al.*, 2017; Klinger and Russman, 2017), though more longitudinal studies also have their place in the literature (Larsson, 2016; Heiss *et al.*, 2019; Serra-Silva *et al.*, 2018).

A common dispute underlining some studies is related to the normalization vs. equalization hypotheses. The first (Klinger, 2013; Ramos-Serrano *et al.*, 2018) states that

major parties already rely on a stronger presence through mainstream media. This allows them to be more conservative in the usage of social media without compromising their primacy. The conflicting “equalization hypothesis” argues that smaller or newer parties can (and do) take advantage of these platforms to counterbalance and eventually overcome an otherwise unfavorable position (Gibson and McAllister, 2014; Samuel-Azran, 2015; Bene, 2021). Another research trend, keeping up with the perceived ascension of populist and “alt-right” movements (Engesser *et al.*, 2017; Alonso-Muñoz and Casero-Ripollés, 2020), tries to make sense of the phenomenon by identifying features in the communicative style that may potentiate this rise.

The rationale for our research was that the communicative interchange that takes place in social media, however asymmetrical it may be, needs to consider both the posts and their content (“supply-side”) and the engagement affordances allowed to the users (“demand-side”). Our main objective was thus to investigate how this communicative exchange took place on Facebook, to have a more accurate description of what happened both on the supply and the demand sides before, during, and after the European and Legislative Elections in Portugal in 2019. The activity on the supply side can be measured by the number of posts from the page owners, while from the demand side it can be measured by the number of comments, shares, and reactions (at the time, *Like, Love, Haha, Wow, Sad or Angry*) from users. These interactions, commonly explained and analyzed as user engagement (Effing *et al.*, 2011, Stetka *et al.*, 2019; Sobacı and Hatipoğlu, 2020), are the major metrics used in the literature.

While these allow for a quantitative assessment of how common users interact with the posts, they may fail in providing a more in-depth explanation of those behaviors, which calls for some additional strategical and methodological approaches. Another objective was thus to investigate what this interchange *actually* means, targeting the content of posts and comments through diverse methodological strategies.

In this study, we evaluate how official parties tried to engage with Facebook users during the two major 2019 elections in Portugal, and which type of content received the most attention, two complementary sides of the communicational exchange which are commonly referred to in the literature respectively as “supply-side” and “demand-side” (Norris, 2003; Gibson, 2012; Theocaris *et al.*, 2016; Xenos *et al.*, 2017). At the same time, it will serve as a testimony about the singularities and learnings of a research process built upon the constraints of a data sprint.

The main research questions that this paper seeks to answer are the following:

RQ1: How differently do official parties try to engage with Facebook users in the two major 2019 elections periods in Portugal?

RQ2: What type of content receives the most attention and engagement from the users?

2. Methodology

2.1. Data context

In the broader context of a research project (PINBook PT: Political Interest Networks in Facebook Portugal) about political participation on Facebook during the Portuguese electoral year of 2019 – for the European Parliament in May, and for the National Parliament in October – a total of more than 9,000 posts and corresponding reactions and comments had been previously retrieved using Facepager. This dataset, spanning a period of 18 months (from July 2018 until the end of 2019), contains the posts from the official pages of the Portuguese parties with parliamentary representation during the legislature that was sworn in after the latter electoral act.

Such an amount of data was indispensable for the project's main objectives – the analysis of political participation through Facebook, from official and non-official actors –, but also inspired more specific research questions. Participating in a data sprint came up as an opportunity to explore some of those questions in an intensive and time-bound context, while at the same time connecting with people outside the project and from different backgrounds, constituting an *ad-hoc* group gathering disparate researchers and skills towards a common goal.

2.2. Corpus

To attain both objectives, while at the same time keeping in mind the tenets of current research (which are far from exhausting the field), we have opted to analyze the two electoral periods of 2019 (European and Legislative campaigns) and select two samples of posts for each period: those with the highest level of interactions, and those with the highest level of “other reactions” (i.e., reactions other than *Like* or *Love*). This latter approach takes advantage of Facebook's current availability of multiple reactions beyond the *Like* button, which, to a certain extent, attenuate the inherent ambiguity of the original reaction and allow a more nuanced expression of meanings and emotions. Though the connection between emotions and reactions is present in the literature (Hughes and van Kessel, 2018; Eberl *et al.*, 2020), this procedure was, as far as the authors have examined, never attempted in previous studies.



Figure 1. Mapping of political parties (political spectrum and foundation time).

For the comparative analyses, the political parties were mapped (Figure 1) according to their foundation time (“older” if founded before 2000 vs. “newer” if founded after that date), and political position (left vs. right). As the sprint developed, other specific data – namely the images embedded in the posts and the emojis within posts and comments – were gathered.

2.3. Query Design and Visual Protocol

For this study, posts from the Portuguese political parties' public pages were collected using Facepager, a tool for fetching publicly available data from various social media platforms (Jünger and Keyling, 2019).

The complete dataset, spanning 18 months of posts from the pages of the Portuguese parties represented in the parliament elected in 2019 (cf. Appendix I: List of political parties) and corresponding comment trees, consists of more than 9,000 posts and more than 100,000 comments, along with their engagement metrics. From that main dataset, the subsets of 100 posts each were retrieved for the periods of the two major

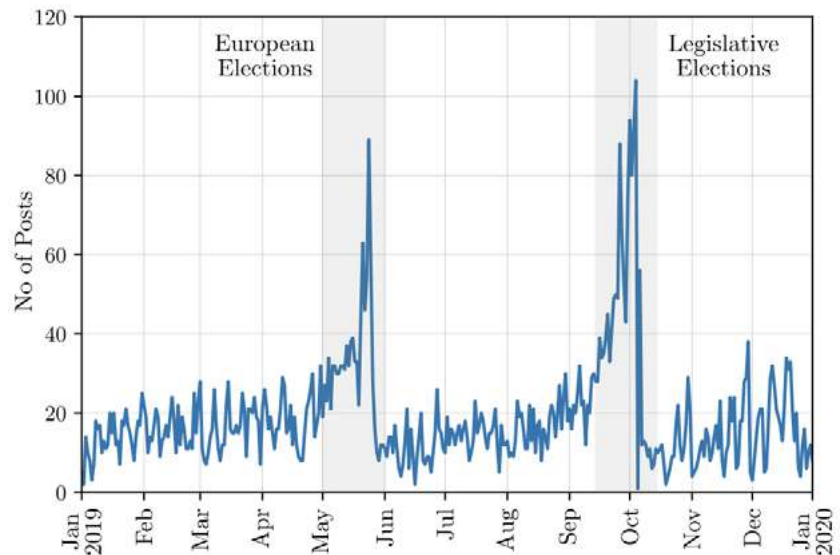


Figure 2. Daily post activity over time (all pages/parties).

elections and related electoral campaigns. Two periods (April to June and September to November 2019) were chosen, to cover two months before each election, which include both the official campaign time and the “pre-campaign” time, and the month after. These periods also correspond to those with higher activity from the supply side, i.e., the higher values of published posts per day by the political parties on their Facebook official pages, as shown in Figure 2.

Within these periods, we then selected the top 100 posts for each with 2 filters, as summarized in Table 1. A similar approach was taken for the top 100 comments.

	European Elections (2 months before, including 15 days of campaign + 1 month after)	Legislative Elections (2 months before, including 15 days of campaign + 1 month after)
Posts ranked by “other reactions” (<i>Haha, Wow, Sad, Angry</i>) [Used for all analysis]	Top 100 posts + Top 100 comments within those posts	Top 100 posts + Top 100 comments within those posts
Posts ranked by total interactions (reactions + comments + shares) [Used only for visual and sentiment analysis]	Top 100 posts + Top 100 comments within those posts	Top 100 posts + Top 100 comments within those posts

Table 1. Datasets.

Figure 3 gives a graphical representation of the visual protocol used for the data sprint.

Study Design – Visual Protocol


Platform	Dataset/Sub-datasets	Extraction	Analysis	Visualization
Facebook 	Posts and comments	Facepager	<ul style="list-style-type: none"> • Text Analysis Manual Coding • Text Sentiment analysis NLTK Vader 	RAWGraphs
	Images within the posts	Python urllib library	<ul style="list-style-type: none"> • Python • Memespector GUI • Image Content analysis 	<ul style="list-style-type: none"> • Imagesorter • Gephi
	Emojis within posts and comments	SpaCy		RAWGraphs

Figure 3. Visual Protocol.

The data sprint lasted a total of five days, leaving approximately two and a half days for actual research, plus the latter half of the last day for the presentation of results. The team was at the outset split into two groups according to their skills and previous background: the first dealing with the task of codifying the posts' content, following a previously established set of categories (cf. Appendix II); the second dealing with quantitative dimensions such as engagement metrics and automated sentiment analysis. About halfway through the process, both groups regathered to crosscheck and discuss the respective findings, aiming at a global analysis and interpretation of the data. For the third and last phase of the project, the team prepared a set of visualizations that synthesized the main results.

2.4. Implementation of the Methodology

To answer RQ1 (How parties try to engage with Facebook users?), we looked at the posts' content: text, images and emojis.

The 100 posts with more "other reactions" of each subset (i.e., European and Legislative elections), were considered for manual coding of categories, which were then analyzed and ran through several visualization techniques. "Other reactions", i.e., all reactions besides *Like* and *Love*, were considered instead of common engagement metrics (total reactions, total interactions, number of shares, etc.) not because these could be considered more representative, but rather because these more diverse responses from the "demand side" could be helpful in showing a wider spectrum of engagement strategies from the "supply side". As a complement, word frequencies were also considered, both for these posts and for the similar corpora of the 100 posts with more interactions from each period. Images and emojis on posts were also subjected to analysis, as described in more detail below. Finally, posts and comments, after being translated to English, were subjected to sentiment analysis with NLTK.VADER¹.

To answer RQ2 (What type of content receives more attention and engagement?), we

additionally analyzed comments to the posts, and particularly emojis within those comments, using SpaCy² for retrieval and RawGraphs³ for visualization. Given that the subsets of data were already selected due to being the top-rated regarding interactions or “other reactions”, the techniques used to answer RQ1 were also applicable for RQ2.

2.4.1. Data processing

Text

1. Posts and comments were translated from Portuguese to English with Google Translator⁴.
2. Text sentiment analysis was processed with NLTK.VADER (Hutto and Gilbert, 2014), a model used for text sentiment analysis that is sensitive to both polarity (positive/negative) and intensity (strength) of emotion.
3. Emoji extraction, through text description, was carried out with SpaCy and the addon Spacymoji⁵.

Images

After being retrieved with the Python urllib library⁶, the images belonging to the corpora were processed with:

1. Google Vision API⁷, with a particular focus on the Safe features detection, and visualized with the help of heatmaps.
2. Image Sorter⁸, which allows the grouping of images according to several criteria, e.g. by color resemblance.

3. Results

3.1. Basic descriptive statistics

Tables from 2 to 5 give a basic overview of the sampled data, from which it can already be observed that none of the subcorpora includes all the initial 11 official party pages (cf. Appendix I). Although every page appears in at least one of the samples, only seven are present in all of them. From left to right in the political spectrum, these are: PCP (Communist Party), L (Free; left-wing), PAN (People Animals Nature, an ecologically and animal rights oriented party), PPD/PSD (Social Democrat Party, center-right), IL (Liberal

2. An open source software library for advanced natural language processing (<https://spacy.io/>).

3. <https://www.rawgraphs.io/>

4. <https://pypi.org/project/googletrans/>

5. A community maintained addon (<https://pypi.org/project/spacymoji/>)

6. <https://docs.python.org/3/library/urllib.html>

7. <https://cloud.google.com/vision>

8. <https://imagesorter.software.informer.com/>

Initiative, neoliberal/libertarian), CDS-PP (Popular Party, conservative), CH (Enough, extreme right populist).

This disproportionality is amplified if we look at the number of posts of each page in the dataset, with PAN and IL always overrepresented, and CH also overrepresented, except in one of the samples. The averages of “other reactions” (Tables 2 and 3) and of total interactions (Tables 4 and 5) allow to compensate for this overrepresentation, showing less intense fluctuations between higher and lower averaged parties.

Page	Number of posts in dataset	Total of “other reactions”	Average of “other reactions” per post
PCP	8	535	66.88
CDU (PCP+PEV)	1	25	25.00
PEV	1	45	45.00
BE	13	765	58.85
L	5	161	32.20
PS	5	518	103.60
PAN	24	2,229	92.88
PPD/PSD	2	59	29.50
IL	23	1,974	85.83
CDS-PP	4	218	54.50
CH	14	728	52.00
<i>TOTAL</i>	<i>100</i>	<i>7,257</i>	<i>72.57</i>

Table 2. Dataset with Top 100 “Other Reactions” for European Elections.

Page	Number of posts in dataset	Total of “other reactions”	Average of “other reactions” per post
PCP	1	128	128.00
CDU (PCP+PEV)	1	79	79.00
BE	6	985	164.17
L	9	2,882	320.22
PAN	17	4,679	275.24
PPD/PSD	3	266	88.67
IL	24	5,230	217.92
CDS-PP	4	691	172.75
CH	35	13,653	390.09
<i>TOTAL</i>	<i>100</i>	<i>28,593</i>	<i>285.93</i>

Table 3. Dataset with Top 100 “Other Reactions” for Legislative Elections.

Page	Number of posts in dataset	Total of interactions	Average of interactions per post
PCP	5	3,750	750.00
CDU (PCP+PEV)	5	3,865	773.00
BE	3	2,086	695.33
L	1	937	937.00
PS	11	13,901	1,263.73
PAN	37	58,814	1,589.57
PPD/PSD	3	3673	1,224.33
IL	31	36,014	1,161.74
CDS-PP	1	1,088	1,088.00
CH	3	3,124	1,041.33
<i>TOTAL</i>	<i>100</i>	<i>127,252</i>	<i>1,272.52</i>

Table 4. Dataset with Top 100 Interactions for European Elections.

Page	Number of posts in dataset	Total of interactions	Average of interactions per post
PCP	2	7,572	3,786.00
L	8	36,991	4,623.88
PS	10	24,645	2,464.50
PAN	13	33,986	2,614.31
PPD/PSD	4	10,233	2,558.25
IL	33	81,261	2,462.45
CDS-PP	1	1,805	1,805.00
CH	29	111,734	3,852.90
<i>TOTAL</i>	<i>100</i>	<i>308,227</i>	<i>3,082.27</i>

Table 5. Dataset with Top 100 Interactions for Legislative Elections.

3.2. Text

3.2.1. Categorization

The two sets of 100 posts with the highest rate of “other reactions” were analyzed and classified manually, using a list of categories and subcategories from a preexisting codebook conceived by the team.

The classification was done by four coders, which independently classified 50 posts each. For this project, and although some posts may have benefited from a multiple categorization approach, the team agreed to have only one code per post to facilitate the statistical treatment and to keep the task more manageable within the time limits of the data sprint. Intercoder reliability was assessed by the categorization and comparison of a small sample, done by all 4 coders, until consensus was reached. The total of 200 posts was subject to a further round of mutual verification, to strengthen the reliability.

Both in the first and second rounds, some limitations in the original codebook were detected, prompting the members involved in this task to rectify the coding scheme, readjusting some of the initial categories and also creating new ones, in order to fit the data more thoroughly.

The data for each period, with added categories and subcategories, were sorted and analyzed according to several criteria, of which 1) the frequency of each *category* and *subcategory*, and 2) the frequency and ratios of categories for each *party* yielded the most relevant interpretations.

The category distributions for European and Legislative Elections were visualized using the RAWGraphs tool and are shown in Figures 4 and 5.

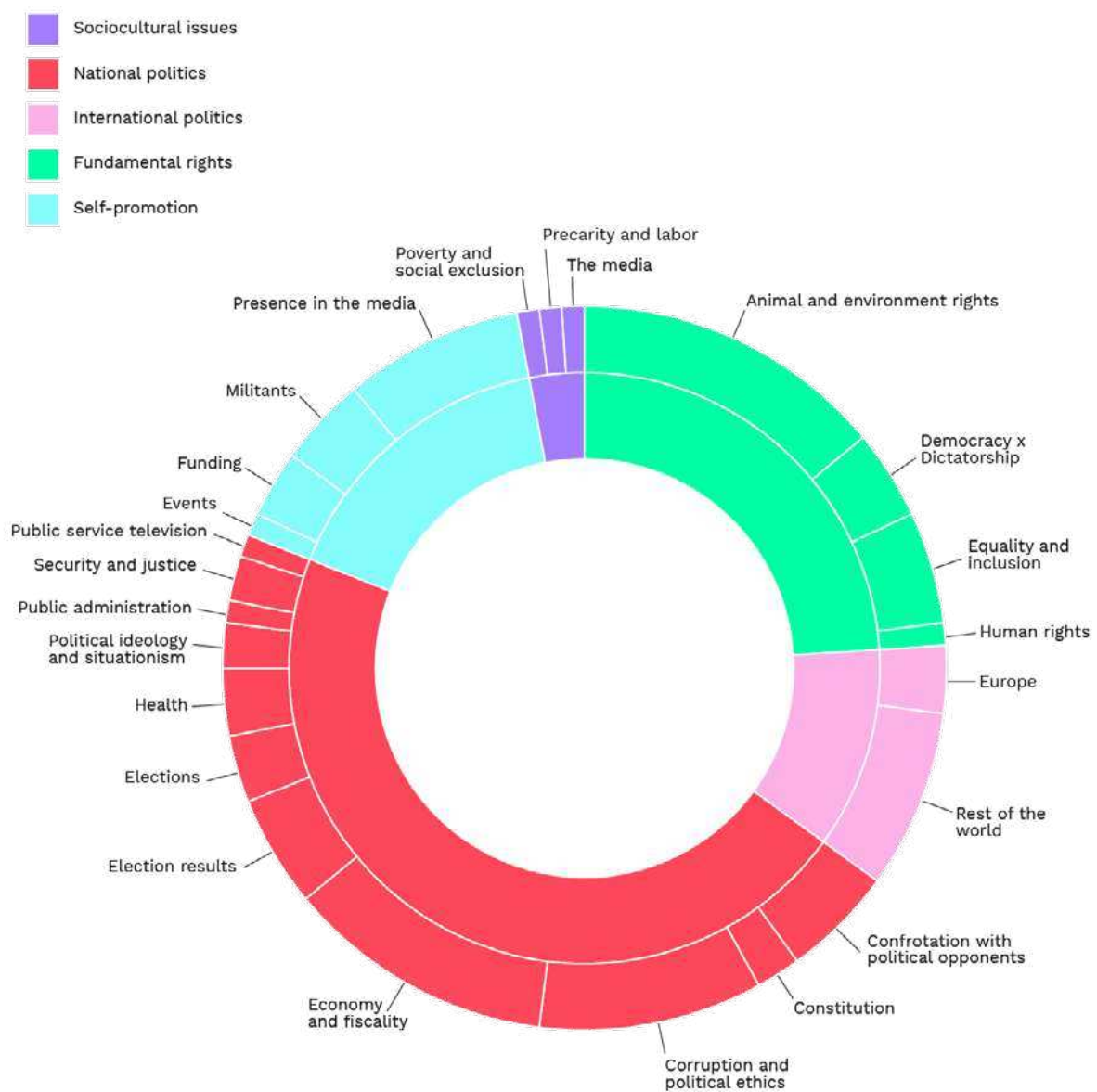


Figure 4. European Elections categories (all parties).

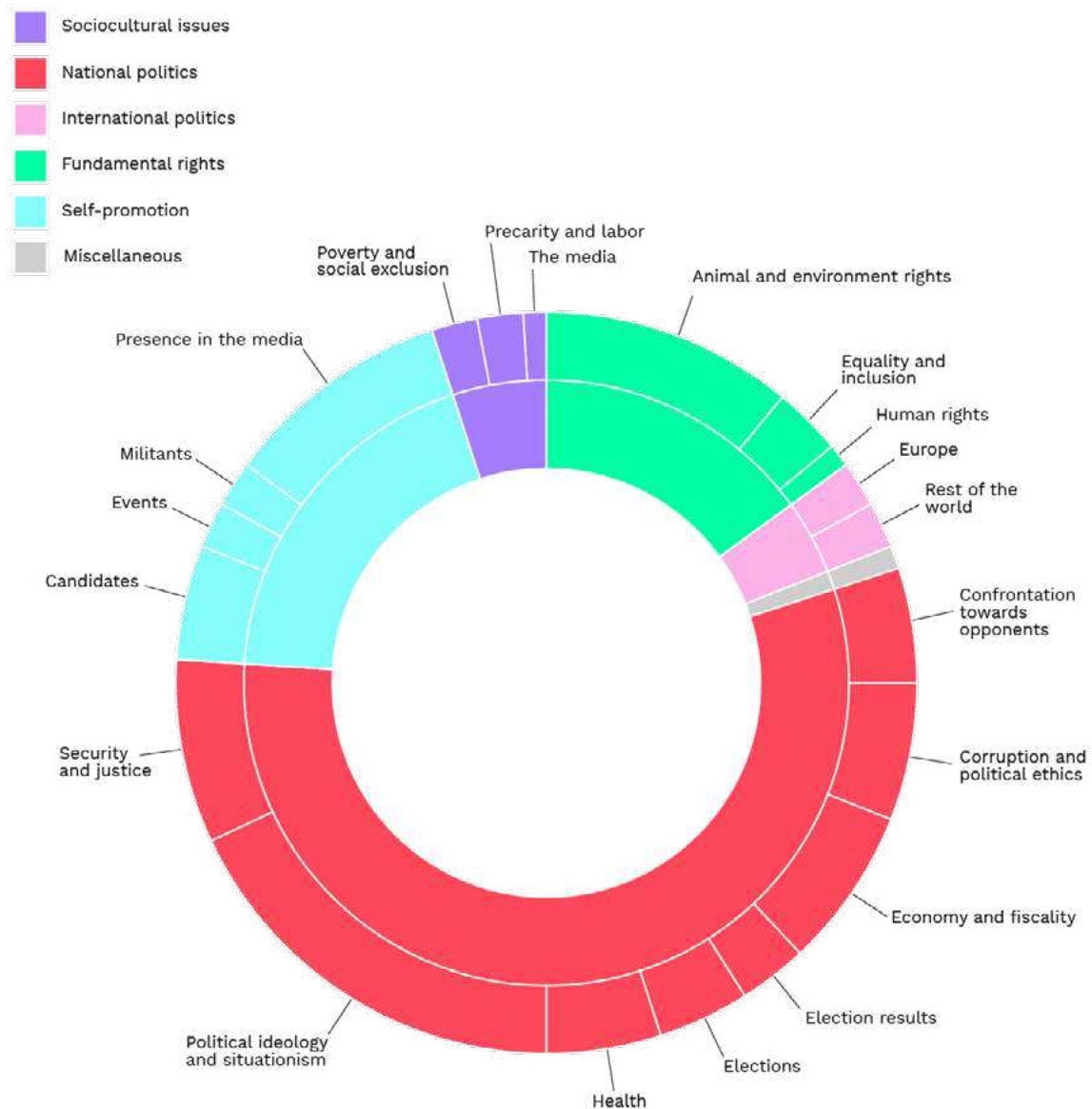


Figure 5. Legislative Elections categories (all parties).

The top 100 European Elections posts with more “other reactions” (Figure 4) highlighted the following results:

- The most frequent category is “National Politics”, with “Economy and Fiscality” and “Corruption and Political Ethics” as its most frequent subcategories;
- “Fundamental Rights” is the second most frequent main category, although this result may be biased by the predominance of the “Animal and Environmental Rights” subcategory in the posts by PAN;
- Surprisingly, “International Politics” is not only the second less-frequent category during European Elections, but “Rest of the World” surpasses “Europe” as a subcategory.

The top 100 Legislative Elections posts with more “other reactions” (Figure 5) highlighted the following results:

- The weight of “National Politics” is overwhelming, with the subcategories “Security and Justice” and “Political Ideology” being more frequent than in the European Elections;
- “Self-promotion” is also more frequent than in the European Elections, with “Presence in the Media” and “Candidates” as the most frequent subcategories;
- The “Security and Justice” subcategory grows in importance.



Figure 6. European Elections category cloud (all parties).

The clouds for most frequent categories (Figures 6 and 7) confirm the previous analysis, showing how some subcategories shift in importance from the European to the Legislative Elections. Such is the case of “Economy and Fiscality” and “Animal and Environmental Rights” decreasing, and “Political Ideology and Situationism” increasing.

A similar analysis was applied at the party level (Figure 8), this time considering the 200 posts (i.e., both European and Legislative Elections), in order to assess individual party strategies.

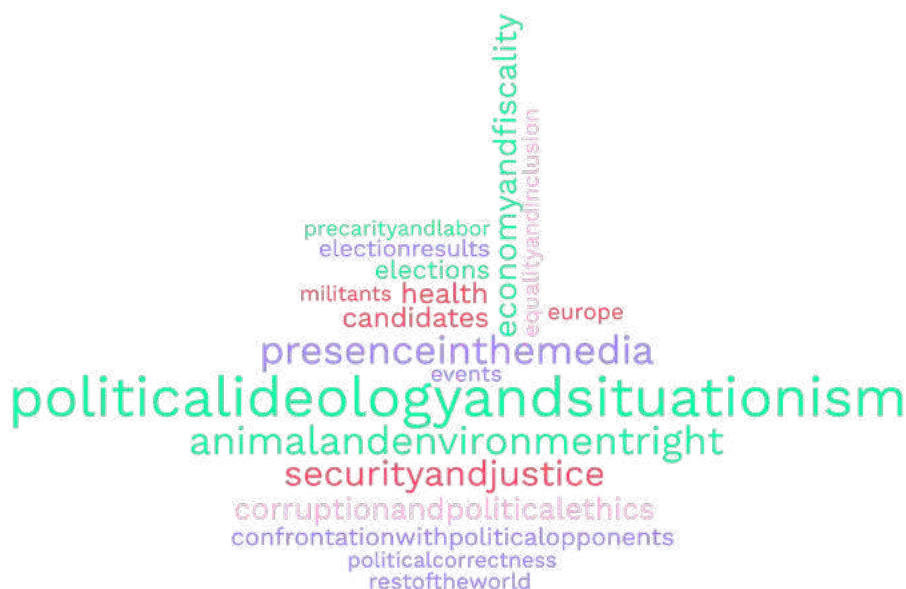


Figure 7. Legislative Elections category cloud (all parties).

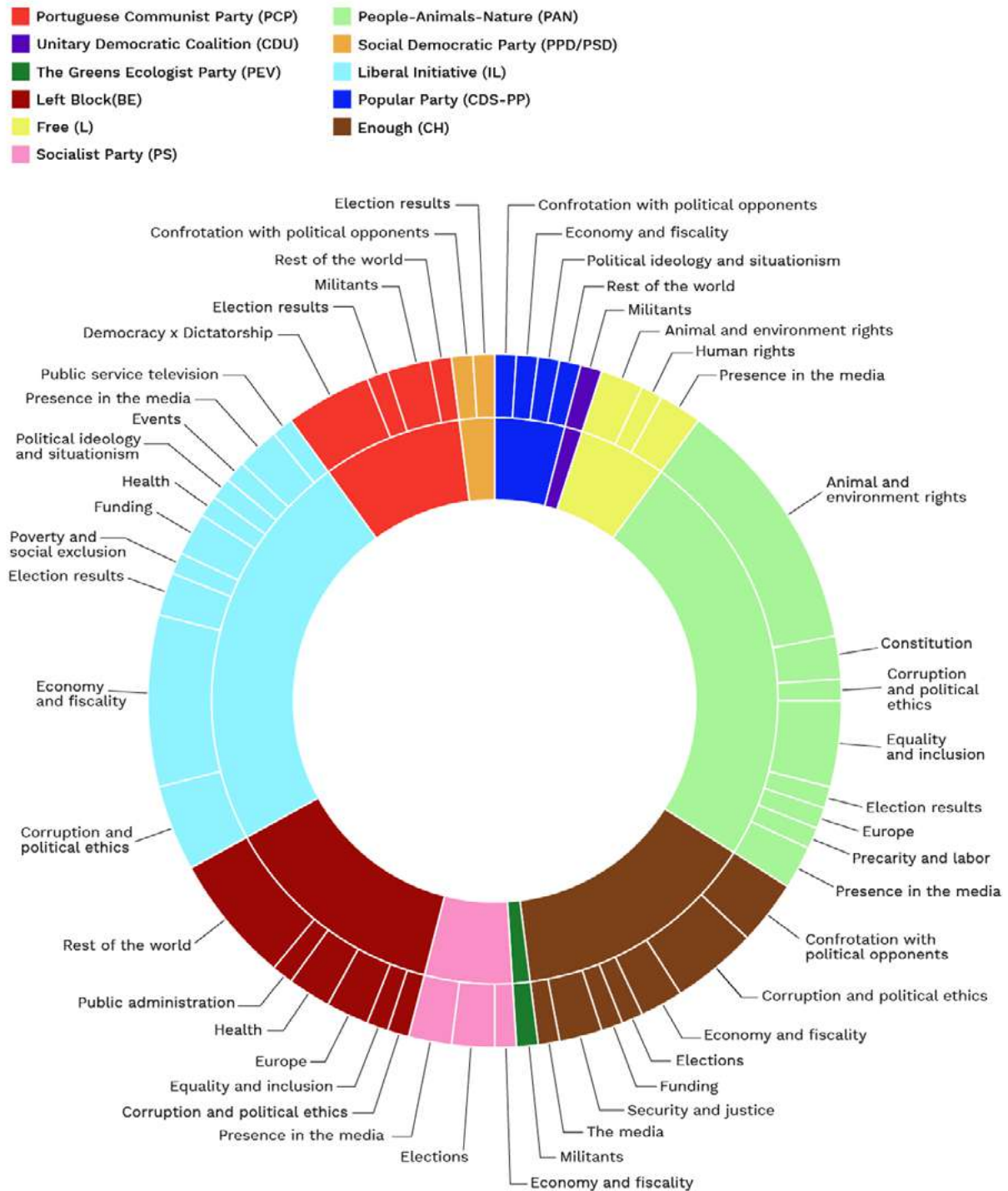


Figure 8. Most frequent categories by political party.

Although the corpus only considers posts that achieved higher levels of “other reactions”, which may disallow a generalization, some evidence can be highlighted:

- The three right-wing parties (CDS-PP, IL and CH) have “Confrontation with Political Opponents” as a more frequent subcategory than other (left, center-left and center-right) parties;
- Left-wing parties BE (Left Block) and CDU (Unitary Democratic Coalition between the Communist and the “Greens”) relate more to “International Politics”;
- Some subcategories are almost monopolized by one single party, as is the case of IL for “Economy and Fiscality” and PAN for “Animal and Environmental Rights”.

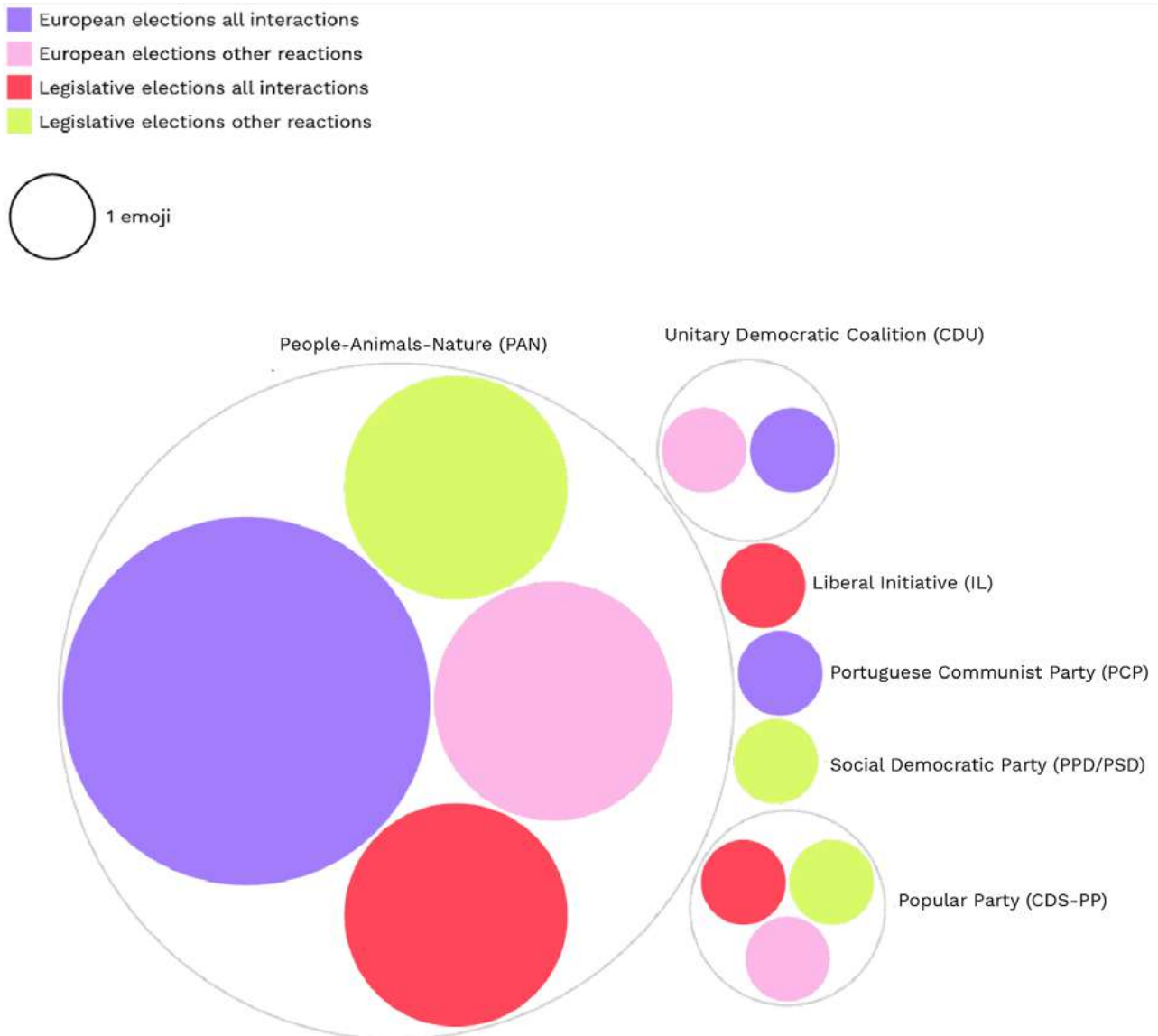


Figure 11. Emojis on posts (divided by parties).



Figure 12. Emojis on comments (divided by parties).

3.2.4. Sentiment analysis

Sentiment analysis techniques were applied with VADER to the automatically translated contents of the posts and of the top 100 comments within those posts. Although this analysis yielded sentiment scores, the results were inconclusive. This task could not be fully carried out due to time constraints – and thus the preliminary results and visualizations that were elaborated were not considered suitable to be included, though that methodological approach remains a goal for further research.

3.3. Images

3.3.1. Image grouping

The images retrieved from the posts, totaling 187, were subjected to some basic sorting techniques with the software Image Sorter. Grouping images by color proved the most interesting approach, as those groupings served as guides for the team to focus on recurrent patterns, and as a crosscheck with the image features described in the next subsection.

Images with red, blue, and white, all found in IL's posts, were the most common, denoting not only consistent use of this template in the party's posts but also confirming their prevalence within posts with more interactions and more "other reactions", which, we may remember, were the primary criteria for corpus selection.

Images with green hues appear only in the European elections data, and these are mostly used for "green" content – i.e., posts with the category "Animal and Environmental Rights", which, as mentioned above, is the dominant subcategory in PAN's posts.

3.3.2. Image features

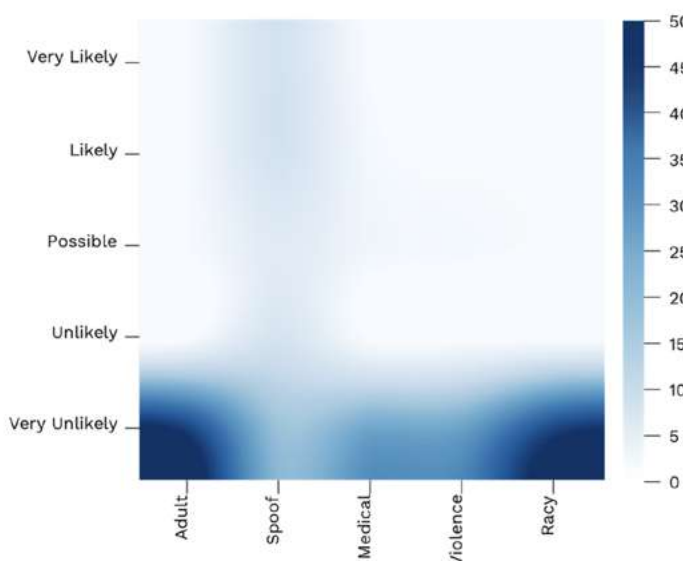


Figure 15. Heatmap for Safe Search features on images (European Elections).

The following heatmap visualizations show, according to Google Vision API's "Safe Search" module, the probability of having, in the images belonging to the corpora, features that may classify them respectively as "adult-oriented", "spoof or parody" (image modified, often to become offensive), "medical", "violent" or "obscene or provocative"⁹. As before, individual analyses were made for the European Elections (Figure 15) and the Legislative Elections (Figure 16).

9. Cf. <https://cloud.google.com/vision/docs/reference> and <https://cloud.google.com/vision/docs/reference/rpc/google.cloud.vision.v1#google.cloud.vision.v1.SafeSearchAnnotation>

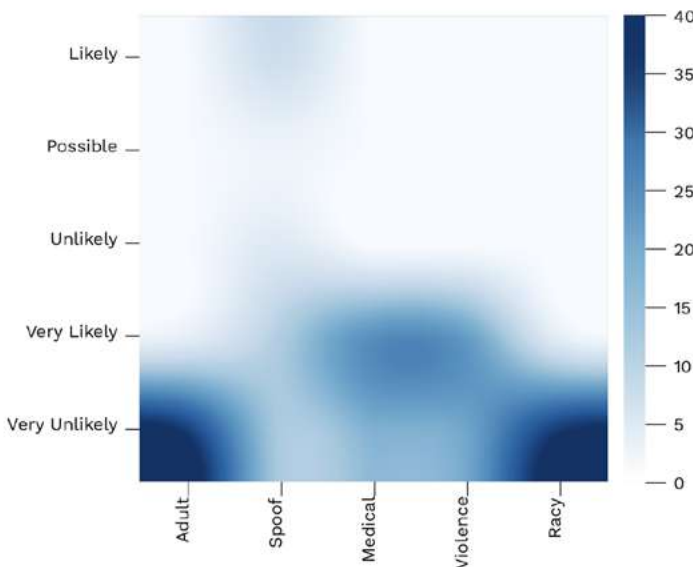


Figure 16. Heatmap for Safe Search features on images (Legislative Elections).

From what can be read in these visualizations, the overwhelming majority of the images was marked as “very unlikely” having any of the above mentioned features, though with some likelihood of spoofing for images both in the data sets from the European and from the Legislative Elections. In the case of those belonging to the latter set, the likelihood of having images with medical or violent content increases slightly, but still far from being unequivocal positives.

One image (Figure 17) that drew the attention of the team is indeed a spoof, a photoshopped image of

PAN’s leader and main candidate André Silva in a post by this party. Both the image and the post content show that it was intentional humor – self-deprecating humor, if the expression suits – as an announcement for a forthcoming interview in a TV program hosted by a well-known comedian in Portugal.



Figure 17. Example of an image with “Spoof” features.

4. Key Findings and Discussion

As mentioned in the Introduction, the answers to our Research Questions should also consider the mapping of the parties according to their foundation time and ideological position, as represented above in Figure 1. Given that the selected corpora filter the posts from the two electoral periods according to the criteria of the highest amount of “other reactions” and interactions, this already shows that newer parties, except L (Free, left-wing), were overrepresented in our data subsets, a clue that we followed across each specific analysis. In the next paragraphs, we will address RQ1 and RQ2 with this in mind.

RQ1: How differently official parties try to engage with Facebook users in the two major 2019 elections periods in Portugal?

Regardless of the party and its specific positioning on the map, and also both for European and Legislative Elections, posts are predominantly from the main categories “National Politics” (the most frequent), “Fundamental Rights”, and “Self-Promotion”. “European Politics” was, surprisingly, barely a topic in the posts with more “other reactions”, even during the European Elections.

However, differences in the communication strategies become more manifest when considering the subcategories. IL favors “Economy and Fiscality” (a subcategory of “National Politics”) and PAN “Animal and Environmental Rights” (a subcategory of “Fundamental Rights”), thus allowing for classifying them as “single-issue parties”. And while left (and especially more left-wing) parties slightly favor broader concepts like “Democracy x Dictatorship” (a subcategory of “Fundamental Rights”), right-wing parties, especially newer and fringe ones (IL and CH), resort to themes that have a higher confrontational potential, such as “Corruption and Political Ethics” or “Political Ideology and Situationism”, if not outright “Confrontation with Political Opponents”. That agonistic approach seems to increase during the campaign periods.

The analysis of the most frequent words in the posts shows a common agreement regarding “freedom” (notably in the Legislative Elections), though only a deeper look at the content may reveal if that concept is understood consensually or not – i.e., whether political positioning or other differentiating features also correspond to diverse appropriations of “freedom”. Other concepts are too spread out (both considering posts and parties) and thus with frequencies too low to draw conclusions.

Emojis are rarely used as part of the parties’ communication strategies, with PAN being the exception (almost always using symbols that confirm their focus with “Animal and Environmental Rights”). However, two other newer parties (IL and CH) showed an increase in emoji usage for the Legislative Elections, which took place later in the year.

Using images is also a transversal strategy across all parties, with an over-representation of posts by IL, one of the newer right-wing parties. Though this may be due to a sampling bias, this party is very consistent regarding visual aesthetics, using their colors (white, red, and especially cyan) in almost all image posts.

For other parties, there is more chromatic diversity, but “risky” images – i.e., those that may be flagged as less appropriate in Google Vision API Safe Search detection algorithms – are for the most part avoided. Spoofing for humorous or parodic objectives, either to mock the opponents or as lighthearted forms of self-deprecation, was however found in a few images, even if Google Vision did not flag these with a high degree of certainty. Again, this particular approach to images tends to be more common with newer parties.

Globally, the newer and smaller parties PAN, IL, and CH are not only among the most represented in the corpora but also reveal more unconventional approaches in their posts. This happens either through specific themes – e.g., IL with posts about “Economy and Fiscality”, CH with “Corruption and Political Ethics” –, cues in image posts (PAN, IL), or emojis (PAN), authorizing the conjecture of a more informal style of communication favoring the “equalization hypothesis”, in line with the claim that minor parties make stronger use of social media (Gibson and McAllister, 2014), even surpassing major ones in engagement levels (going beyond Samuel-Azran, 2015). As such, it may be the case that “Facebook is an appropriate campaign tool for minor political actors to decrease their structural disadvantages” and “it may have contributed to the remarkable transformation of European political systems” (Bene, 2021: 17).

The more confrontational stance in the posts by CH, and to a lesser extent by IL, is also in line with the literature about the ascension of right-wing populism, namely that these political actors “seek a fast, direct, and unmediated connection to the people” (Engesser *et al.*, 2017: 1282), even if we could not fully explore to what extent were present the “three major dimensions of populist style: simplification, emotionalization, and negativity” (Engesser *et al.*, 2017: 1285). The almost non-existence of posts about European politics did now allow us to assess if Euroscepticism is part of their discourse, as proposed by Alonso-Muñoz and Casero-Ripollés (2020) for similar parties in the continent.

RQ2: What type of content receives the most attention and engagement from the users?

If solely considered from the perspective of interactions as a whole, some of the items and interpretations above-identified for RQ1 also apply to RQ2 if we switch from the “supply-side” to the “demand-side” and may inclusively be one of the explanations for the success in the voter turnout of those same newer parties. Comparing the Legislative Elections of 2019 to the previous in 2015, PAN rose from 1.39% to 3.49% and from 1 to 4 MPs, and the newcomers IL and CH (each with 1.35% of the votes), had just enough to earn an MP each.

PAN’s success has some nuances if confronted with those other parties: on the one hand, they are positioned at the center-left rather than on the right of the spectrum; on the other hand, they are, relatively speaking, the “oldest” and most established of the newer parties, with representation in the national parliament since 2015. The engagement with the audience seems to be achieved with their most common topic (“Animal and Environmental Rights”). This is also the party that receives more emojis in the comments, which may be due to a “mirror” effect: emojis in the posts may be read as an incentive to mimic this approach in the comments.

The engagement with image posts in these three parties seems to be particularly significant when considering “other reactions” (*Haha, Wow, Sad, Angry*), i.e., there may be a stronger correlation between image posts and these more polarized reactions than in the cases of *Like* or *Love*, a conjecture that can only be confirmed through a more thorough analysis which could not be done in the data sprint context.

No relevant fluctuations in the sentiments (in intensity or from positive to negative and vice-versa) could be detected. The team members that were more closely involved with this approach conjecture that, due to the idiosyncrasies of the Portuguese language and the lack of a toolkit specifically aimed at this language, the automatic translation into English required the complement of human-assisted verification. Though the time constraints did not allow it, that seems to be a rich path for future analysis.

Overall, and while acknowledging again the caveat that these results cannot be fully generalized due to the relative size of the corpora compared to the full original dataset, also the analysis of the “demand-side” favors the “equalization hypothesis”, i.e., newer and smaller parties tend to receive more engagement from the users, which may result from either a more ludic or more confrontational communication strategy (“supply-side”).

Some approaches contemplated at the outset of this research would have required additional time to be properly adapted. Such was not only the case of sentiment analysis but also of image network analysis, which was decisively discarded due to the lack of a more coherent strategy to make sense of the results.

Regardless of these shortcomings, the data sprint enabled the team to gather very useful results in the scope of a wider project, and both research questions could be at least partially answered, providing relevant clues for future work. We believe that despite the time limit and small sets of data points, the application of less common approaches in computational communication sciences to contents such as texts, emoticons, and images may result in a valuable contribution for communication researchers, giving a better and more interdisciplinary understanding of the exchanges between parties and citizens in social media platforms. That is particularly relevant in the Portuguese context, which, to the authors’ knowledge, is still short of academic work on the political presence online. The methodologies here applied to Facebook can be adapted to other platforms, thus providing further data and insights for engagement analysis and metrics.

Doing research in the context of a data sprint also proved to be an enlightening learning experience for the team, which was for the most part unfamiliar with this intensive and time-bound environment. Most of all, the diverse backgrounds (communication and other social sciences, humanities, data science, design) and levels of research practice promoted the multidisciplinary that is expected in this approach.

Funding

FCT: Fundação para a Ciência e a Tecnologia, Ministério do Ensino Superior, Ciência e Tecnologia, Portugal (Projeto de Investigação “Redes de Participação Política no Facebook em Portugal” – PTDC/COM-CSS/28269/2017).

References

- Alonso-Muñoz, L. and Casero-Ripollés, A. (2020). Populism Against Europe in Social Media: The Eurosceptic Discourse on Twitter in Spain, Italy, France, and the United Kingdom During the Campaign of the 2019 European Parliament Election. *Frontiers in Communication* 5(54): 1-12. <https://doi.org/10.3389/fcomm.2020.00054>
- Bene, M. (2021). Who Reaps the Benefits? A Cross-Country Investigation of the Absolute and Relative Normalization and Equalization Theses in the 2019 European Parliament Elections. *New Media & Society*: 1-20. <https://doi.org/10.1177/14614448211019688>
- Bennett, W.L. (2012). The Personalization of Politics: Political Identity, Social Media, and Changing Patterns of Participation. *The Annals of the American Academy of Political and Social Science* 644(1): 20–39. <https://doi.org/10.1177/0002716212451428>
- Eberl, J.M., Tolochko, P., Jost, P., Heidenreich, T. and Boomgaarden, H.G. (2020). What's in a Post? How Sentiment and Issue Salience Affect Users' Emotional Reactions on Facebook. *Journal of Information Technology & Politics* 17(1): 48-65. <https://doi.org/10.1080/19331681.2019.1710318>
- Effing, R., Jos, V.H. and Huibers, T. (2011). Social Media and Political Participation: Are Facebook, Twitter and YouTube Democratizing our Political Systems? In: *International Conference on Electronic Participation* (pp. 25-35). Springer. https://doi.org/10.1007/978-3-642-23333-3_3
- Enli, G.S. and Skogerbø, E. (2013). Personalized Campaigns in Party-Centred Politics: Twitter and Facebook as Arenas for Political Communication. *Information, Communication & Society* 16(5): 757–774. <https://doi.org/10.1080/1369118X.2013.782330>
- Engesser, S., Fawzi, N. and Larsson, A. O. (2017). Populist Online Communication: Introduction to the Special Issue. *Information, Communication & Society*, 20(9): 1279-1292 <https://doi.org/10.1080/1369118X.2017.1328525>
- Gibson, R. (2012). From Brochureware to 'MyBo': An Overview of Online Elections and Campaigning. *Politics*, 32(2): 77–84. <https://doi.org/10.1111/j.1467-9256.2012.01429.x>
- Gibson, R. K. and McAllister, I. (2014). Normalising or Equalising Party Competition? Assessing the Impact of the Web on Election Campaigning. *Political Studies*, 63(3): 529-547. <https://doi.org/10.1111/2F1467-9248.12107>
- Heiss, R., Schmuck, D. and Matthes, J. (2019). What Drives Interaction in Political Actors' Facebook Posts? Profile and Content Predictors of User Engagement and Political Actors' Reactions. *Information, Communication & Society*, 22(10): 1497–1513. <https://doi.org/10.1080/1369118X.2018.1445273>
- Hughes, A. and van Kessel, P. (2018). 'Anger' Topped 'Love' when Facebook Users Reacted to Lawmakers' Posts after 2016 Election, *PEW Research Center*, 18 July 2018. <https://www.pewresearch.org/fact-tank/2018/07/18/anger-topped-love-facebook-after-2016-election/>
- Hutto, C., and Gilbert, E. (2014). VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. *Proceedings of the International AAAI Conference on Web and Social Media*, 8(1), 216-225. <https://ojs.aaai.org/index.php/ICWSM/article/view/14550/14399>

- Junger, J. and Keyling, T. (2019). Facepager. An application for Automated Data Retrieval on the Web. Source code and releases available at <https://github.com/strohne/Facepager/>
- Kalsnes, B. (2016). The Social Media Paradox Explained: Comparing Political Parties' Facebook Strategy versus Practice. *Social Media + Society*, 2(2): 1-11. <https://doi.org/10.1177%2F2056305116644616>
- Kapoor, K.K., Tamilmani, K., Rana, N.P., Patil, P, Dwivedi, Y.K. and Nerur, S. (2018). Advances in Social Media Research: Past, Present and Future. *Information Systems Frontiers*, 20(3): 531–558. <https://doi.org/10.1007/s10796-017-9810-y>
- Klinger, U. (2013) Mastering the Art of Social Media. *Information, Communication & Society*, 16(5): 717-736. <https://doi.org/10.1080/1369118X.2013.782329>
- Klinger, U. and Russmann, U. (2017) Beer is more Efficient than Social Media: Political Parties and Strategic Communication in Austrian and Swiss National Elections. *Journal of Information Technology & Politics*, 14(4): 299-313. <https://doi.org/10.1080/19331681.2017.1369919>
- Larsson, A.O. (2016). Online, all the Time? A Quantitative Assessment of the Permanent Campaign on Facebook. *New Media & Society*, 18(2): 274–292. <https://doi.org/10.1177%2F1461444814538798>
- Larsson, A.O. and Kalsnes, B (2014). 'Of Course we are on Facebook': Use and Non-use of Social Media among Swedish and Norwegian Politicians. *European Journal of Communication*, 29(6): 653–667. <https://doi.org/10.1177%2F0267323114531383>
- Lilleker, D.G., Tenscher, J. and Stetka, V. (2015). Towards Hypermedia Campaigning? Perceptions of New Media's Importance of Campaigning by Party Strategists in Comparative Perspective. *Information, Communication & Society*, 18(7): 747–765. <https://doi.org/10.1080/1369118X.2014.993679>
- Macnamara, J. (2011). Pre and Post-election 2010 Online: What Happened to the Conversation? *Communication, Politics, Culture*, 44(2): 18–36. <https://search.informit.org/doi/10.3316/informit.627292905802447>
- Magin, M., Podschuweit, N., Haßler, J. and Russmann, U. (2017). Campaigning in the Fourth Age of Political Communication. A Multi-method Study on the Use of Facebook by German and Austrian Parties in the 2013 National Election Campaigns. *Information, Communication & Society*, 20(11): 1698–1719. <https://doi.org/10.1080/1369118X.2016.1254269>
- Nahon, K. (2015). Where There Is Social Media There Is Politics. In A. Bruns *et al.* (eds.), *The Routledge Companion to Social Media and Politics* (pp. 39-55) Routledge. <https://doi.org/10.4324/9781315716299-5>
- Norris, P. (2003). Preaching to the Converted? Pluralism, Participation and Party Websites. *Party Politics*, 9(1): 21-45. <https://doi.org/10.1177/135406880391003>
- Ramos-Serrano, M., Fernández Gómez, J.D. and Pineda, A. (2018). 'Follow the Closing of the Campaign on Streaming': The Use of Twitter by Spanish Political Parties during the 2014 European Elections. *New Media & Society*, 20(1): 122-140. <https://doi.org/10.1177/1461444816660730>

- Samuel-Azran, T., Yarchi, M. and Wolfsfeld, G. (2015). Equalization versus Normalization: Facebook and the 2013 Israeli Elections. *Social Media + Society*, 1(2): 1-9. <https://doi.org/10.1177/2056305115605861>
- Serra-Silva, S., Dias Carvalho, D. and Fazendeiro, J. (2018). Party-citizen Online Challenges: Portuguese Parties' Facebook Usage and Audience Engagement. In: Lobo, M.C., Silva, F.C. and Zúquete, J.P. (eds.), *Changing Societies: Legacies and Challenges. Vol. 2. Citizenship in Crisis* (pp. 185-214). ICS. <https://www.ics.ulisboa.pt/books/book2/ch08.pdf>
- Sobacı, M. Z. and Hatipoğlu, İ. (2020). Measuring the Engagement Level of Political Parties with Public on Facebook: The Case of Turkey. *2017 Conference for E-Democracy and Open Government (CeDEM)* (pp. 209-216). <https://doi.org/10.1109/CeDEM.2017.15>
- Stetka, V., Surowiec, P. and Mazák, J. (2019). Facebook as an Instrument of Election Campaigning and Voters' Engagement: Comparing Czechia and Poland. *European Journal of Communication*, 34(2): 121-141. <https://doi.org/10.1177/0267323118810884>
- Theocharis, Y., Barberá, P., Fazekas, Z., Adrian Popa, S. and Parnet, O. (2016). A Bad Workman Blames His Tweets: The Consequences of Citizens' Uncivil Twitter Use When Interacting With Party Candidates. *Journal of Communication*, 66(6): 1007-1031. <https://doi.org/10.1111/jcom.12259>
- Vogels, E. A., Auxier, B. and Anderson, M. (2021). Partisan Differences in Social Media Use Show up for Some Platforms, but not Facebook, PEW Research Center, 7 April 2021. <https://www.pewresearch.org/fact-tank/2021/04/07/partisan-differences-in-social-media-use-show-up-for-some-platforms-but-not-facebook/>
- Xenos, M. A., MacAfee, T. and Pole, A. (2017). Understanding Variations in User Response to Social Media Campaigns: A Study of Facebook Posts in the 2010 US Elections. *New Media and Society*, 19(6): 826-842. <https://doi.org/10.1177/1461444815616617>

Appendix I: List of political parties

The list below shows the 10 political parties and one coalition, from left to right in the political spectrum, and their hyperlinked Facebook pages:

1. PCP (Portuguese Communist Party): <https://www.facebook.com/pcp.pt/>
2. CDU (Unitary Democratic Coalition, between PCP and PEV): <https://www.facebook.com/CDUPCPPEV/>
3. PEV (The Greens Ecologist Party): <https://www.facebook.com/PartidoEcologistaOsVerdes/>
4. BE (Left Block) [changed page in November 2021; this page remains as the page of their official media outlet, Esquerda.Net]: <https://www.facebook.com/esquerda.net/>
5. L (Free): <https://www.facebook.com/LIVREoficialpt/>
6. PS (Socialist Party): <https://www.facebook.com/SedeNacionalPartidoSocialista/>
7. PAN (People-Animals-Nature): <https://www.facebook.com/PANpartido/>
8. PPD/PSD (Social Democratic Party): <https://www.facebook.com/ppdpsd/>
9. IL (Liberal Initiative): <https://www.facebook.com/iniciativoliberal/>
10. CDS-PP (Popular Party): <https://www.facebook.com/CDSPP/>
11. CH (Enough): <https://www.facebook.com/PartidoChegaOficial/>


Appendix II: List of categories

MAIN CATEGORY	SUBCATEGORIES
FUNDAMENTAL RIGHTS	Animal and Environmental rights
	Democracy x Dictatorship
	Digital Rights
	Equality and Inclusion
	Human Rights
INTERNATIONAL POLITICS	Europe
	Rest of the World
MISCELLANEOUS	Miscellaneous
NATIONAL POLITICS	Confrontation with Political Opponents
	Constitution
	Corruption and Political Ethics
	Culture
	Decentralization
	Economy and fiscality
	Education
	Elections
	Election Results
	Health
	Nativity
	Political Ideology and Situationism
	Public Administration
	Public Service Television
	Security and Justice
Yellow Jackets	
SELF-PROMOTION	Candidates
	Events
	Funding
	Internal Organizational Affairs
	Militants
	Online Presence
	Presence in the Media
	The Party in Europe
SOCIOCULTURAL ISSUES	Entrepreneurship
	The Media
	Political Correctness
	Poverty and Social Exclusion
	Precarity and Labor
	Religion



Visual Communication strategy of populist leaders on Instagram in 2020

Estrategia de comunicación visual de los líderes populistas en Instagram en 2020

 Ilya Lavrov
irlavrov@gmail.com
Universidade NOVA de Lisboa

Chiara Miozzo
miozzochiara@gmail.com
Universidade NOVA de Lisboa

Franziska Schranz
franziska.schranz@gmx.at
Universidade NOVA de Lisboa

Marie Palaffre
mpalaffre@yahoo.fr
Universidade NOVA de Lisboa

Received: 14/01/2022 | Accepted: 30/05/2022

ABSTRACT In recent years social media platforms, especially Instagram, have become one of the most crucial elements of visual political communication strategies, assisting politicians with creating their vision in the minds of citizens. At the same time, the strengthening of the populist trend in the modern political landscape is noted in many countries worldwide. The present study uses digital methods to examine the virtual visual strategy of 12 populist leaders in democratic states in 2020. This study describes the type of content populists publish on their Instagram profiles. A comparative analysis between populists in different countries is conducted, as well as the engagement and textual analysis focusing on the COVID-19 rhetoric. It was revealed that populist leaders post pictures that communicate personal and national power. While showing their strength and reliability, they also portray themselves as ordinary citizens.

KEYWORDS Instagram, Populism, digital methods, Vision API, communication strategy

RESUMEN En los últimos años, las plataformas de redes sociales, especialmente Instagram, se han convertido en uno de los elementos más cruciales de las estrategias de comunicación política visual, ayudando a los políticos a crear su imagen en la mente de los ciudadanos. Al mismo tiempo, se observa el fortalecimiento de la tendencia populista en el panorama político moderno en muchos países del mundo. El presente estudio utiliza métodos digitales para examinar la estrategia visual virtual de 12 líderes populistas en estados democráticos en 2020. Este estudio describe el tipo de contenido que los populistas publican en sus perfiles de Instagram. Se realiza un análisis comparativo entre populistas de diferentes países, así como el análisis de engagement y textual centrado en la retórica COVID-19. Se reveló que los líderes populistas publican fotos que comunican el poder personal y nacional. A la vez que muestran su fuerza y fiabilidad, también se retratan como ciudadanos de a pie.

PALABRAS CLAVE Instagram, populismo, métodos digitales, Vision API, estrategia de comunicación

How to cite this article:

Lavrov, I.; Schranz, F.; Miozzo, C. & Palaffre, M. (2022). Visual Communication strategy of populist leaders on Instagram in 2020. *Dígitos. Revista de Comunicación Digital*, 8: 155-177. DOI: 10.7203/drdcd.v1i8.238



Visual Communication strategy of populist leaders on Instagram in 2020

1. Introduction and problem statement

Political communication has changed drastically over the last several decades: politicians still need to use 'traditional' means of expression to win elections; however, they should also have digital media tools at their disposal. Indeed, digital media tools are becoming impactful in influencing the results of an election race (Tambini, 2018). Moreover, with 'the professionalization' of political communication (Maarek, 2014), digital media has unquestionably become a tool that politicians cannot ignore and need to use as effectively as possible. Nowadays, politicians are free to create a different persona of themselves on the Net, and it is unclear which persona is more important to voters. This trend is supported by the 'hyper individualization' and 'personalization' of political communication in the modern political media landscape (Russmann et al., 2019).

Visuals have become even more crucial with the evolution of technology (Farkas & Bene, 2021): "political communication today is built on a visual foundation" (Schill, 2012). The modern World can even be described as the age of "ubiquitous photography" (Hand, 2012), with more than 300 million photos uploaded to Facebook and more than 95 million to Instagram (IG) every day (Stout, 2021). Nowadays, Instagram is one of the leading platforms for sharing visual content; it is becoming a vital tool for shaping a holistic identity (Gordillo-Rodríguez & Bellido-Pérez, 2021) mainly because "impression management" on this platform is possible. Thus, politicians, specifically political leaders, find it extremely helpful for political communication. In an 'image-centric' social network like Instagram, colours, graphic elements, and image style are used as visual marketing tools to draw images of politicians comparable to brand identities (Zailskaitė-Jakštė et al., 2017). Evidence shows that visual communication can create an emotional resonance that conveys as strong a message as a literal one (Lilleker, 2019). Digital tools, in addition to assisting politicians with projecting power, trustworthiness, and prestige, essentially establish a more humanized and relatable version of a politician by displaying certain carefully chosen parts of their private life (Lalancette & Raynaud, 2017; Page & Duffy, 2016).

According to several studies, the populist trend in the contemporary political landscape has been consolidating. However, there is no consensus on determining which regime can be named a 'populist regime' (Mudde & Kaltwasser, 2018). Three main conceptual approaches are suggested in the Oxford Handbook of Populism: an ideational, a political-strategic, and a sociocultural approach (Miller-Idriss, 2019). Most scholars tend to use *an ideational approach* that revolves around the idea of a strong division between "the pure people" and "the corrupt elite." This approach also claims that "politics is about respecting popular sovereignty at any cost" (Mudde & Kaltwasser, 2018). A less common

variant of conceptualization of populism is represented by a *political-strategic approach*. In particular, K. Weyland defines populism as a political strategy of a personalist leader based on direct non-institutionalized support of followers from different *strata* of society (Weyland, 2020). According to P. Kenny, populism is the mobilization of the masses directed by charisma in order to gain power (Kenny, 2021). Finally, representatives of a *sociocultural approach* consider populism as a “popular style” of politics, manifested in the “wrong” behaviour of a leader who goes beyond the accepted cultural framework to get closer to voters. According to P. Ostiguy, populists “prefer the authentic (pure) lower culture of the people over the inauthentic (corrupt) higher culture of the elite” (Ostiguy, 2009). As B. Moffitt notes, populist leaders are extraordinary personalities who carry out the “presidentialization” and glorification of politics, which allows them to establish such a type of socio-political interaction as performance (Moffitt, 2016).

Regardless of the approach considered, it is evident that many populist leaders widely use social networks to gain power and mobilize people (Dittrich & the Jacques Delors Institut, 2017). It has been happening since ‘social media changes the communication structure of the public sphere’ (Spiekermann, 2020). Yet, despite populism being widely discussed in the scientific community, their visual communication strategies are not studied extensively, rarely on a small, manually collected sample of images.

This study is intended to reveal the dynamics of visual political communication of populist leaders. Specific objectives are 1) to determine the type of content populists publish on their Instagram profiles; 2) to conduct a comparative analysis between populists in different states; 3) to discover what type of content produces the most engagement on populists’ Instagram profiles; 4) to conduct textual analysis of the populists’ posts on Instagram.

The Tony Blair Institute methodology has been used (Kyle & Meyer, 2020) to determine a pool of populist political leaders. First of all, all populists have two claims in common: “(1) that a country’s ‘true people’ are locked into a moral conflict with ‘outsiders’ and (2) that nothing should constrain the will of the ‘true people’” (Kyle & Meyer, 2020). The classification of populist leaders studied in this paper was made by the same institute and was based on three aspects of a populist style. “(1) A high degree of anti-elite rhetoric, (2) the evidence of efforts to delegitimize political opposition, and (3) the evidence that the leader cultivated a ‘cult of personality or emphasized that they alone could serve as the voice of the people’” (Kyle & Meyer, 2020).

For this paper, only free and partly free countries (Freedom House, 2020) and leaders with Instagram accounts were considered, and the study has been limited to the year 2020. As a result, a sample of Instagram profiles of 12 populist political leaders, who were in power in 2020, was conducted.

While several studies analyse the visual communication of specific political leaders (Strand & Schill, 2019; Lalancette & Raynauld, 2017) or some groups of politicians in a particular region (Ekman & Widholm, 2017), an investigation of a sample of populists — especially with Vision API tools —, can produce insights on how populists gain and maintain their power on social media.

1.1. Research questions

RQ1: What kind of content do populist leaders publish?

- a. What kind of pictures do they use?
- b. Which colours do they use?
- c. Are their accounts more personal / politician-oriented or institutional?

RQ2: Is there a common communication strategy among populist leaders around the World? Do they refer to common themes?

RQ3: Which content produces the highest and lowest engagement?

RQ4: What vocabulary do populists use to transfer their message, and what topics do they post about, with a special emphasis on the Corona crisis (a topic that affected all of the chosen countries)?

1.2. Research visual protocol

The initial dataset consisted of 3388 Instagram posts (excluding videos, IGTV, and stories) from 12 Instagram accounts of populists from the 1st of January 2020 until the 31st of December 2020. They comprised 3388 image *files* as only the first picture in the carousel was chosen for the analysis using PhantomBuster (<https://phantombuster.com>). Each of the 3388 images obtained from the Instagram posts was processed through Google Cloud Vision API (Omena, J. J. et al., 2021) and Microsoft Azure Computer Vision API with Memespector (Chao, 2021) to identify: GV_Label_Descriptions, GV_Web_Entity_Descriptions, GV_Web_BestGuessLabels (<https://cloud.google.com/vision/docs/detecting-web>), MA_Tags, MA_Description_Captions, MA_Objects (<https://azure.microsoft.com/en-us/services/cognitive-services/computer-vision/#features>) which were used for further analysis.

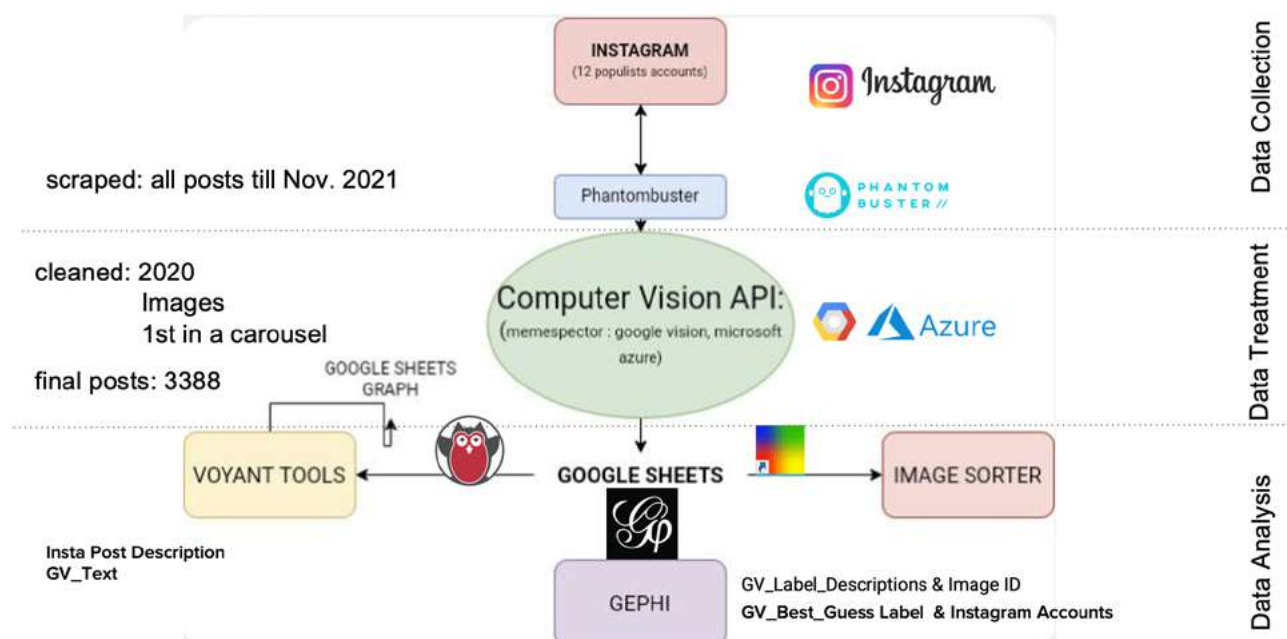


Figure 1. Visual protocol of the project. Source: Own construction

2. Methodology

2.1. Data selection and list of chosen leaders

The 12 populist leaders selected (Kyle & Meyer, 2020) used to govern, in 2020, democratic countries in three parts of the World: America, Asia, and Europe. In America, the leaders are Jair Bolsonaro for Brazil, Andrés Manuel López Obrador for Mexico, and Donald Trump for the USA. The Asian continent is represented by Narendra Modi for India, Benjamin Netanyahu for Israel, and Mahinda Rajapaksa for Sri Lanka. In Europe, the populist leaders are Bojko Borissov for Bulgaria, Giuseppe Conte for Italy Conte, Viktor Orbán for Hungary, Andrzej Sebastian Duda for Poland, Janez Janša for Slovenia, and Aleksandar Vučić for Serbia.

The work began with exploring how these populist leaders appropriate their Instagram profiles. Along with an account overview, politicians were grouped, for the study purpose, according to the number of their followers. Since there is no academic agreement about this group division, the Influencer Marketing Hub (2021) differentiation was used as an authority. This Hub divides the categories of IG influencers being

Micro-influencer:	less than 15,000 followers
Regular-influencer:	between 15,000 and 50,000 followers
Rising-influencer:	between 50,000 and 100,000 followers
Mid-influencer:	between 100,000 and 500,000 followers
Macro-influencer:	between 500,000 and 1,000,000 followers
Mega-influencer:	more than 1,000,000 followers

Thus, in this study, the largest category is the “Mega-Influencers,” including half of the profiles. It shows a massive gap between the smallest accounts (@lopezobrador from Mexico and @b.netanyahu from Israel) and the biggest one, namely the Indian president. Indeed with his 63,5 million followers, Narendra Modi’s account (39,8M followers) exceeds the account of the second biggest one: Donald Trump.

Table 1. Populists in power in democratic countries and their Instagram accounts in 2020. Source: Own elaboration

Name	Username (@)	Country	Followers	Number of posts in 2020	Influencer classification
Narendra Damodardas Modi	narendramodi	India	63,5 m	39	Mega-influencer
Donald J. Trump	realdonaldtrump	USA	23,7 m	459	Mega-influencer

Table 1. (cont.)

Name	Username (@)	Country	Followers	Number of posts in 2020	Influencer classification
Jair M. Bolsonaro	jairmessiasbolsonaro	Brazil	19 m	556	Mega-influencer
Giuseppe Conte	giuseppeconte_ufficiale	Italy	1.8 m	225	Mega-influencer
Andrés Manuel López Obrador	lopezobrador	Mexico	1 m	35	Mega-influencer
Benjamin Netanyahu	b.netanyahu	Israel	1 m	625	Mega-influencer
Gotabaya Rajapaksa	gotabayar	Sri Lanka	248 k	111	Mid-influencer
Aleksandar Vučić	buducnostsrbihev	Serbia	156 k	721	Mid-influencer
Andrzej Duda	prezydent_pl	Poland	147 k	381	Mid-influencer
Viktor Orbán	orbanviktor	Hungary	124 k	134	Mid-influencer
Boyko Borissov	boyko.borissov	Bulgaria	38.2 k	61	Regular-Influencer
Ivan Janša	jjansasds	Slovenia	14.9 k	41	Micro-Influencer

2.2. Visual communication

A bipartite graph has been designed to get an overview of the image corpus and connect pictures with their accounts. The analysis focused on the Images described by labels given by Google Vision API. Using an image-label network allows the study of the visual image content through a literal description. It should be noted that this method results in a network according to the image classification by Google Vision through a pre-trained machine learning algorithm. Still, the resulting network facilitates a qualitative analysis of the image content and interpretation of the strategic image use of populists (Omena et al., 2021). The approach of computer vision-based analysis was chosen due to the sample size, as it has already proved helpful in previous work to examine patterns at a “macro-level” (Pearce & Gaetano, 2021). The network was rendered by Gephi displaying images and their assigned labels as nodes and the edges as relations between both node types. ForceAtlas2, as proposed by Omena et al. (2021) and Venturini et al. (2018) was used to detect clusters among the labels and contiguous images. This particular layout is a force-driven algorithm that spatially arranges nodes by repulsion and attraction, with edges acting as attractive forces and nodes acting as repulsive forces between nodes, drawing nodes with the same or similar labels closer together (Venturini et al., 2018). Since no unique clusters could be mapped through spatialisation, the *Modularity*

Class was then applied to identify the most densely connected communities (Blondel et al., 2008). In Figure 2, the resulting graph, coloured according to the detected Modularity Classes, is shown.

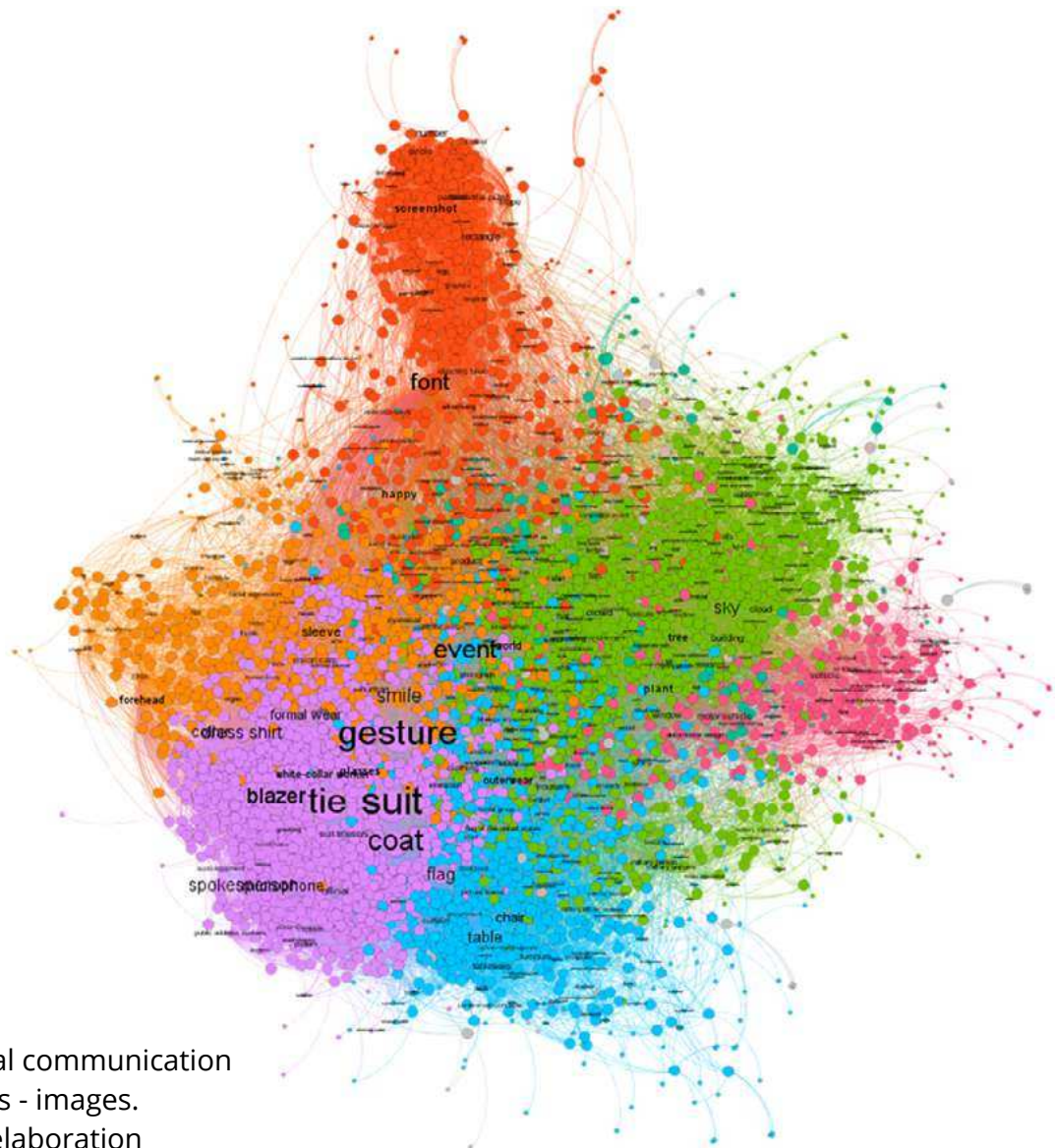


Figure 2. Visual communication strategy: labels - images.
Source: Own elaboration

Six main clusters were detected. To analyse clusters, labels were filtered by Partition (Modularity Class) to see the used labels per cluster and their occurrence. Table 2 shows all six clusters and their top labels ranked by occurrence count.

Table 2. Visual communication strategy: clusters and detected labels. Source: Own elaboration

purple	blue	rose	green	red	orange
Suit (1477)	Event (915)	Motor vehicle (185)	Sky (526)	Font (832)	Smile (629)
Gesture (1343)	Table (459)	Vehicle (185)	Plant (327)	Screenshot (304)	Sleeve (387)
Tie (1317)	Chair (366)	Automotive design (115)	Tree (271)	Number (236)	Happy (310)
Coat (1072)	Outerwear (277)	Tire (88)	Building (240)	Circle (233)	Forehead (244)
Blazer (752)	Furniture (187)	Wheel (88)	Cloud (240)	Product (228)	Human (213)

with Depth 1 to colour the two politicians in the periphery and their connected best guess labels. Due to this procedure, the common visual communication elements could be detected, and the specialties of the politicians in the periphery were investigated. The results can be found in the findings section.

Regarding colour analysis, the tool ImageSorter was used for the image dataset of every politician to detect noticeable patterns in their communication, which can be compared among other politicians.

With the use of such Vision APIs as Google Vision API and Microsoft Azure (GV_Web_Entity_Description and MA_Description_Captions respectively), we were able to determine whether the politician is depicted in a picture or some other people can be found there. That information was then used to identify the type of Instagram accounts.

According to Lilleker, in political communication, an “image must convey an idea that is consistent with the audience’s learned values” (Lilleker, 2019). As this study aimed to investigate the communication style of politicians from 12 different countries, taking into account the cultural background of each would have been beyond the scope of this study. For this reason, the study focused on describing the visible features of their communication strategy. However, for a deeper investigation, further work is suggested to focus on understanding their strategy in the light of cultural and social background.

2.3. Content engagement

A sequence of steps was made in order to depict the most and least engaged-with images on each of the populists’ Instagram pages. Engagement is calculated by adding the number of likes to the number of comments on each post extracted using PhantomBuster. Then, tables on Google Spreadsheets allowed for ranging and numbering.

2.4. Written communication

To understand populist leaders’ communication strategies, every word they used in 2020 on Instagram was analysed with the Voyant Tool (<https://voyant-tools.org/>) and the Google Vision API. The latter identified the words written in the pictures. The collected text corpus was then translated into English for adequate results. Both the description of the posts and the posts with text were qualitatively analysed as some of the leaders post on Instagram screenshots of their own tweets such as Jair Bolsonaro.

3. Findings

3.1. Visual communication

The two most significant detected label clusters were “Personal power” and “National pride.” “Personal power” is conveyed by pictures that show the politicians in an official setting, usually in the centre, wearing suits and ties, holding a speech, pictured alone or with other political leaders. The “National pride” cluster contains pictures that show national monuments like the White House of the USA or the Triglav Mountain in

Slovenia and pictures with the military or the president in military clothes. This indicates that populist leaders use the strategy of posting pictures that show them as a strong central personality who is very connected with the country and its traditions. However, almost as a counterpole to this unattainable figure, the two clusters, “Everyday life” and “ordinary people”, who together account for nearly 30% of the total network, convey a sense of the politician being accessible and a man just like everyone else. Some dominant pictures show their daily life at work, eating with their families, or at social events. Also, close-ups, where the political leader is still in the middle, are found in this cluster but with a positive and friendly facial expression and often in leisure clothing. Figure 4 provides a stereotypical picture of every cluster.



Figure 4. Visual communication strategy: Cluster Descriptions. Source: Own elaboration

Since the cluster graphic elements account for 15% of the network and contain mainly screenshots showing pure text and small amounts of pictures with slogans, this cluster was closely investigated in point 3, “Written content”.

The detected visual strategy is also mirrored in the analysis of the Best Guess Label - Instagram Account Network. Nodes with the highest in-degree (Figure 5) shown by size and colour in the centre of the network can also be found in the above described dominant clusters:

- Personal power: public speaking, suit, official, business
- National pride: tree, sky, photograph, crowd
- Everyday life: social group, furniture

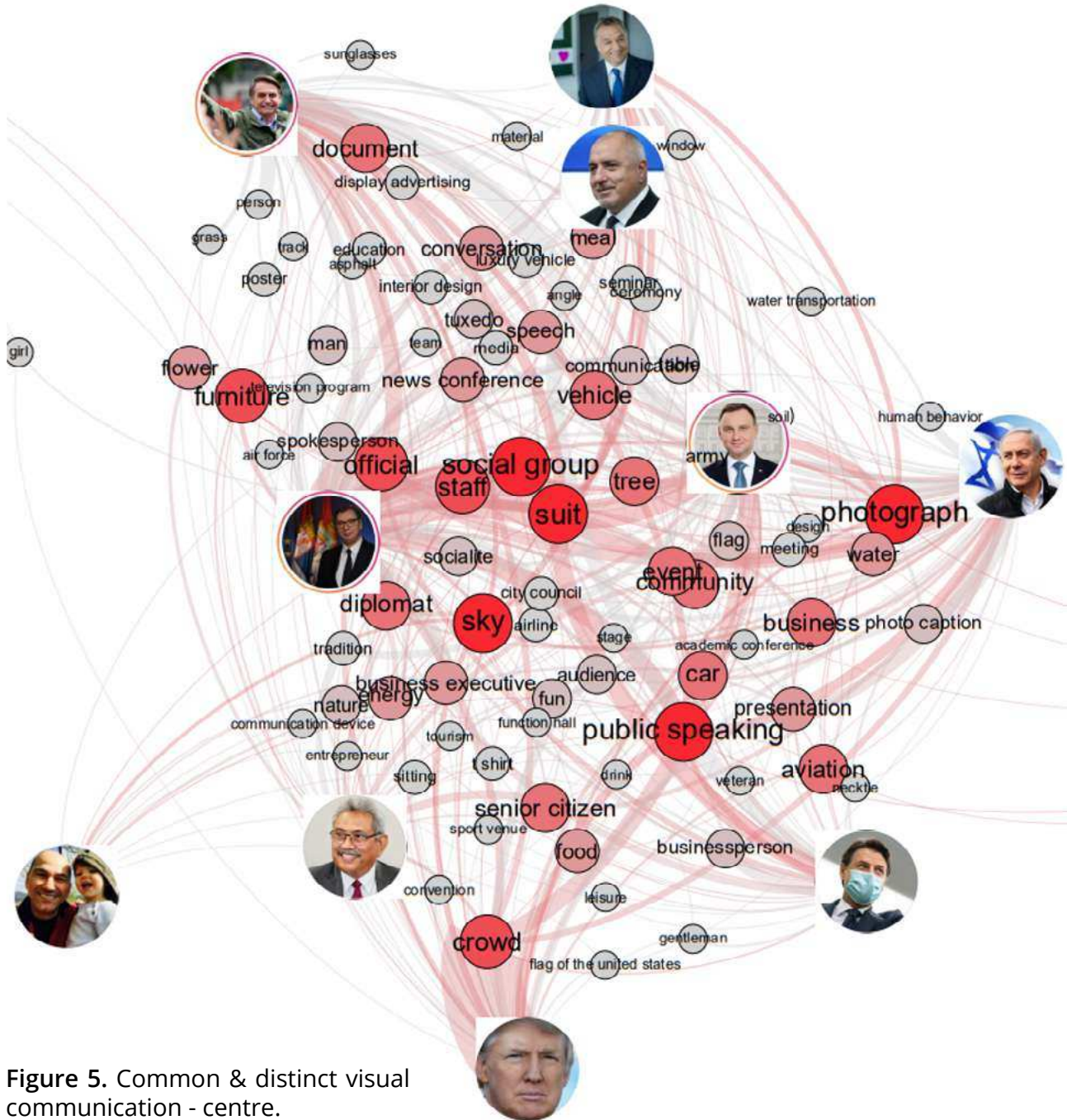


Figure 5. Common & distinct visual communication - centre.
 Source: Own elaboration

When looking at the centre, we can see that all politicians except two are very close to the centre, which indicates that their visual content posted contains similar topics. The best-guessed terms associated only with one Instagram account usually have names or descriptions of the specific politicians and their countries, which is why this analysis focused only on the more general words. By closely investigating the connected labels to Narendra Modi and Andrés Manuel López Obrador in the periphery, one can see that

the labelling algorithm of Google Vision influenced their spatial position. In the case of Andrés Manuel López Obrador, Google Vision extracted the labels in Spanish. It often used his initials “amlo”, but he is using similar topics as detected in the centre: e.g. official clothing. Another influence on his spatialisation is the low amount of pictures that were extracted from his Instagram account; therefore, fewer connections could be made which is also true for Narendra Modi. However, in the case of the Indian leader specific general descriptions were labelled differently by Google Vision due to the country’s specifics: Orange, Festival, Writing. The pictures corresponding to those descriptions could have also been labelled as Orange = official, business, Festival = crowd, and Writing = document, which would draw Narendra Modi’s node closer to the centre. Therefore, it can be argued that similar visual content is used by all of the 12 investigated populist leaders. They convey power through posting pictures where they are shown in official settings, alone in the centre or with other political leaders and showing solidarity with one’s own country by posting monuments, traditional celebrations but also pictures from their everyday lives.

3.2. Colour analysis

Having performed the colour analysis, we can divide politicians into three main groups:

- a) with a specific pattern of action;
- b) with the will to transmit a mood;
- c) without a detectable style.

The first category (Figure 6) contains politicians whose pictures overall transmit a mood with homogeneous colour palettes despite defined clusters. Orbán’s pictures, for instance, are dark and shady, with colours just related to flags and his figure, which are highlighted. Rajapaksa’s images have a prevalence of the orange colour, which is highly associated with religious ceremonies. Furthermore, the Bulgarian Borissov has a majority of blue, with images representing huge infrastructures. Finally, Modi’s dataset has a prevalence of the orange colour as leader but images related to military power are also noticeable. The predominance of orange in the images of Rajapaksa and Modi is linked to cultural meanings that are also associated with the religion of Buddhism and Hinduism (Kudrya-Marais & Olalere, 2022).

The second category can group politicians that use specific patterns (Figure 7). For instance, Bolsonaro, Netanyahu, and Conte usually share Twitter screenshots. The Brazilian president shows two main clusters: the national flag and close-up photos with a blue background. Moreover, he shared the same photo many times, resizing it. Netanyahu’s profile shares many titles of articles related to the Covid vaccine and images with the Israeli flag. Conte’s dataset presents a group of close-up photos with a specific graphic containing his quotes, pictures taken while signing documents, and the country’s flag. Images of Trump’s profile present many graphics as well, related in particular to his election campaign. Moreover, he has many photos with crowds, mainly associated with the 2020 elections.



Orbán

Rajapaksa



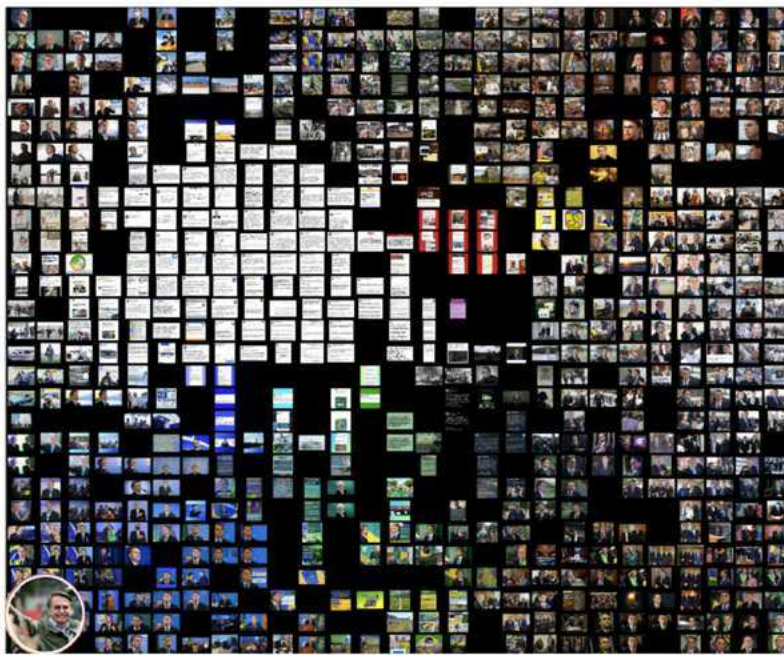
Borissov

Modi

Figure 6. Distribution of photos by colour. Source: Own elaboration

Even though the group includes politicians with a similar communication pattern, an oversight of the profiles shows Bolsonaro as an outsider due to his lower image quality and a “less refined” style. The remaining three politicians seem to have well-defined communication, with recurring graphic elements used for their posts. While this latter group seems to be guided by professional communication decisions made by social media managers, the “less refined” style of Bolsonaro’s profile may also result from a professional, intentional choice.

The overview through Image Sorter of the remaining politicians does not present noticeable outlines.



Bolsonaro



Netanyahu



Conte

Trump

Figure 7. Distribution of photos by colour. Source: Own elaboration

3.4. Personal / politician-oriented vs. institutional accounts

The accounts of Mega-influencers can be grouped either into:

- 1) *personal or politician-oriented* if more than 60% of posts include politicians themselves;
- 2) *mixed accounts* with circa an equal distribution of pictures with the politician and other content;
- 3) *institutional accounts* with more than 50% of posts portraying other people, events, text content, or/and architecture.

The accounts of Narendra Damodardas Modi (@narendramodi) and Jair M. Bolsonaro (@jairmessiasbolsonaro) were identified as *personal accounts*. Indeed, they use Instagram as a diary, posting pictures with their portraits in different environments. The accounts of Donald J. Trump (@realdonaldtrump), Giuseppe Conte (@giuseppeconte_ufficiale), and Andrés Manuel López Obrador (@lopezobrador) are *balanced in terms of content*. These politicians publish pictures of them and other people and events. Finally, the Instagram account of Benjamin Netanyahu (@b.netanyahu) is the only account that can be characterized as an *institutional one* because the majority of posts include content that does not have photos of Benjamin Netanyahu himself.

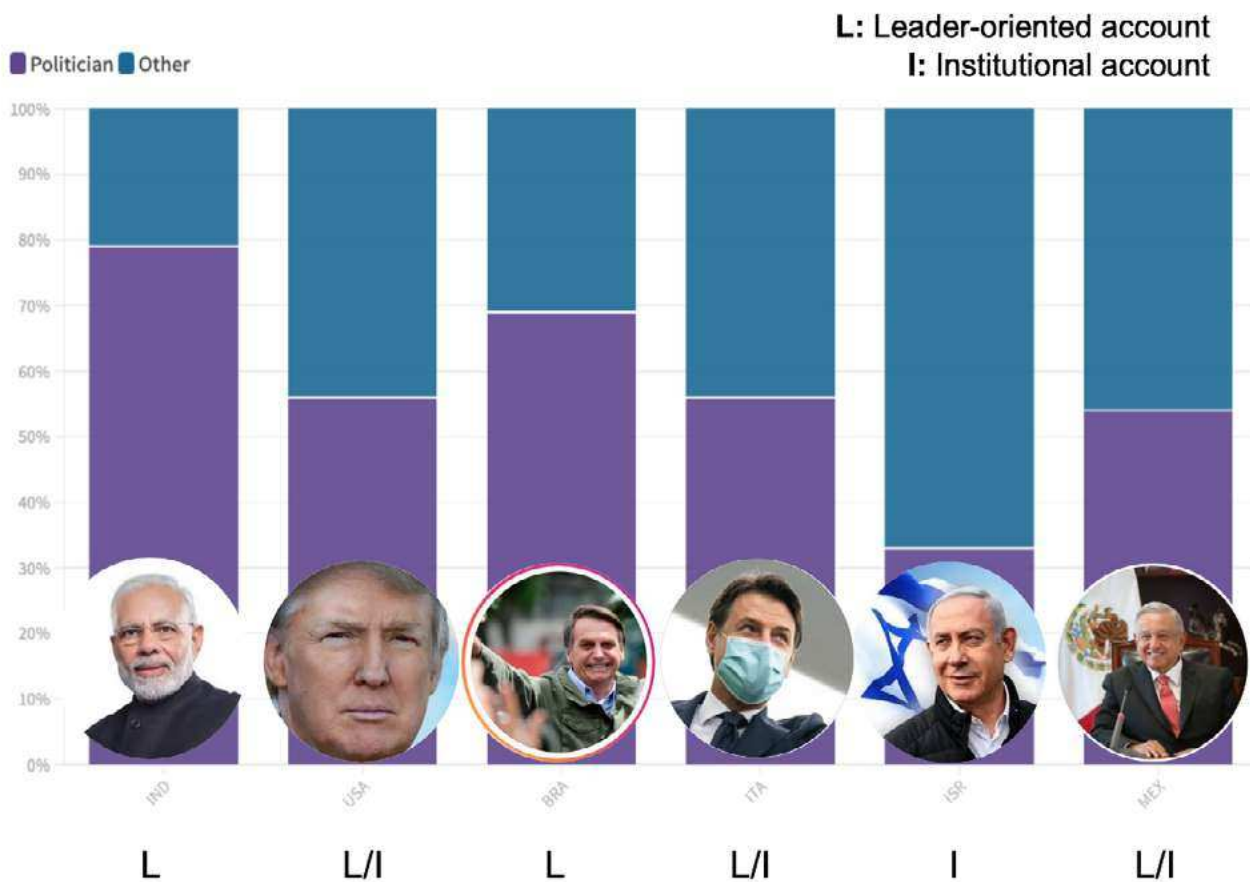


Figure 8. Shares of photos with Mega-influencers, present on them.
Source: Own elaboration

3.5. Content engagement

Analysing the engagement level of Instagram posts, several insights were found. First, in accordance with common patterns, Instagram accounts were grouped into three tables.

Insight 1: Close-up portraits and images depicting 'ordinary life' increase engagement, while formal settings and posters decrease the number of likes and comments (Figure 9).



Figure 9. Insight 1 illustration. Own elaboration.

Aleksandar Vučić (@buducnostsrbiyeav) is recognized in the five pictures from the list of the most engaged-with publications as they are his portraits. At the same time, it is pretty challenging to find him on images from the list of least engaged with. As for Andrés Manuel López Obrador (@lopezobrador), the President of Mexico is usually easily recognized in the pictures on his Instagram page; however, it is evident that his portraits taken outside are more engaged-with than the photos taken in his office. Posts with portraits of Giuseppe Conte (@giuseppeconte_ufficiale) also produce more likes and comments. Donald J. Trump's Instagram (@realdonaldtrump) shows that events sometimes are more important than the pictures chosen to depict them; however, it is clear that pictures with the face of Trump are pretty popular and that 'professionally designed' posts are not as engaging.

Insight 2: Posts with family members and children are very popular, while official meetings do not produce so many likes and comments (Figure 10).

For example, Ivan Janša (@jjansasds) publishes pictures with his wife and children in informal settings, and these pictures are more engaged-with than the so-called 'official publications.' Posts by Boyko Borissov (@boyko.borissov) and Orbán Viktor (@orbanviktor) with children produce significantly more likes and comments than pictures depicting their working environment. The President of Poland, Andrzej Duda (@prezydent_pl), quite often publishes pictures with his wife, Agata Kornhauser-Duda, the current first lady of Poland, and these pictures produce much more engagement than other posts.



Figure 10. Insight 2 illustration. Source: Own elaboration

Insight 3: Posts with a lot of text on pictures are significantly less popular than others (Figure 11).

Instagram accounts of Gotabaya Rajapaksa (@gotabayar), Benjamin Netanyahu (@b.netanyahu), Jair M. Bolsonaro (@jairmessiasbolsonaro), and Narendra Damodardas Modi (@narendramodi) prove that Instagram is not a platform for posting text content if it is not a screenshot of a tweet. Posts with politicians themselves in them are substantially more engaging than posts with their quotes, screenshots of articles, and posters.



Figure 11. Insight 3 illustration. Source: Own elaboration

3.6. Written communication

Some of the most used words detected by the text analysis are: “Day”, “People”, “President”, “Health”, “Citizens”, “Government”, “National”, “State”, “Forces” or “Measures”.

"In honor of Serbian heroes who lost their lives in Kosovo and Metohija. 38 The 38 were killed in the battle of baskets. Serbia remembers!" - The President of the Republic of Serbia, Aleksandar Vučić, laid a wreath on the monument to the father-in-laws 1998-1999. Years."

Aleksandar Vučić

Populist leaders are very devoted to promoting their national culture to gather the people around the nation. Pictures present many national buildings, flags, and traditional events, a national communication reflected in the language. All the leaders use words that refer to their country and allow them to promote their patriotic sentiment. They also use terms that evoke power and strength, talking about decisions they make for the country and demonstrating the power of the government they rule. The effect is the emphasis on their power and strength, a strategy reflected in the pictures, where the leaders show themselves in situations where they demonstrate control and influence.

"I'm not fighting for me, I'm fighting for the 74.000.000 million people (not including the Trump ballots that were "tossed"), a record for a sitting President who voted for me !"

Donald Trump

The analysis shows that they also use trivial language, which evokes routine and everyday life. For example, they write about their family, and some, like Modi, use religious and spiritual lexicon. People can relate to this vocabulary, making them think that populist leaders come from the people, live like them, and use a simple language that everybody can understand and relate to.

Populist leaders are cautious about the type of language and vocabulary they use on Instagram. They use words about power and the nation to promote their political leadership and trivial vocabulary to get close to the people and gain popularity.

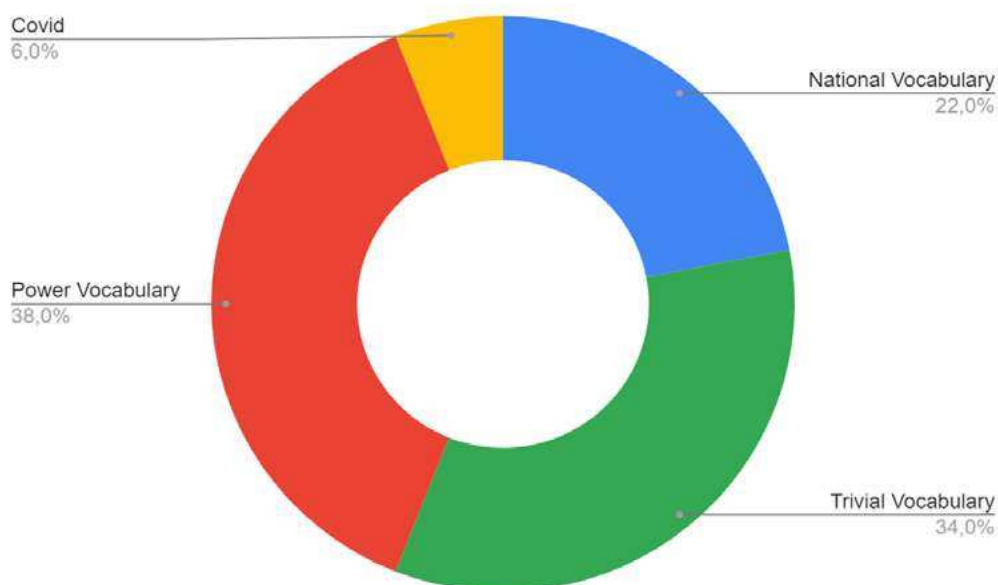


Figure 12. Distribution of different types of vocabulary across politician's posts.

Source: Own elaboration

Coronavirus represents only 6 % of the words used by all the populists. The frequency of mention of COVID-19 (and relatable words) was also analysed. It was calculated country by country (e.g., 25% in Rajapaksa's post of posts in 2020, while 2% in Duda's communication), but since the frequency is based on a %, it varies a lot due to considerable differences in the total quantity of posts between the leaders. But there is a difference between text and images about COVID-19. It was noticed that Benjamin Netanyahu posted a lot of pictures with vaccines because he wanted to promote the scientific progress and economy of his country. Still, in the text analysis, we only found 2% of posts mentioning covid-19. In the end, the populist leaders did not mention covid a lot in their written communication and preferred to underline the messages of the pictures by using an impactful vocabulary. Nevertheless, it could be seen that political leaders of countries critical of the existence of Corona or downplaying its impact rarely addressed the issue in written or visual form.

"The Chinese vaccine of João Dória.

For my government, any vaccine, before being made available to the population, should be scientifically proven by the Ministry of Health and Certified by ANVISA.

The Brazilian people will not be a guinea pig.

No billionaire is justified for a financial contribution in a medicine that has even surpassed its testing phase.

In view of the foregoing, my decision is not to acquire the said vaccine."

Jair Bolsonaro

4. Conclusion and future studies

Connecting our findings with the theoretical framework described in the introduction, we observe that the main clusters detected during the network analysis may represent different characteristics of populism. The cluster "Personal power" relates to populist leaders always being portrayed as extraordinary personalities, a characteristic described in Moffitt (2016). Besides, one of the strategies such politicians apply is to address those "pure" people they supposedly represent and show that they all have one mission (Mudde & Kaltwasser, 2018). This explains why there is a cluster of "Ordinary people". Populists also strive to get closer to the "pure" lower culture of the people rather than the higher culture of the elite (Ostiguy, 2009). Hence, the cluster of images "Everyday life" emerges. Furthermore, as a result of the engagement analysis, we discovered that images of "ordinary life", unlike images in official settings, and photos with family members and children produce the most likes and comments.

There are two strategies of populists on social media that stand out: 1) either a politician heavily relies on their charisma and, thus, the account is more politician-oriented and, thus, may even seem like a personal account (for instance, Narendra D. Modi's and Jair M. Bolsonaro's accounts); 2) or a politician posts images where he is not depicted, so he is not the central figure of the publication, and such an account may be characterized

as an institutional one (for example, Benjamin Netanyahu). However, it is true that in order to detect the distinctive features of populists, there is a need to compare the content of these populists with the content of other politicians who are not considered to be such, and that may become the basis for future studies.

While the use of digital methods and the computer vision API has allowed us to look at the issue of communication strategy from different angles, there are still some limitations that need to be kept in mind. As far as the labels attributed by Google Vision are concerned, they follow a certain hierarchy in interpreting the data that must not be disregarded. In this paper, a recording, like already suggested by D'Andréa & Mintz (2021), was done by the authors based on the clusters displayed through Google Vision labels and Gephi, as described in the Methodology section. Another interesting finding concerning the labelling algorithms was the output of specific images not in English but in the national language of the country where the image was taken. While no conclusion can be drawn as to when this is the case, it is something we should consider for future studies

Regardless of the limitations, the use of digital methods, computer vision APIs, and the representation of networks made it possible to answer all research questions to a certain extent. The findings indicated that populist leaders post images that convey personal and national power. They convey the message that they are the embodiment of national power, strong leaders on whom the people can rely, but also show that they are close to the people, that they follow traditions and that they can be referred to and relied upon. Through Instagram, the image conveys that they are at the same time ordinary citizens and strong men who are loyal to their country and act only in the best interests of it, as the textual *issuefication* of the global topic, COVID-19, has shown.

As for future studies, this study can be enriched by emotion analysis (using APIs) and sentiment analysis (using Python) of the text in the images and the post descriptions. Thus, an overview can be given of the range of emotions that populist leaders use for their communication, which seems most attractive to their audience/voters. To get a holistic picture of their communication strategies, we suggest that in further research, to also examine posts of these politicians on other well-known platforms such as Twitter and Facebook and extend the time frame of the analysed data. This would provide insights into the dynamics of their visual communication and show how and if it has changed over time.

Finally, to better understand the differences and similarities between the Instagram communication of populist and non-populist leaders, a comparative analysis of the visual communication strategies of both camps would be an interesting topic for future research.

Author contributions

Ilya Lavrov (IL), Chiara Miozzo (CM), Marie Palaffre (MP), Franziska Schranz (FS)

Project pitch, the initial research and design: IL

Project introduction and contextualization: IL

Research design: IL, FS, CM, MP

Dataset building and research visual protocol: IL, FS, CM, MP

Data analysis (and respective visualisations): image-label network analysis (FS), engagement analysis (IL), determining the type of accounts (IL), textual analysis (MP), image grouping/colour analysis (CM), Instagram-Profile appropriation (CM, FS, IL, MP)

References

- Blondel, V.D., Guillaume, J.-L., Lambiotte, R. & Lefebvre, E. (2008). Fast unfolding of communities in large networks, *J. Stat. Mech.*, P10008.
- Chao, J. (2021). "Memespector GUI: Graphical User Interface Client for Computer Vision APIs [Computer Software] (Version 0.1)" [C#] (Original work published 2021). <https://github.com/jason-chao/memespector-gui>
- Google Cloud.(2022). Features List | Cloud Vision API. Google Cloud. <https://cloud.google.com/vision/docs/features-list>
- D'Andréa, C. and Mintz, A. (2021). "Investigating cross-platform visual issuefication." #SMARTDataSprint. <https://smart.inovamedialab.org/past-editions/2021-platformisation/project-reports/investigating-cross-platform/>
- Dittrich, P. J. and the Jacques Delors Institut. (2017). "Social networks and populism in the EU. Four things you should know." <https://institutdelors.eu/wp-content/uploads/2020/08/socialnetworksandpopulism-dittrich-jdib-april17-2.pdf>
- Ekman, M. and Widholm, A. (2017). "Political Communication in an age of visual connectivity: Exploring Instagram practices among Swedish politicians." *Northern Lights: Film & Media Studies Yearbook*, 15(1), 15–32. https://doi.org/10.1386/nl.15.1.15_1
- Farkas, X. and Bene, M. (2021). "Images, Politicians, and Social Media: Patterns and Effects of Politicians' Image-Based Political Communication Strategies on Social Media." *The International Journal of Press/Politics*, 26(1), 119–142. <https://doi.org/10.1177/1940161220959553>
- Freedom House. (2020): "Freedom in the World 2020 A Leaderless Struggle for Democracy". Available on the Internet: https://freedomhouse.org/sites/default/files/2020-02/FIW_2020_REPORT_BOOKLET_Final.pdf
- Gordillo-Rodríguez M.T. and Bellido-Pérez, E. (2021). "Politicians self-representation on Instagram: the professional and the humanized candidate during 2019 Spanish elections." *Observatorio*, 15(1), 109-136.
- Hand, M. (2012). "Ubiquitous photography." UK: Polity Press.

- Influencer Marketing Hub. (2021). Influencer Marketing Benchmark Report https://influencermarketinghub.com/ebooks/influencer_marketing_benchmark_report_2021.pdf
- Kenny, P. D. (2021). "The strategic approach to populism". In D. B. Subedi, T. Lynch, A. Scott, & H. Brasted (Eds.), *Handbook on populism in the Asia Pacific*. Routledge.
- Kyle, J. and Meyer, B. (2020). "High Tide? Populism in Power, 1990–2020". Institute for Global Change. <https://institute.global/policy/high-tide-populism-power-1990-2020>
- Kudrya-Marais, E. & Olalere, F. (2022). The Spiritual Language of Colour: A Case Study of Colour Symbolism in Hindu Community in South Africa. *Pharos Journal of Theology*. 103. 10.46222/pharosjot.10321.
- Lalancette, M. and Raynauld, V. (2017). "The power of political image: Justin Trudeau, Instagram, and celebrity politics." *American Behavioral Scientist*, 63(7), 888–924. <https://doi.org/10.1177/0002764217744838>
- Lilleker, D. (2019). The Power of Visual Political Communication: Pictorial Politics Through the Lens of Communication Psychology. 10.1007/978-3-030-18729-3_3.
- Maarek, P. J. (2014). "Politics 2.0: New Forms of Digital Political Marketing and Political Communication." *Tripodos*, 1(34), 13-22. http://www.tripodos.com/index.php/Facultat_Comunicacio_Blanquerna/article/view/163
- Miller-Idriss, C. (2019). "The Oxford Handbook of Populism", edited by Cristóbal Rovira Kaltwasser, Paul Taggart, Paulina Ochoa Espejo, and Pierre Ostiguy. *Populism*, 2(2), 247–251. <https://doi.org/10.1163/25888072-02021036>
- Mudde, C. and Rovira Kaltwasser, C. (2018). "Studying Populism in Comparative Perspective: Reflections on the Contemporary and Future Research Agenda." *Comparative Political Studies*, 51(13), 1667–1693. <https://doi.org/10.1177/0010414018789490>
- Omena, J. J. and Amaral, I. (2019). "Sistema de leitura de redes digitais multiplataforma." In Janna Joceli Omena (Ed.), *Métodos Digitais: Teoria-Prática-Crítica* (pp. 121–140). Lisbon: ICNOVA.
- Omena, J. J., Elena, P., Gobbo, B., & Jason, C. (2021). "The Potentials of Google Vision API-based Networks to Study Natively Digital Images". *Revista Diseña*, 19. <https://doi.org/10.7764/disena.19.article.1>
- Ostiguy, P. (2009). "The high and the low in politics: a two-dimensional political space for comparative analysis and electoral studies". Kellogg Institute, https://kellogg.nd.edu/sites/default/files/old_files/documents/360_0.pdf
- Page, J. T. and Duffy, M. E. (2016). "What does credibility look like? tweets and walls in U.S. presidential candidates' visual storytelling." *Journal of Political Marketing*, 17(1), 3–31. <https://doi.org/10.1080/15377857.2016.1171819>
- Pearce, W., & De Gaetano, C. (2021). Google Images, Climate Change, and the Disappearance of Humans. *Diseña*, (19), Article.3. <https://doi.org/10.7764/disena.19.Article.3>
- Russmann, U., Svensson, J., & Larsson, A. O. (2019). "Political parties and their pictures: Visual Communication on Instagram in Swedish and Norwegian election campaigns." *Visual Political Communication*, 119–144. https://doi.org/10.1007/978-3-030-18729-3_7

- Schill, D. (2012). "The Visual Image and the Political Image: A Review of Visual Communication Research in the Field of Political Communication." *Review of Communication*, 12(2), 118–142. <https://doi.org/10.1080/15358593.2011.653504>
- Spiekermann, Kai (2020). "Why Populists Do Well on Social Networks." *Global Justice: Theory Practice Rhetoric* 12 (2):50-71.
- Stout, D. W. (2021). "Social Media Statistics 2022: Top Networks By the Numbers. Dustin Stout." <https://dustinstout.com/social-media-statistics/>
- Strand, R. T. and Schill, D. (2019). "The Visual Presidency of Donald Trump's First Hundred Days: Political Image-Making and Digital Media." In A. Veneti, D. Jackson, & D. G. Lilleker (Eds.), *Visual Political Communication* (1st ed. 2019 ed., pp. 167–187). Springer International Publishing. <https://doi.org/10.1007/978-3-030-18729-3>
- Tambini, Damian (2018) Social media power and election legitimacy. In: Tambini, Damian and Moore, Martin, (eds.) *Digital dominance: the power of Google, Amazon, Facebook, and Apple*. Oxford University Press, New York, NY, pp. 265-293. ISBN 9780190845117
- Venturini, T., Jacomy, M., & Jensen, P. (2019). "What Do We See When We Look at Networks. An Introduction to Visual Network Analysis and Force-Directed Layouts." *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.3378438>
- Zailskaitė-Jakštė, L., Ostreika, A., Jakštas, A., Staneviciene, E. & Damasevicius, R. (2017). Brand communication in social media: The use of image colours in popular posts. 1373-1378. 10.23919/MIPRO.2017.7973636.
- Weyland, K. (2020). "Populism's Threat to Democracy: Comparative Lessons for the United States." *Perspectives on Politics*, 18(2), 389–406. <https://doi.org/10.1017/s1537592719003955>


OPEN

ZONE



Game frames y populismo en las elecciones al Parlament de Catalunya de 2021. Actividad, temas y discursos de los candidatos en Twitter

Game frames and populism in the elections to the Catalan Parliament in 2021. Candidates' activity, issues and speeches on Twitter

 **Miguel Ibáñez-Cuquerella**
miguel.ibanez.cuquerella@gmail.com
Universitat de València

RESUMEN Este trabajo analiza la actividad en Twitter de los candidatos a la presidencia de la Generalitat en las elecciones al Parlament de Catalunya del 14 de febrero de 2021. El objetivo de la investigación es conocer qué uso han dado los candidatos a sus perfiles en Twitter, de qué temas han hablado y qué estrategias discursivas populistas han empleado. A través de una triangulación metodológica, el estudio aplica un análisis cuantitativo de la actividad en Twitter, un análisis temático de contenido y un análisis del discurso. Los resultados obtenidos muestran que los candidatos utilizaron Twitter de una forma unidireccional y que solo interactuaron con sus esferas partidistas, que usaron más los game frames –enmarcados para obtener votos– que los issue frames –propuestas sectoriales concretas– y que mostraron diversos rasgos de comunicación populista en sus discursos.

PALABRAS CLAVE Comunicación política; campaña electoral; Cataluña; strategic game frames; populismo; Twitter

ABSTRACT This paper analyses the Twitter activity of the candidates for the presidency of the Generalitat in the elections to the Parliament of Catalonia on 14 February 2021. The aim of the research is to find out how the candidates have used their Twitter profiles, what topics they have talked about and what populist discursive strategies they have employed. Through a methodological triangulation, the study applies a quantitative analysis of Twitter activity, a thematic content analysis and a discourse analysis. The results obtained show that the candidates used Twitter in a unidirectional way and that they only interacted with their party spheres, that they used game frames –framed to obtain votes– more than issue frames –specific sectorial proposals– and that they showed various features of populist communication in their speeches.

KEYWORDS Political communication; electoral campaign; Catalonia; strategic game frames; populism; Twitter

Edita: Vicente Fenoll

Recibido: 03/09/2021 | Aceptado: 21/10/2021

Cómo citar este artículo:

Ibáñez-Cuquerella, M. (2022). *Game frames y populismo en las elecciones al Parlament de Catalunya de 2021. Actividad, temas y discursos de los candidatos en Twitter*. *Dígitos. Revista de Comunicación Digital*, 8: 179-200. DOI: 10.7203/rd.v1i8.225



Game frames y populismo en las elecciones al Parlament de Catalunya de 2021. Actividad, temas y discursos de los candidatos en Twitter.

1. Introducción

Las redes sociales se han convertido en una herramienta muy importante de comunicación y marketing político. Twitter es una de las redes sociales más estudiadas en los procesos de comunicación política (Campos-Domínguez, 2017), especialmente en lo referido a la actividad de los actores políticos y los temas de sus agendas (Alonso-Muñoz et al., 2016b; Alonso-Muñoz & Casero-Ripollés, 2018a; Alonso-Muñoz & Casero-Ripollés, 2018b; Enli & Skobergø, 2013; García-Ortega & Zugasti-Azagra, 2014; García-Ortega & Zugasti-Azagra, 2018; López-García, 2016; López-Meri, 2016; Verger, 2015; Zamora & Zurutuza, 2014; Zugasti-Azagra & Sabés, 2015). Redes sociales como Twitter tienen un largo recorrido en contextos electorales complejos y muy competitivos como el catalán, marcado por la cuestión independentista. Desde que empezó el “procés” en septiembre de 2012, Catalunya ha registrado cuatro citas electorales y no ha agotado ningún ciclo político. La decimosegunda legislatura, previa a las elecciones autonómicas del 14 de febrero, estuvo marcada por la efímera presidencia de Quim Torra tras las elecciones convocadas por el gobierno de España a raíz de la aplicación del artículo 155 en Catalunya y la suspensión de la autonomía por el referéndum del 1 de octubre.

Las pugnas electorales y las posteriores negociaciones entre los partidos independentistas, Junts per Catalunya (JuntsxCat), Esquerra Republicana de Catalunya (ERC) y la Candidatura d'Unitat Popular (CUP), ha provocado constantes cambios de liderazgos y de aritmética parlamentaria en los últimos años, como la renuncia de Artur Mas en 2012 o la sustitución de Convergència por el Partit Demòcrata Europeu Català (PDeCAT) y su escisión liderada por Puigdemont en Junts per Catalunya (Junts). Entre los partidos contrarios a la independencia, el Partit dels Socialistes de Catalunya (PSC), Ciutadans (C's) y el Partit Popular Català (PPC), la pugna ha sido diferente, al haber sido C's el partido que desbancó al PSC como líder de la oposición y ser el PPC un partido residual. A este bloque se añade un competidor más, VOX, un partido de extrema derecha que irrumpió en la arena electoral española tras obtener doce diputados en las elecciones autonómicas de Andalucía en 2018, y que se iba a presentar en Catalunya por primera vez desde entonces. Entre estos dos bloques –referidos siempre a su posición respecto de la independencia– se encuentra En Comú Podem (ECP), el partido político formado en Catalunya por Podemos y los restos del espacio ecosocialista de Iniciativa per Catalunya-Verds (ICV), partidarios de un referéndum pactado con el Estado.

Finalizada la legislatura, las elecciones al Parlament de Catalunya se celebraron el 14 de febrero de 2021 en medio de la tercera ola de la pandemia de COVID-19 en España. Pese a su primera convocatoria en estas fechas y su aplazamiento debido a la situación sanitaria, el Tribunal Superior de Justicia de Cataluña falló –primero de forma cautelar y después de forma definitiva– que las elecciones debían mantenerse y la campaña electoral debía comenzar el 29 de enero.

Este trabajo tiene como objeto de estudio la actividad de los candidatos a las elecciones a la presidencia de la Generalitat durante la campaña electoral en Twitter. Los tres elementos en los que se centra este trabajo son:

1. El estudio de la actividad de las cuentas de Twitter para ver cómo interactúan con los demás usuarios, cuánto contenido generaron y de qué cuentas compartieron contenido, entre otros.
2. El estudio de los temas de los candidatos buscando los *strategic game frames* (Aalberg et al., 2011).
3. El estudio de los rasgos y el estilo de la comunicación populista en sus estrategias discursivas.

2. Marco teórico

2.1 Usos de Twitter durante las campañas electorales

Las Elecciones Generales de 2011 supusieron un punto de inflexión en la comunicación política digital en España (Gamir-Ríos, 2016); a partir de entonces, el uso de Twitter por parte de los actores políticos y especialmente durante las campañas electorales es una realidad plenamente consolidada y estudiada.

Superado el debate entre *ciberoptimistas* y *ciberpesimistas* sobre la potencialidad dialógica y de discusión pública de Twitter (Campos-Domínguez, 2017; López-García et al., 2018), la evidencia mostró que esta red social presenta un alto grado de polarización, pues facilita la creación de comunidades que comparten contenido común. En correlación con la “balcanización” del espacio público, ya que los partidos solo interactuaban con sus cuadros, estructuras, representantes y militantes, pero no con otros partidos, actores del sistema político o ciudadanos, lo que permitía que estas comunidades funcionaran como “cámaras de eco” (Aragón et al., 2013; Verger, 2015).

Larsson (2015), que identificó cuatro formas de los actores políticos para comunicarse a través de Twitter, apuntó a la difusión como la forma más utilizada por los políticos. Este punto ha sido evidenciado, también, por muchos otros autores (Alonso-Muñoz et al., 2016a; García-Ortega & Zugasti-Azagra, 2018; Jungherr, 2016; López-García, 2016; Segado-Boj et al., 2016; Suau-Gomila et al., 2020) lo que ha supuesto una forma de afrontar la comunicación con la ciudadanía poco interactiva. Para otros autores, la mera presencia en las redes sociales fue vista como un fin en sí misma, como un canal en el que amplificar su discurso sin tener que pasar por el filtro de los medios tradicionales y como un canal de autopromoción (Marcos-García et al., 2021; Segado-Boj et al., 2016). El perfil de los candidatos ha sido, generalmente, muy institucional y poco personal, centrado en su agenda propia y sin interés por conversar con la ciudadanía (Zamora & Zurutuza, 2014).

El retuit y la mención cumplen una función autorreferencial o de interacción. Así, se observó durante las campañas electorales la autorreferencialidad de los candidatos y las estrategias de comunicación híbrida, por medio de las cuales se compartía material y se interactuaba con las cuentas de los medios de comunicación tradicionales y de periodistas, y se compartían hipervínculos a sitios de web en los que aparecían los candidatos para, de esta manera, amplificar su aparición y presencia en los medios (García-Ortega & Zugasti-Azagra, 2018; Jungherr, 2016; Marcos-García et al., 2021; Miquel-Segarra et al., 2017). Esto reforzaba la lógica de la hibridación de los medios ya

que compartían contenidos e interactuaban con perfiles de los medios de comunicación tradicionales, pero en un nuevo medio como Twitter. Se relacionaba una mayor presencia de los actores políticos en Twitter con una mayor presencia de los políticos en los medios tradicionales –programas, entrevistas o debates electorales– (Alonso-Muñoz et al., 2016a; Alonso-Muñoz et al., 2016b; Jungherr, 2016).

Por todo ello, este trabajo plantea la primera pregunta de investigación:

- P1: ¿Qué tipo de actividad presentan los candidatos durante la campaña?

2.2 Los *game frames*: la política como juego

Concluyó López-García (2016), en su análisis sobre las elecciones generales de 2015, que el factor novedoso era que los “nuevos” partidos –Ciudadanos y Podemos– utilizaban un estilo de comunicación diferente en el que primaba la forma sobre el fondo, que hablaban de pronósticos, expectativas y sensaciones antes que de propuestas y que, en suma, su comunicación representaba la conversión de la política en una “carrera de caballos” (Aalberg et al., 2011; Alonso-Muñoz & Casero-Ripollés, 2018b; López-García, 2016).

Este fenómeno también ha sido observado por Aalberg et al. (2011) en su revisión sobre los *game frames* y los *strategy frames* (Aalberg et al., 2017). Estos autores plantearon el enmarcado de la política como un juego de estrategia de manera que los asuntos sustanciales y los *issues* quedaban desplazados. Este encuadramiento estaba fuertemente influido por el periodismo de “carrera de caballos” (Aalberg et al., 2011; Alonso-Muñoz & Casero-Ripollés, 2018b; López-García, 2016), que se centraba en el enfoque ganadores-perdedores, en las encuestas y los resultados electorales, y que usaba constantemente lenguaje bélico y/o de juego.

Alonso-Muñoz y Casero-Ripollés (2018a) propusieron una síntesis que aplicaba las categorías de Aalberg et al. (2011) al estudio de los discursos políticos en Twitter durante las campañas electorales. Este enmarcado de las noticias tiene un correlato temático en las agendas de partidos y candidatos. Así, las formaciones y los líderes tratan con mayor frecuencia temas relacionados con las encuestas electorales, los llamamientos al voto, su aparición en los medios de comunicación, los pactos de gobierno o los discursos movilizadores, que las propuestas sectoriales (Alonso-Muñoz & Casero-Ripollés, 2018a; Alonso-Muñoz & Casero-Ripollés, 2018b). Como concluyen Alonso-Muñoz y Casero-Ripollés:

“Los actores políticos no conciben esta red social como un instrumento desde el cual es posible lanzar propuestas concretas, sino como una red social como instrumento mediante el que pueden persuadir al electorado para que voten por ellos y contribuir a su victoria electoral.” (2018b: 22)

Estos enmarcados han sido estudiados, aunque no de forma idéntica, desde la perspectiva de la Agenda Setting. Estas investigaciones mostraron que en la gran mayoría de ocasiones el primer tema del que hablaban partidos y candidatos era sobre la misma campaña, esto es, utilizaban Twitter como tablón de anuncios de sus propios actos (Alonso-Muñoz et al., 2016b; Enli & Skobergø, 2013; García-Ortega & Zugasti-Azagra, 2014; García-Ortega & Zugasti-Azagra, 2018; López-García, 2016; López-Meri, 2016; Verger, 2015; Zamora & Zurutuza, 2014; Zugasti-Azagra & Sabés, 2015), lo que suponía desperdiciar su potencial dialógico para convertirse en un mero altavoz. Otros estudios apuntaron hacia una dirección similar: los temas más tratados

eran las cuestiones políticas generales y las propuestas sectoriales, si bien también eran recurrentes los temas relacionados con el desarrollo de las campañas electorales (Baviera et al., 2019; López-García et al., 2016), como también lo eran las cuestiones políticas e ideológicas genéricas (López-García, 2016; Suau-Gomila et al., 2020). Hay algunas excepciones, como Alonso-Muñoz & Casero-Ripollés (2018b) que encontraron que las agendas de varios líderes populistas estaban más centradas en propuestas concretas que en los enmarcados de juego.

Por todo ello, este trabajo plantea la segunda pregunta de investigación:

- P2: ¿Qué diferencias existen en los temas de los candidatos en función de su posición respecto a la independencia de Catalunya?

2.3 Redes polarizadas y élites populistas

La polarización es un fenómeno que se ha producido en mayor o menor medida a lo largo de la historia de las democracias liberales, si bien ha cobrado mayor relevancia en las últimas décadas. Países como España o Italia, con modelos mediáticos pluralistas polarizados (Hallin & Mancini, 2004; López-García & Gamir-Ríos, 2011), han experimentado más este fenómeno que otros estados. Hay estudios que relacionaron el grado de polarización con el modelo mediático del país, pese a haber dos países de modelo liberal -como Estados Unidos y Reino Unido- que la han acusado más (Fletcher et al., 2019; Urman, 2009).

Este trabajo entiende la polarización como un estadio de gran distancia ideológica entre los partidos pero de gran homogeneidad en su seno -los seguidores de los partidos adoptan con mayor frecuencia los discursos de sus formaciones, aunque sean contradictorios- (Druckman et al. 2013), así como el proceso mediante el cual las múltiples diferencias políticas se alinean en una única dimensión (Waisbord, 2020), de manera que los afectos por el partido y los líderes tienen mayor importancia que la ideología (Iyengar et al., 2012).

Algunos autores relacionan la polarización con los procesos comunicativos y políticos que vienen determinados por las élites políticas y mediáticas por lo que se considera al populismo como uno de los factores que la explican (Dubois & Blanck, 2018; McCoy et al., 2018; Tucker et al., 2018; Waisbord, 2020). La producción y difusión de los discursos populistas tiene una mayor adherencia en la comunicación online (Engesser et al. 2017a) por lo que en un contexto mediático híbrido como el actual, los actores populistas intervienen tanto en los medios tradicionales como en los digitales, pero de formas diferentes. En el caso de las redes sociales, la función de intermediación de los periodistas no se realiza, por lo que los mensajes pueden ser más personales y sensacionalistas, además de permitir centrar sus mensajes en grupos específicos (Engesser et al., 2017a; Ernst et al., 2017b; Mazzoleni & Bracciale, 2018). Las redes sociales son el espacio preferido para la comunicación populista, siendo Facebook la red social preferida por los actores populistas (Engesser et al., 2017a; Ernst et al., 2019; Hamelers & Schmuck, 2017).

El populismo es una ideología delgada compuesta por tres ideas centrales: la centralidad del pueblo, el antielitismo y la restauración de la soberanía, que, además, cuenta con una lógica comunicativa propia (Engesser et al., 2017a; Ernst et al., 2017a; Zulinello et al., 2018). La comunicación populista es el resultado de una estrategia política que usa mensajes clave ideológicos y unos elementos de estilo concretos (Ernst et al., 2019).

Diferentes estudios encontraron rasgos populistas en los discursos de los partidos tradicionales que no lo eran y apuntaron que los partidos en la oposición o los que tenían posicionamientos más extremos –normalmente de extrema derecha– eran más propensos a utilizar una comunicación populista (Engesser et al., 2017b; Mazzoleni & Bracciale, 2018).

Por todo ello, este trabajo presenta la siguiente pregunta de investigación:

- P3: ¿Qué estrategias discursivas populistas utilizan los candidatos?

3. Metodología

3.1 Corpus

La investigación aplica un análisis cuantitativo, de contenido y del discurso a los tuits publicados durante la campaña electoral por los candidatos a la presidencia de la Generalitat del PSC, ERC, Junts, VOX, la CUP, ECP, C's y el PPC, las formaciones para las que las encuestas preelectorales preveían una intención de voto superior al 3% (CIS, 2021; CEO, 2021). El corpus está formado por 3.357 tuits recogidos en los 18 días desde el inicio de la campaña electoral –29 de enero– hasta el día siguiente de las votaciones –15 de febrero–, esta delimitación temporal está en la línea de estudios anteriores (Alonso-Muñoz & Casero-Ripollés, 2018a; Alonso-Muñoz et al., 2016b; Gamir-Ríos, 2016; Gamir-Ríos et al., 2017; Gamir-Ríos, 2020; García-Ortega & Zugasti-Azagra, 2018; López-García, 2016; López-García et al., 2015; López-García et al., 2016; Marcos-García et al., 2020).

Los tuits y los datos referidos a seguidores, “me gusta” y retuits han sido extraídos a través de la aplicación *Twitonomy* entre las 00:00 del 29 de enero de 2021 y las 23:59 del 15 de febrero de 2021. El resto de cálculos, clasificaciones y codificaciones ha sido realizado de forma manual a partir de estos datos.

Este trabajo plantea tres estrategias complementarias para responder a las preguntas de investigación: un análisis cuantitativo de la actividad para la P1, un análisis de contenido para la P2 y un análisis del discurso para la P3.

3.2 Análisis cuantitativo

En primer lugar, el análisis cuantitativo ha abordado cuatro variables de las cuentas de los candidatos:

1. La actividad. El número de publicaciones de las cuentas durante el periodo de días estipulado, expresado a través de la frecuencia diaria de actualización obtenida de la división del número total de publicaciones de la cuenta de cada candidato entre los 18 días estipulados.
2. Su interacción con los demás usuarios, por lo que se tiene en cuenta la relación de retuits y respuestas sobre el total de publicaciones en esos 18 días y el origen de los retuits. Para la primera subvariable se desglosan los tuits propios, las respuestas y los retuits. Para la segunda se utilizan las siguientes categorías: a) la cuenta del propio candidato; b) cuentas oficiales del partido, cargos públicos, políticos condenados por el 1 de octubre y residentes fuera de España y pertenecientes a la esfera partidista; c) cuentas de medios de comunicación, periodistas, opinadores, profesionales de los medios, profesores, expertos y referentes del sector cultural; d) cuenta trol; e) instituciones; f) movimientos sociales; g) otras cuentas.

La esfera partidista está formada, también, por las cuentas que contaban con referencias al partido, candidato o a la campaña, tanto en sus elementos fijos como en el apartado multimedia. Estas cuentas se clasificaron de forma manual accediendo a las 2.423 cuentas retuiteadas.

La categoría de cuenta trol se inspira en el concepto de “trol independiente” (Tucker et al., 2018), referido a los usuarios que buscan fomentar la discordia y causar angustia emocional al resto.

3. La notoriedad. Se toma en consideración el tamaño de las comunidades de seguidores de los candidatos a través del aumento o la disminución de las mismas, tomando como referencia el día que empieza la campaña electoral y el día después de las elecciones.

4. El impacto. Se relacionan los “me gusta” y los retuits recibidos únicamente por los tuits propios y por las respuestas.

3.3 Análisis de contenido

El análisis de contenido (Krippendorff, 1990) se ha desarrollado a partir de la metodología diseñada por Alonso-Muñoz & Casero-Ripollés (2018a), por lo que se trabaja con las categorías de *game frame* e *issue frame*, obteniendo veinte variables –consultables en la Tabla 1–. De forma excepcional se crea una variable para englobar las políticas relacionadas con la gestión de la pandemia de la COVID-19; además, se crean variables específicas como trabajo y empleo, justicia, derechos y libertades, sanidad, seguridad ciudadana y medios de comunicación.

Tabla 1. Libro de códigos. Fuente: Elaboración propia a partir de Alonso-Muñoz & Casero-Ripollés (2018a).

Libro de códigos	
A - Game Frame	
1. Gobernanza	
1.1 Estrategia y pactos	Intención de construir un cierto tipo de gobierno o pactos.
2. Carrera de caballos	
2.1 Votaciones y resultados electorales	Encuestas, información, análisis y estimaciones de los resultados electorales. También los tuits relacionados con la acción de votar.
3. Política como individuo	
3.1 Temas personales	Vida personal de los actores políticos (gustos, preferencias, perfil...).
4. Estrategia política	
4.1 Función de organización y campaña	Operaciones de campaña y eventos de la organización, como <i>rallies</i> , mítines y eventos políticos y debate sobre los debates.
4.2 Arengas políticas	Se destaca la importancia de la unión del partido y se emplaza a los seguidores a unirse al partido y ganar. También llamamientos al voto y a la participación.

5. Gestión de apariciones en medios	
5.1 Relación con medios	Apariciones de los candidatos en los medios de comunicación (entrevistas, artículos de opinión, declaraciones a medios...).
B - Issue Frame	
6. Gestión de la pandemia	Gestión de la pandemia, ayudas, confinamientos, vacunación, medidas restrictivas, gestión del gobierno central o de la Generalitat, impacto en la economía...
7. Economía y hacienda	Propuestas económicas que no tengan que ver con la crisis de la COVID-19. Economía sostenible, modelo productivo, fiscalidad, presupuestos, sectores productivos, desarrollo tecnológico...
8. Trabajo y empleo	Reforma laboral, condiciones laborales, diálogo social, pensiones y seguridad social, políticas activas de empleo, formación para el empleo, autónomos y conflictos laborales.
9. Educación, cultura, ciencia y deportes	Cultura catalana, catalán, modelo educativo, política lingüística, colegios, institutos, universidades, deporte e industrias culturales.
10. Medioambiente y cambio climático	Energía, renovables, nuclear, cambio climático...
11. Infraestructuras y transportes	Rodalies, AVE, autopistas, conexiones con otros países, inversión pública (autonómica y estatal)...
12. Corrupción	Casos de corrupción y medidas para combatirla.
13. Justicia, derechos y libertades	Represión, tribunales, juicios del "procés", amnistía, euroórdenes, despolitización de la justicia, transparencia, ley electoral, libertades públicas, memoria histórica, relación Iglesia-Estado y regeneración democrática.
14. Modelo territorial	Nuevo referéndum, DUI, supresión del estado de las autonomías, cambios constitucionales, negociaciones con España y financiación.
15. Sanidad	Propuestas sobre sanidad fuera de la pandemia
16. Seguridad ciudadana	Mossos d'Esquadra, guardias urbanas, GC, CNP, modelo de seguridad, terrorismo, seguridad/inseguridad.
17. Igualdad y políticas sociales	Políticas de igualdad, LGTBI, juventud, derechos sociales, servicios públicos, dependencia, vivienda, okupación e inmigración.
18. Asuntos exteriores	Relaciones con otros países, con la UE...
19. Medios de comunicación	Medios de comunicación y CCMA
20. Otros	No cabe en las anteriores

El análisis se ha realizado sobre un total de 513 tuits de elaboración propia de los candidatos. Para validar la codificación se ha elaborado un test de intercodificadores sobre 50 unidades del total de la muestra para la que se ha obtenido un coeficiente de Alfa de Krippendorff de 0.797 (Igartua & Humanes, 2006) para la categoría del enmarcado utilizado en cada uno de los tuits.

3.4 Análisis del discurso

Para desgranar las estrategias discursivas populistas utilizadas por los candidatos durante la campaña, este trabajo toma como referencia los cinco rasgos del populismo –acentuación de la soberanía popular, defensa del pueblo, ataque a las élites, exclusión de los otros y apelación a la patria (Engesser et al., 2017b)– y las tres dimensiones del estilo populista –simplificación, emocionalización y negatividad (Engesser et al., 2017a).

4. Resultados

4.1 Análisis cuantitativo

La actividad de los candidatos durante la campaña fue muy diferente. La candidata que destacó por haber sido la más activa en sus publicaciones diarias –como puede consultarse en la Tabla 2– fue Laura Borràs con una media de 105,2 tuits al día. Esta gran actividad se explica porque dedicó gran parte de sus tuits a compartir contenido de otros usuarios y, en menor medida, a responderlos. Después de Borràs, con mucha diferencia, se encontraron el resto de candidatos. Illa destaca, en último lugar, porque no estuvo, prácticamente, presente en Twitter durante la campaña electoral.

El candidato con más tuits de contenido propio respecto al total fue Carrizosa, 50,5%. En cuanto a la difusión de contenido de otros usuarios –retuits–, la candidata que más lo hizo –con una clara diferencia en términos absolutos– fue Borràs con un 84,4%.

Respecto a los orígenes del contenido retuiteado –disponibles en la Tabla 3– hubo varios aspectos reseñables. En primer lugar, las cuentas pertenecientes a la esfera partidista eran la más presente en los retuits de todos los candidatos en unos porcentajes muy altos. En segundo lugar, destaca el contenido compartido de cuentas de medios de comunicación, periodistas o similares, entre los que destacaron Fernández (17,2%), Sabater (12%) y Borràs (12%). En el caso de esta última candidata resaltó que, por un lado, compartió un 10,5% de contenido de ciudadanía anónima, esto es, de parte de la ciudadanía que no mostraba ninguna afiliación partidista clara en sus perfiles de Twitter. Por otro, fue la candidata que más contenido compartió –en términos absolutos– de usuarios troles, seguida por Fernández y por Carrizosa. También destacó que Sabater (6,7%), Albiach (4,2%) y Borràs (0,9%) fueron las candidatas que más contenido compartieron de cuentas pertenecientes a movimientos sociales.

En último lugar, el candidato que más respondió a los usuarios de forma proporcional fue Illa con un 19,6%, pese a que fue Borràs la candidata que más lo hizo en términos absolutos. Este análisis destacó que los candidatos utilizaron muy poco la opción de responder a otros usuarios durante la campaña, si bien es verdad que hay dos casos en los que los candidatos respondiendo más que contenido propio compartieron en esta red social. Estos fueron los casos de Borràs –lo hizo de forma más acusada– y de Fernández –prácticamente respondió igual que generó contenido propio.

Tabla 2. Resultados análisis cuantitativo. Fuente: Elaboración propia.

	Illa (PSC)	Aragonès (ERC)	Borràs (Junts)	Garriga (Vox)	Sabater (CUP)	Albiach (ECP)	Carrizosa (C's)	Fernández (PPC)
Tuits	56	189	1.894	279	145	347	224	223
Propios	19	53	48	47	51	147	113	35
Retuits	26	116	1.598	208	75	143	106	151
Respuestas	11	20	248	24	19	57	5	37
Frecuencia diaria	3,1	10,5	105,2	15,5	8,1	19,3	12,4	12,4
Seguidores 00:00 29/01/2021	101.164	53.417	123.576	88.816	29.742	21.254	39.763	48.486
Seguidores 23:59 15/02/2021	102.385	54.539	125.553	95.982	32.324	26.556	40.549	51.700
Aumento de seguidores	1.221	1.122	1.977	7.166	2.582	5.302	786	3.214
"me gusta" recibidos	22.881	52.495	1.129.686	555.997	31.582	130.349	52.823	158.133
"me gusta"/ tuit propio o respuesta	16.424	12.260	86.401	10.6211	14.625	46.101	30.612	75.496
Retuits recibidos	8.703	21.633	39.4354	21.9733	10.271	46.507	26.897	47.556
Retuits/tuit propio o respuesta	5.718	1.986	21.420	37.665	3.605	20.314	15.350	19.755
Tuits a 00:00 29/01/2021	3.711	2.6049	16.3001	15.452	18.590	8.570	6.315	5.991
Tuits a 23:59 15/02/2021	3.768	2.6237	16.4822	15.731	18.736	8.916	6.536	6.219

Tabla 3. Cuentas de origen de los retuits. Fuente: Elaboración propia.

	Illa	Aragonès	Borràs	Garriga	Sabater	Albiach	Carrizosa	Fernández
Cuenta del candidato	0%	0,9%	0,1%	3,9%	2,7%	0%	0%	2,7%
Esfera partidista	88,5%	89,7%	75,3%	90,9%	77,3%	82,5%	94,3%	75,5%
Medios de comunicación	7,7%	5,1%	12%	5,3%	12%	7,7%	0,9%	17,2%
Instituciones	3,9%	4,3%	0,1%	0%	0%	0%	1,9%	0%
Trol	0%	0%	1,3%	0%	0%	0%	0,9%	3,3%
Movimientos Sociales	0%	0%	0,89%	0%	6,8%	4,2%	1,9%	0,7%
Otras cuentas	0%	0%	10,5%	0%	1,3%	5,6%	0%	0,7%

Todas las cuentas aumentaron seguidores, pero de forma desigual. La candidata que más seguidores consiguió fue Albiach, con un aumento del 25%. Después, en proporciones similares, Sabater, 8,7% y Garriga, 8,1% –es el candidato que más seguidores totales tiene–. El cuarto candidato que aumentó en seguidores fue Fernández, con un 6,6%. Se observó que los candidatos menos conocidos fueron los candidatos que más seguidores obtuvieron a lo largo de la campaña electoral. Por contra, candidatos como Aragonès –vicepresidente de la Generalitat–, Carrizosa –líder de la oposición–, Borràs –había sido diputada en el Congreso– e Illa –había sido ministro de Sanidad– lo hicieron un 2,1%; 2%; 1,6%; y 1,2%, respectivamente.

El candidato que guardó una proporción mayor de “me gusta” y retuits de sus propios tuits frente al total de mensajes emitidos fue Illa, con un 71,8% y 65,7% respectivamente. Los candidatos guardaron una proporción similar entre el número de retuits y el de “me gusta” recibidos, solo destacaron Illa y Sabater al guardar ambos una mayor proporción de “me gusta” que de retuits. También destacó el poco impacto de las publicaciones de Borràs, por debajo del diez por ciento y muy por detrás del resto de candidatos.

4.2 Análisis de contenido

Los encuadres más utilizados fueron los *game frames* con un 71,7% del total frente al 27,3% de los *issue frame*. En primer lugar, estaban los referidos a la estrategia política –campañas y discursos movilizadores– con un 34,9%, seguido de los de gobernanza (11,3%) y los candidatos (11,3%). Después estaban los referidos a la presencia de los candidatos en los medios de comunicación (9,9%). En cuanto a los *issue frames*, en primer lugar, se encontraron las políticas sociales y de igualdad (6,6%), seguidas de las referidas a la gestión de la pandemia (3,7%), economía y hacienda (2,7%) y justicia, derechos y libertades (2,5%).

Por bloques, los candidatos del bloque independentista hablaron más de la estrategia política. En segundo lugar, se situaron las propias candidatas como tema de sus tuits con el objetivo de crear un perfil más cercano con el resto de usuarios –destacan los tuits de Borràs con citas de escritores catalanes y de la literatura universal– mientras que Aragonès habló más de los pactos y de ERC como garantía de gobierno independentista y progresista. En las propuestas sectoriales los temas fueron variados: Aragonès habló mucho de gestión de la pandemia –el segundo candidato que más lo hizo por detrás de Carrizosa–, Borràs del modelo territorial y las relaciones entre Cataluña y España, y Sabater de justicia, derechos y libertades.

Albiach trató también los temas relacionados con la campaña y los discursos movilizadores en primer lugar. Por detrás de estos temas, habló sobre los pactos de gobierno, insistiendo en su partido como garantía de un tripartito progresista. La candidata destacó por ser la que más tuiteó sobre políticas sociales y de igualdad.

Los candidatos de los partidos contrarios a la independencia también utilizaron el desarrollo de la campaña y sus discursos en primer lugar durante la campaña. Carrizosa y Garriga fueron quienes más tuitearon sus apariciones en medios de comunicación –tradicionales y digitales–. Carrizosa y Fernández se centraron en los futuros pactos insistiendo ambos en que el PSC mentía cuando decía que no pactaría con ERC. Illa y Fernández fueron quienes más hablaron de ellos mismos siendo este último el que más lo hizo. Illa fue el que más tuiteó sobre educación, cultura, ciencia y deporte, así como el que más propuestas sobre sanidad –que no estaban relacionadas con la pandemia– hizo, justo por delante de Albiach. Garriga fue quien más habló sobre seguridad ciudadana –prácticamente el único–, centrando sus tuits en la supuesta inseguridad

ciudadana que existe en Catalunya debido, según él, a la inmigración. El resto de resultados se pueden consultar en la Tabla 4.

Tabla 4. Resultado análisis temático de contenido. Fuente: Elaboración propia.

	Illa	Aragonès	Borràs	Garriga	Sabater	Albiach	Carrizosa	Fernández
Game frames	42,1%	79,23%	79,2%	83%	62,8%	65,3%	73,5%	85,7%
Gobernanza	5,3%	20,8%	2,1%	0%	0%	22,5%	8%	8,6%
Carrera de caballos	5,3%	5,7%	6,3%	2,1%	0%	4,8%	5,3%	2,9%
Política como individuo	10,5%	5,7%	54,2%	6,4%	11,8%	6,1%	0,9%	22,9%
Estrategia política	21,1%	35,9%	16,7%	57,5%	49%	23,1%	41,6%	42,9%
Función de organización y campaña	0%	18,9%	10,4%	42,6%	23,5%	10,2%	23,0%	31,4%
Arengas políticas	21,1%	17%	6,3%	14,9%	25,5%	12,9%	18,6%	14,3%
Gestión de apariciones medios	0%	11,3%	0%	17%	2%	8,8%	17,7%	8,6%
Issue frames	42,1%	18,9%	18,8%	17%	37,3%	33,3%	26,6%	14,3%
Gestión de la pandemia	5,3%	9,4%	0%	2,1%	3,9%	2,0%	5,3%	2,9%
Economía y hacienda	0%	0%	0%	0%	2%	4,1%	6,2%	0%
Trabajo y empleo	5,3%	0%	0%	0%	0%	0,7%	0%	0%
Educación, cultura, ciencia y deporte	15,8%	0%	2,1%	0%	2%	1,4%	3,5%	0%
Medioambiente y cambio climático	5,3%	3,8%	4,2%	0%	2%	2,7%	0,9%	0%
Infraestructuras y transporte	5,3%	0%	0%	0%	0%	0,7%	0%	0%
Corrupción	0%	0%	0%	0%	0%	1,4%	4,4%	0%
Justicia, derechos y libertades	0%	1,9%	0%	0%	15,7%	0,7%	1,8%	2,9%
Modelo territorial	0%	1,9%	8,3%	0%	0%	0%	0%	2,9%
Sanidad	10,5%	0%	0%	0%	3,9%	4,1%	0,9%	0%
Seguridad ciudadana	0%	0%	4,1%	8,5%	2%	0,7%	0%	0%
Igualdad y políticas sociales	5,3%	1,9%	0%	4,3%	5,9%	14,3%	3,5%	5,7%
Asuntos exteriores	0%	0%	0%	0%	0%	0%	0%	0%
Medios de comunicación	0%	0%	0%	2,1%	0%	0,7%	0%	0%
Otros	5,3%	1,9%	2,1%	0%	0%	1,4%	0%	0%

4.3 Análisis del discurso

Salvador Illa

Illa, quien fue ministro de Sanidad hasta que empezó la campaña electoral y había ganado popularidad por su gestión de la pandemia de la COVID-19, fue un caso particular. Tuiteó poco durante la campaña por lo que resulta complicado hablar de la presencia del populismo en su discurso. Pese a esto, atacó a las élites políticas de Catalunya – su gobierno– cuando dijo que quería formar un gobierno que “genere confianza, que recupere el prestigio de las instituciones. Lo contrario de lo que ha estado sucediendo los últimos diez años”. Planteó también la defensa del pueblo, votar al PSC era “el voto útil para un nuevo Gobierno progresista que deje atrás diez años de decadencia”, y se mostró como la persona capaz de mejorar la vida de la gente y asegurar estabilidad política y económica.

Se observaron la simplificación y la emocionalización en su discurso. Simplificación, pues hubo una única causa que explicara la situación de Catalunya, gobernaban partidos independentistas desde hacía diez años, y la solución era un gobierno liderado por el PSC. La emocionalización en la medida en que se habló de la esperanza de que todo fuera a mejor, tanto política como económicamente, si el PSC gobernaba Catalunya.

Pere Aragonès

Aragonès, que era el vicepresidente de la Generalitat desde las elecciones de 2017 y había tomado relevancia como responsable de la Generalitat desde la inhabilitación del presidente Torra en septiembre de 2020, utilizó la reivindicación de la soberanía popular y la apelación a la patria como ejes de su discurso. La apelación a la patria – como una futura Catalunya independiente y progresista– fue una constante. Reivindicó la soberanía popular y se mostró como el líder que cumpliría los deseos de la gente, él estaba para que Catalunya “se comprometiera desde la izquierda con la república catalana”.

Atacó a las élites –el Estado español– cuando dijo que había mostrado su “odio y rencor” hacia la libertad –al recurrir la Fiscalía el tercer grado de los presos– o cuando lo reconoció como el enemigo a batir.

Se sirvió de la simplificación y la emocionalización. Simplificación, porque todos los problemas que tenía Catalunya se solucionarían con la independencia y emocionalización porque utilizó el miedo para construir su discurso, por ejemplo, cuando dijo: “Los jóvenes de este país también tienen en sus manos parar a la extrema derecha llenando las urnas de voto antifascista, republicano y progresista. La historia os necesita, movilizaos”. También utilizó la esperanza como una emoción para la transformación y para llamar al voto.

Laura Borràs

Borràs, que se convirtió en la candidata del partido de Puigdemont después de su escisión del PDeCAT, presentó un discurso con tres rasgos del estilo populista. Con la independencia de Cataluña como telón de fondo fueron numerosos los tuits en los que reivindicó la soberanía popular y la democracia como dos elementos ligados a este proceso. Se observó cómo puso al pueblo en la centralidad y cómo de él emergía el mandato que ella garantizó que cumpliría. Atacó a las élites –el Estado español, su gobierno e instituciones– refiriéndose al “club España” –PSOE, PP, ECP, C’s y VOX– que

ataca a Catalunya sistemáticamente y destacó que el Gobierno de España no dudó en pactar con la extrema derecha para atacarlos. Las apelaciones a la patria fueron un elemento transversal a todos los demás rasgos.

Redujo todos los problemas que tenía Catalunya a una única solución, su independencia, como cuando publicó: "La mejor política lingüística es la independencia. Dignidad y Resistencia. Independencia". No planteó los problemas como elementos complejos. Destacó, por el contrario, que utilizó mucho las emociones, concretamente la esperanza, como un rasgo asociado siempre al futuro de la independencia de Cataluña, como un futuro próspero para todos.

Ignacio Garriga

Garriga, fue diputado en el Congreso y participó de los debates al entender la Junta Electoral que era un partido representativo, planteó cuatro de los cinco rasgos del populismo y destacó por no reivindicar la soberanía popular. Las apelaciones a la patria fueron constantes. Las amenazas del pueblo fueron la inmigración y los separatistas. Esta inmigración estaba permitida por las élites políticas de Catalunya y del Gobierno de España, y por el "buenismo progre". La "mafia separatista" y la izquierda "abren sus chiringuitos" mientras "cierran nuestros comercios". Ante toda esta situación, Garriga se presentó como salvador: "daré voz a la Cataluña silenciada y harta de consecuencias de la mafia separatista y la izquierda".

La simplificación estuvo muy presente. La solución para los problemas de Catalunya era expulsar a los independentistas de la Generalitat. El miedo fue una constante, ya que reiteró que los inmigrantes cometían delitos –robos, agresiones, terrorismo y violaciones– y que eran un peligro real para las personas, aunque hubo algo de esperanza depositada en que VOX ganara las elecciones y pudiera expulsarlos. La negatividad recorrió todo el discurso de forma reiterada a partir del miedo, representando una Catalunya en la que a nadie le gustaría vivir.

Dolors Sabater

Sabater, que había sido alcaldesa de Badalona y fundadora de Guanyem Catalunya, la plataforma que se presentó conjuntamente con la CUP a las elecciones, utilizó la defensa del pueblo y el ataque a las élites. Planteó un pueblo amenazado por unas élites políticas represoras y represivas –el Estado español, la Generalitat y sus maquinarias judiciales– que tenían una "causa general contra el independentismo" y que era la extrema derecha la que "sigue dictando las órdenes", por lo que el pueblo estaba en peligro. Relacionó la represión que sufrió Rosa Parks en Estados Unidos con la conocida como ley mordaza. Sabater planteó la CUP y a ella misma como las soluciones y los actores que protegerían al pueblo de estas amenazas.

Esto también estuvo presente en el estilo de su discurso. Hubo simplificación porque redujo los problemas y sus causantes de Catalunya a la represión del Estado, la extrema derecha o los partidos unionistas. Negatividad porque insistió mucho en la idea del miedo, miedo ante la represión y ante la extrema derecha. Aunque frente a este miedo propuso, eso sí, una cierta dosis de esperanza tanto en la CUP y en ella como candidata como en la sociedad civil organizada.

Jèssica Albiach

Albiach, que era diputada desde 2015, desplegó tres rasgos del populismo en su discurso. Reivindicó la soberanía popular pidiendo el voto para su formación para que “ningún voto republicano, fraterno, feminista y verde se quede en casa” o cuando se erigió como la única capaz de garantizar un gobierno de izquierdas, “o desgobierno o un gobierno de izquierdas. Y os aseguro que el único voto que garantiza esta alternativa [...] es el voto a ECP”. Atacó a las élites e identificó a dos enemigos. Las élites económicas y políticas y la formación de extrema derecha VOX cuando dijo que eran “violencia en sí mismo” al ser negacionistas de “la violencia machista, el coronavirus y los derechos humanos”. Ella defendió al pueblo formado por “las mujeres de la limpieza, las cajas de supermercado, las polígonas y las “kellys”, un pueblo de “gente común”.

Uso la simplificación cuando se presentó como la garantía de un gobierno entre ERC y el PSC. En cuanto a la emocionalización, habló del amor y la esperanza como el voto a ECP en contraposición a la resignación y al odio de los votos a otros partidos.

Carlos Carrizosa

Carrizosa, que había sido líder de la oposición en Cataluña tras la elección de Arrimadas como lideresa del partido en el ámbito estatal, utilizó cuatro rasgos del populismo. La soberanía popular estuvo presente cuando planteó que los constitucionalistas eran los que tenían la capacidad de elegir, de votar a C's y así “acabar con el nacionalismo” y “ganar cuatro años de futuro”. Se planteó como defensa del pueblo, de la gente normal y constitucionalista. Atacó a las élites –Generalitat–, que eran una amenaza para la economía y los catalanes, “si no se supera el procés no habrá ni recuperación económica ni lucha contra la corrupción”. También excluyó a una parte de los catalanes que tienen una ideología diferente a la suya que era totalitaria –separatistas, nacionalistas–.

Simplificó la situación al plantear que o ganaban los “constitucionalistas” o seguía la “ruina”, también cuando acusó al nacionalismo de ser el culpable de todo lo malo que pasa en Catalunya. La emocionalización se mostró a través de él como esperanza del cambio. La negatividad estuvo presente al presentar una situación crítica para todos, insistió en asociar al gobierno de la Generalitat y a un posible tripartito como gobiernos de “ruina” para Catalunya.

Alejandro Fernández

Fernández, líder del partido desde que en 2018 dimitió García Albiol, presentó un discurso peculiar. Mientras las tres dimensiones del estilo populista estuvieron presentes, casi no lo estuvieron los rasgos de comunicación populista. La reivindicación de la soberanía popular quedó patente cuando puso en valor la participación electoral para poder decidir y que era él quien había tenido “el honor de liderar el mejor proyecto para Cataluña”. Atacó a las élites, –Generalitat y Gobierno Central– y denunció que el candidato del PSC gobernaría con los independentistas, aunque gracias al PPC esto no pasaría, por lo que también defendió al pueblo.

Alejandro Fernández destacó en el uso de las tres dimensiones del estilo populista, aunque la simplificación fue la menos utilizada. Fue el candidato que más habló de las emociones. Una fotografía con Álvarez de Toledo en el balcón del hemiciclo del Parlament hablando con Alejandro Fernández cuyo título rezaba “Mi foto en San Valentín” o un video con Martínez Almeida con el título “Almeida es cojonudo”, son

algunos ejemplos. Utilizó, otra vez, el “todo mal”, con una Catalunya que “nunca había estado peor” y que lideraba “las listas del paro, la delincuencia y la ocupación ilegal”.

5. Discusión y conclusiones

En lo referido a la P1 los candidatos, de forma general, utilizaron la difusión de contenido de otras cuentas –retuit– durante la campaña electoral, en consonancia con lo estudiado previamente (Alonso-Muñoz et al., 2016a; García-Ortega & Zugasti-Azagra, 2018; Jungherr, 2016; Larsson, 2015; López-García, 2016; Segado-Boj et al., 2016; Suau-Gomila et al., 2020), es decir, que prevalece compartir contenido de otras cuentas por encima de generar contenido propio. Este contenido es compartido a partir de cuentas de la esfera partidista, entre miembros de una misma esfera ideológica –incluso de cuentas troles–, lo que refuerza las tesis de la polarización de Twitter (Aragón et al., 2013; Verger, 2015). Otra parte del contenido compartido proviene de los medios de comunicación lo que refuerza la lógica de la hibridación de los medios, además también generaron contenido propio cuya fuente eran los medios de comunicación –tradicionales y digitales, especialmente de entrevistas– buscando amplificar su presencia en Twitter a través de contenidos mediáticos (Jungherr, 2014; Jungherr, 2016; Marcos-García et al., 2021; Miquel-Segarra et al., 2017).

Los candidatos no contestaron a los usuarios que les interpellaron, prefirieron no explotar el potencial dialógico de Twitter para conversar con la ciudadanía, tanto durante como después de las campañas electorales (Marcos-García et al., 2021; Segado-Boj et al., 2016).

En respuesta a la P2 el tema preferido de los candidatos fue la propia campaña, que generaron un metadiscurso en el que los candidatos compartían, fundamentalmente, convocatorias de sus actos de campaña y enlaces para asistir de forma online –lo que fue frecuente debido a la situación sanitaria–. Además, dentro de esta categoría –*political strategy frames*– también se encuentran todos los tuits que llamaban a la movilización, al voto, o que emitían algún tipo de discurso movilizador y cohesionador para con el partido/candidato y con la ciudadanía. Estos temas son los más frecuentes durante las campañas electorales (Alonso-Muñoz et al., 2016b; Enli & Skobergø, 2013; García-Ortega & Zugasti-Azagra, 2014; García-Ortega & Zugasti-Azagra, 2018; López-García, 2016; López-Meri, 2016; Verger, 2015; Zamora & Zurutuza, 2014; Zugasti-Azagra & Sabés, 2015) de esta manera sus perfiles en Twitter se convierten en tablones de anuncios sobre sus actividades de campaña. Por el contrario, no hablaron casi de políticas sectoriales concretas – pese a estar viviendo una situación general de crisis pandémica, económica y política– y no se encontraron diferencias en los temas entre los partidos independentistas y los contrarios a la independencia.

La construcción de la imagen de los candidatos también ha estado presente durante la campaña. En muy diversas formas –desde las más cercanas a la emocionalización de Fernández a las más relacionadas con el conocimiento literario de Borràs–, algunos candidatos hablaron de ellos mismos antes que, por ejemplo, de la gestión de la pandemia (Suau-Gomila et al., 2020; Zamora & Zurutuza, 2014).

En estos temas es más importante la forma que el contenido, los candidatos han enfocado la campaña como una “carrera de caballos” en la que lo importante es ganar y pactar antes que proponer y hablar de gestión (Aalberg et al., 2011; Alonso-Muñoz & Casero-Ripollés, 2018a; López-García, 2016).

En cuanto a la P3 los candidatos presentan, en mayor o menor medida, rasgos y

características estilísticas de la comunicación populista –lo que no quiere decir que lo sean–, indistintamente de su ideología o de la posición que ocupen. No hay diferencias claras entre partidos de izquierdas ni de derechas, ni entre partidos en el gobierno ni en la oposición, ni entre los nuevos ni los tradicionales (Engesser et al., 2017b; Mazzoleni & Bracciale, 2018).

Los rasgos más presentes son la defensa del pueblo y el ataque a las élites, la dualidad “nosotros” contra “ellos” –fundamento del populismo– está presente en sus diferentes formas, desde la “mafia separatista” de Garriga al “club de España” de Borràs (Engesser et al., 2017a; Ernst et al., 2017a; Ernst et al., 2019; Zulinanello et al., 2018). Pese a esto, la exclusión de los otros únicamente está presente en los candidatos de C’s y VOX. Las patrias también son un elemento muy utilizado en sus discursos, aunque cada candidato lo enfoca de una manera diferente –la España mítica de VOX o la Catalunya libre de Junts–.

La simplificación es la dimensión del estilo populista más presente entre los candidatos, lo que encaja con la política como “carrera de caballos” (Aalberg et al., 2011; Alonso-Muñoz & Casero-Ripollés, 2018b; López-García, 2016). El uso de la emocionalización y la negatividad en su contraposición esperanza-miedo también ha sido muy utilizada. Como indican Engesser et al. (2017a) la esperanza es una emoción válida que puede ayudar a la movilización. La estrategia que utilizan es siempre la misma, plantean una situación muy negativa y frente a la que solo existe esperanza de cambio si es el propio candidato quien gana las elecciones.

La principal limitación de este trabajo es que el estudio se circunscribe únicamente al análisis de los *game frame* y la comunicación populista durante la campaña electoral, por lo que habría que realizar análisis posteriores tanto fuera de campaña como en las futuras campañas para poder concluir sobre tendencias y procesos más profundos.

Referencias

- Aalberg, T., de Vreese, C. y Strömbäck, J. (2017). “Strategy and game framing”. En: de Vreese, F., Esser, F. y Hopmann, D. N. (eds): *Comparing Political Journalism*. Londres: Routledge, pp. 33-49. <https://doi.org/10.4324/9781315622286>
- Aalberg, T., Strömäck, J. y de Vreese, C. H. (2011). “The framing of politics as strategy and game: A review of concepts, operationalizations and key findings”, en *Journalism*, vol. 13, nº 2, pp 162-178. <https://doi.org/10.1177/1464884911427799>
- Alonso-Muñoz, L. y Casero-Ripollés, A. (2018a). “Political agenda on Twitter during the 2016 Spanish elections: Issues, strategies, and users’ responses”, en *Communication & Society*, vol. 31, nº 3, pp. 7-25. <http://dx.doi.org/10.15581/003.31.3.7-25>
- Alonso-Muñoz, L. y Casero-Ripollés, A. (2018b). “Communication of European populist leader son Twitter: Agenda setting and the “more is less” effect”, en *El profesional de la información*, vol. 27, nº6, pp. 1193-1202. <https://doi.org/10.3145/epi.2018.nov.03>
- Alonso-Muñoz, L., Marcos-García, S. y Casero-Ripollés, A. (2016b). “Political leaders in (inter)action: Twitter as a strategic communication tool in electoral campaigns”, en *Tripodos*, vol. 39, pp. 71-90. <https://www.raco.cat/index.php/Tripodos/article/view/335035>
- Alonso-Muñoz, L., Miquel-Segarra, S. y Casero-Ripollés, A. (2016a). “Un potencial

comunicativo desaprovechado. Twitter como mecanismo generador de diálogo en campaña electoral”, en *Obra Digital*, vol. 11, pp. 39-59. <http://www.raco.cat/index.php/ObraDigital/article/view/315496>

Aragón, P., Kappler, K. E., Kaltenbrunner, A., Laniado, D., y Volkovich, Y. (2013). “Communication dynamics in twitter during political campaigns: The case of the 2011 Spanish national election”, en *Policy & Internet*, vol. 5, nº 2, pp. 183-206. <https://doi.org/10.1002/1944-2866.poi327>

Baviera, T., Calvo, D. y Llorca-Abad, G. (2019). “Mediatization in Twitter: an explanatory analysis of the 2015 Spanish general election”, en *The Journal of International Communication*, pp. 1-26. <https://doi.org/10.1080/13216597.2019.1634619>

Campos-Domínguez, E. (2017). “Twitter y la comunicación política”, en *El profesional de la información*, vol. 26, nº 5, pp. 785-1793. <https://doi.org/10.3145/epi.2017.sep.01>

Casero-Ripollés, A., Feenstra, R. A., y Tormey, S. (2016). “Old and New Media Logics in an Electoral Campaign: The Case of Podemos and the Two-Way Street Mediatization of Politics”, en *The International Journal of Press/Politics*, pp. 1-20. <https://doi.org/10.1177/1940161216645340>

Centre d'Estudis d'Opinió. (2021). “Encuesta electoral al Parlamento de Cataluña. 2021”. Disponible en Internet: <http://upceo.ceo.gencat.cat/wsceop/7908/Resumen%20en%20espa%C3%B1ol%20-992.pdf> [Consulta: 15 de marzo de 2021]

Centro de Investigaciones Sociológicas. (2021). “Preelectoral de Cataluña. Elecciones autonómicas de 2021”. Disponible en Internet: http://www.cis.es/cis/export/sites/default/-Archivos/Marginales/3300_3319/3306/es3306mar.pdf [Consulta: 15 de marzo de 2021]

Druckman, J. N., Peterson, E. y Slothuus, R. (2013). “How elite partisan polarization affects public opinion formation”, en *American Political Science Review*, vol. 107, nº 1, pp. 57-79. <https://doi.org/10.1017/S0003055412000500>

Dubois, E. y Blank, G. (2018). “The echo chamber is overstated: the moderating effect of political interest and diverse media”, en *Information, Communication & Society*, vol. 21, nº 5, pp. 729-745. <https://doi.org/10.1080/1369118X.2018.1428656>

Engesser, S., Ernst, N., Esser, F. y Büchel, F. (2017b). “Populism and social media: how politicians spread a fragmented ideology”, en *Information, Communication & Society*, vol. 20, nº 8, pp. 1109-1126. <https://doi.org/10.1080/1369118X.2016.1207697>

Engesser, S., Fawzi, N. y Larsson, A. (2017a). “Populist online communication: introduction to the special issue”, en *Information, Communication & Society*, pp. 1-14. <http://dx.doi.org/10.1080/1369118X.2017.1328525>

Enli, G. S. y Skobergø, E. (2013). “Personalized campaigns in party-centred politics. Twitter and Facebook as arenas for political communication”, en *Information, Communication & Society*, pp. 757-774. <https://doi.org/10.1080/1369118X.2013.782330>

Ernst, N., Blassnig, S., Engesser, S., Büchel, F. y Esser, F. (2019). “Populist Prefer Social Media Over Talk Shows: An Analysis of Populist Messages and Stylistic Elements Across Six Countries”, en *Social Media + Society*, January-March, pp. 1-14. <https://doi.org/10.1177/2056305118823358>

Ernst, N., Engesser, S y Esser, F. (2017b). “Bipolar Populism? The use of Anti-Elitism and People-Centrism by Swiss Parties on Social Media”, en *Swiss Political Science Review*, vol. 23, nº3, pp. 253-261. <https://doi.org/10.1111/spsr.12264>

- Ernst, N., Engesser, S., Büchel, F., Blassnig, S. y Esser, F. (2017a). "Extreme parties and populism: an analysis of Facebook and Twitter across six countries", en *Information, Communication & Society*, vol. 20, nº 9, pp. 1347-1364. <https://doi.org/10.1080/1369118X.2017.1329333>
- Fletcher, R., Alessio, C. y Nielsen, K. (2019). "How Polarized Are Online and Offline News Audiences? A Comparative Analysis of Twelve Countries", en *The International Journal of Press/Politics*, pp. 1-27. <https://doi.org/10.1177/1940161219892768>
- Gamir-Ríos, J. (2016). "Blogs, Facebook y Twitter en las Elecciones Generales de 2011. Estudio cuantitativo del uso de la web 2.0 por parte de los cabezas de lista del PP y del PSOE", en *Revista Dígitos*, nº 2, pp. 101-120. <https://doi.org/10.7203/rd.v0i2.53>
- Gamir-Ríos, J., Cano-Orón, L. y Calvo, D. (2017). "La campaña electoral de 2015 en cifras. La presencia en la blogosfera, Facebook y Twitter de los cabezas de lista provinciales de PP, PSOE, Podemos y Ciudadanos". En: López-García, G. y Valera-Ordaz, L. (eds): *Pantallas electorales. El discurso de partidos, medios y ciudadanos en la campaña de 2015*. Barcelona: Editorial UOC, pp. 41-58.
- Gamir-Ríos, J. (2020). "La transmisión digital de las agendas políticas en las Elecciones Locales de 2015 de la ciudad de Valencia", en *Revista de Comunicación Política*, vol. 2, pp. 55-79. <https://doi.org/10.29105/rcp2-3>
- García-Ortega, C., y Zugasti-Azagra, R. (2014). "La campaña virtual en Twitter: análisis de las cuentas de Rajoy y de Rubalcaba en las Elecciones Generales de 2011", en *Historia y Comunicación Social*, vol. 19, pp. 299-311. https://doi.org/10.5209/rev_HICS.2014.v19.45029
- García-Ortega, C., y Zugasti-Azagra, R. (2018). "Gestión de la campaña de las elecciones generales de 2016 en las cuentas de Twitter de los candidatos: entre la autorreferencialidad y la hibridación mediática", en *El Profesional de la Información*, vol. 27, nº 6, pp. 1215-1225. <https://doi.org/10.3145/epi.2018.nov.05>
- Hallin, D. y Mancini, P. (2004). *Comparing Media Systems. Three Models of Media and Politics*. Cambridge: Cambridge University Press.
- Hameleers, M. y Schmuck, D. (2017). "It's us against them: a comparative experiment on the effects of populist messages communicated via social media", en *Information, Communication & Society*, vol. 20, nº 9, pp. 1425-1444. <https://doi.org/10.1080/1369118X.2017.1328523>
- Igartua, J. J. y Humanes, M., L. (2004). *Teoría e investigación en comunicación social*. Madrid: Síntesis.
- Iyengar, S., Sood, G. y Lelkes, Y. (2012). "Affect, not ideology: A social identity perspective on polarization", en *Public Opinion Quarterly*, vol. 76, nº 3, pp. 405-431. <https://doi.org/10.1093/poq/nfs038>
- Jungherr, A. (2016). "Twitter use in election campaigns: A systematic literature review", en *Journal of Information Technology & Politics*, vol. 13, nº 1, pp. 72-91. <https://doi.org/10.1080/19331681.2015.1132401>
- Krippendorff, K. (1990). *Metodología de análisis de contenido*. Barcelona: Paidós.
- Larsson, O. (2015). "Comparing to Prepare: Suggesting Ways to Study Social Media Today-and Tomorrow", en *Social Media + Society*, abril-junio, pp. 1-2. <https://doi.org/10.1177/2056305115578680>


- López-García, G. (2016). "'Nuevos' y 'viejos' liderazgos: la campaña de las elecciones generales españolas de 2015 en Twitter", en *Communication & society*, vol. 29, nº 3, pp. 149-168. <https://doi.org/10.15581/003.29.3.149-168>
- López-García, G., Cano-Orón, L. y Argilés-Martínez, L. (2016). "Circulación de los mensajes y establecimiento de la agenda en Twitter: el caso de las elecciones autonómicas de 2015 en la Comunidad Valenciana", en *Trípodos*, nº 39, pp. 163-183. http://www.tripodos.com/index.php/Facultat_Comunicacio_Blanquerna/article/view/387
- López García, G. y Gamir Ríos, J. (2011). "Los medios españoles en la campaña de 2008". En: López-García, G. (Ed.): *Política binaria y SPAM electoral. Elecciones Generales de 2008: Nuevas herramientas, viejas actitudes*. Valencia: Tirant lo Blanch, pp. 87-114.
- López-García, G., Gamir-Rios, J. y Valera-Ordaz, L. (2018). *Comunicación política*. Madrid: Síntesis.
- López-García, G., Gamir-Ríos, J., García-Ull, F., J., Llorca-Abad, G., Cano-Orón, L. y González-Esteban, J. L. (2015). "El debate sobre Europa en Twitter. Discursos y estrategias de los candidatos de las elecciones al Parlamento Europeo de 2014 en España", en *Revista de Estudios Políticos*, vol. 170, pp. 213-246. <https://doi.org/10.18042/cepc/rep.170.07>
- López-Meri, A. (2016). "Twitter-retórica para captar votos en campaña electoral. El caso de las elecciones de Cataluña de 2015", en *Comunicación y Hombre*, vol. 12, pp. 97-118. <http://www.comunicacionyhombre.com/articulo.php?articulo=177>
- Marcos-García, S., Alonso-Muñoz, L. y López-Meri, A. (2021). "Campañas electorales y Twitter. La difusión de contenidos mediáticos en el entorno digital", en *Cuadernos. info*, nº 48, pp. 27-47. <https://doi.org/10.7764/cdi.48.27679>
- Marcos-García, S., Viunnikoff-Benet, N. y Casero-Ripollés, A. (2020). "What is There in a 'Like'? Political Content in Facebook and Instagram in The 2019 Valencian Regional Election", en *Debats. Revista de cultura, poder y sociedad*, nº 5, pp. 85-109. <https://doi.org/10.28939/iam.debats-en.2020-5>
- Mazzoleni, G. y Bracciale, R. (2018). "Socially mediated populism: the communicative strategies of political leaders on Facebook" en *Palgrave Communications*, vol. 4, nº 50, pp. 1-10. <https://doi.org/10.1057/s41599-018-0104-x>
- McCoy, J., Rahman, T. y Somer, M. (2018). "Polarization and the Global Crisis of Democracy: Common Patterns, Dynamics, and Pernicious Consequences for Democratic Polities", en *American Behavioral Scientist*, vol. 32, nº 1, pp. 16-42. <https://doi.org/10.1177/0002764218759576>
- Miquel-Segarra, S., Alonso-Muñoz, L. y Marcos-García, S. (2017). "Buscando la interacción. Partidos y candidatos en Twitter durante las Elecciones Generales de 2015", en *Prisma Social*, nº 18, pp. 34-54. <https://revistaprismasocial.es/article/view/1353>
- Scherpereel, J. A., Wohlgemuth, J., y Schmelzinger, M. (2017). "The adoption and use of Twitter as a representational tool among members of the European Parliament", en *European politics and society*, vol. 18, nº 2, pp. 111-127. <https://doi.org/10.1080/23745118.2016.1151125>
- Segado-Boj, F., Díaz-Campo, J. y Lloves Sobrado, B. (2016). "Objetivos y estrategias de los políticos españoles en Twitter", en *Index Comunicación*, vol. 6, nº 1, pp. 77-98. <https://journals.sfu.ca/indexcomunicacion/index.php/indexcomunicacion/article/view/244>


- Suau-Gomila, G., Pont-Sorribes, C., y Pedraza-Jiménez, R. (2020). "Politicians or influencers? Twitter profiles of Pablo Iglesias and Albert Rivera in the Spanish general elections of 20-D and 26-J", en *Communication & Society*, vol. 33, nº 2, pp. 209-225. <https://doi.org/10.15581/003.33.2.209-225>
- Tucker, J. A., Guess, A., Barberá, P., Vaccari, C., Siegel, A., Sanovich, S., Stukal, D. y Nyhan, B. (2018). "Social Media, Political Polarization, and Political Disinformation: A Review of the Scientific Literature", en *SSRN Electronic Journal*, March 2018, pp. 1-95. <https://doi.org/10.2139/ssrn.3144139>
- Urman, A. (2019). "Context matters: political polarization on Twitter from a comparative perspective", en *Media, Culture & Society*, pp. 1-23. <https://doi.org/10.1177/0163443719876541>
- Vergeer, M. (2015). "Twitter and Political Campaigning", en *Sociology Compass*, vol. 9, nº 9, pp. 745-760. <https://doi.org/10.1111/soc4.12294>
- Waisbord, S. (2020). "¿Es válido atribuir la polarización política a la comunicación digital? Sobre burbujas, plataformas y polarización afectiva", en *Revista SAAP*, vol. 14, nº 2, pp. 249-279. <https://doi.org/10.46468/rsaap.14.2.A1>
- Zamora, R. y Zurutuza, C. (2014). "Campaigning on Twitter: Towards the 'Personal Style' Campaign to Activate the Political Engagement During the 2011 Spanish General Elections", en *Communication & Society*, vol. 27, nº 1, pp. 83-106. <https://doi.org/10.15581/003.27.1.83-106>
- Zugasti, R. y Pérez, J. (2016). "Los temas de campaña en Twitter de @PPopular y @ahorapodemos para las elecciones europeas de 2014", en *adComunica Revista Científica de Estrategias, Tendencia e Innovación en Comunicación*, nº 12, pp. 205-2024. <https://raco.cat/index.php/adComunica/article/view/312866>
- Zugasti-Azagra, R. y Sabés, F. (2015). "Los issues de los candidatos en Twitter durante la campaña de las elecciones generales de 2011", en *Zer*, vol. 20, nº 38, pp. 161-178. <https://ojs.ehu.eus/index.php/Zer/article/view/14792>
- Zulianello, M., Albertini, A. y Ceccobelli, D. (2018). "A Populist Zeitgeist? The Communication Strategies of Western and Latin American Political Leader on Facebook", en *The International Journal of Press/Politics*, vol. 23, nº 4, pp. 439-457. <https://doi.org/10.1177/1940161218783836>



Feminist discourses of Spanish digital influencers: An analysis of their activity on social media

Discursos feministas de las *influencers* digitales españolas: Análisis de su actividad en redes sociales

 **Celina Navarro**
celina.navarro@uab.cat
Universitat Autònoma de Barcelona

 **Isabel Villegas-Simón**
isabelmaria.villegas@upf.edu
Universitat Pompeu Fabra

ABSTRACT With the rise of digital feminism, this article explores the role of Spanish digital influencers considered as feminist referents and the topics addressed by them on social networks from a gender perspective. An online survey was used to create a ranking of the twelve most popular influencers by social media users. Firstly, an analysis of their professional and sociodemographic profiles was made. Secondly, a cross-platform content analysis of the posts published by these influencers on Facebook, Twitter and Instagram was undertaken. The results show the homogenous profile of the most popular feminist influencers. The main topics addressed in their social networks are linked to gender violence, especially on Twitter, which is the social network where they publish the largest number of posts with a gender perspective despite having more followers on Instagram. Finally, the study concludes that online feminist discourses replicate the issues in the traditional media agenda and omit other topics that involve women belonging to minorities or related to the private sphere.

KEYWORDS feminism; social media; influencers; digital activism

RESUMEN Este artículo explora el rol de las *influencers* digitales españolas consideradas referentes feministas y los temas que abordan con perspectiva de género en sus redes sociales. A partir de la realización de una encuesta, se ha establecido un *ranking* con las doce *influencers* digitales más populares consideradas feministas. En primer lugar, se analizaron sus perfiles sociodemográficos junto con su actividad profesional. En segundo lugar, se llevó a cabo un análisis de contenido de sus publicaciones en Facebook, Twitter e Instagram. Los resultados muestran una homogeneidad de los perfiles de las *influencers* feministas más populares. Los principales temas abordados en sus redes sociales están vinculados a la violencia de género, siendo Twitter la red social donde publican mayor cantidad de *posts* con perspectiva de género, a pesar de que en Instagram tienen más seguidores. Finalmente, el estudio concluye que los discursos feministas *online* replican los temas y asuntos de la agenda de los medios tradicionales y dejan de lado otros temas de la agenda feminista que implican a mujeres que pertenecen a minorías y temas relacionados con la esfera privada.

PALABRAS CLAVE feminismo; redes sociales; *influencers*; activismo digital

Edita: Irene Liberia Vayá

Recibido: 30/06/2021 | Aceptado: 12/12/2021

Cómo citar este artículo:

Navarro, C. y Villegas-Simón, I. (2022). Feminist discourses of Spanish digital influencers: An analysis of their activity on social media. *Dígitos. Revista de Comunicación Digital*, 8: 201-214. DOI: 10.7203/rd.v1i8.223



Feminist discourses of Spanish digital influencers: An analysis of their activity on social media

1. Introduction

Over the last decade, Spain has experienced a period of mobilization by different social segments such as pensioners, the health sector, educational workers or, at the core of this paper, the feminist movement. The global economic crisis of 2008 greatly damaged the economy and successive governments have responded with public cuts that continue to be questioned and protested against by a significant portion of Spanish society.

The increased presence of feminism in the international and national public spheres is undeniable. In recent years, there have been significant mobilizations in Spain particularly illustrated by the first feminist labour strike on March the 8th 2018, which coincided with International Women's Day. There have also been large demonstrations against judicial rulings considered to be part of a patriarchal system such as the trial known as "La Manada" [The Pack in English] (Navarro and Coromina, 2020).

Social networks are currently one of the spaces where popular culture is more visible and have facilitated the spread of feminism within the public sphere. The rise of hashtag feminism internationally is one example of this (Myles, 2019). From this perspective, the hashtags: #YesAllWomen in the US and Germany (Baer, 2015), #EverydaySexism in the UK (Bowles Eagle, 2015) or #Niunamenos [#Notoneless in English] in Latin America (Giraldo-Luque et al., 2018) are clear examples of the interest in analysing how these movements are presented and the impact that they have. The importance of this new way of commenting on, debating and organizing the movement has even been considered as the fourth wave of the feminist movement (Baumgardner, 2011; Zimmerman, 2017).

One of the main aspects praised about hashtag feminism is the possibility of creating a community from the individual experiences and opinions of Twitter users (Thrift, 2014; Myles, 2019). However, the role of influencers is also key for the diffusion of feminism online. In general, we can define influencers as individuals who have the potential "to lead others to engage in a certain act" (Cha et al., 2010: 2). Therefore, the activity of the feminist influencers is important when understanding the evolution of feminism in popular culture.

The objective of this paper is to determine the Spanish feminist influencers on social networks, to describe their profile and to map the issues related to feminism or from the gender perspective that they discuss according to the social network on which they publish: Twitter, Facebook or Instagram. In addition, the role that these feminist influencers have in creating new topics beyond the agenda setting of traditional media is investigated and how these topics correlate to the interests of social media users. Regarding the methodology, firstly, a ranking of the most popular Spanish influencers

was created through a survey of 743 participants that has allowed, in the second part of our analysis, the creation of a cross-platform content analysis of social networks to establish the current feminist discourses in the Spanish digital sphere.

In this paper, first of all in the theoretical framework, we discuss the concept and evolution of digital feminism and its consolidation on social media. Afterwards, the different aspects of the implementation of the survey and the content analysis of social media posts are explained in the methodology section. The results are divided into three sections focusing on the results of the survey, the profile of the Spanish feminist influencers and their presence on social media. To conclude, a discussion of the results contextualizing them within different theoretical paradigms is conducted to close the main conclusions.

2. Theoretical framework

2.1. Digital Feminism

The popularization of feminism has had a major impact on the digital sphere (Caro-Castaño, 2015; Hogsnes and Steenbjerg Hansen, 2018). In fact, the 4th wave of feminism has been defined as the spread of the 3rd wave into cyberspace (Zimmerman, 2017). Authors who support this idea argue that young women can find a space on the internet where they can express their experiences and views with more freedom (Munroe, 2013). Thus, the 4th wave is characterized as being sex-positive, trans-accepting and body-positive (Baumgardner, 2011). Furthermore, this wave has a strong influence on companies and their commercial strategies (Hogsnes and Steenbjerg Hansen, 2018).

The pervasiveness of feminism into cyberspace has been observed from different perspectives. Cyberfeminism claims that the Internet offers the possibility of creating new identities and relationships out of gender dichotomy. Also, cyberfeminism believes that new technologies can help to reduce gender inequality through online activism and by spreading global connections around the world (Paasonen, 2011). Consequently, cyberfeminism authors affirm that social media platforms movements such as #MeToo or #Niunamenos or first feminist strike on March the 8th illustrate how social networks can give visibility and support women's struggle across the whole world and in a transversal way. Thus, from a utopian and optimistic point of view, the Information and Communications Technologies (ITC) encourage broader feminist strategies and promotes activist discourse from the public digital sphere (Núñez-Puente et al., 2016).

On the other hand, more negative perspectives state that the patriarchal structures of the offline society are duplicated in cyberspace. For example, Judy Wajcman (2000), one of the main authors who criticizes cyberfeminism, argues that technology does not guarantee a genuine change in eradicating social injustices but it can have a variety of consequences. For this reason, Wajcman (2000) queries the technological determinism of cyberfeminism. Previous studies conclude that the digital sphere is not conquered by the feminist groups (Sánchez-Duarte, 2015). Also, they said that cyberspace may not be useful in dismantling the patriarchy and sexism despite it being a good space to spread feminist ideas (Giraldo-Luque et al., 2018). In conclusion, the possibilities that the digital sphere offers to women are in dispute (Zafra, 2011).

2.2. Feminism on social networks

For Catherine Rottenberg (2014) and Rosalind Gill (2007), the popularization of feminism and influencers on social networks constitutes an individualistic feminism of neoliberal consumer culture, which is characterized by the visibility of messages concerning self-making, self-love and self-care. These kinds of messages are also found in the advertising campaigns of major companies and are aimed at white, privileged women. Furthermore, these messages are related to the neoliberal culture of self-employment and DIY culture among young women who self-identify as feminists (Banet-Weiser and Portwood-Stacer, 2017).

From this perspective, two terms have been coined. The first is 'Commodity Feminism' which defines the relationship between feminism and consumerism in a symbolic and commercial approach to the digital sphere (Banet-Weiser and Portwood-Stacer, 2017). The second is 'Accidental Feminist' by Rose Maloney (2017). In her research, Maloney explores the Kardashian family and she concludes that they have become role models for women because of their unapologetic sex and body positivity and their support for the transsexual movement in spite of never having dedicated their life to gender equality explicitly.

Some authors consider that these kinds of feminism have a negative effect on the real fight for equality. In particular, Banet-Weiser and Portwood-Stacer (2017) state that the rise of individualistic feminist neoliberal consumer culture is only for privileged women to gain visibility and reward in the market, at the expense of converting feminism into a commodity. This reinforces neoliberal individualism and white supremacy, while the millions of women who belong to a different race, sexual identity and or class continue to be oppressed.

Alternatively, Caro-Castaño (2015) raises some challenges for feminism on social networks. Initially, the author claims that feminism on these digital platforms is light and far from radical, and therefore considers that a more inclusive and radical method is necessary. Secondly, the author points out the obligation in overcoming the individualistic neoliberal commercial heteronormativity. In addition, social media are also public spaces where misogynistic messages and attacks are common, especially for users defending feminist discourses (Jane, 2016). This makes safe participation on these platforms difficult for users who openly define themselves as feminists and hinders productive discussion within the movement.

3. Method

Taking under consideration the theoretical and empirical preceding works, the main objective of this paper is the identification and analysis of the digital influencers related to feminism in Spain. In this research paper, a social media influencer is classed as a feminist if other users consider them to be, whether self-identified as a feminist or not. With this premise, the specific objectives of this paper are:

1. To describe the sociodemographic and professional profile of the Spanish feminist digital influencers.
2. To discover if the topics published by the influencers correlate to the interests of Spanish social media users and the agenda of traditional media.

3. To analyse whether the social networks (Facebook, Twitter or Instagram) determines the topics published by the Spanish feminist influencers.

To address these objectives a preliminary online survey has firstly been used to detect the main Spanish feminist digital influencers and the topics Spaniards are interested in regarding gender inequality. Second, a cross-platform content analysis of the posts by the feminist referents on Facebook, Twitter and Instagram has been used to detect their activity on each platform in relation to gender perspective and whether there is a relationship between the results by the respondents in the survey.

3.1. Survey

A total of 743 valid responses from citizens across the country were received using the online survey distributed by Google forms, 448 women and 295 men (users who were not Spanish were excluded). All valid responses were analysed with the open source programming software *R Studio*. Respondents were predominantly aged between 18 and 35 years old and needing to be recurrent users of Facebook, Twitter or Instagram, with the latter being the most used by 41.28% of the respondents. The answers of by participants not fulfilling these requirements were not considered. The survey was spread across social media platforms and was answered through an online form during the month of October, 2018.

The participants were firstly asked about their sociodemographic characteristics (gender, age, residence and level of studies). In the second section, participants were asked if they considered themselves as feminists. In the cases which answered positively, two main aspects were further investigated: the themes they were interested in relating to feminism or gender inequality and mentioning a maximum of three Spanish feminist influencers. The topics regarding gender inequality included in the survey were grouped from a selected range of precedent feminist literature, including theoretical works and research relating to feminism and popular culture (Baer, 2015; Banet-Weiser and Portwood-Stacer, 2017; Dixon, 2014; Jane, 2016; McRobbie, 2009; Munroe, 2013). They are the following: a) roles and gender inequalities (from an educational and divulgation approach), b) news related to gender inequalities, c) history and knowledge about feminism, d) cultural recommendations, e) questioning of beauty standards, f) information on the inequalities of lesbian, bisexual and transgender women, information on the inequalities of racialized women, g) female sexuality and h) new dynamics in romantic relationships.

3.2. Cross-platform social media content analysis

After analysing the survey, the results were used to create a ranking of the top 12 digital feminist influencers. Once these were selected, their sociodemographic profile and their professional activities were analysed. For this analysis, information was sought through their description on their social networks and their appearance in the media. Furthermore, a cross-platform analysis of their posts was carried out using a quantitative content analysis. The sample included all of their entries posted on Facebook, Twitter and Instagram during November 2018 with a total of 3670 units of analysis compiled. Posts were firstly divided between those including a gender point of view and those without. The first were further classified according to their topic, using the same categories as for those presented in the survey, and the subtopic. The subtopics were filled using an inductive approach based on the main feminist events in Spain during

the period analysed.

The intercoder reliability test, conducted by two coders from a subsample of 367 posts (10% of the sample), was measured for each variable by Krippendorff's (2004) alpha. All variables reached the Krippendorff's standards of unconditionally reliability. The coefficients for gender perspective posts and topic were 0.89 and 0.93. In the survey results, to verify statistical significance relationships in the contingency tables, chi-square was used, considering 0.05 as the mean value of significance.

4. Results

4.1. Feminist influencers by social media users

The analysis of the responses to the survey shows that a significant number of the survey respondents consider themselves to be feminists (84.93%, n=631), however, this percentage is statistically higher in the case of women (88.20%, n=395) than in the responses from men (80%, n=236, p-value = 0.003277). Therefore, from the results of our survey, Spanish women are more likely to view themselves as feminists than men do.

From a gender perspective, the respondents reported that the main topics they are interested in on social media are the role and inequality of genders, news related to gender inequality and the history and knowledge of feminism. In addition, it is interesting to point out that men and women agree on their interest in these issues. However, men show greater interest in the inequality of lesbian, bisexual and transgender women and in the information related to the new dynamics of romantic relationships than women, while women show greater interest in female sexuality and the questioning of beauty standards (Table 1).

Table 1. Topics interested in on social media related to with a gender perspective

Topics	Women	Men	Total
Roles and gender inequalities (wage gap, sexism in the media, etc)	24.92%	25.75%	25.21%
News related to gender inequalities	17.44%	20.51%	18.50%
History and knowledge about feminism	11.70%	10.63%	11.33%
Cultural recommendations with a gender perspective (TV series, movies, books, conferences, etc)	11.31%	7.78%	10.08%
Questioning of beauty standards	9.79%	8.23%	9.25%
Information on inequalities of lesbian, bisexual and transgender women	7.01%	8.98%	7.69%
Information on inequalities of racialized women	6.21%	6.89%	6.44%
Female sexuality	6.93%	4.79%	6.19%
New dynamics in romantic relationships	4.70%	6.44%	5.31%

Based on the answers by the respondents, a ranking of the 12 most popular Spanish feminist influencers on social media has been created (Table 2). In general, there is

a firm consensus on the first two influencers on the list by a significant number of participants. The following 10 influencers were mentioned a similar amount of times. However, major differences are observed between men and women. Firstly, among men there is less consensus about their feminist influencers since no names occurred predominantly. The influencer that most appeared in their responses, Amaia from a music reality show, was only referenced by 4,75% of the male participants in the survey. In contrast, women agreed more on their feminist influencers.

Table 2. Ranking of Spanish feminist referents on social media

	Women	Men	Total
Leticia Dolera	19.42%	2.37%	12.65%
Barbijaputa	11.16%	3.73%	8.21%
La vecina rubia	3.79%	3.73%	3.77%
Amaia Romero	2.46%	4.75%	3.36%
Moderna de Pueblo	4.46%	1.02%	3.10%
Srta Bebi	3.79%	1.69%	2.96%
Ana Isabel Bernal Triviño	3.13%	0.00%	1.88%
Devermut	2.23%	1.36%	1.88%
Irantzu Varela	2.90%	0.34%	1.88%
Cristina Fallarás	2.23%	1.36%	1.88%
Lola Vendetta	2.46%	0.68%	1.75%
Amarna Miller	0.67%	2.71%	1.48%

4.2. Profile of Spanish Feminist influencers

The analysis of the 12 Spanish digital influencers reveals that the users considered to be feminist influencers have similar sociodemographic and economic backgrounds. Firstly, all are white, belong to a similar social class and the vast majority are heterosexual. Also, all of them are cis women and only two mention having female partners on their posts (Irantzu Varela y Devermut). In addition, another similarity is that all of their work is related to the media and cultural industry, and is recognized in their professional field. Specifically, seven of these influencers started to gain popularity due to their activity on social networks (Barbijaputa, La vecina rubia, Moderna de Pueblo, Srta. Bebi, Devermut, Irantzu Valera, and Lola Vendetta) although all of them have developed part of their profession outside the digital platforms. The remaining five were already popular for their professional work before joining social media (Leticia Dolera, Amaia Romero, Ana Isabel Bernal Triviño, Cristina Fallarás and Amarna Miller).

Regarding their professions, two are journalists working on digital newspapers (Cristina Fallarás and Ana Isabel Bernal Triviño), two actresses and writer-actress, (Leticia Dolera and Amarna Miller), two illustrators (Moderna de Pueblo and Lola Vendetta), one a singer that participated in a hit television talent show (Amaia Romero), two are activists that have a relationship with offline and online media (Barbijaputa and Irantzu Valera) and three have their main professional activity on social media platforms (La vecina rubia, Srta. Bebi and Devermut).

Within this analysis of their profiles, it is important to highlight that only half reveal their real identity on social networks, some use a pseudonym on their profiles and three hide all data on their real identity. All of the profiles correspond to only one person except Devermut, which has two people using this profile. In addition, only three define themselves as feminists on their social network profile.

4.3. Spanish feminist influencers on social media

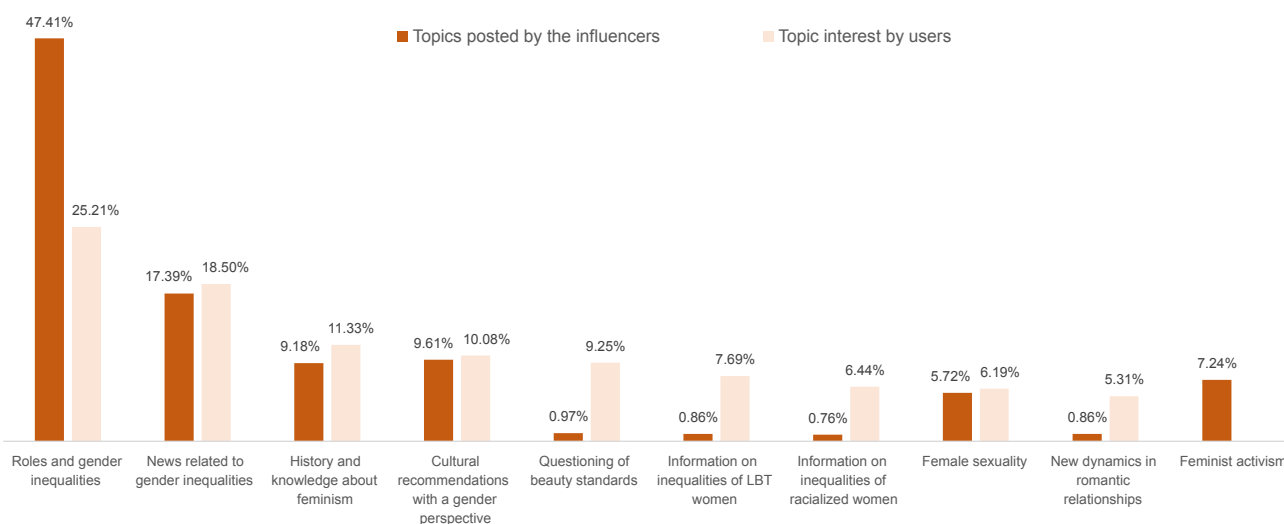
Strikingly, all 12 of the influencers are present on Instagram, Twitter and Facebook. This order is the same for the number of followers on each social network with Instagram having the most, followed by Twitter and finally Facebook. The social network most used by the influencers analysed is Twitter, since it is the network where they post the most publications, followed by Facebook, and finally Instagram, with the fewest publications (Table 3). If we look at the publications with a gender perspective, Facebook is proportionally where the most publications are found, since a little more than 30% of publications are linked to feminism, followed by Instagram and finally by Twitter. However, in absolute terms, the largest number of publications with a gender perspective can be found on Twitter.

Table 3. Posts with gender perspective topics by social network

	Gender perspective topics		Others		Total	
Facebook	102	30.27%	235	69.73%	337	9.18%
Instagram	70	28.23%	178	71.77%	248	6.76%
Twitter	754	24.44%	2331	75.56%	3085	84.06%
Total	926	25.23%	2744	74.77%	3670	100.00%

Posts with a gender perspective are linked to two main topics: first, roles and gender inequality and, second, news with a gender perspective, as can be seen in Figure 1. The topics addressed less often are: information on the inequalities of racialized, lesbians, transsexuals and bisexuals and the new dynamic in romantic relationships. Also, posts related to feminist activism account for 7.24%. This category was added during the codification process to group all posts that were calls for offline mobilisation.

Figure 1. Topics of gender perspective posts



The more common subtopics are gender-based violence, patriarchy justice and prostitution. Each one of these subtopics represents over 10% of the total of gender perspective posts, especially the first, since 22.24% of the posts are related to gender-based violence. The subtopics with a greater frequency are also related to violence toward women: feminicides, sexual harassment and allegations of sexism, all slightly above 3%. The subtopics pornography, surrogate mothers, sexism in the media and women and politics are also present, at slightly more than 2%. Otherwise, the subtopics less represented are female poverty, commodification of the female body and humour limits, which have less than 0.3%.

Finally, during our analysis we detected posts with a gender perspective linked to media events that took place during the sample period. They were mainly related with the 25th of November, the day against gender violence, cases of patriarchal justice and feminicides. Other media controversies were found. Such as, the case of Leticia Dolera, director of a television series who was publicly questioned when she fired a pregnant actress. As well as surrogacy and the legislation against unionised prostitution in Spain.

5. Discussion

In recent years, issues relating to feminism have grown strongly in the media and social debate. In this context, the present research seeks to discover who the most relevant feminist influencers on social networks are and what their discourses are since they have tremendous popularity. Results have to be explored considering the sociocultural, economic and political context of 2018 when the survey was implemented since both feminism and social media evolve rapidly.

The 12 feminist influencers identified have several common characteristics since they are all white, mostly heterosexual (except for two) cis women and are recognized within their professional field. Also all of them are professionally linked to the culture and media sector. This result is related to the postulates of Banet-Weiser and Portwood-Stacer (2017) and Rottenberg (2014) who assert that the most popular feminism found on social networks is led by privileged white women who manage to gain greater visibility thanks to their activity on those social networks. However, according to the results, they can be further classified into three groups based on their professional activity and the type of publications with a gender perspective posted.

In the first group are those who have an explicit feminist activity on their social networks, inasmuch as they dedicate a major part of their posts to aspects related to a wide gender perspective and openly claim to be feminists. Barbijaputa, Iranztu Valera, Cristina Fallarás, Leticia Dolera, Devermut and Ana I. Bernal Triviño are found in our sample. Furthermore, all of them carry out a professional activity that is not linked to social networks but they do use their accounts to further their fight against sexism.

The second group is formed by those whose professional activity is based on their activity on social networks and do not have an activity directly related to the defence of feminist issues. On the contrary, all of them publish posts related to women, but they do not explicitly define themselves as feminists. This is the case of La vecina rubia, Moderna de Pueblo and Srta. Bebi. These influencers match the concept of "accidental feminism" by Maloney (2017). Additionally, these influencers may be connected with the ideas suggested by Rottenberg (2014) and Banet-Weiser and Portwood-Stacer (2017) about individualistic feminism and commodity feminism, since they are users who tie

their commercial and professional activity to feminist discourses on social networks or, at least, content that the participants of the survey perceive as feminist.

Finally, there is one profile, Amaia Romero, that is considered to be feminist by social media users in the survey performed because of her activity in other spaces but is barely active on the social networks analysed. This point shows the importance of other media beyond social networks and its ability to influence a topic, in this case, feminism. As Fotopoulou (2016) and Van Dijck (2011) point out, social networks play a major part in creating cultural imaginaries, but they do not act alone. Future studies must explore the relationship between all kinds of media to construct feminist discourses on social media.

Another interesting finding is the use of pseudonyms by half of the digital influencers. This leads to two thoughts: the first is how a user without a real identity can become an influencer of authority on social networks, and the second is related to how hiding one's identity may be a protection mechanism against hate on digital networks, specifically towards women who defend feminist ideas (Jane, 2016).

The results of the content analysis on posts with a gender perspective show a strong consensus on the condemnation of gender violence: feminicides, sexual harassment or patriarchal justice. As Menéndez Menéndez (2014) has pointed out, sexist violence is one of the topics that is of most concern to society and greatly influences the setting of the agenda. The relationship between digital gender perspective discourses and the agenda setting of traditional media is also observed because the second topic is news related to gender inequalities. Therefore, this result may support the positions of López-García, Cano-Orón and Argilés-Martínez (2016) and Van Dijck (2011) who affirm that media setting transfers into the digital sphere. Alternatively, the potential for social networks to promote offline political mobilisation is reflected in this research since posts related to offline feminist activism are often found. Also, this result may be explained by the rise of the feminist struggle in Spain. However, posts analysed for a gender perspective do not discuss racial or the non-mainstream sexual identities of women.

Therefore, the media agenda has a great influence on the online feminist discourse identified. On one hand, the analysed influencers use news and media events in their publications to speak about feminism and gender perspective issues. Consequently, influencers repeat the topics from traditional media on their social media profiles. On the other hand, the topics mostly addressed by influencers match the main gender perspective topics in traditional media agenda: issues related to sexist violence (Casero-Ripollés, 2010). This fact also has been pointed out by Fernández-Rovira & Villegas-Simón (2019), who highlight the relationship between Twitter and the media agenda, since most tweets with gender perspective published by Spanish politicians are related to news events and feminicides.

In addition, the topics in which the survey respondents show the greatest interest coincide with the topics most discussed by influencers on social networks. The sample highlights a greater interest in topics with regard to gender roles in the media and public sphere, especially those which concern income differences between men and women, sexism in the media, cases of gender violence or media controversy around feminism, while they show less interest in aspects related to minority groups within feminism (such as racism, transsexuals, lesbianism, etc.) and the personal sphere (such

as romantic relationships or beauty standards). This result suggests two main ideas. The first is that the interests of the sample as well as the topics most covered in social networks by the influencers analysed coincide with a light (Caro-Castaño, 2015) and liberal feminism (Banet-Weiser and Portwood-Stacer, 2017; Rottenberg, 2014), based on the defence of women in public and the most visible issues of feminism (gender violence and visibility of women). The second is that it is a biased feminism focused on white and privileged women (McRobbie, 2009). It is important to note that the sample used is homogeneous in terms of the educational level, economic level and sexual orientation of the participants. Therefore, future studies should observe more heterogeneous samples and identify differences in the choice of topics of interest and digital profiles based on more varied sociodemographic categories, such as sexual orientation, ethnicity or educational level.

This research also reveals how the nature and structure of each social network influences the type of posts and activity on each platform. Despite the fact that Twitter is the social network with the largest number of posts, those related to feminism and gender perspective represent the smallest percentage. This digital platform has been explored as an instrument for social movements, but authors have pointed out its limited capacity to cause real social change (Anduiza et al., 2014; Castells, 2012; Giraldo-Luque, 2015; Fuchs, 2014). Instagram is the social network in which the selected influencers have the greatest numbers of followers and it is the most used platform by the participants of the survey. This has been central to the consolidation of digital influencers (Hogsnes and Steenbjerg Hansen, 2018). Notwithstanding, Instagram has the fewest number of posts, which can be explained by the visual nature that characterises this particular social network (Thelwall et al., 2015). However the percentage of posts about gender perspective is slightly greater than those found on Twitter.

These results may link to more critical discourses related to the relationship between the feminist struggle and social networks. On one hand, the low number of messages on gender perspective on the most popular social network is not favourable for the diffusion of feminist ideas. In addition, the prominence of images can lead to a simplification of the feminist discourse in the digital sphere. Both effects have been stated by authors such as Banet-Weiser and Portwood-Stacer (2017) or McRobbie (2009), who consider that feminism can dissolve on social media and be harmful to gender equality.

6. Conclusion

This paper contributes to the understanding of digital feminism and the role of influencers in building and influencing the popular feminist discourses. With our results it can be concluded that while feminism is present at high levels in the Spanish digital public sphere and a large number of people consider themselves to be feminists, the influencers considered as feminist referents have a very homogenous sociodemographic profile. Nevertheless, with our analysis, we can conclude that there are three types of influencers on social media: the ones that focus their professional work on the movement and are mainly from the journalism and media sectors, those that due to their work have become accidental feminists whose presence is mainly online, and the third group that are known for their presence on other platforms but rarely use social media.

Furthermore, this analysis also concludes that while Instagram is the social network mostly used and where the feminist influencers have most followers, it is rarely used to debate these issues. Therefore, the most viewed posts created by feminist influencers do not have a gender perspective. The popularity of Instagram, an image-oriented platform, may simplify the feminist ideas on social media.

In conclusion, there is a homogenization in the topics related to gender inequalities which are addressed by Spanish feminist digital influencers. It may be stated that they replicate the hegemonic discourse of the agenda setting of traditional media focused on gender violence and discourse while the realities of women's minority groups and topics related to the private sphere are left out. However, it needs to be highlighted that some of the aforementioned influencers have lost popularity since 2018 (for example, Leticia Dolera or Devermut) and others have increased (such as Bernal Triviño). In addition, feminist topics on both social media and in the mass media have evolved rapidly. This highlights the need to further study the topic at different periods and in socioeconomic contexts to understand the role and evolution of feminist influencers.

References

- Anduiza, E.; Cristancho, C. and Sabucedo, J. (2014): "Mobilization through online social networks: the political protest of the indignados in Spain", in *Information, Communication & Society*, vol. 17, nº 6, pp. 750-764.
- Baer, H. (2015): "Redoing feminism: digital activism, body politics, and neoliberalism", in *Feminist media studies*, vol. 16, nº 1, pp 17-34.
- Banet-Weiser, S. and Portwood-Stacer, L. (2017): "The traffic in feminism: an introduction to the commentary and criticism on popular feminism", in *Feminist Media Studies*, vol. 17, nº 5, pp. 884-888.
- Baumgardner, J. (2011): *F´em!: Goo Goo, Gaga, and Some Thoughts on Balls*. Berkeley: Seal Press.
- Bowles Eagle, R. (2015): "Loitering, lingering, hashtagging: Women reclaiming public space via #BoardtheBus, #StopStreetHarassment, and the #Everydaysexism Project", in *Feminist Media Studies*, vol, 15, nº 2, pp. 350-353.
- Caro-Castaño, L. (2015): "Construir y comunicar un 'nosotras' feminista desde los medios sociales. Una reflexión acerca del 'feminismo del hashtag'", in *Commons: Revista de comunicación y ciudadanía digital*, vol. 2, nº 4, pp. 124-154.
- Casero-Ripollés, A. (2010): "Prensa en internet: nuevos modelos de negocio en el escenario de la convergencia", in *El profesional de la información*, vol. 19, nº 6, pp. 595-601.
- Castells, M. (2012): *Redes de indignación y esperanza: los movimientos sociales en la era de Internet*. Madrid: Alianza.
- Cha, M.; Haddadi, H.; Benevenuto, F. and Gummadi, P.K. (2010): "Measuring user influence in Twitter: the million follower fallacy", in *4th AAAI International Conference on Weblogs and Social media (ICWSM)*.
- Dixon, K. (2014): "Feminist online identity: Analyzing the presence of hashtag feminism", in *Journal of Arts and Humanities*, vol. 3, nº 7, pp. 34-40.

- Fernández-Rovira, C. and Villegas-Simón, I. (2019): "Estudi comparat de les posicions a Twitter dels polítics espanyols sobre el feminisme", in *Anàlisi: quaderns de comunicació i cultura*, nº 61, pp. 77-92.
- Fotopoulou, A. (2016): "Digital and networked by default? Women's organisations and the social imaginary of networked feminism", in *New Media & Society*, vol. 18, nº 6, pp. 989-1005.
- Fuchs, C. (2014): *Social media. A critical introduction*. Thousand Oaks: Sage.
- Gill, R. (2007). Postfeminist media culture: Elements of a sensibility, in *European Journal of Cultural Studies*, vol. 10, nº 2, pp. 147-166.
- Giraldo-Luque, S. (2015): *Més enllà de Twitter. De l'expressió indignada a l'acció política*. Vic: Eumo Editorial.
- Girado-Luque, S.; Fernández-García, N. and Pérez-Arce, J. (2018): "La centralidad temática de la movilización #NiUnaMenos en Twitter", in *El Profesional de la Información*, vol. 27, nº 1, pp. 96-105.
- Hogsnes, M. and Steenbjerg Hansen, K. (2018): "Negotiating Consumerism and Feminism on Instagram". Paper presented at the *11th IEEE International Conference on Social Computing and Networking. SocialCom*. Melbourne, 11-13 December 2018.
- Jane, E. A. (2016): "Online misogyny and feminist digitalism", in *Continuum*, num. 30, vº 3, pp. 284-297.
- Krippendorff, K. (2004): "Measuring the reliability of qualitative text analysis data", in *Quality and Quantity*, vol. 38, pp. 787-800.
- López-García, G.; Cano-Orón, L. and Argilés-Martínez, L. (2016): "Circulación de los mensajes y establecimiento de la agenda Twitter: el caso de las elecciones autonómicas de 2015 en la Comunidad Valenciana" in *Tripodos*, vol. 39, pp. 163-183.
- Maloney, A. R. (2017): "Influence of the Kardashian-Jenners on Fourth Wave Feminism", in *Elon Journal of Undergraduate Research in Communications*, vol. 8, nº 2, pp. 48-59.
- McRobbie, A. (2009): *The Aftermath of Feminism: Gender, Culture and Social Change*. Thousand Oaks: SAGE.
- Menéndez Menéndez, M.I. (2014): "Retos periodísticos ante la violencia de género. El caso de la prensa local en España", in *Comunicación y Sociedad*, vol. 22, pp. 53-77.
- Myles, D. (2019): "'Anne goes rogue for abortion rights!': Hashtag feminism and the polyphonic nature of activist discourse", in *New Media & Society*, vol. 21, nº 2, pp. 507-527.
- Munroe, E. (2013): "Feminism: A Fourth Wave?", in *Political Insight*, vol. 4, pp. 22-25.
- Navarro, C. and Coromina, Ò (2020): "Discussion and mediation of social outrage on Twitter: The reaction to the judicial sentence of "La Manada"", in *Communication & Society*, vol. 33, nº 1, pp. 93-106.
- Núñez-Puente, S.; Vázquez-Cupeiro, S. and Fernández-Romero, D. (2016): "Ciberfeminismo contra la violencia de género: análisis del activismo online-offline y de la representación discursiva de la víctima", in *Estudios sobre el mensaje periodístico*, vol. 22, nº 2, pp. 861-877.

- Paasonen, S. (2011): "Revisiting cyberfeminism", in *Communications*, vol. 36, pp. 335-352.
- Rottenberg, C. (2014): "The rise of neoliberal feminism", in *Cultural Studies*, vol. 28, nº 3, pp. 418-437.
- Sánchez-Duarte, J.M. (2015): "Participación digital en los partidos políticos. Autonomía y prácticas de militancia en red", in *Dígitos*, vol. 1, pp. 59-69.
- Thelwall, M; Vis, F.; Faulkner, S.; Burns, A.; Aulich, J; Mas-Bleda, A.; Stuart, E. and D'orazio, F. (2015): "Chatting Through Pictures? A Classification of Images in One Week in the UK and USA", in *Journal of the Association for Information Science and Technology*, vol. 67, nº 11, pp. 2575-2586.
- Thrift, S.C. (2014): "#YesAllWomen as feminist meme event", in *Feminist Media Studies*, vol. 14, nº 6, pp. 1090-1092.
- Van Dijck, J. (2011): "Facebook as a tool for producing sociality and connectivity", in *Television & New Media*, vol. 2, nº 13, pp. 1-17
- Wajcman, J. (2000): "Reflections on gender and technology studies: In what state is the art?", in *Social studies of science*, vol. 30, nº 3, pp. 447-464.
- Zafra, R. (2011). "Un cuarto propio conectado. Feminismo y creación desde la esfera público-privada online", in *Asparkía*, vol. 22, pp. 115-129.
- Zimmerman, T. (2017): "#Intersectionality: The Fourth Wave Feminist Twitter Community", in *Atlantis*, vol. 38, nº 1, pp. 54-70



El tratamiento periodístico de la privacidad en las aplicaciones de rastreo de COVID-19 en España y Reino Unido

The journalistic treatment of privacy in COVID-19 tracking applications in Spain and the United Kingdom

Sofiya Kamalova
so.kamalova@gmail.com
Universidad de Alicante

Alfredo Moralejo
alfredo.moralejo@gmail.com
Universidad de Alicante

RESUMEN

Para gestionar la pandemia del Covid-19 se planteó el uso de la tecnología para facilitar la detección y control de nuevos contagios. El uso de estas aplicaciones en teléfonos móviles para solucionar esta crisis sanitaria inició un debate sobre el impacto de estas aplicaciones en la privacidad de los usuarios. El objetivo de este estudio es identificar cómo se plantea esta cuestión en España y Reino Unido, a través del análisis de contenido de las comunicaciones periodísticas sobre estas aplicaciones de rastreo en ambos países. En esta investigación se han estudiado 249 artículos de periódicos británicos y españoles publicados en el periodo desde el 11 de marzo de 2020 hasta el 31 de octubre 2020. Con la información de los artículos analizados se ha realizado un análisis cuantitativo con el objetivo de encontrar y explicar las tendencias, similitudes y diferencias en el tratamiento periodístico de las aplicaciones de rastreo en los dos países. Los autores piensan que las diferencias en el planteamiento del debate de privacidad es debido a diferentes contextos socio culturales.

PALABRAS CLAVE Aplicaciones de rastreo, Radar Covid, NHS Covid App, España, Reino Unido

ABSTRACT

To manage the Covid-19 pandemic, the use of technology was proposed to facilitate the detection and control of new infection cases. The use of tracing applications on mobile phones to solve this health crisis started a debate on the impact of these applications on the privacy of users. The objective of this study is to identify, by applying content analysis of journalistic communications about these tracking applications in Spain and the United Kingdom, how this question arises in both countries. In this research, 249 articles from British and Spanish newspapers published in the period from March 11, 2020 to October 31, 2020 have been studied. With the information from the analysed articles, a quantitative analysis has been carried out to find and explain the trends, similarities and differences in the journalistic treatment of the tracking applications in the two countries. The authors believe that the differences in the approach to privacy debate is a result of different socio-cultural contexts.

KEYWORDS Tracing apps, Radar Covid, NHS Covid App, Spain, United Kingdom

Edita: Dafne Calvo

Recibido: 27/06/2021 | Aceptado: 25/11/2021

Cómo citar este artículo:

Kamalova, S. y Moralejo, A. (2022). El tratamiento periodístico de la privacidad en las aplicaciones de rastreo de COVID-19 en España y Reino Unido. *Dígitos. Revista de Comunicación Digital*, 8: 215-230. DOI: 10.7203/rd.v1i8.222



El tratamiento periodístico de la privacidad en las aplicaciones de rastreo de COVID-19 en España y Reino Unido

1. Introducción

Desde el inicio de la pandemia mundial del Covid-19 numerosos gobiernos crearon diversas iniciativas para aplicar las nuevas tecnologías para la gestión de la crisis. Una de las áreas propuestas desde el principio fue el uso de aplicaciones en dispositivos móviles para identificar contactos y potenciales contagios del virus. El planteamiento del uso de las aplicaciones para “nivelar la curva” de los contagios (Chan, 2021) ha generado preocupación sobre el impacto que éstas podrían tener sobre el derecho a la intimidad de los ciudadanos (Rudgard, 2020; Sabbagh et al. 2020a), lo que abre el debate sobre el compromiso entre la privacidad y la salud pública (Chan, 2021). Existen estudios que han utilizado una aproximación similar para analizar la cobertura periodística de las aplicaciones de rastreo en Alemania, Austria y Suiza (Amann et al. 2021) y en Suiza (Von Wyl 2021). La novedad de este estudio consiste en un análisis de la cobertura periodística en el contexto de España y Reino Unido enfocado en cuestiones de privacidad.

El objetivo de este estudio es identificar, a través del análisis comparado de las comunicaciones periodísticas sobre estas aplicaciones de rastreo en España y Reino Unido, las diferencias en percepción de la privacidad como factor clave y los enfoques utilizados en este debate. El análisis trata de determinar si las publicaciones periodísticas plantean la cuestión de privacidad y si lo plantean de una forma positiva (asegurando que se respeta el derecho a privacidad), negativa (criticando el uso inadecuado de información personal que se comparte) o neutra (sin evaluar). Para complementar el análisis, se determinan cambios en los enfoques de los artículos periodísticos que se refieren a las aplicaciones de rastreo durante el periodo estudiado.

El uso de aplicaciones de rastreo para la pandemia del COVID-19, así como el debate entre privacidad y salud pública está sujeto a los aspectos culturales y sociológicos en cada sociedad. Estos aspectos también influyen en los enfoques dados a las noticias de los medios de comunicación. Los autores piensan que las noticias referentes a los riesgos relacionados con la privacidad en las aplicaciones de rastreo en España y Reino Unido son una fuente para estudiar la opinión pública y mostrarán diferencias en el enfoque propias de las diferencias sociales y culturales. Para demostrarlo, se plantea la cuestión de privacidad con la introducción y el uso de las aplicaciones de rastreo de Covid-19 en una selección de periódicos en España y Reino Unido desde el 11 de marzo de 2020 hasta el 31 de octubre de 2020.

2. El uso de las aplicaciones en salud pública

El rastreo de casos es el proceso para el control de epidemias por el cual se identifican los contactos directos de personas infectadas y se corta la cadena de infección a través de la aplicación de cuarentenas. Tradicionalmente, este proceso se ha realizado

de manera manual por rastreadores que contactan con personas potencialmente infectadas y les solicitan que tomen medidas para evitar su transmisión. Este proceso tiene una serie de limitaciones, como la dificultad para conseguir datos completos de los pacientes infectados. Además, requiere una importante cantidad de tiempo y recursos personales para llevarlo a cabo en una pandemia como la del Covid-19. Estas limitaciones, junto con la velocidad de transmisión de este virus, condujeron a la idea del uso de las tecnologías para automatizar total o parcialmente el proceso de rastreo de contactos utilizando aplicaciones en teléfonos móviles (Ferreti et al., 2020). Existen varios diseños para las aplicaciones de rastreo: centralizado, descentralizado y una combinación entre ambos (Nadeem et al., 2020).

El uso de estas técnicas para el rastreo tiene una serie de consideraciones y riesgos tanto prácticos, técnicos, legales y éticos que introducen una serie de dificultades para su implementación y efectividad (Braithwaite et al. 2020). Entre ellas, se encuentra la posible invasión de la privacidad de los ciudadanos por parte de organizaciones públicas y privadas. Son muchos los organismos que han publicado recomendaciones o guías de buenas prácticas sobre el proceso de diseño e implantación de dichas aplicaciones de rastreo para mitigar estos riesgos. Algunos ejemplos de estos organismos son el Information Commissioner's Office de Reino Unido (2020) y organizaciones como Amnistía Internacional (2020), los cuales aprueban su uso como herramienta de control de la pandemia bajo determinadas recomendaciones.

El Gobierno de España anunció el 20 de mayo su iniciativa para la implantación de su aplicación de rastreo llamada *Radar Covid*. El objetivo era ponerla en marcha un mes más tarde en las Islas Canarias. Desde el primer momento se anunció que se seguirían las recomendaciones realizadas por organismos internacionales en cuanto a privacidad y se aplicarían las lecciones aprendidas por otros países que habían comenzado el desarrollo anteriormente. Por un lado, se utilizaría la tecnología creada por Apple y Google que no se apoya en la geolocalización del móvil para rastrear, sino que hace uso del bluetooth del mismo dispositivo para detectar la proximidad con otros dispositivos. Por otro lado, sería una aplicación descentralizada, lo que significa que la información de cada usuario se almacena en su propio dispositivo y no en un centro de datos de algún organismo o empresa (Vega, 2020a). La aplicación inició una fase de prueba en la isla de la Gomera el 23 de junio de 2020 y a finales de julio se anunció que las pruebas habían terminado con resultados positivos y se procedería a su puesta en marcha en todas las Comunidades Autónomas (Vega, 2020b). A mediados de agosto de 2020, las comunidades comenzaron a unirse a la aplicación que necesita la participación activa de las consejerías de sanidad para el mecanismo por el que los usuarios que dan positivo en los test PCR proporcionan esta información a la aplicación. La fase de puesta en marcha en todas las comunidades se extendió hasta octubre de 2020, cuando Madrid y Cataluña se unieron a *Radar Covid* (Vega, 2020c).

El Reino Unido anunció su iniciativa de crear una aplicación antes que el Gobierno de España, pero tardó más en ponerla en funcionamiento. El 31 de marzo de 2020 NHS (El sistema nacional de salud) empezó a trabajar en su propia aplicación de rastreo para detectar los nuevos casos de Covid-19. Al contrario de las recomendaciones de los expertos en privacidad (Lapolla et al. 2019) y preferencias de los países europeos, optaron por una aplicación que funciona de una forma centralizada, porque según el gobierno proporcionará datos más precisos. Al principio rechazaron la colaboración con Apple y Google para centrarse en una aplicación propia (Sabbagh et al. 2020b). En

mayo de 2020 empezó el periodo de pruebas en la isla Wight (Murphy, 2020a). Como resultado de este periodo de prueba se detectaron muchos fallos técnicos que impedían el funcionamiento de la aplicación. Al no poder crear una aplicación propia, optaron por una aplicación con Apple y Google que funciona de una manera descentralizada (Murphy et al., 2020b). Con mucho retraso lanzaron la aplicación a finales de septiembre de 2020 (Hern, 2020).

De acuerdo con diferentes estudios realizados, la privacidad es una de las principales preocupaciones de los ciudadanos en cuanto a la adopción de la aplicación (Lucy Simko et. al 2020). Tal y como se menciona en Shavneet Sharma et al., (2020), la actitud de los posibles usuarios ante la instalación y uso de las aplicaciones de rastreo es subjetiva y viene dada por la percepción personal de riesgo y beneficio, así como de las experiencias previas.

Además, estas percepciones vienen condicionadas por una serie de valores culturales que influyen de diferentes maneras en distintos países. Los valores culturales son un conjunto de creencias que guían las actitudes y el comportamiento. Tienden a perdurar incluso cuando otras diferencias entre países desaparecen por cambios en la economía, la política, la tecnología y otros factores externos (Bellman, 2004).

De acuerdo al mencionado estudio, y aplicando la teoría de las dimensiones culturales de Hofstede, hay dos dimensiones especialmente aplicables a este caso, el *Individualismo* y la *Aversión a la Incertidumbre* (Krasnova et al. 2010). Por ello, diferentes sociedades pueden (y de hecho presentan) diferentes actitudes ante el debate entre privacidad y salud pública. Una diferencia cultural en la dimensión *individualismo/colectivismo* sugiere que los usuarios de Internet de culturas que premian ser un individuo en lugar de ser parte de una comunidad pueden ser más egocéntricos (Hofstede, 1991). Como resultado, se centran menos en las gratificaciones sociales que puede ofrecer el hecho de compartir datos. Una diferencia entre países en la dimensión cultural de *aversión a la incertidumbre* sugiere que algunos usuarios pueden intentar evitar situaciones de incertidumbre, como por ejemplo enfrentar riesgos de privacidad (Hofstede, 1991). Por lo tanto, las personas de culturas que ocupan un lugar destacado en la prevención de la incertidumbre deben tener mayores preocupaciones sobre la privacidad. (Trepte et al., 2017)

El siguiente gráfico muestra los valores de las dimensiones de Hofstede en España y Reino Unido y las diferencias en cuanto al individualismo (donde Reino Unido se muestra como un país más individualista que España) mientras que la sociedad española destaca por una aversión al riesgo mucho más acentuada que Reino Unido.

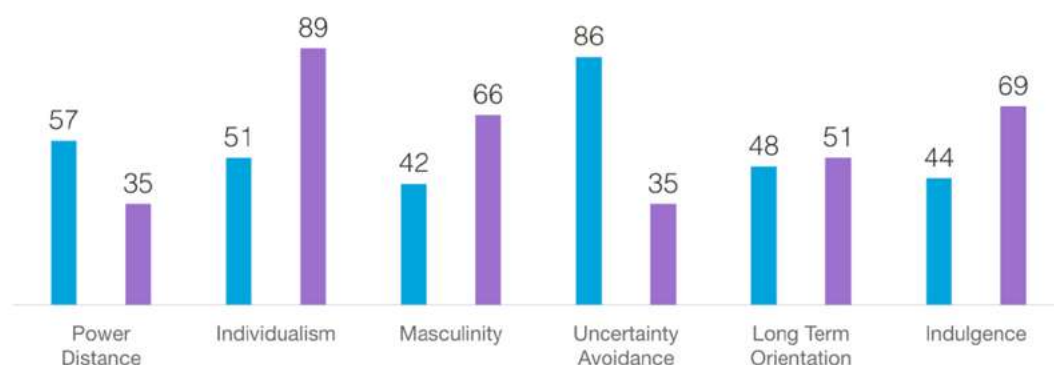


Gráfico 1. Dimensiones culturales de Hofstede de España (azul) y Reino Unido (morado)

Fuente: <https://www.hofstede-insights.com/product/compare-countries/>

La interacción entre los medios de comunicación y la sociedad es estudiada por teoría del encuadre o *framing*. Esta interacción se produce a los diferentes niveles de la comunicación, incluyendo la elaboración de las noticias. El individuo-periodista y toda su ideología, carga valorativa, criterios e intereses tienen un efecto en el enfoque que se aplica a las noticias cómo se interpretan y cómo se construye la narrativa alrededor de cada tema. Además, la sociología de las redacciones y, por tanto, su entorno influye en la elaboración de las noticias y sus enfoques (Amadeo, 2002). Son factores que pueden influir en la opinión pública, que resulta significativa a la hora de introducir la aplicación y asegurar que el uso de la aplicación forma parte del día a día de los ciudadanos. El éxito de las aplicaciones de rastreo de Covid-19 depende de la extensión que obtendrá en la población (Colizz, 2021).

3. Aproximación metodológica

Dado que el objetivo de este estudio es identificar las diferencias de percepción de la privacidad en relación a las aplicaciones de rastreo en diferentes países se ha realizado un análisis cuantitativo del contenido de los artículos periodísticos sobre estas aplicaciones. Se ha diseñado una muestra intencional de periódicos digitales generalistas basada en los datos de audiencia en cada país. Tanto para España como para Reino Unido se ha utilizado la información del “Reuters Institute Digital News Report” de 2019. Este estudio proporciona una lista de medios digitales que la población utiliza para informarse sobre las noticias. De esta lista, se seleccionaron versiones digitales de periódicos de cada país. Todos los medios en ambos países son privados y hay un equilibrio entre la ideología de los medios, para prevenir así que el debate de la privacidad se asocie al discurso de los partidos de la derecha o partidos de la izquierda.

La búsqueda se hizo por el nombre de la aplicación (“Radar Covid” en España y “NHS Covid-19” en Reino Unido) de cada país en los medios seleccionados en la base de datos MyNews, desde el 11 de marzo de 2020 hasta 31 de octubre 2020. Se decidió comenzar la búsqueda el 11 de marzo de 2020, cuando la OMS oficialmente anunció que Covid-19 era una pandemia. El 31 de octubre de 2020, por su parte, se asocia al fin de la segunda ola. En la primera búsqueda se obtuvieron 344 artículos en los periódicos españoles y 120 artículos en los periódicos británicos.

Sobre los artículos encontrados se ha realizado primer filtrado en base de una lectura inicial de los títulos de artículos y los párrafos introductorios para eliminar posibles noticias duplicadas, no relacionadas con el tema analizado o que estén escritas en una lengua que no sea español o inglés.

Se seleccionaron artículos con el tema de “Radar Covid” o “NHS Covid-19” como tema principal o un tema relevante. Se eliminaron los artículos que hacían una mención de la aplicación sin desarrollar el tema con más profundidad. Todas estas noticias junto con sus datos principales (fecha, titular, medio) se han añadido a una ficha básica (Anexo) de noticias atribuyendo un código para identificar cada noticia. Después del primer filtrado, el corpus se compuso de 252 artículos (Tabla 1).

Se creó una ficha para analizar individualmente cada artículo recogido (disponible en anexo). La ficha incluye información bibliométrica, enfoque de privacidad y otros enfoques, como el enfoque agenda, social, económico, político y tecnológico, para identificar qué enfoque suelen tener las noticias que abordan el tema de la privacidad. Para identificar estos enfoques se ha realizado una lectura de una muestra aleatoria, pero que incluye artículos de cada medio seleccionados.

Tabla 1. Número de artículos recogidos organizado por medios en España y en Reino Unido

Periódicos españoles	Número de artículos	Periódicos de Reino Unido	Número de artículos
<i>El País</i>	43	<i>The Guardian</i>	23
<i>El Mundo</i>	17	<i>Telegraph</i>	6
<i>ElDiario.es</i>	54	<i>Times</i>	1
<i>20minutos</i>	84	<i>Independent</i>	24
Total	198	Total	54

3.1. La evaluación del tratamiento de la privacidad

Se acordó la definición de cada enfoque por consenso entre los autores. Como objeto principal del estudio, se analizó el *tratamiento sobre la privacidad* de las aplicaciones de rastreo. En primer lugar, se identificó si este tema es tratado y en caso de que esté presente si se habla de la gestión de la privacidad de una forma positiva (la aplicación respeta la privacidad de los usuarios), negativa (la aplicación pone en riesgo la privacidad de los usuarios) o neutra (simplemente se menciona la cuestión sin evaluar). Es importante señalar que el tratamiento de la privacidad se puede deducir implícitamente de los términos usados para describir el funcionamiento de las aplicaciones, por ejemplo:

- Se entiende un tratamiento positivo cuando las noticias mencionan que las notificaciones e información generadas por la aplicación son “anónimas”.
- Las noticias que especifican que la aplicación no puede acceder a datos personales o de manera muy limitada dan un tratamiento positivo a la privacidad de la aplicación ya que se infiere que, si no acceden a los datos, no pueden hacerlos públicos a otros usuarios.
- Describir que una aplicación funciona de manera “descentralizada” implica un tratamiento positivo, ya que se considera que las aplicaciones descentralizadas tienen un mejor tratamiento de la privacidad que las centralizadas.

Para asegurar la unicidad de criterios entre los evaluadores se ha realizado un pre-test (Vilches, 2020) en el que ambos analizaron un grupo de artículos, se discutieron los términos que identifican enfoques específicos y se compararon los resultados obtenidos para aclarar las posibles diferencias de criterios. Como resultado se ha detectado que el enfoque social y el enfoque político requiere una definición más específica. Se revisaron las definiciones de los dos enfoques que causaron complicaciones para poder continuar con el análisis de los artículos.

Además, del enfoque de privacidad, se evalúa si el enfoque de la noticia es de agenda, social, económico, político o tecnológico, pues se asume que una noticia puede asociarse a uno o varios enfoques al mismo tiempo.

- El *enfoque agenda* incluye las noticias que informan sobre la cantidad de usuarios, el número de descargas, comparación del número de descargas entre diferentes regiones o diferentes países, las regiones administrativas donde se implanta la aplicación, información sobre las fechas de superación de prueba o los progresos en su implantación.

- El *enfoque social* cubre un abanico variado de noticias que tratan el tema de la aplicación y su uso integrando la sociedad y enfatizando el papel que tienen los ciudadanos a la hora de utilizar la aplicación. Se incluye un artículo dentro del enfoque social cuando se considera a la sociedad como participante activa para que la aplicación tenga éxito. También el enfoque social cubre las historias humanas.
- El *enfoque económico* es la categoría que cubre las noticias que se refieren al impacto económico de la pandemia o a las acciones encaminadas a la recuperación económica, así como a aspectos económicos del desarrollo o implantación de las aplicaciones de rastreo.
- El *enfoque político* se refiere a noticias que tratan la implantación de Covid-19 desde un punto de vista de los conflictos o alianzas que ésta genera entre diferentes partidos o instituciones políticas. No se consideraron políticas las noticias simplemente por la mención de instituciones o políticos, era necesario que cubrieran un aspecto politizado de su gestión.
- El *enfoque tecnológico* cubre las noticias tecnológicas, instrucciones de uso (cómo y dónde se puede descargar la aplicación y qué información hay que proporcionar para usarla), descripción de cómo funciona (sistema centralizado o descentralizado, tecnología Bluetooth) y de qué otros usos puede tener la aplicación.

Con la información de los artículos analizados, se ha realizado un análisis cuantitativo de los datos cualitativos con el objetivo de identificar diferencias de enfoque en las noticias sobre aplicaciones de rastreo entre los medios españoles y británicos, variaciones en formas de abordar la privacidad entre ambos países, cambios en el tratamiento de la privacidad a lo largo del tiempo, relación entre el tratamiento de la privacidad de las noticias y otros enfoques encontrados en los artículos analizados.

4. Resultados

Los resultados de esta investigación muestran datos de cada país centrados en el enfoque de la privacidad. La separación de los datos por países permite hacer una comparación entre España y Reino Unido. Hay una diferencia significativa entre el número de artículos que abordan la privacidad. En España, como lo muestra (Tabla 2) se recogieron 96 noticias que tratan el tema de privacidad mientras en Reino Unido se han obtenido 21 artículos. Una posible explicación del mayor número de noticias españolas con el enfoque de la privacidad es la sección regional del periódico *20 minutos*, donde se obtuvieron 84 noticias (Tabla 2).

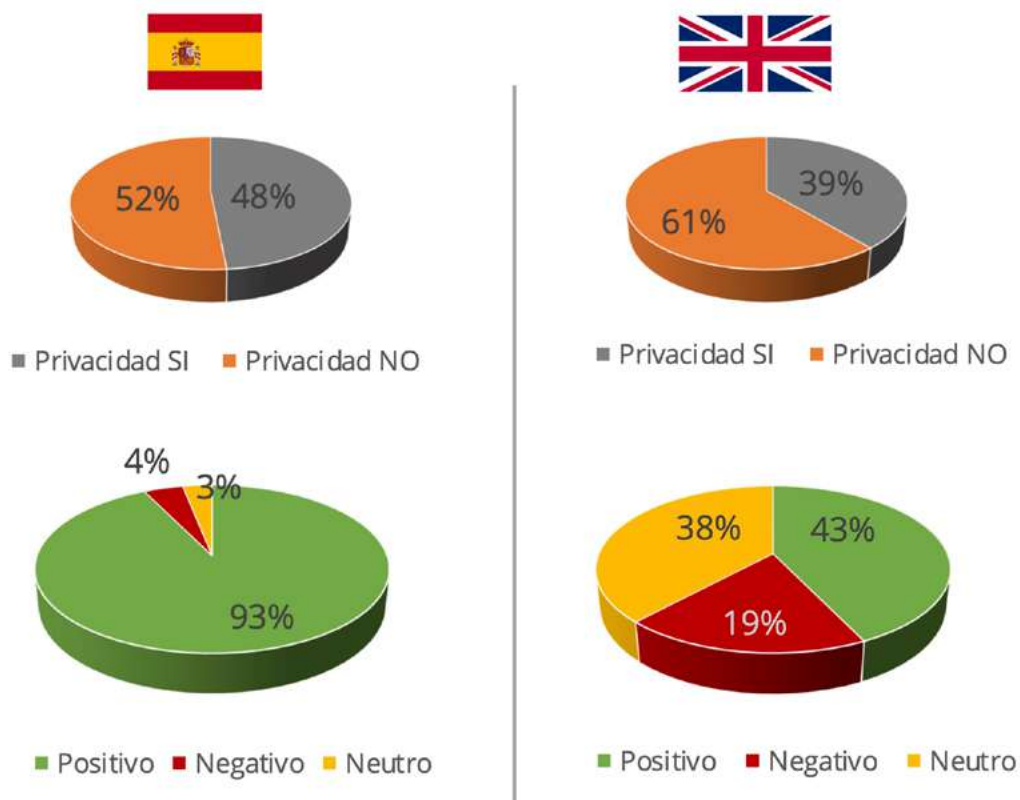
El análisis del tratamiento de la privacidad en las noticias sobre Radar Covid en España Gráfica 2 muestra que aproximadamente la mitad de las noticias sobre este tema tratan de alguna manera su impacto en la privacidad de los usuarios. De estas noticias una gran mayoría (el 92,7%) lo hacen de una manera positiva, mientras que solamente un 4,2% lo hacen de manera desfavorable. El mismo análisis de las noticias en los medios británicos muestra que la privacidad es tratada por un porcentaje importante de la muestra, aunque ligeramente inferior (38,9%). El porcentaje de noticias que lo hacen de manera positiva es mucho más bajo que en las españolas (42,9%) y es muy similar a las noticias que lo hacen de una manera neutra (38,1%). Casi una de cada cinco noticias muestra una visión negativa de la privacidad de *Radar Covid*.

Tabla 2. El total de noticias de cada país organizado por los enfoques de privacidad

	España		Reino Unido	
Tratan privacidad	96	48,5%	21	38,9%
No tratan privacidad	102	51,5%	33	61,1%
Enfoque positivo de privacidad	89	92,7%	9	42,9%
Enfoque negativo de privacidad	4	4,2%	4	19,0%
Enfoque neutro de privacidad	3	3,1%	8	38,1%
Total noticias	198		54	

El siguiente gráfico muestra el porcentaje de noticias que tratan la privacidad y el enfoque de la privacidad de estas en España y en Reino Unido.

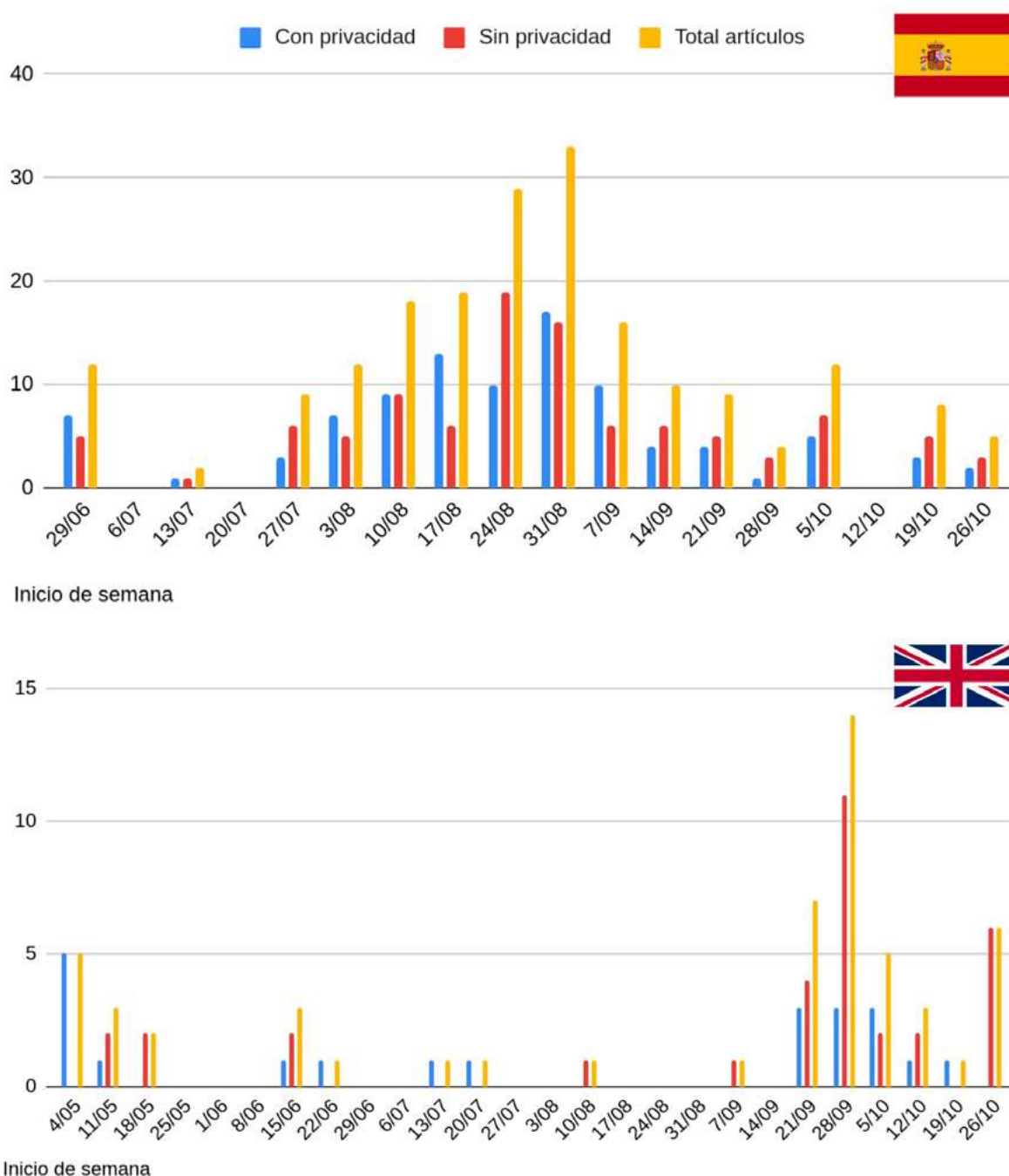
Gráfico 2. Tratamiento de privacidad en las noticias sobre Radar Covid en medios de España



En el gráfico 3 se pueden observar el número de artículos sobre *Radar Covid* y su tratamiento de privacidad por semana en los medios españoles. A finales de junio existe un primer pico de artículos que se corresponden con el lanzamiento de las pruebas de la aplicación en la isla de La Gomera. Se trata la privacidad en más de la mitad de estos artículos. Durante julio se encuentra un número bajo de artículos, hasta que a

finales del mes se publicaron los resultados de la prueba. A partir de aquí se observa una subida continuada del número artículos sobre la aplicación durante todo el mes de agosto, que se corresponde con informaciones sobre el avance del despliegue en las distintas comunidades autónomas (enfoque agenda), sobre su funcionamiento y los potenciales beneficios de su implantación. El mayor número de noticias se produce en la última semana de agosto, a partir de ahí comienza a bajar el número de publicaciones que sobre todo se refieren al progreso en las últimas comunidades que lo adoptaron y ciertas controversias políticas relacionadas con estas comunidades. En cuanto al porcentaje de noticias con o sin privacidad, no se aprecian grandes cambios en la evolución a lo largo del tiempo.

Gráfico 3. Artículos por semana sobre Radar Covid en medios españoles y británicos



En el mismo gráfico 3 también se muestra el seguimiento cronológico en Reino Unido. Se puede ver como el número de los artículos aumenta en fechas claves del desarrollo e implantación de la aplicación. En la primera semana de mayo, NHS lanza su primera versión de la aplicación con un sistema centralizado. Las pruebas se realizaron en la isla de Wight. Debido a que un sistema centralizado se considera menos seguro en cuanto a la privacidad de los usuarios, todos los artículos abordan la privacidad.

El 21 de mayo Apple y Google lanzan su aplicación mientras NHS aún no puede anunciar ninguna fecha. La razón por la que la privacidad no se aborda en las noticias de la segunda y tercera semana de mayo podría ser porque se centran en la crítica de poca eficiencia de gobierno que supone el retraso de lanzamiento de aplicación.

Hay una pausa de tres semanas. El 18 de junio el gobierno anuncia que su intento de crear una aplicación propia con sistema centralizado no tuvo éxito y que optan por una colaboración con Apple y Google para lanzar una aplicación de NHS. Las noticias de la tercera y cuarta semana hablan sobre el nuevo anuncio del gobierno y el sistema descentralizado que utiliza Apple y Google.

El mayor número se produce en la segunda y tercera semana de septiembre debido a que el 24 de septiembre se publica la nueva aplicación, fruto de la colaboración entre Apple y Google y NHS. Las noticias que abordan privacidad explican cómo funciona el sistema descentralizado y qué datos requiere la aplicación.

El índice de noticias a finales de octubre no incluye aquellas que tratan privacidad, pues se centra en los problemas técnicos que aparecieron cuando se lanzó la aplicación y las soluciones que el gobierno ha propuesto.

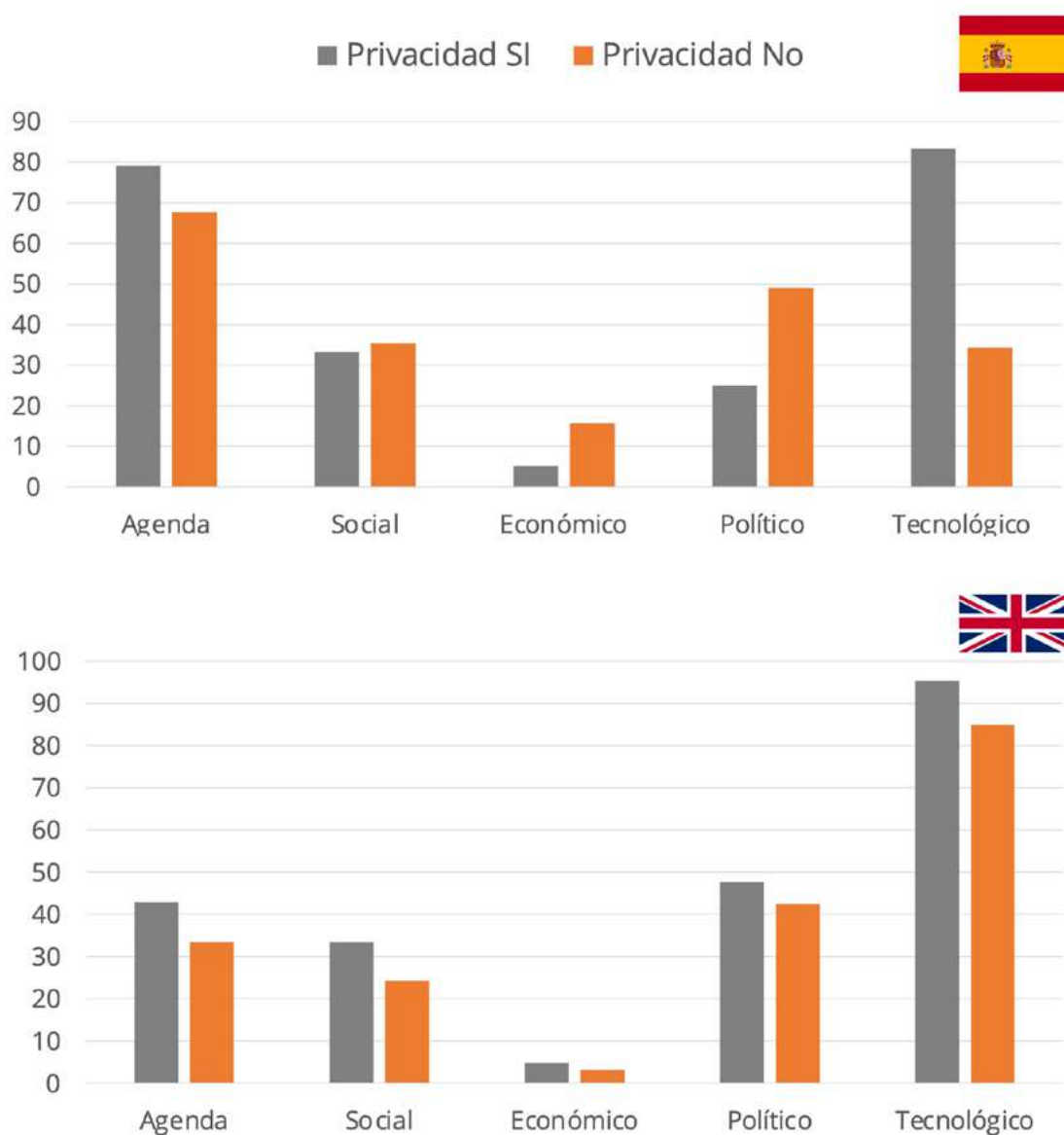
En el gráfico 4 se muestran los porcentajes de presencia de enfoque por tratamiento de privacidad en España y en Reino Unido. En España hay muchas más noticias de agenda. Como ya se ha mencionado anteriormente, esto es debido a la importancia de las noticias regionales que documentan el proceso de implantación en cada comunidad. Mientras en Reino Unido, excluyendo Escocia, la implantación de la aplicación era uniforme para todo el país.

En España, la mayoría de las noticias con un enfoque económico no suelen tratar la privacidad. Mientras en Reino Unido el número de las noticias con un enfoque económico es demasiado bajo para sacar conclusiones. Además, la cuestión de privacidad en España no es tan politizada como en Reino Unido. Incluso en Reino Unido hay más noticias con enfoque político que abordan la privacidad que noticias que no la abordan. Tanto en España como en Reino Unido, la mayoría de las que lo abordan tiene un enfoque tecnológico. Hay mucha más información tecnológica en Reino Unido que en España, sobre todo en las noticias que no abordan privacidad, debido a un gran número de artículos que explican cómo se utiliza y se instala la aplicación.

5. Discusión

Según la teoría del *framing*, los medios de comunicación tienen la capacidad de dirigir la mirada de gran parte de los ciudadanos hacia unos temas u otros. Como consecuencia, la cobertura periodística influye en las prioridades y preocupaciones de los ciudadanos (Núñez Gómez et al., 2020). El análisis presentado en este trabajo muestra diferentes planteamientos en el debate sobre la privacidad de las aplicaciones de rastreo. Mientras que en España los medios de comunicación enfatizan la idea de seguridad, los

Gráfico 4. Porcentajes de presencia de enfoques por tratamiento de privacidad en España y en Reino Unido



medios de comunicación británicos proporcionan un debate crítico sobre la privacidad. Esta diferencia entre los medios de comunicación a la hora de enfocar la cuestión de privacidad puede interpretarse como el resultado de factores socio culturales de cada país.

De las seis características que señala Hofstede (1991), *aversión a la incertidumbre* e *individualismo* son las que más relevancia tienen en las cuestiones de privacidad (Trepte et al. 2017). Por un lado, en la sociedad española la aversión a la incertidumbre está a 86% mientras en Reino Unido está a 35% (Gráfica 1). Sin embargo, la sociedad británica es más individualista (89%) mientras que la española se caracteriza como una sociedad colectiva (51%). Por ello, las dos sociedades tienen predisposición a proteger sus datos, pero desarrollan el debate de la privacidad de formas diferentes.

Debido a que uno de los factores críticos para el éxito de la aplicación era la percepción de riesgo y beneficio potencial, dar seguridad a la ciudadanía sobre el tratamiento de sus datos privados es un elemento importante. Como resultado, los medios españoles

han tenido un papel no sólo informativo, sino que se han involucrado en la promoción y concienciación ciudadana a favor de la adopción de Radar Covid como herramienta de rastreo. Las conexiones sociales y las gratificaciones que uno recibe al utilizar la aplicación aumentan el sentimiento de seguridad (Trepte et al. 2017). En esta campaña, los medios han enfatizado los beneficios de la aplicación y una visión positiva de su tratamiento de la privacidad, lo que ha dejado poco espacio al juicio propio de los ciudadanos o a un debate público sobre la privacidad.

A diferencia de los medios españoles, los británicos presentan rasgos más individualistas y han optado por realizar un debate en mayor profundidad en cuanto a posibles consecuencias en la privacidad de los usuarios (también en algunos aspectos sociales). Además, el hecho de que Reino Unido ha lanzado una primera versión de NHS Covid-19 siguiendo el modelo centralizado en lugar del descentralizado (recomendado por todos los organismos que analizaron el tema) redundaba en una percepción más negativa. La decisión de compartir datos es resultado de un cálculo racional de costo-beneficio en el que los riesgos de compartir datos se comparan con los beneficios obtenidos (ayuda a gestión de la pandemia y prevención de nuevos contagios). Sin embargo, en una sociedad individualista, las gratificaciones sociales como la seguridad colectiva, la protección ante el contagio o el compromiso con la sociedad tienen menos valor que en una sociedad colectiva.

Hoy en día, los valores y las normas culturales son decisivos para la forma en que se percibe la privacidad del espacio virtual (Trepte et al. 2017). En este trabajo se demuestra cómo estas diferencias culturales afectan la cobertura de la cuestión de privacidad en el uso de las aplicaciones de rastreo de Covid-19 en los medios de comunicación en España y en Reino Unido. En el caso de estas aplicaciones, la percepción de los ciudadanos sobre sus riesgos, era de vital importancia ya que su éxito dependía de una implantación masiva entre la población.

6. Conclusiones

Con los datos que se han obtenido en esta investigación, podemos constatar que la privacidad es un tema al que se atribuye importancia a la hora de hablar de las aplicaciones de rastreo tanto en España como en Reino Unido. El análisis del tratamiento de la privacidad en las noticias sobre *Radar Covid* en España (Gráfica 2) muestra que aproximadamente la mitad de las noticias sobre este tema tratan de alguna manera su impacto en la privacidad de los usuarios. El mismo análisis de las noticias en los medios británicos indica que la privacidad es tratada por un porcentaje ligeramente inferior, aunque también relevante (38,9%).

Como resultado, se ha mostrado que las noticias españolas tienden a hacer un tratamiento más positivo de la privacidad, mientras los medios británicos proporcionan un debate más profundo donde 19% de noticias tienen un enfoque negativo y 38,1% de noticias tienen un enfoque neutro. El tratamiento positivo de la privacidad (92,7% de las noticias que abordan privacidad tienen un enfoque positivo) es resultado de una campaña de concienciación cuyo objetivo era promover el uso de la aplicación en la sociedad. Por otra parte, los medios británicos afectados por un desarrollo de la implantación de la aplicación más complejo y situados en un contexto social y cultural distinto, son más críticos con el tratamiento del tema de la privacidad.

Se recogieron más artículos de los periódicos españoles que de los periódicos

británicos. Por lo tanto, hay más artículos que abordan la privacidad en España que en Reino Unido. La cronología muestra como el número de noticias aumenta en los momentos clave, como el anuncio del proyecto, el periodo de pruebas y el lanzamiento de la aplicación. La diferencia tan drástica en el número de noticias (198 en periódicos españoles y 54 en británicos) supone una limitación a la hora de comparar estos dos países. Por otra parte, no se ha analizado la población a la que llegaron estos artículos, por lo que no es posible obtener conclusiones en cuanto a la capacidad de influencia sobre la percepción y el uso de las aplicaciones en estos países.

En futuras investigaciones, sería interesante hacer un análisis discursivo de los artículos que tienen el tema de la privacidad como temática principal para observar los argumentos que se utilizan, a fin de asegurar que la aplicación respeta la privacidad, y explorar los argumentos se utilizan para constatar que la aplicación abusa de la privacidad.

Referencias

- Amadeo, B. (2002): "La teoría del Framing. Los medios de comunicación y la transmisión de significados", en *Revista de Comunicación*, vol. 1, pp. 6-32. Disponible en: <https://revistadecomunicacion.com/es/articulos/2002/Art006-32.html> [Consulta: 3 de mayo de 2021]
- Amann, J., Sleight, J., y Vayena, E. (2021): "Digital contact-tracing during the Covid-19 pandemic: an analysis of newspaper coverage in Germany, Austria, and Switzerland", en *Plos one*, vol. 16, nº 2. Disponible en: <https://doi.org/10.1371/journal.pone.0246524> [Consulta: 11 de noviembre 2021]
- Amnistía Internacional (2020): "10 bulos sobre COVID-19 y derechos humanos". Disponible en: <https://www.es.amnesty.org/en-que-estamos/blog/historia/articulo/10-bulos-covid19-derechos-humanos/> 10 bulos sobre COVID-19 y derechos humanos [Consulta: 1 de mayo de 2021]
- Ahmed, N., Michelin, R. A., Xue, W., Ru, S., Malaney, R., Kanhere, S. S., Seneviratne, A., Hu, W., Janicke, H. y Jha, S. K. (2020): "A Survey of COVID-19 Contact Tracing Apps", en *IEEE Access*, vol. 8, pp. 134577-134601. Disponible en: <https://ieeexplore.ieee.org/abstract/document/9144194> [Consulta: 1 de mayo de 2021]
- Bacciarelli, A. (2020): "Contact tracing apps: a test for privacy in Europe". Disponible en: <https://www.amnesty.org/en/latest/news/2020/04/contact-tracing-apps-privacy-in-europe/> [Consulta: 1 de mayo de 2021]
- Bellman, S., Johnson, E. J., Kobrin, S. J. y Gerald L. Lohsem G. L. (2004): "International Differences in Information Privacy Concerns: A Global Survey of Consumers", en *The Information Society* vol. 20, nº 5, pp. 313-324. Disponible en <https://doi.org/10.1080/01972240490507956> [Consulta: 2 de mayo de 2021]
- Braithwaite I., Callender T., Bullock M. y Aldridge R. W. (2020): "Automated and partly automated contact tracing: a systematic review to inform the control of COVID-19", en *The Lancet Digital Health*, vol. 2, nº 11, pp. E607-E621. Disponible en: [https://www.thelancet.com/journals/landig/article/PIIS2589-7500\(20\)30184-9/fulltext](https://www.thelancet.com/journals/landig/article/PIIS2589-7500(20)30184-9/fulltext)
- Chan, E. Y. y Saqib, N. U. (2021): "Privacy concerns can explain unwillingness to download and use contact tracing apps when COVID-19 concerns are high", en *Computers*

- in Human Behavior*, vol. 119, nº 106718, pp. 1-12. Disponible en: <https://www.sciencedirect.com/science/article/pii/S0747563221000406> [Consulta: 2 de mayo de 2021]
- Colizza, V., Grill, E., Mikolajczyk, R., Cattuto, C., Kucharski, A., Riley, S., Kendall, M., Lythgoe, K., Bonsall, D., Wymant, C., Abeler-Dörner, L., Ferretti, L. y Fraser, C. (2021): "Time to evaluate COVID-19 contact-tracing apps", en *Nat Med* vol. 27, pp. 361-362. Disponible en: <https://www.nature.com/articles/s41591-021-01236-6> [Consulta: 1 de mayo de 2021]
- Ferretti, L., Wymant, C., Kendall, M., Zhao, L., Nurtay, A. y Abeler-Dörner, L. (2020): "Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. *Science*", en *Science*, vol. 368, nº 6491, pp. 1-13. Disponible en: <https://www.sinobiological.com/research/virus/sars-cov-reagent> [Consulta: 3 de mayo de 2021]
- Hern, A. (2020): "NHS coronavirus app to launch in England and Wales on 24 September". Disponible en: <https://www.theguardian.com/world/2020/sep/11/nhs-coronavirus-app-launch-england-wales-24-september> [Consulta: 1 de mayo de 2021]
- Hofstede, G. H. (1991): *Cultures and organizations: Software of the mind*. New York, NY: McGraw-Hill.
- Information Commission Office (2020): "COVID-19 contact tracing: data protection expectations on app development". Disponible en: <https://ico.org.uk/media/for-organisations/documents/2617676/ico-contact-tracing-recommendations.pdf> [Consulta: 1 de mayo de 2021]
- Klenk, M. y Duijf, H. (2020): "Ethics of Digital Contact Tracing and COVID-19: Who Is (Not) Free to Go?", en *Ethics and Information Technology*, vol. 2, nº 10. Disponible en: <https://link.springer.com/article/10.1007/s10676-020-09544-0> [Consulta: 3 de mayo de 2021]
- Krasnova, H., Spiekermann, S., Koroleva, K. y Hildebrand, T. (2010): "Online Social Networks: Why We Disclose", en *Journal of Information Technology*, vol. 25, nº 2, pp. 109-125. Disponible en: <https://doi.org/10.1057/jit.2010.6> [Consulta: 3 de mayo de 2021]
- Lapolla, P. y Lee, R. (2020): "Privacy versus safety in contact-tracing apps for coronavirus disease 2019", en *Sage Journal*, vol. 6, nº 2055207620941673, pp. 1-2. Disponible en: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7364796/> [Consulta: 3 de mayo de 2021]
- Morley, J, Cowls, J., Taddeo, M. y Floridi, L (2020): "Ethical guidelines for COVID-19 tracing apps: Protect privacy, equality and fairness in digital contact tracing with these key questions", en *Nature*, vol. 582, pp. 29-31. Disponible en: <https://www.nature.com/articles/d41586-020-01578-0> [Consulta: 1 de mayo de 2021]
- Murphy, S. (2020): "What have we got to lose?: Isle of Wight responds to NHS app". Disponible en: <https://www.theguardian.com/world/2020/may/08/what-have-we-got-to-lose-isle-of-wight-responds-to-nhs-covid-19-contact-tracing-app-> [Consulta: 1 de mayo de 2021]
- Murphy, S., Sabbagh, D. y Hern, A. (2020): "Piloted in May, ditched in June: the failure of England's Covid-19 app". Disponible en: <https://www.theguardian.com/world/2020/>

[jun/18/piloted-in-may-ditched-in-june-the-failure-of-englands-covid-19-app](#)
[Consulta: 1 de mayo de 2021]

Newman, N., Fletcher, R., Kalogeropoulos, A. & Nielsen, R. K. (2019): *Reuters Institute Digital News Report 2019*, https://reutersinstitute.politics.ox.ac.uk/sites/default/files/inline-files/DNR_2019_FINAL.pdf

Núñez Gómez, P., Abuín Vences, N., Sierra Sánchez, J. y Mañas Viniegra, L. (2020): "El enfoque de la prensa española durante la crisis del Covid-19. Un análisis del framing a través de las portadas de los principales diarios de tirada nacional.", en *Revista Latina de Comunicación Social*, vol. 78, pp. 41-63.

Rudgard, O. (2020): "Will a contact tracing app actually stop the coronavirus spread? The Telegraph". Disponible en: <https://www.telegraph.co.uk/technology/0/contact-tracing-app-how-do-actually-work/> [Consulta: 3 de mayo de 2021]

Sabbagh, D., Alex Hern, A. y Proctor, K. (2020): "UK racing to improve contact-tracing app's privacy safeguards". Disponible en: <https://www.theguardian.com/technology/2020/may/05/uk-racing-to-improve-contact-tracing-apps-privacy-safeguards> [Consulta: 2 de mayo de 2021]

Sabbagh, D. y Hern, A. (2020): "UK coronavirus contact tracing app: is it safe and will people use it?". Disponible en: <https://www.theguardian.com/world/2020/may/05/uk-coronavirus-contact-tracing-app-is-it-safe-and-will-people-use-it> [Consulta: 1 de mayo de 2021]

Sharma, S., Singh, G., Sharma, R., Jones, P., Kraus, S. y Dwivedi, Y. K. (2020): "Digital Health Innovation: Exploring Adoption of COVID-19 Digital Contact Tracing Apps" en *IEEE Transactions on Engineering Management*, pp. 1-17. Disponible en: <https://covid19.elsevierpure.com/es/publications/digital-health-innovation-exploring-adoption-of-covid-19-digital-> [Consulta: 3 de mayo de 2021]

Simko, L., Calo, R., Roesner, F., y Kohno, T. (2020): "COVID-19 Contact Tracing and Privacy: Studying Opinion and Preferences", en *Cornell University*. Disponible en: <https://arxiv.org/abs/2005.06056> [Consulta: 3 de mayo de 2021]

Trepte, S., Reinecke, L., Ellison, N.B., Quiring, O., Yao, M.Z., Ziegele, M. A. (2017): "Cross-Cultural Perspective on the Privacy Calculus", en *Social Media + Society*. Disponible en <https://doi.org/10.1177/2056305116688035> [Consulta: 3 de mayo de 2021]

Vega, G. (2020): "Así funciona Radar Covid, la app de rastreo de contagiados de coronavirus: La aplicación ya está disponible en la Google Play Store y próximamente lo estará en la de Apple". Disponible en: <https://elpais.com/tecnologia/2020-07-01/asi-funciona-radar-covid-la-app-de-rastreo-de-contagiados-de-coronavirus.html> [Consulta: 1 de mayo de 2021]

Vega, G. (2020): "El Gobierno lanza la primera fase de Radar Covid, la 'app' de alerta de contagios del coronavirus: La aplicación funcionará en Android e iOS, y ya ha sido enviada a Google y Apple para que supere sus procesos de validación y se publique en sus respectivas tiendas". Disponible en <https://elpais.com/tecnologia/2020-06-29/el-gobierno-lanza-la-primera-fase-de-radar-covid-la-app-de-alerta-de-contagios-del-coronavirus.html> [Consulta: 1 de mayo de 2021]

Vega, G. (2020): "Trece comunidades autónomas están implantando la aplicación Radar

Covid: Fernando Simón anuncia que “tres o cuatro” regiones están en disposición de ponerla en marcha antes de septiembre”. Disponible en <https://elpais.com/tecnologia/2020-08-17/trece-comunidades-autonomas-estan-implantando-la-aplicacion-radar-covid.html> [Consulta: 1 de mayo de 2021]

Vilches, L. (2020): *La investigación en comunicación: Métodos y técnicas en la era digital*. Barcelona: Editorial GEDISA

Von Wyl, V. (2021): “Challenges for nontechnical implementation of digital proximity tracing during the COVID-19 pandemic: media analysis of the SwissCovid app”, en *JMIR mHealth and uHealth*, vol. 9, nº 2. Disponible en: <https://mhealth.jmir.org/2021/2/e25345/> [Consulta: 11 de noviembre de 2021]

Zens, M., Brammertz, A., Herpich, J., Südkamp, N. y Hinterseer M. (2020): “App-Based Tracking of Self-Reported COVID-19 Symptoms: Analysis of Questionnaire Data”, en *J Med Internet Res*, vol. 22, nº 9. Disponible en: <https://www.jmir.org/2020/9/e21956/> [Consulta: 3 de mayo de 2021]


ANEXO: La ficha utilizada para análisis de los artículos recogidos


Código ID	Número que identifica el artículo, se usa ES-X para noticias españolas y UK-X para noticias británicas.
Medio	Nombre del medio de comunicación
Sección	Sección dentro del medio donde se encuentra el artículo
Fecha de publicación	Los artículos están ordenados cronológicamente empezando con los artículos más recientes.
Título del artículo	
¿Aborda privacidad?	Si o No
Enfoque de privacidad	Positivo, negativo o neutro
Enfoques	Los enfoques posibles son: agenda, social, económico, político y tecnológico
Fuente de información	




El consumo de alcohol, tabaco y drogas en los jóvenes: un estudio sobre las teen series españolas recientes (2015-2021)

Alcohol, tobacco and drug use among young people: a study of recent Spanish teen series (2015-2021)

 Patricia Palomares-Sánchez
patriciapsanchez15@gmail.com
Universidad de Alicante

 Tatiana Hidalgo-Marí
tatiana.hidalgo@ua.es
Universidad de Alicante

 Jesús Segarra-Saavedra
jesus.segarra@ua.es
Universidad de Alicante

RESUMEN

El presente trabajo analiza la representación de los adolescentes y el consumo de alcohol, tabaco y otras drogas en las *teen series* españolas recientes (2015-2021) creadas para la televisión convencional y las plataformas de *video on demand* (VOD). Resulta esencial estudiar el discurso que la ficción televisiva ofrece sobre el consumo de sustancias en la adolescencia para conocer los modelos identitarios y las representaciones sociales que se presentan, puesto que las personas jóvenes suelen recurrir a las series de televisión como fuente de información sobre temas en los que son más inexpertos, como pueden ser algunas cuestiones y problemáticas específicas en torno al consumo de sustancias psicoactivas. Para ello, se ha trabajado sobre una metodología mixta, que conjuga datos cuantitativos con cualitativos desde un enfoque socio-semiótico, que analiza desde dos escenarios distintos un total de seis ficciones: tres creadas para la televisión generalista y tres para las plataformas *in streaming*, con el objetivo de identificar las tendencias de representación más recientes a partir de la llegada de las distribuidoras de contenido televisivo digital de pago a España en 2015. El estudio evidencia que, a pesar de que la ficción juvenil incorpora narrativas que abordan las posibles consecuencias negativas del consumo de sustancias, las *teen series* españolas reflejan un consumo extendido y normalizado entre los adolescentes que se asocia a la diversión y a la socialización, en lugar de apostar por relatos críticos y reflexivos.

PALABRAS CLAVE

Ficción *online*; ficción televisiva; representación juvenil; series de televisión; televisión española

FINANCIACIÓN

Artículo realizado en el marco del proyecto de investigación emergente del Vicerrectorado de Investigación y Transferencia de Conocimiento de la Universidad de Alicante, "Ficción *online* a la carta: Producción, contenido e interacción en las series españolas de las plataformas de TV *in streaming* (2016-2019)" (GRE 19-10).

Edita: Silvia Guillamón

Recibido: 15/11/2021 | Aceptado: 25/02/2022

Cómo citar este artículo:

Palomares-Sánchez, P.; Hidalgo-Marí, T. y Segarra-Saavedra, J. (2022). El consumo de alcohol, tabaco y drogas en los jóvenes: un estudio sobre las teen series españolas recientes (2015-2021). *Dígitos. Revista de Comunicación Digital*, 8: 231-250. DOI: 10.7203/rd.v1i8.227

ABSTRACT

This paper analyses the representation of teenagers and the consumption of alcohol, tobacco and other drugs in recent Spanish teen series (2015-2021) created for conventional television and video on demand (VOD) platforms. It is essential to study the discourse that television fiction offers on the consumption of substances in adolescence in order to understand the identity models and social representations that are presented, since young people often turn to television series as a source of information on topics in which they are more inexperienced, such as some specific issues and problems related to the consumption of psychoactive substances. For this purpose, we have worked on a mixed methodology, which combines quantitative and qualitative data from a socio-semiotic approach, analysing a total of six fictions from two different scenarios: three created for general television and three for streaming platforms, with the aim of identifying the most recent trends in representation since the arrival of digital pay TV content distributors in Spain in 2015. The study shows that, despite the fact that youth fiction incorporates narratives that address the possible negative consequences of substance use, Spanish teen series reflect a widespread and normalised consumption among adolescents that is associated with fun and socialisation, instead of focusing on critical and reflective narratives.

KEYWORDS

Online fiction; television fiction; youth representation; television series; Spanish television



El consumo de alcohol, tabaco y drogas en los jóvenes: un estudio sobre las teen series españolas recientes (2015-2021)

1. Introducción

La ficción televisiva debe ser entendida como un retrato de la sociedad que adquiere un valor significativo en la construcción del conocimiento y la comprensión de la realidad debido al gran impacto de su discurso. Su capacidad como agente socializador ejerce un rol determinante en la configuración de la identidad colectiva, puesto que constituye un dispositivo cultural que presenta modelos de identificación que reproducen valores, creencias, comportamientos y actitudes que son adoptados por una audiencia amplia y heterogénea (Galán, 2007; Tous-Rovirosa, 2015).

En este sentido, la ficción incorpora un conjunto de símbolos y valores con los que las personas receptoras se vinculan en la configuración de su propia identidad, llegando a contemplarlos como espacios simbólicos de gran riqueza a través de los que construir numerosos significados que pueden emplear en sus experiencias personales (Pindado, 2006). No obstante, la identidad no es algo único e inmutable, sino que se mantiene en continua evolución y hace uso del potencial de los discursos televisivos como recurso idóneo para construirse, deconstruirse y modificarse (Barker, 2003).

En los casos en los que la identidad se encuentra en un periodo de formación, como es la adolescencia, el efecto socializador de la ficción incide en mayor medida. Es durante

esta etapa de transición entre la niñez y la edad adulta cuando, como individuos, vivimos una transformación física, psicológica y social que constituye un proceso de maduración y construcción de la personalidad (Luzón, Figueras y Jiménez, 2009). Tal y como afirma Gabelas:

La adolescencia es el período narcisista por excelencia, el niño deja de serlo y en la entrada al adulto precisa una buena imagen que le proporcione seguridad, confianza. (...) La televisión agiganta y multiplica las dimensiones de este espejo, poblando de reflejos y espejismos el universo icónico juvenil en sus modelos y paraísos publicitarios (Gabelas, 2005: 143).

Los adolescentes son espectadores muy susceptibles de ser influenciados por los contenidos vehiculados por los medios, tanto por su necesidad de exploración y aprendizaje de la realidad como por su capacidad para comprender e interpretar aquello que consumen (Fedele y García-Muñoz, 2010). Mediante las series televisivas, los jóvenes pueden identificarse con los personajes (Montero, 2006; Igartua, 2010), adquirir una percepción de similitud con estos (Hoffner & Buchanan, 2005), comparar y reflexionar acerca de las situaciones representadas y los comportamientos de los personajes con los propios (Von Feilitzen, 2014), e identificar modelos sobre cómo debería ser la adolescencia y los adolescentes (Masanet y Fedele, 2019).

Por tanto, analizar el discurso que la ficción televisiva ofrece sobre la adolescencia resulta esencial para conocer los modelos identitarios y aspiracionales que se presentan y que son susceptibles de ser puestos a prueba, modificados y contrastados por los espectadores (Livingstone, 1998).

En la adolescencia existen diversas fuentes de información que se potencian u oponen: la familia, las instituciones educativas y los grupos de pares son las más significativas. Debido a la precariedad de la esfera profesional y la incorporación de la mujer al mercado laboral, a medida que las instituciones tradicionales han perdido peso en la carga educativa, la construcción de la identidad se ha desplazado hacia otras actividades que se desarrollan en el tiempo libre (Luzón, Figueras y Jiménez, 2009). En dicha dirección, las series adquieren una importancia significativa al constituirse como uno de los contenidos televisivos favoritos de los adolescentes, tanto a través de la televisión como desde el ordenador (Funes, 2008; López y Gómez, 2012; Fedele, 2021).

Estos productos audiovisuales ofrecen una transcripción simbólica de la realidad capaz de complementar, potenciar o anular la influencia de otros agentes sociales. Durante la etapa adolescente, caracterizada por la construcción de la identidad y la pertenencia del individuo, aumenta la posibilidad de generar relaciones parasociales con los personajes de ficción, de forma que estos puedan llegar a ser percibidos como parte de su colectivo familiar, amistoso o social (Livingstone, 1998). Estas relaciones merecen especial atención en el caso de las *teen series*, puesto que están dirigidas específicamente a un *target* juvenil y adquieren una mayor verosimilitud al estar protagonizadas por personas de edades similares a las de su público (Harwood, 1997; Meyer, 2009).

Siguiendo las conclusiones de Buckingham (2003), se puede afirmar que los medios de comunicación no ofrecen una ventana translúcida del mundo, pero sí configuran nuestra visión de la realidad y nos proporcionan las herramientas necesarias para interpretarla. Más allá del entretenimiento, los jóvenes pueden atribuir a los productos de ficción seriada diferentes funciones, entre las que destacan la informativa y la social (Von Feilitzen, 2004).

La función informativa hace referencia a la posibilidad de conocer y adquirir nuevos conocimientos sobre temas en los que son más inexpertos: roles sociales, relaciones interpersonales y afectivas, problemáticas específicas de esta fase clave como el consumo de alcohol o de drogas, y otras cuestiones de carácter amoroso y sexual (Fedele y García-Muñoz, 2010). Esta capacidad se ve potenciada cuanto menor sea la experiencia directa del espectador sobre un aspecto concreto de la realidad ya que, tal y como apunta Chicharro (2011: 183), la ficción televisiva “hace de instrumento de educación informal de los espectadores”. Por otro lado, la función social permite a los adolescentes incrementar las interacciones con el grupo de iguales, buscando la aceptación a través de vivencias compartidas y convirtiéndolas en un tema de conversación a través del que dialogar y aportar opiniones propias (França, 2011).

El papel atribuido a la ficción televisiva difiere de unos estudios a otros en función de la perspectiva a partir de la cual se abordan. Tradicionalmente, ha prevalecido una visión centrada en el mensaje que sitúa al receptor como un ser pasivo, concluyendo que los efectos de los medios son nocivos y ejercen un poder sustitutorio de las experiencias reales (Thompson, 1998), tal y como afirma Bindig (2008: 5), “aunque sería ridículo pensar que los espectadores imitan exactamente lo que se retrata en los medios que les rodean, sería igualmente ingenuo creer que estos mensajes no tienen trascendencia”. La ficción puede ser mucho más eficaz a la hora de ejercer influencia sobre opiniones y actitudes que los contenidos informativos o publicitarios, donde la audiencia es más consciente de la posible manipulación o persuasión existente (Guarinos, 2009). No obstante, las investigaciones más recientes que contemplan una perspectiva sociocultural y constructivista apuntan que los adolescentes son telespectadores activos y con capacidad crítica para incorporar la información televisiva de forma complementaria a las experiencias reales (Pindado, 2006; Medrano, Aierbe y Orejudo, 2009; Luzón, Figueras y Jiménez, 2009; Masanet y Buckingham, 2015).

De este modo, se puede afirmar que los mensajes televisivos no son unívocos ni cerrados, puesto que existe una implicación por parte de las personas jóvenes en la producción de significado que puede ser guiada por distintos mediadores con el fin de reconstruirlo (Rodríguez, Medrano, Aierbe y Martínez, 2013). En este sentido, parece haber consenso en la necesidad de enseñar a leer, analizar y decodificar críticamente los mensajes que contienen las series de televisión para contrarrestar el impacto negativo que puedan tener en el desarrollo de la identidad adolescente (Carducci y Rhoads, 2005; Gabelas, 2005; Medrano y Cortés, 2007). Por tanto, se puede afirmar que quien realmente tiene la capacidad de entender, decidir e interpretar de forma activa y rutinaria si la influencia de la ficción es positiva o negativa es el propio espectador y su relación con el entorno.

1.1. Antecedentes del estudio de las *teen series*

El estudio de la construcción social de los adolescentes en el seno de los trabajos sobre ficción televisiva ha suscitado un notable interés en el ámbito académico. Las investigaciones sobre esta temática se inician en Estados Unidos y Reino Unido a partir de la década de los ochenta, momento en el que los adolescentes emergen como audiencia potencial del género (Luzón, Figueras y Jiménez, 2009) y se establecen conclusiones acerca de la identificación de estos con los personajes y las tramas (Lemish, 1985; Rubin, 1985; Allen, 1989; Arnett, 1995).

A partir de los noventa, autores estadounidenses defienden estudios que confirman el uso instrumental de los medios como fuente de aprendizaje en la configuración de la identidad juvenil (Barker y Andre, 1996) y se incrementan las investigaciones de las *teen series* desde la metodología del estudio de caso, tanto en el ámbito norteamericano, con el análisis de ficciones como *Sensación de vivir* (FOX, 1990-2000) de Povlsen (1996); Dawson Crece (WB, 1998-2003) de Meyer (2003); o *Smallville* (WB, 2001-2011) de Banks (2004); como en países de habla no inglesa, con casos de estudio como la comedia francesa *Hélène et les garçons* (TF1, 1992-1994) de Pasquier (1998).

En el contexto español, el estudio de las *teen series* como subgénero de la ficción no había suscitado tanto interés como en el mundo anglosajón, a pesar de que algunos trabajos que han analizado la representación social en la ficción televisiva han contemplado la figura de los adolescentes (Pindado, 2006; Guarinos, 2009; Lacalle, 2013). Los estudios más relevantes en lo que a las *teen series* se refiere residen en los análisis de caso, como ocurriría a finales de los noventa en la investigación de Estados Unidos y Reino Unido. En concreto, se han realizado estudios sobre ficciones juveniles paradigmáticas como *Compañeros* (França, 2001), *El Internado* (Guarinos, Gordillo, Ramírez y Hermida, 2010), *Física o Química* (Masanet, Medina y Ferrés, 2012; Falcón y Díaz-Aguado, 2013), *Los protegidos*, *El Barco* (Masanet y Fedele, 2019), *SKAM España* (Sabina, de la Fuente y Martínez, 2019) o *Merlí* (Mateos-Pérez, 2021), entre otros. Recientemente, un trabajo de Lacalle, Gómez e Hidalgo (2021) ha puesto sobre la mesa el fenómeno de la *teen serie* española con un enfoque sociosemiótico que aporta una radiografía del estado del subgénero en el panorama español.

Además, se constata el interés por las investigaciones sobre los personajes jóvenes desde una perspectiva de género, incidiendo en la representación de los estereotipos (Guarinos, 2009), y el tratamiento de la sexualidad (García-Manso, 2013; Figueras, Tortajada y Araüna, 2014; Lacalle y Castro, 2017). Por otro lado, también se observan estudios enfocados a la investigación de los discursos sociales y educativos sobre la juventud presentes en la ficción española (Belmonte, 2018), así como al análisis del perfil y las preferencias de consumo televisivo juvenil (Medrano, Aierbe y Orejudo, 2009; Korres y Elexpuru, 2016; Menéndez, Figueras-Maz y Núñez, 2017).

Es evidente que las series se han constituido como un espacio simbólico e ideológico significativo en la construcción de la identidad adolescente y el “sentimiento de pertenencia a una cultura con modas, hábitos, estilos de vida, valores, preocupaciones e inquietudes propios” (Montero, 2006: 18). Por ello, resulta fundamental considerar que estos productos pueden articular metáforas sobre los cambios sociales juveniles (Belmonte, 2017) y ser conductores de posibles estereotipos que perpetúan creencias sociales e imaginarios colectivos ya asentados en la sociedad, al mismo tiempo que repiten y reafirman las normas y modelos que imitan (García-Manso, 2003).

Otros estudios se han centrado en el análisis de la percepción de los jóvenes respecto a la ficción, y concluyen que existe un mayor interés por parte del género femenino, pese a que ambos los siguen con continuidad (Medrano y Cortés, 2007). Además, Menéndez, Figueras-Maz y Núñez (2017) han revelado que la selección de títulos por parte de los adolescentes refleja un consumo estereotipado y con roles de género muy marcados, puesto que los chicos tienden a elegir series de acción o ciencia ficción, mientras que las chicas prefieren las de corte romántico y/o familiar. Por otro lado, Alberó (2006) constata un prominente interés por las teleseries de producción nacional que integran tramas de protagonismo adolescente, tales como *Los Serrano* (Telecinco, 2003-2008)

o *Cuéntame cómo pasó* (TVE, 2001-). Respecto a los personajes, se concluye que los adolescentes tienden a identificarse con aquellos que tienen su misma edad y género (Montero, 2006), aunque diversos autores manifiestan que tanto chicos como chicas prefieren a los protagonistas masculinos (Medrano, Aierbe y Orejudo, 2009; Bermejo, 2012).

2. Objetivos y metodología

El presente trabajo pretende realizar un estudio sobre la representación que se realiza en las *teen series* recientes españolas (2015-2021) sobre el consumo de alcohol, tabaco y otras drogas por parte de la población joven. Teniendo en cuenta los distintos canales de consumo de ficción audiovisual a los que tiene acceso la juventud, se han seleccionado productos de ficción de dos medios diferenciados, como son la televisión convencional y las plataformas de vídeo bajo demanda, con el objetivo de poder abordar diferencias y similitudes en lo que al tratamiento sobre el consumo de sustancias psicoactivas por parte de los jóvenes se refiere.

De este objetivo general surgen nuevos objetivos específicos que pretenden:

- OE1. Analizar el tratamiento del consumo de alcohol de los jóvenes en las *teen series*.
- OE2. Analizar el tratamiento del consumo de tabaco de los jóvenes en las *teen series*.
- OE3. Analizar el tratamiento del consumo de otras drogas y/o adicciones de los jóvenes en las *teen series*.
- OE4. Establecer diferencias y similitudes en la representación del consumo de sustancias en las series de la televisión convencional y en las series creadas para el VOD.

Para la realización de la investigación se ha empleado una metodología cualitativa, que se apoya puntualmente en datos descriptivos cuantitativos. Esta metodología se ha desarrollado, en primer lugar, seleccionando los personajes adolescentes que aparecen representados en las ficciones televisivas que conforman la muestra, entendidos como personas que se encuentran en una franja de edad comprendida entre los 14 y los 19 años¹, esto es, la etapa de adolescencia mediana (14-16 años) y tardía (17-19 años), según la segmentación establecida por UNICEF.

La muestra comprende una selección de seis series de ficción, tres de ellas emitidas originalmente en cadenas españolas en abierto y tres creadas para plataformas de vídeo bajo demanda, con el objetivo de poder ofrecer una comparación sólida sobre los modelos adolescentes representados en sendos medios.

Se han descartado, por tanto, ficciones creadas exclusivamente para el medio digital de cadenas convencionales, como es el caso de la *webserie Si fueras tú* (Playz, 2017). También aquellas cuyos protagonistas preadolescentes no se encuentran en la franja de edad señalada (14-19 años), como la serie de corte infantil-juvenil *Los espabilados* (Movistar+, 2021-).

Todas las series han sido emitidas a partir de 2015, criterio que se ha aplicado con

1. El presente estudio recopila a los personajes que se encuentran en una etapa de adolescencia mediana (14-16 años) y tardía (17-19 años), según la clasificación adoptada por [UNICEF](#).

el objetivo de identificar las tendencias más recientes a partir de la transformación del sector audiovisual con la llegada a España de Netflix, la primera distribuidora de contenido televisivo digital de pago. Se ha considerado este año como el punto de inflexión de la rápida expansión de nuevas plataformas de distribución de contenidos audiovisuales, como HBO, Amazon Prime Video o Movistar+, que han revolucionado el modelo convergente de la televisión con internet y se han configurado como un rival de la televisión generalista nacional. Las series sobre las que se ha trabajado se pueden consultar en la Tabla 1.

Medio	Muestra	Cadena	Año	Temporadas analizadas	Episodios	Duración	N.º personajes analizados
TV	<i>Merlí</i>	TV3	2015-2018	3	40	60'	12
	<i>Las del hockey</i>	TV3	2019-2020	2	26	50'	13
	<i>HIT</i>	TVE	2021-	1	10	60'	9
VOD	<i>Élite</i>	Netflix	2018-	3	24	45'	15
	<i>SKAM España</i>	Movistar+	2018-2021	4	39	30'	18
	<i>El Internado: Las Cumbres</i>	Amazon Prime Video	2021-	1	8	50'	9

Tabla 1. Muestra de análisis. Fuente: Elaboración propia.

Teniendo en cuenta que algunas de las series escogidas no han finalizado su período de emisión, es conveniente señalar que el análisis realizado comprende a los personajes que han participado en las temporadas emitidas desde la creación de las producciones hasta el primer trimestre del año 2021.

Para la elaboración del trabajo se ha llevado a cabo el visionado completo de las ficciones que componen la muestra, así como la recopilación de los personajes adolescentes de las series escogidas y su categorización dentro de su contexto narrativo. Posteriormente, se ha realizado una tabla de recogida de datos (Tablas 2 y 3) en la que se detalla información acerca de los 76 personajes que comprende el corpus de análisis.

Dado que la finalidad del presente estudio reside en ofrecer una visión global de la representación y el tratamiento narrativo sobre el consumo de sustancias psicoactivas (alcohol, tabaco y otras drogas) por parte de los jóvenes, se han establecido tres posibles etiquetas para cada una de las variables específicas (consumo social, habitual o no consumo) con el fin de determinar de forma más concreta la frecuencia y el modo en el que los adolescentes se representan consumiendo. De este modo, se ha analizado si se trata de un consumidor social, es decir, no consume en solitario, pero sí lo hace de manera esporádica cuando se encuentra con compañeros y/o amigos; de un consumidor habitual, entendido como alguien que consume cotidianamente, sea en soledad o en grupo; o de un personaje que no se representa consumiendo a lo largo de la narración. Finalmente, resulta necesario señalar que a pesar de que encontramos personajes en los que se infiere una adicción a alguna de las sustancias mencionadas, se ha prescindido de esta etiqueta debido a la dificultad para identificarla en ausencia

de un diagnóstico profesional que lo determine.

Información general sobre el personaje	
Variable	Etiqueta
Género	Masculino, femenino u otro.
Tipo de personaje	Principal, secundario o recurrente.
Protagonismo	Individual, en pareja o coral.
Procedencia	País de origen.
Religión	Católico, musulmán, judío, ateo u otro.
Rendimiento académico	Malo, regular, bueno o excelente.
Clase social	Nivel socioeconómico del personaje: baja, media o alta.

* En todas las variables se ha contemplado la etiqueta N/S para hacer referencia a aquellos casos en los que no hay datos o se ha planteado otra circunstancia ajena a los parámetros de la investigación.

Tabla 2. Variables generales de análisis. Fuente: elaboración propia.

Información específica sobre el personaje	
Variable	Etiqueta
Alcohol	Indica si el personaje es bebedor en un entorno social, es consumidor habitual o no consume alcohol.
Tabaco	Indica si el personaje es fumador en un entorno social, es consumidor habitual o no consume tabaco.
Drogas	Indica si el personaje se droga en un entorno social, es consumidor habitual o no consume drogas.
Otras adicciones	Variable abierta

Tabla 3. Variables específicas de análisis. Fuente: elaboración propia.

3. Resultados

3.1. Rasgos descriptivos de los jóvenes de las *teen series*

Los datos descriptivos acerca de los 76 personajes que conforman la selección muestral permiten detectar una presencia global equilibrada entre el género masculino (54%, 41 casos) y el femenino (46%, 35 casos). En relación con el tipo de personaje, se observa que 50 (60%) de los adolescentes representados cuentan con un papel principal en la trama, 20 (26%) son personajes secundarios, y 6 (8%) aparecen de forma recurrente. Además, se detecta el predominio de un reparto coral en todas las *teen series* analizadas, ya que el 95% (72 casos) de los personajes cuenta con una relevancia similar en la narración, a excepción de Bruno y Pol en *Merlí* y Amaia y Paul en *El Internado: Las Cumbres* que obtienen un protagonismo individual.

En cuanto a la procedencia de los personajes, se observa que España es el país predominante con un 79% (51 casos), seguido por una representación residual de Marruecos (7%, 3 casos), Palestina (3%, 2 casos) y Francia (3%, 2 casos). Por otro lado, si se analiza la variable que identifica la religión, se detecta que 8 (10%) jóvenes manifiestan ser practicantes del islam, seguidos por los 6 (8%) que mencionan ser católicos y los 2 (3%) que son ateos. Por último, destaca una notable mayoría de 60 (79%) personajes

que no manifiestan en la narración sus creencias religiosas, en caso de que las tengan.

Según los datos obtenidos acerca de la clase social, se detecta que la clase media es la más representada con un 49% (37 casos), seguida por la clase alta con un 24% (18 casos) y la baja con un 18% (14 casos). Con respecto al rendimiento académico, destacan los jóvenes con un buen nivel (37%, 28 casos), regular (21%, 16 casos) y malo (20%, 15 casos), otorgando una escasa representación a aquellos cuyo rendimiento es excelente (13%, 10 casos).

Por otro lado, si comentamos las variables específicas de análisis que hacen referencia al consumo sustancias de los personajes que conforman la muestra nos ha permitido identificar que, en términos generales, el alcohol es la sustancia más representativa al ser detectada en 71 de los casos, seguida por el tabaco en 59 casos, y otras drogas, entre las que destacan el cannabis o la cocaína, también detectadas en 59 casos (ver gráfico 1). Del total de casos analizados, se observa que únicamente 4 personajes (5%) no consumen ninguna de las sustancias analizadas, 10 consumen solo una de ellas (13%) y 62 (82%) se representan consumiendo al menos dos.

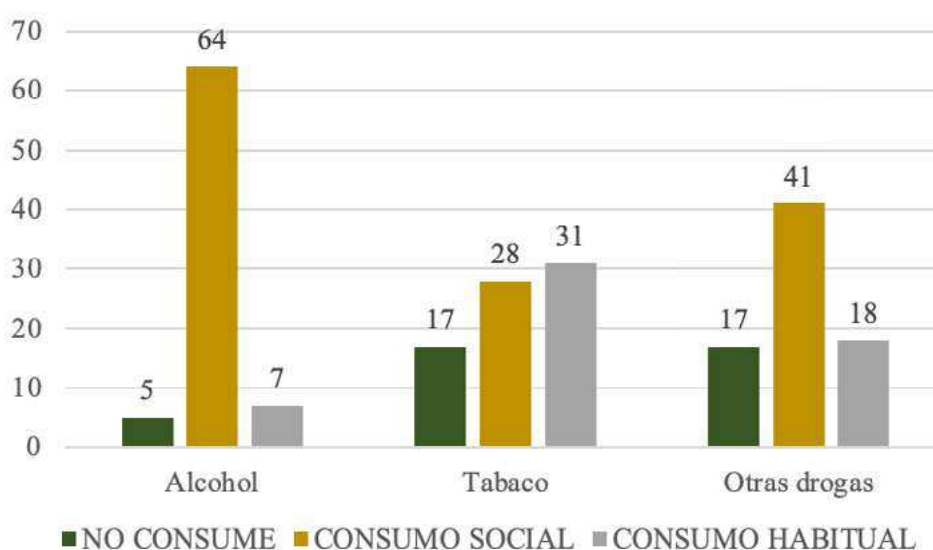


Gráfico 1. Consumo de sustancias por parte de los jóvenes. Fuente: Elaboración propia

3.2. El consumo de alcohol en los jóvenes de las *teen series*

En cuanto a las distintas formas de consumo establecidas, se observa que los jóvenes que consumen alcohol lo hacen principalmente en un entorno social (84%, 64 casos). Se detecta una tendencia notable en las *teen series* españolas a representar a los adolescentes bebiendo en situaciones sociales que pueden ir desde una reunión tranquila hasta entornos festivos en los que se ingiere una cantidad abundante de alcohol mediante la práctica del botellón, tal y como se refleja en todas las ficciones analizadas. Por otra parte, podemos encontrar a 7 jóvenes (9%) que beben alcohol habitualmente, tanto en grupo como individualmente, y dejan entrever una alta posibilidad de que esta práctica les conduzca a una adicción. Si atendemos al tratamiento narrativo ofrecido acerca de los jóvenes que beben alcohol de forma habitual podemos encontrar dos perfiles diferentes. Por un lado, se observan personajes que utilizan el alcohol como medio para hacer frente a sus problemas pero que, cuando el conflicto que les había empujado a beber se resuelve, recuperan el control y frenan el consumo, como podemos ver en Raquel en *Las del hockey* o Guzmán en *Élite*. Por otro lado, encontramos casos

en los que el consumo de alcohol es tan elevado que interfiere de forma significativa en el desarrollo de diferentes aspectos de las vidas de los personajes, tal y como se observa en Óscar en *Las del hockey* o Valerio en *Élite*, quienes se muestran incapaces de controlar la cantidad de alcohol que ingieren.

Resulta relevante poner en relación la variable que analiza el consumo de alcohol por parte de los jóvenes y la información obtenida en las variables generales con el fin de identificar las tendencias de representación y profundizar en el tratamiento narrativo ofrecido por las *teen series* españolas. De este modo, si analizamos el tipo de personaje observamos que del total de jóvenes que consumen alcohol (93%, 71 casos), ya sea de forma social o habitual, el 66% (47 casos) cuenta con un papel principal en la trama, el 27% (19 casos) son personajes secundarios y el 7% (5 casos) aparece de forma recurrente. En cuanto al protagonismo, se observa el predominio de un reparto coral (94%, 67 casos), algo lógico si tenemos en cuenta que en las ficciones estudiadas destaca la presencia de personajes principales con una relevancia similar. No obstante, es evidente el vínculo existente entre el elevado consumo de alcohol por parte de los jóvenes, quienes beben principalmente acompañados en un entorno social, y el retrato coral, es decir, grupal, que las *teen series* ofrecen sobre los adolescentes.

Por otra parte, observamos que el perfil del consumidor de alcohol se detecta de forma ligeramente superior en el género masculino (58%, 41 casos) que en el femenino (42%, 30 casos). En términos relativos, cabe destacar que el 100% de los hombres representados beben alcohol, frente al 86% (30/35 casos) de las mujeres. A pesar de que las diferencias no son demasiado significativas, dado que el relato no incide en este aspecto, destaca el hecho de que los escasos 5 personajes que no consumen alcohol sean mujeres. Un ejemplo de ello lo encontramos en Amira y Dounia en *SKAM España*, quienes aclaran no beber por respeto a sus creencias religiosas, y en Inés, Adele y Rita en *El Internado: Las Cumbres*, cuyo consumo de alcohol en la narración es inexistente. Si atendemos a los datos obtenidos sobre la procedencia, se observa que de los personajes que consumen alcohol, el 80% (57 casos) procede de España, lo que supone un 97% (57/59 casos) de los jóvenes españoles. Por su parte, los personajes procedentes de otros países, representan un 20% (14 casos) del total de consumidores, y un 82% (14/17 casos) entre los jóvenes de origen extranjero representados.

No parece que exista una relación directa que ponga de relieve el rendimiento académico con el consumo de alcohol, puesto que se detecta un equilibrio entre los jóvenes buenos o excelentes estudiantes que beben alcohol (48%, 34 casos) y los que tienen un perfil académico regular o malo (42%, 30 casos). En cuanto a la clase social, se observa un mayor consumo entre los jóvenes con un nivel socioeconómico medio con un 49% (35 casos), seguido por el alto y el bajo con un 23% (16 casos) y un 20% (14 casos), respectivamente. Dado que los datos obtenidos resultan acordes al número de personajes representados en las diferentes clases sociales, se ha considerado relevante profundizar en el grado de consumo que presentan con el fin de obtener resultados más concretos. De este modo, se observa que los personajes de clase media que consumen alcohol son principalmente bebedores sociales (94%, 33 casos), aunque también se detectan 2 casos (6%) cuyo consumo es habitual, como vemos en Erika en *HIT* y Óscar en *Las del hockey*. Por su parte, en la clase alta encontramos 11 (69%) personajes que consumen en un entorno social y 5 (31%) que lo hacen habitualmente, mientras que todos los personajes de clase baja son bebedores sociales (100%, 14 casos).

Si realizamos una comparación entre la representación del consumo de alcohol en las series de la televisión convencional y en las series creadas para el VOD, se detecta que en las producciones para la televisión generalista todos los personajes beben, bien sea de forma social (91%, 31 casos) o habitual (9%, 3 casos). En cambio, en las plataformas *in streaming* se observa a 5 (12%) adolescentes que no consumen alcohol, aunque lo cierto es que la mayoría sí lo hace, ya sea como bebedores sociales (79%, 34 casos) o habituales (9%, 4 casos). A pesar de que en ambos medios se representa un elevado número de personajes que beben alcohol, las ficciones del VOD contemplan una mayor normalización del no consumo, algo inexistente en la televisión generalista (Gráfico 2).

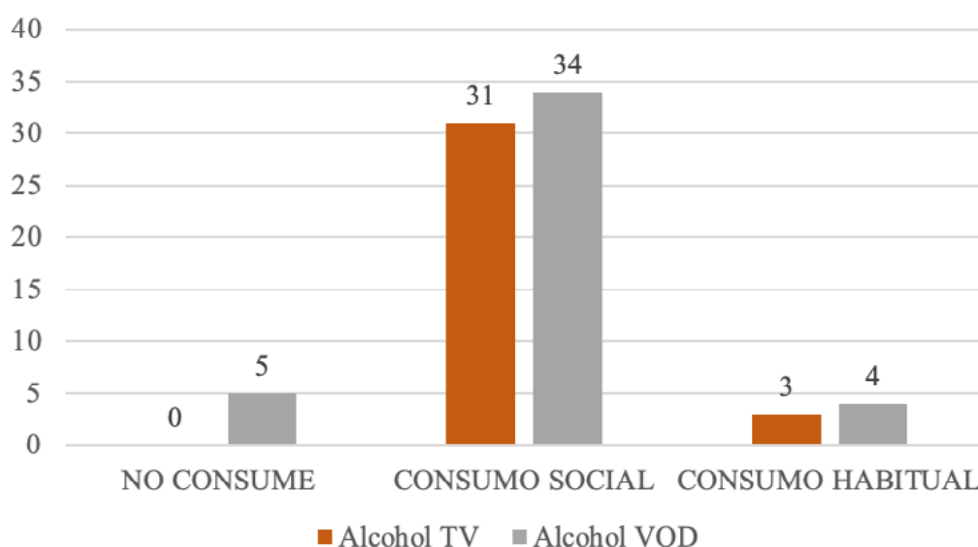


Gráfico 2. Comparativa de la representación del consumo de alcohol en la televisión convencional y las plataformas de VOD. Fuente: Elaboración propia

Si bien es cierto que los datos apuntan a un consumo levemente menor en las *teen series* para el VOD, el tratamiento narrativo que se presenta dista mucho entre un medio u otro. En las series de la televisión generalista *Merlí* y *Las del hockey* se representa a los jóvenes bebiendo alcohol de forma puntual en reuniones sociales o fiestas, y aunque en *HIT* se muestra un consumo más elevado, todas las narrativas inciden en algún momento de la serie en las consecuencias negativas que puede producir un consumo abusivo: Joan en *Merlí* sufre un coma etílico, Óscar en *Las del hockey* atropella a una persona por conducir ebrio, y en *HIT* es el propio adulto protagonista el que experimenta una recaída en el alcohol que sirve de ejemplo para el grupo de adolescentes. Por su parte, las ficciones de las plataformas presentan diversos escenarios: mientras que en *SKAM España* se muestra un consumo moderado asociado al ocio y en *El Internado: Las Cumbres* el alcohol no tiene apenas protagonismo, en *Élite* los jóvenes aparecen bebiendo alcohol en todos los capítulos de forma desenfrenada, sin ningún tipo de reflexión o crítica al respecto.

Por todo lo expuesto, podemos decir que los adolescentes representados en las *teen series* españolas reflejan un consumo de alcohol extendido y normalizado. Teniendo en cuenta que el alcohol es, con diferencia, la sustancia psicoactiva más consumida entre los adolescentes de 14 a 18 años en España (OEDA, 2020), se constata la capacidad de la ficción televisiva para configurarse como un retrato de la sociedad en la que se encuentra inmersa.

3.3. El consumo de tabaco en los personajes de las *teen series*

En cuanto al consumo de tabaco, se observa que el 41% (31 casos) de los personajes son fumadores habituales, el 37% (28 casos) son fumadores sociales, y el 22% (17 casos) no se muestra fumando tabaco en ninguna ocasión. Algunos ejemplos los podemos encontrar en los personajes de Cris y Dani en *SKAM España* o Andrés y Marga en *HIT* (fumadores habituales), Emma y Raquel en *Las del hockey* (fumadoras sociales); y Mònica e Iván en *Merlí* o Eric, Paul y Julio en *El Internado: Las Cumbres* (no fumadores).

Si relacionamos el consumo de tabaco con el tipo de personaje observamos que, al igual que sucede con el alcohol, son los personajes principales los que más se relacionan con el tabaco con un 68% (38 casos). En relación con el protagonismo, el 95% (56 casos) de los personajes que consume tabaco forma parte de un reparto coral, frente al 5% (3 casos) de protagonistas individuales, tal y como vemos en Amaia en *El Internado: Las Cumbres* y Pol y Bruno en *Merlí*.

En lo que al perfil psicosocial del personaje que consume tabaco se refiere, detectamos que el 44% (26 casos) de ellos son mujeres y el 56% (33 casos) son hombres. De entre las mujeres representadas en las ficciones analizadas, se detecta que el 74% (26/35 casos) son fumadoras, principalmente de forma social (16 casos), aunque también encontramos personajes que fuman habitualmente (10 casos). Por su parte, de todos los hombres que componen la muestra, los que fuman tabaco suponen el 80% (33/41 casos), de los cuales el 36% (12 casos) lo hace en un entorno social y el 64% (21 casos) de forma habitual. En este sentido, cabe señalar que los resultados obtenidos no coinciden con los datos estadísticos del Observatorio Español de las Drogas y Adicciones (OEDA, 2020) para el año 2020 en España, según los cuales el consumo de tabaco es más frecuente en el grupo femenino que en el masculino, a pesar de que la diferencia se ha reducido progresivamente en los últimos años.

En relación con la procedencia, se observa un consumo de tabaco mayor por parte de los jóvenes españoles (82%, 49 casos) que de los personajes de origen extranjero (18%, 11 casos), lo que implica que el 83% de los jóvenes españoles que aparecen en las *teen series* fuma, en contraste con el 65% de los personajes procedentes de otros países. Por otra parte, no parece que la religión adquiriera un peso significativo en este aspecto, puesto que entre los jóvenes que manifiestan ser católicos encontramos casos en los que no son fumadores, como Mònica en *Merlí* o Rita en *El Internado: Las Cumbres*, y casos en los que sí fuman, como Lucrecia en *Élite* o Janina en *Las del hockey*. Del mismo modo, algunos personajes musulmanes como Nadia o Malick en *Élite* son no fumadores, y otros como Nourdin en *HIT* o Kassim en *SKAM España* son fumadores habituales.

En cuanto al rendimiento académico de los personajes fumadores, se detecta que el 45% (25 casos) responde al perfil de buen estudiante y el 50% (28 casos) al de estudiante con un nivel inferior. Si profundizamos en el análisis, destaca el hecho de que de los 31 jóvenes etiquetados como mal o regular estudiante, 28 consumen tabaco (90%), mientras que los buenos o excelentes estudiantes solo fuman en el 64% de los casos (25/39 casos). Por otra parte, si atendemos a la clase social se observa que la clase media es la que más consume tabaco con un 56% (33 casos), seguida por la clase baja con un 20% (12 casos) y la clase alta con un 17% (10 casos). Teniendo en cuenta que la muestra está compuesta por 18 personajes de clase alta y 14 de clase baja, destaca que el porcentaje de jóvenes fumadores con un nivel socioeconómico bajo sea mayor.

A pesar de que el consumo de tabaco se representa como una práctica normalizada

que pasa desapercibida, la acción de fumar se asocia a los personajes que tienden a ocupar una posición superior en el grupo de pares, caracterizados por ser los rebeldes y “malotes” de la serie. Un ejemplo de ello se observa en Amaia y Paz en *El Internado: Las Cumbres*, Pol y Berta en *Merlí*, Rebeka, Omar y Valerio en *Élite* o cualquiera de los personajes de *HIT*. Se trata de una paradoja de la ficción televisiva juvenil en la que el tabaco se representa como un símbolo de irreverencia y empoderamiento que contribuye a normalizar la figura del fumador, en contraste con la mala imagen que tiene actualmente el tabaco en la sociedad.

En cuanto a la presencia del tabaco entre los jóvenes de las ficciones para la televisión convencional y las plataformas *in streaming*, se detecta un menor consumo en las series para el *VOD*, donde encontramos la ausencia de personajes que fumen tabaco en 13 casos (31%), en contraste con los 4 casos (12%) de la ficción generalista. En este sentido, podemos afirmar que las series para el *VOD* apuestan por una representación más realista, acorde a los datos obtenidos por el OEDA (2020)², al mostrar una cantidad menor de jóvenes que no fuman, frente a la escasez de personajes que no consumen tabaco de la televisión generalista (ver gráfico 3). En relación con el consumo social y el habitual la diferencia detectada es mínima, ya que en ambos medios se observa un consumo ampliamente extendido.

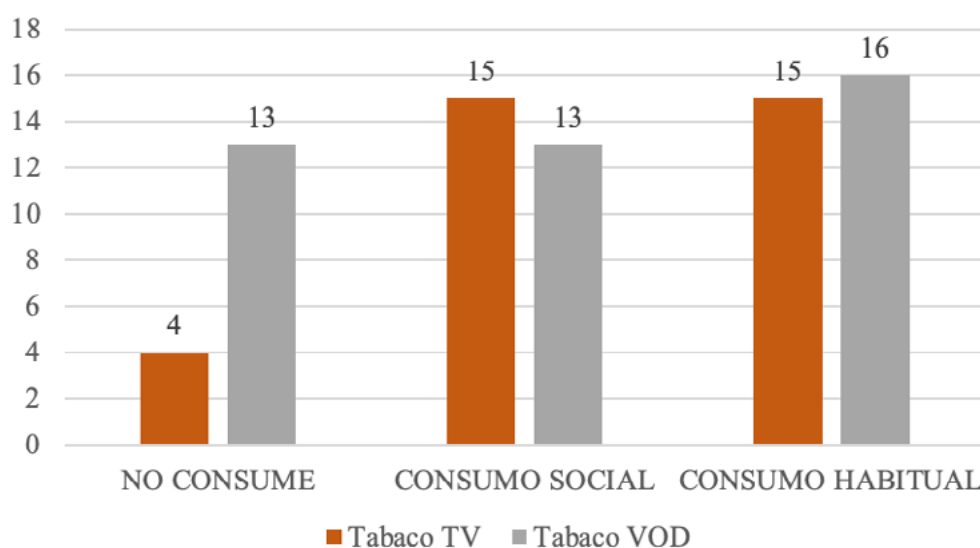


Gráfico 3. Comparativa de la representación del consumo de tabaco en la televisión convencional y las plataformas de *VOD*. Fuente: Elaboración propia

3.4. Consumo de otras drogas y otras adicciones

En cuanto al consumo de otras sustancias psicoactivas, entre las que destaca el cannabis como la sustancia con mayor prevalencia, se ha detectado el consumo en un entorno social en un 57% (43 casos) de los personajes, seguido por un 21% (16 casos) que hacen uso de forma habitual, y 17 que no consumen (22%). Teniendo en cuenta el protagonismo que tienen las drogas en *Élite*, se puede encontrar una representación de ambas formas de consumo entre sus personajes: un ejemplo de ello se observa en Samuel y Lucrecia, quienes no se drogan con frecuencia, pero sí en entornos festivos (consumidores sociales), o Ander, Omar, Valerio y Carla (consumidores habituales), quienes se representan fumando cannabis a diario y, en el caso de Valerio y Carla, consumiendo cocaína sin control en múltiples ocasiones. Por otro lado, resulta relevante el caso de Inés en *El Internado: Las Cumbres*, dado que sufre una adicción a las drogas

que no es voluntaria, si no que está provocada por la fuerte medicación que su padre le obliga a tomar desde hace años y que le lleva incluso a padecer amnesia.

En términos generales el consumo de drogas en la ficción tiende a reflejarse como una práctica común, sobre todo en relación con el cannabis, que se representa como una sustancia de bajo riesgo y de fácil disponibilidad, algo que se corresponde con los datos del OEDA (2020), según los cuales el cannabis es la sustancia psicoactiva ilegal con mayor prevalencia entre los jóvenes de 14 a 18 años en España. Otra de las drogas que cuenta con una presencia significativa en las *teen series* españolas es la cocaína, cuyo consumo tiende a representarse con normalidad asociado a entornos festivos y el descontrol, tal y como se observa en *Élite*, donde el consumo de cocaína está presente de forma evidente en la mayoría de los encuentros.

No obstante, debemos resaltar que la ficción televisiva también contempla tramas que tratan de incidir en las consecuencias negativas que se pueden derivar del consumo de drogas, tales como los brotes psicóticos (Gerard en *Merlí*), la dificultad para tomar consciencia bajo los efectos de las drogas (Carla en *Élite*) o la recurrencia a estas para aliviar el malestar (Erika en *HIT*). Además, en este último caso, el abuso que hace Erika de las drogas adquiere una relevancia significativa en la narración, ya que se representa como el conflicto central del personaje y se constituye como su característica definitoria: Erika es una joven adicta al alcohol y a las drogas que comienza a consumir para tratar de ocultar su dolor tras la muerte de su madre por sobredosis. El negocio de su padre, una farmacia, le permite acceder a determinadas sustancias y robarlas, tanto para venderlas a terceros como para su propio consumo. En este contexto, *HIT* apuesta por mostrar a lo largo de la narración los efectos que tienen las drogas en Erika, quien adopta una actitud pasiva frente a todo lo que le rodea, se presenta en clase colocada diariamente y llega a sufrir un desmayo como consecuencia de un fuerte abuso de sustancias durante una fiesta. Finalmente, la ficción plasma la necesidad de tomar consciencia del problema, contar con el apoyo del entorno más cercano y buscar ayuda profesional.

Por otro lado, si atendemos al género de los personajes que consumen otras drogas diferentes al alcohol y el tabaco no detectamos diferencias destacables, ya que aunque el 46% (27 casos) son mujeres y el 54% (32 casos) son hombres, estos casos suponen el 77% (27/35 casos) y 78% (32/41 casos) del total de mujeres y hombres representados en las ficciones analizadas. No obstante, si relacionamos el género con la forma de consumo de sustancias podemos observar que el género masculino tiene una mayor tendencia a consumir drogas de forma habitual (85%, 14 casos) que el femenino (15%, 4 casos). En contraste, el consumo social tiene mayor prevalencia entre las mujeres (55%, 23 casos), frente al 45% (18 casos) de los hombres.

En el consumo de drogas no se observan diferencias significativas en lo que al rendimiento académico se refiere, ya que los personajes que consumen drogas se representan en el mismo número de casos, ya sean buenos (50%, 27 casos) o malos estudiantes (50%, 27 casos). Sin embargo, en proporción a los casos totales detectados como jóvenes con un mal rendimiento académico, se detecta que el 87% (27/31 casos) ha consumido sustancias, mientras que los buenos estudiantes se representan consumiendo en el 67% (27/39 casos) de los casos. Además, si relacionamos el consumo de drogas con la clase social podemos decir que la clase media es la más representada con un 56% (33 casos), seguida por la clase alta con un 22% (13 casos) y la baja con un 19% (11 casos).

Si realizamos una comparativa entre las plataformas de VOD y la televisión convencional,

se detecta una mayor presencia de jóvenes que consumen sustancias en la televisión generalista, puesto que de los 34 personajes que conforman la muestra, 31 de ellos aparecen consumiendo drogas en algún momento de la narración (91%). En cambio, en las series de las plataformas el número de casos desciende a 28 (67%) de los 42 personajes analizados. A pesar de que es un porcentaje notable en ambos casos, dado que se trata de más de la mitad de los personajes, debemos destacar que en las ficciones para las plataformas su ausencia es del 14 (33%) de los casos, mientras que en la televisión convencional únicamente de 3 (9%) (ver gráfico 4). Finalmente, si atendemos al tratamiento narrativo que se lleva a cabo en sendos medios, podemos afirmar que las ficciones para la televisión generalista plantean un discurso más educativo vinculado a las consecuencias del consumo que las series para el *VOD*, donde el consumo de drogas se representa ampliamente aceptado y se asocia al ocio.

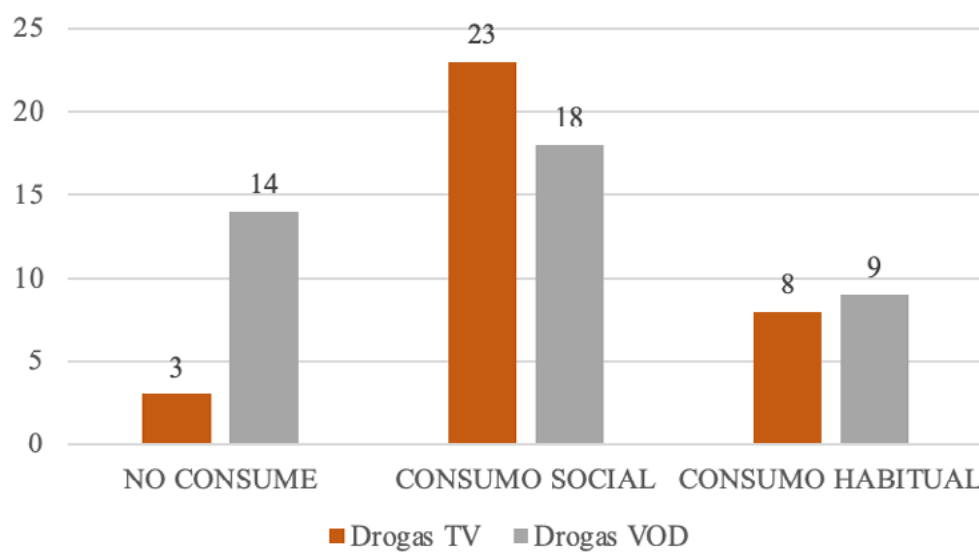


Gráfico 4. Comparativa de la representación del consumo de drogas en la televisión convencional y las plataformas de *VOD*. Fuente: Elaboración propia

Por otro lado, si se atiende a la representación de otros consumos que puedan llevar a una adicción distinta a las ya mencionadas podemos observar que en las plataformas de *VOD* no se representa a ningún personaje que merezca especial atención. Los contenidos para la televisión convencional, por su parte, reflejan un total de 3 casos (9%), en los que se abordan problemáticas relacionadas con la tecnología móvil, la obsesión por el deporte o la adicción a los juegos de apuestas *online*. Un ejemplo de ello lo encontramos en Ricard de *Las del hockey*, quien se dopa para mejorar su rendimiento físico y está obsesionado con estar al máximo para ser el mejor en el deporte que practica. Del mismo modo, su compañero Óscar, además del ya mencionado consumo habitual de alcohol, sufre un trastorno adictivo por las apuestas en internet que le llevan a perder todo su dinero e incluso a robar a su madre para saciar la necesidad incontrolable de jugar. En ambos casos las ficciones representan dichas adicciones como un problema grave que padecen los personajes y del que necesitan ser conscientes para dar el primer paso hacia la recuperación. Por otro lado, podemos encontrar una adicción distinta en Darío de *HIT*, un chico solitario que sufre una dependencia a su teléfono móvil. La serie ahonda en el aislamiento que provoca esta adicción en el personaje, quien vive constantemente pegado a su dispositivo y es incapaz de relacionarse socialmente con los demás como consecuencia del mundo interno que el mismo ha creado.

4. Discusión y conclusiones

Tras el análisis presentado, podemos decir que el consumo de alcohol, tabaco y otras sustancias psicoactivas por parte de los jóvenes en la ficción televisiva es evidente, algo lógico si tenemos en cuenta que se trata de un hábito social y cultural ampliamente extendido en nuestra sociedad. Aunque el alcohol y el tabaco son las drogas más toleradas y socialmente aceptadas en las ficciones analizadas, lo cierto es que no son las únicas que cuentan con la aprobación de los jóvenes, puesto que otras sustancias como el cannabis o la cocaína están igualmente presentes asociadas al ocio.

Por otra parte, a pesar de que el consumo de tabaco parece ser un hábito en declive en la sociedad, puesto que la tendencia en los últimos años es decreciente (Guerrero, 2021), podemos afirmar que se trata de una sustancia que sigue estando muy presente en la vida de los jóvenes representados en la ficción televisiva juvenil, lo que se corresponde con los datos del OEDA (2020, según los cuales el tabaco continúa siendo la segunda droga más consumida entre los adolescentes de entre 14 y 18 años.

Como se ha constatado en la exposición de los resultados, se evidencia un consumo extendido y normalizado de sustancias psicoactivas entre las personas adolescentes representadas en la ficción juvenil. En las *teen series* españolas prácticamente todos los jóvenes beben alcohol, una gran parte fuma tabaco, y se observa un elevado número de personajes que consume drogas con frecuencia. Se constata un hecho significativo en las narrativas y es la tendencia a reflejar un consumo de sustancias asociado a la socialización y la diversión, lo que provoca que los adolescentes tengan una baja percepción de riesgo del consumo de drogas. Si bien es cierto que se presentan distintos personajes que padecen los efectos de un consumo de sustancias abusivo, en general no se apuesta por relatos críticos que reflejen las consecuencias negativas que se pueden derivar, tanto en el propio individuo que consume habitualmente (riesgo de consumo de otras drogas y adicciones, problemas de salud mental, enfermedades respiratorias, etc.) como en su entorno cercano (conflictos familiares y de pareja, aislamiento, problemas laborales, problemas económicos, etc.).

Se han detectado diferencias notables en la representación del consumo entre las producciones para la televisión generalista y las creadas para las plataformas de VOD, ya que se observa un mayor consumo en las ficciones para la televisión generalista, donde prácticamente todos los jóvenes consumen alguna de las sustancias. A pesar de que ambos medios reflejan un elevado consumo por parte de la juventud, el VOD contempla en mayor medida la representación de personajes que no consumen alcohol, tabaco y otras drogas. Sin embargo, mientras que en las ficciones para la televisión el discurso es más moderado y plantea una perspectiva más educativa y crítica en torno a las consecuencias del consumo, las series para las plataformas son las que reflejan un consumo más descontrolado y extendido entre los jóvenes en favor de aportar credibilidad a situaciones requeridas por la narración.

En definitiva, el presente trabajo nos permite concluir que las *teen series* españolas representan un estilo de vida juvenil en el que el alcohol, el tabaco y otras drogas están presentes, probablemente, en un intento por normalizar el discurso realista en las tramas narrativas. No obstante, y a pesar de la representación de tramas próximas al discurso social imperante que trata de incidir en las problemáticas derivadas del consumo de sustancias, la presencia de relatos vinculados directamente con dicho consumo no es un tema principal en los productos analizados, si no que se aborda como una particularidad más del personaje. En este sentido, se hace evidente la ausencia de

discursos que aporten una visibilidad notable a los efectos negativos derivados de las drogas y rompan con los estigmas y la desinformación existente entre los jóvenes y la sociedad en general.

Referencias

- Albero, M. (2006): "Violencia, sexo y televisión: la mirada adolescente", en *Quaderns del CAC*, vol. 22, pp. 81-90. Disponible en Internet: <https://bit.ly/3mhiNXm> [Consulta: 10 de mayo de 2021].
- Allen, R. (1989): "Bursting bubbles: *Soap Opera*, audiences and the limits of genre". En: Seiter, E.; Borchers, H.; Kreutzner, G. y Warth, E. (Eds.), *Remote Control: Television, Audiences and Cultural Power*. London: Routledge, pp. 44-45.
- Arnett, J. (1995): "Adolescents' uses of media for self-socialization", en *Journal of Youth and Adolescence*, vol. 24, nº 5, pp. 519-533. Disponible en Internet: <https://doi.org/fjrd7k> [Consulta: 15 de mayo de 2021].
- Banks, M. (2004): "A Boy for All Planets: *Roswell*, *Smallville* and the Teen Male Melodrama". En: Davis, G. y Dickinson, G. (Eds.), *Teen TV. Genre, Consumption and Identity*. London: BFI, pp. 17-28.
- Barker, C. (2003): *Televisión, globalización e identidades culturales*. Barcelona: Paidós.
- Barker, C., & Andre, J. (1996): "Did you see? Teenage soaps talk and gendered identity", en *Young Nordic Journal of Youth Research*, vol. 4, nº 4, pp. 21-38. Disponible en Internet: <https://doi.org/fw8g74> [Consulta: 2 de junio de 2021].
- Belmonte, J. (2017): "Mutantes, *mutatis mutandis*: de metáfora transcultural e intermedial sobre la juventud a objeto para una coeducación audiovisual", en *EU-topías. Revista de interculturalidad, comunicación y estudios europeos*, vol. 14, pp. 101-120. Disponible en Internet: <https://bit.ly/3eOR4dQ> [Consulta: 3 de enero de 2022].
- Belmonte, J. (2018): "Entre educación y televisión: discursos ideológicos en la representación de la juventud", en *Dilemas Contemporáneos: Educación, Política y Valores*, vol. 5, nº 2, pp. 1-29. Disponible en Internet: <https://bit.ly/3mSeWBJ> [Consulta: 3 de enero de 2022].
- Bermejo, J. (2012): "Los personajes y las series de ficción en la vida de los y las jóvenes", en *Revista de Estudios de Juventud*, nº 96, pp. 31-49. Disponible en Internet: <https://bit.ly/2Rt9sjW> [Consulta: 15 de mayo de 2021].
- Bindig, L. (2008): *Dawson's Creek. A Critical Understanding*. Lexington Books.
- Buckingham, D. (2003): *Media Education. Literacy, Learning and Contemporary Culture*. Cambridge: Polity Press.
- Carducci, R., & Rhoads, M. (2005): "Of Minds and Media: Teaching Critical Citizenship to the Plugged-in Generation", en *About Campus*, vol. 10, nº 5, pp. 2-9. Disponible en Internet: <https://doi.org/b3vjkj> [Consulta: 15 de mayo de 2021].
- Chicharro, M. (2011): "Aprendiendo de la ficción televisiva. La recepción y los efectos socializadores de 'Amar en tiempos revueltos'", en *Comunicar*, vol. XVIII, nº 36, pp. 181-190. Disponible en Internet: <https://doi.org/b2dpdj> [Consulta: 2 de junio de 2021].
- Falcón, L., y Díaz-Aguado, M. J. (2014): "Adolescent Students as Media Fictional Characters", en *Comunicar*, vol. XXI, nº 42, pp. 147-155. Disponible en Internet: <https://doi.org/fd6q> [Consulta: 10 de mayo de 2021].
- Fedele, M. (2021): "La segunda generación de *teen series*: programas estadounidenses, británicos y españoles de los 2000- 2010", en *index.comunicación*, vol. 11, nº 1, pp. 297-

327. Disponible en Internet: <https://bit.ly/3ufDGVH> [Consulta: 10 de mayo de 2021].
- Fedele, M., y García-Muñoz, N. (2010): "El consumo adolescente de la ficción seriada". En *Vivat Academia*, nº 111, pp. 47-64. Disponible en Internet: <https://doi.org/f367> [Consulta: 10 de mayo de 2021].
- Figueras-Maz, M., Tortajada, I., y Araña, N. (2014): "La erótica del malote. Lecturas adolescentes de las series televisivas: atracción, deseo y relaciones sexuales y afectivas". En *Revista de Estudios de Juventud*, nº 106, pp. 49-61. Disponible en Internet: <https://bit.ly/3rCjLhM> [Consulta: 10 de mayo de 2021].
- França, M. E. (2001): *La contribución de las series juveniles de televisión a la formación de la identidad en la adolescencia. Análisis del contenido y de la recepción de la serie "Compañeros" de Antena 3* (Tesis doctoral). Universitat Autònoma de Barcelona, España. Disponible en Internet: <https://bit.ly/3werFBy> [Consulta: 15 de mayo de 2021].
- Funes, M. J. (2008): *Cultura, Política y Sociedad. Informe de la Juventud en España*. Gobierno de España, Ministerio de Igualdad. Disponible en Internet: <https://bit.ly/31wTzuy> [Consulta: 15 de mayo de 2021].
- Gabelas, J. A. (2005): "Televisión y adolescentes: una mítica y controvertida relación", en *Comunicar*, vol. 13, nº 25, pp. 137-146. Disponible en Internet: <https://bit.ly/3dlY1Sb> [Consulta: 15 de mayo de 2021].
- Galán, E. (2007): "Construcción de género y ficción televisiva en España", en *Comunicar*, vol. 14, nº 28, pp. 229-236. Disponible en Internet: <https://bit.ly/39Y8nqG> [Consulta: 2 de junio de 2021].
- García-Manso, A. (2013): "Series de ficción y homosexualidad en España: Un intento por visibilizar la diversidad sexual", en *Revista Iberoamericana de Salud y Ciudadanía*, vol. 2, nº 1, pp. 29-55. Disponible en Internet: <https://bit.ly/3ryFdo8> [Consulta: 15 de mayo de 2021].
- Guarinos, V. (2009): "Fenómenos televisivos «teenagers»: prototipias adolescentes en series vistas en España", en *Comunicar*, vol. 17, nº 33, pp. 203- 211. Disponible en Internet: <https://bit.ly/3walJll> [Consulta: 10 de mayo de 2021].
- Guarinos, V., Gordillo, I., Ramírez, M. D. M., Jiménez-Varea, J., y Hermida, A. (2010): "Masculinidades de ficción televisiva y retroalimentación online. Jóvenes y adolescentes en-red-ados fuera y dentro de Física o química y El Internado". En *Comunicación y desarrollo en la era digital: II Congreso Internacional de la AE-IC (20 p.)*. Asociación Española de Investigadores de la Comunicación.
- Guerrero, J. A. (2021): Cae el número de fumadores en España. *Heraldo*. Disponible en Internet: <https://bit.ly/3p7PJVw> [Consulta: 2 de junio de 2021].
- Harwood, J. (1997): "Viewing age: lifespan identity and television viewing choices", en *Journal of Broadcasting and Electronic Media*, vol. 41, nº 2, pp. 203-213. Disponible en Internet: <https://doi.org/fkprmb> [Consulta: 2 de junio de 2021].
- Hoffner, C., & Buchanan, M. (2005): "Young Adults Wishful Identification with Television Characters: The Role of Perceived Similarity and Character Attributes", en *Media Psychology*, vol. 7, nº 4, pp. 325-351. Disponible en Internet: <https://doi.org/c2dnjm> [Consulta: 15 de mayo de 2021].
- Igartua, J.J. (2010): "Identification with Characters and Narrative Persuasion through Fictional Feature Films", en *Communications: The European Journal of Communication Research*, vol. 35, nº 4, pp. 347-373. Disponible en Internet: <https://doi.org/cj939f> [Consulta: 15 de mayo de 2021].

- Korres, O., y Elexpuru, I. (2016): "Las preferencias de los adolescentes sobre los personajes televisivos de ficción seriada", en *Trípodos*, nº 38, pp. 141-159. Disponible en Internet: <https://bit.ly/3sAL3Xr> [Consulta: 10 de mayo de 2021].
- Lacalle, C. (2013): *Jóvenes y ficción televisiva: Construcción de identidad y transmedialidad*. Barcelona: UOC Editorial.
- Lacalle, C., & Castro, D. (2017): "Representations of female sexuality in Spanish television fiction", en *Convergencia*, vol. 24, nº 75, pp. 45-64. Disponible en Internet: <https://doi.org/drd9> [Consulta: 2 de junio de 2021].
- Lacalle, R., Gómez, B. e Hidalgo-Marí, T. (2021): "Historia de las teen series en España: evolución y características", en *Comunicación y Sociedad*, pp. 1-22. Disponible en Internet: <https://doi.org/10.32870/cys.v2021.7979> [Consulta: 24 de octubre de 2021].
- Lemish, D. (1985): "Soap Opera Viewing in College: A naturalistic inquiry", en *Journal of Broadcasting & Electronic Media*, vol. 29, nº 3, pp. 275-293. Disponible en Internet: <https://doi.org/d3ngtn> [Consulta: 10 de mayo de 2021].
- Livingstone, S. (1998): "Audience research at the crossroads: the 'implied audience' in media and cultural theory", en *European journal of cultural studies*, vol. 1, nº 2, pp. 193-217. Disponible en Internet: <https://doi.org/c69gjm> [Consulta: 15 de mayo de 2021].
- López Vidales, N., y Gómez Rubio, L. (2012): "Géneros, formatos y programas de televisión preferidos por los jóvenes". Análisis comparativo por Comunidades Autónomas", en *Revista de Comunicación y Tecnologías Emergentes*, vol. 10, nº 3, pp. 258-283. Disponible en Internet: <https://doi.org/f37d> [Consulta: 10 de mayo de 2021].
- Luzón, V., Figueras, M., y Jiménez, M. (2009): *La imagen de los y las adolescentes en el prime time televisivo. Transmisión, consumo y recepción*. Universidad Autónoma de Barcelona. Disponible en Internet: <https://bit.ly/3fEuu9t> [Consulta: 10 de mayo de 2021].
- Masanet, M.J., Medina, P., y Ferrés i Prats, J. (2012): "Representación mediática de la sexualidad en la ficción seriada dirigida a los jóvenes: estudio de caso de "Los Protegidos" y "Física o Química"", en *Comunicación*, vol. 10, nº 1, pp. 1537-48. Disponible en Internet: <https://bit.ly/2Ocq6ml> [Consulta: 15 de mayo de 2021].
- Masanet, M. J., y Buckingham, D. (2015): "Advice on life? Online fan forums as a space for peer-to-peer sex and relationships education", en *Sex Education. Sexuality, Society and Learning*, vol. 15, nº 5, pp. 486-499. Disponible en Internet: <https://doi.org/f369> [Consulta: 2 de junio de 2021].
- Masanet, M. J., y Fedele, M. (2019): "El "chico malote" y la "chica responsable": modelos aspiracionales y representaciones juveniles en las teen series españolas", en *Palabra Clave*, vol. 22, nº 2, pp. 1-27. Disponible en Internet: <https://doi.org/f37b> [Consulta: 10 de mayo de 2021].
- Mateos-Pérez, J. (2021): "Modelos de renovación en las series de televisión juveniles de producción española. Estudio de caso de Merlí (TV3, 2015) y Skam España (Movistar, 2018)", en *Doxa Comunicación. Revista Interdisciplinaria De Estudios De Comunicación Y Ciencias Sociales*, vol. 32, pp. 143-157. Disponible en Internet: <https://doi.org/gznb> [Consulta: 15 de julio de 2021].
- Medrano, C., Aierbe, A., y Orejudo, S. (2009): "El perfil de consumo televisivo en adolescentes: diferencias en función del sexo y estereotipos sociales", en *Infancia y Aprendizaje*, vol. 32, nº 3, pp. 293-306. Disponible en Internet: <https://doi.org/d6k3p7> [Consulta: 15 de mayo de 2021].
- Medrano, C., y Cortés, A. (2007): "Teaching and learning of values through television", en *Review International of Education*, vol. 53, nº 1, pp. 5-21. Disponible en Internet: <https://>

doi.org/d2vmmn [Consulta: 15 de mayo de 2021].

- Menéndez, M. I., Figueras-Maz, M., y Núñez, B. (2017): "Consumo y percepción juvenil sobre la ficción seriada televisiva: influencia por sexo y edad", en *OBETS. Revista de Ciencias Sociales*, vol. 12, nº 2, pp. 369-394. Disponible en Internet: <https://doi.org/f37j> [Consulta: 10 de mayo de 2021].
- Meyer, M. (2003): "It's me. I'm it. Defining Adolescent Sexual Identity through Relational Dialectics in Dawson's Creek", en *Communication Quarterly*, vol. 51, nº 3, pp. 262-276. Disponible en Internet: <https://doi.org/d67nzb> [Consulta: 15 de mayo de 2021].
- Meyer, M. (2009): "I'm Just Trying to Find my Way Like Most Kids: Bisexuality, Adolescence and the Drama of One Tree Hill", en *Sexuality & Culture*, vol. 13, nº 4, pp. 237-251. Disponible en Internet: <https://doi.org/ffnww4> [Consulta: 2 de junio de 2021].
- Montero, Y. (2006): *Televisión, valores y adolescencia*. Barcelona: Gedisa.
- Observatorio Español de las Drogas y las Adicciones [OEDA] (2020): *Informe 2020. Alcohol, tabaco y drogas ilegales en España*. Gobierno de España, Ministerio de Sanidad. Disponible en Internet: <https://bit.ly/3D1Mrao> [Consulta: 2 de junio de 2021].
- Pasquier, D. (1998): "Identification au héros et communautés de téléspectateurs: la réception d'«Hélène et les garçons»", en *Hermès*, vol. 1, nº 22. Disponible en Internet: <https://doi.org/b88h65> [Consulta: 10 de mayo de 2021].
- Pindado, J. (2006): "Los medios de comunicación y la construcción de la identidad adolescente", en *Zer, Revista de estudios de Comunicación*, vol. 11, nº 21, pp. 11-22. Disponible en Internet: <https://bit.ly/3m6Ezwl> [Consulta: 15 de mayo de 2021].
- Povlsen, K. K. (1996): "Global teen soaps go local: Beverly Hills 90210 in Denmark", en *Young*, vol. 4, nº 4, pp. 3-20. Disponible en Internet: <https://bit.ly/3a45uEH> [Consulta: 10 de mayo de 2021].
- Rodríguez, A., Medrano, C., Aierbe, A., y Martínez de Morentin, J. I. (2013): "Perfil de consumo televisivo y valores percibidos por los adolescentes: un estudio transcultural", en *Revista de Educación*, vol. 18, nº 2, pp. 513-538. Disponible en Internet: <https://bit.ly/3m6Rpvi> [Consulta: 10 de mayo de 2021].
- Rubin, A. M. (1985): "Uses Daytime Television Soap Operas by College Students", en *Journal of Broadcasting & Electronic Media*, vol. 29, nº 3. Disponible en Internet: <https://doi.org/d8m2qr> [Consulta: 15 de mayo de 2021].
- Sabina, J., de la Fuente, J., y Martínez, R. (2019): "El ecosistema mediático juvenil en España: un estudio de caso sobre el fandom de la serie Skam". En: N. López y E. Medina (Coords.), *Comunicación y Pensamiento. Relatos de la nueva comunicación*. Ediciones Egregius, pp. 33-52.
- Thompson, J. (1998): *Los medios y la modernidad*. Barcelona: Paidós.
- Tous-Rovirosa, A. (2015): *La política en las series de televisión. Entre el cinismo y la utopía*. Barcelona: UOC Editorial.
- Von Feilitzen, C. (2004): *Young People, Soap Operas and Reality TV*. Göteborg: Göteborg University.



La cobertura mediática de la ocupación de viviendas y el pánico moral. Análisis de los magacines matinales de TVE y Telecinco

Media coverage of Squatting and the moral panic. Analysis of morning magazines on TVE and Telecinco

Lourdes García Molina
lourgarciamolina@gmail.com
Universitat de València

RESUMEN Este artículo analiza el tratamiento mediático realizado por dos magacines, La Mañana (TVE) y El Programa de Ana Rosa (Telecinco), sobre las noticias de ocupaciones de viviendas, una cuestión que ha ganado espacio, tanto mediático como político, en los últimos tiempos. El objetivo principal de esta investigación es comprobar si estas informaciones se corresponden con una cobertura generadora de pánico moral, así como determinar si se prioriza un enfoque de suceso criminal o de problema social. Para ello, se estudian veinte programas emitidos en los meses de junio, julio y agosto de 2020 mediante técnicas de análisis cualitativas. El estudio está basado en la teoría *framing* partiendo de cinco *frames* propuestos por Semetko y Valkenburg (2000) y con 21 rasgos establecidos *ad hoc* para detectar los matices de la cobertura. Se concluye que la televisión privada presta una mayor atención a las noticias sobre ocupaciones de viviendas que la televisión pública, así como que el tratamiento mediático se corresponde con una cobertura generadora de pánico moral, con gran presencia del marco de conflicto. Además, prima un enfoque de suceso criminal representado como un atentado violento contra la propiedad privada.

PALABRAS CLAVE ocupación, vivienda, *frames*, televisión, pánico moral

ABSTRACT This article analyzes the media coverage conducted by two morning shows, La Mañana (TVE) and El Programa de Ana Rosa (Telecinco), on news of squatting, an issue that has gained scope, both in the media and in politics, in recent times. The main objective of this research is to verify if this information corresponds to a coverage that generates moral panic, as well as to determine if it is focused as a criminal event or as a social problem. To this end, 20 programs broadcasted between the months of June, July and August 2020 were studied, using qualitative analysis techniques. The study is based on framing theory starting from five frames proposed by Semetko and Valkenburg (2000) and 21 traits created *ad hoc* to detect the nuances of the coverage. It is concluded that private television pays more attention to news about squatting than public television. As well as that the media treatment corresponds to a coverage that generates moral panics, with great prominence for the conflict context. In addition, a criminal event approach prevails, represented as a violent attack against private property.

KEYWORDS squatting, housing, frames, TV, moral panics

Edita: José Gamir-Ríos

Recibido: 10/11/2021 | Aceptado: 18/03/2022

Cómo citar este artículo:

García Molina, L. (2022). La cobertura mediática de la ocupación de viviendas y el pánico moral. Análisis de los magacines matinales de TVE y Telecinco. *Dígitos. Revista de Comunicación Digital*, 8: 251-272. DOI: 10.7203/rd.v1i8.226



La cobertura mediática de la ocupación de viviendas y el pánico moral. Análisis de los magazines matinales de TVE y Telecinco

1. Introducción

La vivienda constituye mucho más que un lugar “para ser habitado por personas” (RAE), “es quizás el factor que determina en mayor medida el acceso que tendrá una persona a los recursos sociales [...] y participa en última instancia en la creación de las relaciones de poder” (Madden y Marcuse, 2016: 107). A pesar de su importancia, la vivienda se ha mercantilizado, es decir, “su función como bien inmueble tiene prioridad sobre su utilidad como lugar en el que vivir” (Ibíd.: 41). Según recogen portales como Fotocasa o InfoJobs, los españoles destinan alrededor de un 40% de su salario al pago del alquiler, y las personas jóvenes, hasta el 92,9% (Alonso, 2021).

Para contextualizar la situación de la vivienda, cabe destacar que, según los datos del Instituto Nacional de Estadística (INE) de 2011, de las 25.208.623 viviendas que hay en España, un 71,7% son viviendas principales, un 14,6% son secundarias, y el resto, un 13,7%, están vacías, las cuales han aumentado un 10,8% en la última década. Además, la vivienda pública en España se encuentra en el 2,5% respecto a las viviendas principales, muy por debajo de la media europea que se sitúa en un 9,3%. Con este panorama, los medios de comunicación y algunos partidos conservadores y ultraconservadores han puesto el foco en la ocupación ilegal de viviendas como un problema de primer orden. Sin embargo, según los datos disponibles, la magnitud del problema no sería de este calibre: un estudio elaborado por el Institut Cerdà en 2017, en cuyo patronato se encuentran miembros que están o han estado vinculados a bancos como La Caixa o Banco Sabadell, señala que existen 87.500 familias ocupando en España. Esta cifra representa un 0,34% del total estatal (Palomera y Virgós, 2019). Las agencias de vivienda de Catalunya y la Comunidad Autónoma de Madrid calculan que de estas ocupaciones ilegales, solo entre un 10% o un 15% pueden calificarse como “conflictivas” (Institut Cerdà, 2017: 18).

Otras entidades como Obra Social Barcelona, en su *Informe sobre la ocupación de vivienda vacía en Catalunya* (2018), apuntan que el 82% de las viviendas ocupadas en Catalunya pertenecen a grandes tenedores y, de estas, el 70% son propiedad de los bancos. Respecto al perfil de las personas ocupadoras, el 75% asegura que es por la carencia de ingresos en relación al coste de la vivienda. De hecho, según este estudio, el 93% de las unidades de convivencia que ocupan ilegalmente se encuentran por debajo del umbral de riesgo de pobreza, y la mayoría de las ocupaciones, un 87%, se llevan a cabo sin intercambio económico.

También es importante contextualizar brevemente la situación jurídica que envuelve

esta cuestión. Es clave diferenciar dos tipos de delitos relacionados con la ocupación de viviendas. Por una parte, el allanamiento de morada, que se refiere a la ocupación de una primera vivienda de un particular, recogido en el artículo 202 del Código Penal. Y, por otra parte, cuando la ocupación se da en viviendas no habituales o que se encuentran vacías, se trataría de un delito de usurpación de vivienda, también regulado a través del artículo 245 del Código Penal, incluso sin necesidad de violencia o intimidación. Por la vía de lo civil también se han introducido reformas dirigidas a agilizar este proceso. Concretamente, la Ley 5/2018, de 11 de junio —conocida como *Ley anti-okupas o del desahucio exprés*—, que modifica la Ley de Enjuiciamiento Civil para lograr “una recuperación inmediata de la vivienda ocupada ilegalmente”. Para ello, se han reducido los trámites y los periodos y se otorgan cinco días a los ocupantes para que “porten un título que justifique su situación posesoria. Si no se aportara justificación suficiente, el tribunal ordenará mediante auto la inmediata entrega de la posesión de la vivienda al demandante, siempre que el título que se hubiere acompañado a la demanda fuere bastante para la acreditación de su derecho a poseer” y, además, no cabrá recurso alguno por parte de los ocupantes ilegales.

2. Objetivos y preguntas

El objetivo principal de esta investigación es comprobar si la información sobre casos de ocupaciones de viviendas emitida en los programas de La Mañana de TVE y El Programa de Ana Rosa de Telecinco en los meses de junio, julio y agosto de 2020 se corresponde con una cobertura generadora de pánicos morales. Además, se establece otro objetivo secundario: determinar si la información sobre estos hechos prioriza un enfoque de suceso criminal, es decir, de atentado contra la propiedad privada, o de problema social, focalizado en la falta de recursos habitacionales de las personas que ocupan.

Las preguntas de investigación de este trabajo son:

1. ¿Qué espacio ocupa la cobertura de la ocupación de viviendas durante los meses que forman parte del periodo de estudio y cuál de los dos programas le presta más atención?
2. ¿Qué *frames* de la clasificación propuesta por Semetko y Valkenburg (2000) resultan preferentes en la cobertura de la ocupación de viviendas en los programas que forman parte del objeto de estudio?
3. ¿Existen diferencias entre los *frames* que se priorizan en Telecinco, de titularidad privada, frente a los que se priorizan en TVE, de titularidad pública?

3. Marco teórico

El concepto de miedo ha sido estudiado por numerosos autores en distintos campos y se entiende como un fenómeno social que experimentan los individuos, pero que está provocado por agentes externos (Morán San Juan, 2018). “Desde el punto de vista sociológico, el miedo es una construcción social y depende en gran medida de la

cultura del ambiente en el que se desarrolle para afrontar las diferentes maneras de interpretarlo. En nuestra sociedad occidental, el miedo posee una connotación negativa" (Morán San Juan, 2018: 18). Se trata, además, de un sentimiento aprendido, como lo es la respuesta hacia él y, en este sentido, los sociólogos Ulrich Beck y Zygmunt Bauman han reflexionado de manera extensa y reconocida acerca del miedo y su percepción social a través del concepto de sociedad del riesgo (Beck, 2006) enmarcado en un contexto de modernidad líquida (Bauman, 2007) en el que predominan la inseguridad y la incertidumbre.

Teniendo en cuenta la dimensión colectiva del miedo, cabe profundizar en el concepto de pánico moral, acuñado en 1972 por el sociólogo Stanley Cohen en su libro *Folk Devils and Moral Panics*. En él explica una dimensión más amplia de la inseguridad entendida como un sentimiento colectivo y describe el pánico moral como un acontecimiento, persona o grupo de personas a quienes se les define como una amenaza a los valores e intereses sociales, presentados por los medios de comunicación de manera estereotipada (Cohen, 2017: 51). Así, para que un caso de pánico moral tenga éxito debe cumplir tres condiciones (Cohen, 2017: 15-16):

- Un enemigo adecuado: alguien con poco poder, escaso acceso a los medios de comunicación, fácilmente denunciado y sin legitimidad.
- Una víctima correcta: alguien con quien uno pueda identificarse.
- Un consenso acerca de que no se trata de un problema aislado sino de una problemática que compete a toda la sociedad, ya que le puede suceder a cualquiera.

Y los elementos que constituyen el sentimiento de pánico moral son (Cohen, 2017: 32):

- Preocupación hacia la amenaza potencial o imaginada.
- Hostilidad hacia los actores definidos como *demonios populares* quienes encarnan el problema. Y también hacia las instituciones públicas, responsables en última instancia.
- Consenso, aunque no es necesario que sea total, de que la amenaza existe, es seria y se debería hacer algo al respecto. Para que funcione, la mayoría de las élites y grupos de influencia, especialmente los medios de comunicación, deben compartirlo.
- Desproporción: exageración del número o la magnitud de los casos, en términos del daño que causan, la ofensa moral y el riesgo potencial. Esto provoca que la preocupación de la opinión pública no sea proporcional al daño objetivo.
- Volatilidad: el pánico moral entra en erupción y se disipa de repente.

James P. Walsh (2016: 645) se refiere al pánico moral como un estado de alerta en el que ciertos acontecimientos o conductas se perciben como un riesgo para los valores y las normas de una sociedad. Por ello, y tomando como referencia el trabajo de Cohen, considera esencial la existencia de los *folk devils* para que tenga lugar una situación de

pánico moral. Estos demonios populares son representados como los responsables de los malos comportamientos y encarnan, a ojos de la sociedad, la antítesis de la estabilidad, el orden y la seguridad. Walsh (2016: 645) advierte de que no se trata de un clima espontáneo, sino que surge de la relación entre los agentes de control social, los medios de comunicación y el público. Así, llega a la conclusión de que los medios de comunicación son los principales impulsores y beneficiarios del pánico moral. Y lo argumenta alegando que la mayoría de la sociedad contagiada por esta alarma colectiva no experimenta directamente el problema, sino que son los medios de comunicación los que publicitan la indignación, ofrecen perspectivas sobre el control social y hacen que existan estos demonios populares.

Amadeo y Aruguete (2012, 2013) y Danieli y Paz García (2020) hacen referencia a la configuración de otros dos fenómenos complementarios al pánico moral: el de inseguridad y el de ola de violencia. Lo más llamativo de la inseguridad es su relativa autonomía respecto del delito efectivo, ya que aumenta a pesar de la disminución de los delitos. Asimismo, consideran que este sentimiento amedrenta la calidad de vida y favorece el apoyo ciudadano a políticas fuertemente punitivas (Danieli y Paz García, 2020: 8). El concepto de ola de violencia, por su parte, se comprende como una agrupación de casos violentos que generan un *efecto cascada* y una amplificación de la percepción de la inseguridad (Danieli y Paz García, 2020: 11). Además, a menudo está descontextualizada dando lugar a expresiones como “otro caso más”, lo que impide conocer las causas de fondo y generar un debate riguroso sobre sus consecuencias y soluciones. Una práctica mediática que, según autores como Van Dijk o Fernández (en Danieli y Paz García, 2020: 12), tendría como resultado prejuicios hacia colectivos vulnerables y una concepción de un *nosotros* y un *ellos*.

4. Metodología

Respecto al material de análisis, se ha escogido los magazines La Mañana de TVE, principal cadena de televisión pública de España, y El Programa de Ana Rosa de Telecinco, de titularidad privada. El motivo reside en que este estudio busca comparar el tratamiento mediático de dos cadenas con diferente titularidad y de cobertura estatal para que abarquen una misma realidad, así como en la accesibilidad al material de estudio que presenta Telecinco frente a otras opciones como Antena 3. La elección del magazine frente a otras opciones como los informativos se debe a la *riqueza*, entendida como variedad, que puede aportar este género donde se tratan más extensamente los temas y tienen mayor cabida las noticias de *sucesos*. El periodo de estudio se corresponde con los meses de junio, julio y agosto de 2020 por la percepción de que podría haber un incremento de noticias que aborden esta cuestión tras leer numerosos artículos en prensa en los que se alertaba de una cobertura desmesurada de casos de ocupaciones de viviendas en verano; así como a que tradicionalmente en los meses estivales suele haber menos actualidad política y, por tanto, más espacio para otras noticias.

Tras delimitar el tema, el objeto y el periodo de estudio, se visualizaron de manera generalizada los programas para preseleccionar aquellos en los que se hablaba sobre ocupaciones de viviendas y contabilizar cuánto tiempo les dedicaban. Tras conseguir

esta información, se escogieron los 10 programas de cada cadena que más tiempo dedicaron a la cuestión para formar un corpus de investigación de 20 programas. A continuación, se seleccionó el análisis del contenido como método para dar respuesta a las preguntas de investigación a partir de cinco *frames* propuestos por Semetko y Valkenburg (2000: 95-96), ya que se consideró que eran marcos amplios que podrían encajar con el discurso observado. Además, se elaboró de manera inductiva un listado de rasgos característicos que encajaban dentro de estos cinco marcos y que, tras visualizar los programas, estaban presentes en mayor o menor grado y ayudaban a dar respuesta a los interrogantes de la investigación. Con estos criterios, se configuró una ficha de análisis del contenido (Tabla 1).

Tabla 1. Ficha de análisis del contenido

MARCO DE CONFLICTO
<ul style="list-style-type: none"> ■ Dan explicaciones reduccionistas que dan respuestas rápidas y sencillas al problema de la ocupación y que, en ocasiones, están descontextualizadas. ■ Legitiman la violencia. ■ Abordan el problema como un choque entre propietarios y ocupadores. ■ Privilegian un léxico ligado a la agresión y la violencia. ■ Representan la ocupación como un ataque y pone el foco en la inseguridad ciudadana.
MARCO DE INTERÉS HUMANO
<ul style="list-style-type: none"> ■ Hablan de un caso en concreto para explicar un problema más amplio. ■ Realzan la vulnerabilidad de la persona a quien le ocupan la vivienda, o de quien ocupa la vivienda. ■ Inciden en el drama de quien ha sufrido una ocupación, o de quien ocupa una vivienda. ■ Utilizan un caso concreto para ilustrar un problema que se asegura que es mucho más general.
MARCO DE CONSECUENCIAS ECONÓMICAS
<ul style="list-style-type: none"> ■ Señalan la situación económica en la que se encuentra la persona a quien le han ocupado una vivienda o de quien la ha ocupado. ■ Inciden en los problemas económicos para quien le han ocupado la vivienda. ■ Aluden al impacto de la ocupación en el mercado inmobiliario. ■ Aluden al impacto de la ocupación en el mercado de la seguridad.
MARCO DE MORALIDAD
<ul style="list-style-type: none"> ■ Hablan en términos morales dicotómicos (bien/mal, legales/delincuentes, trabajadores/vagos, merecedores/privilegiados). ■ Empujan a algún actor de la noticia a que se exprese en términos morales, o incluyen en la pregunta la que esperan que sea la respuesta. ■ Reparten papeles de víctimas y villanos.
MARCO DE ATRIBUCIÓN DE RESPONSABILIDADES
<ul style="list-style-type: none"> ■ Dicen quién es el culpable, tanto de una situación concreta como de la ocupación de viviendas en general. ■ Dicen quién puede solucionar, tanto una situación concreta como la ocupación de viviendas en general. ■ Insinúan que el problema de la ocupación es culpa de los políticos. ■ Insinúan que habría que cambiar las leyes. ■ Insinúan que España está más desprotegida o tiene leyes menos eficaces que otros países.

Por tanto, esta investigación aplica una metodología de análisis de contenido de corte fundamentalmente cualitativo, enmarcado en la teoría del *framing*. Aun así, para dar respuesta a la primera pregunta, se cuantifica la presencia de las noticias sobre ocupaciones de viviendas en los magazines del periodo de estudio. La teoría del *framing* cuenta con unos orígenes multidisciplinares que remiten al campo de la psicología cognitiva, la sociología, la economía y la lingüística cognitiva. Desde un punto de vista comunicativo, "puede definirse como un proceso en el que se seleccionan algunos aspectos de la realidad, a los que les otorgará un mayor énfasis o importancia, de manera que se define el problema, se diagnostican sus causas, se sugieren juicios morales y se proponen soluciones y conductas apropiadas a seguir" (Ardèvol-Abreu, 2015: 424). Por tanto, esta área de investigación se ocupa de analizar la forma en que se presentan las noticias y la interpretación que promueven los medios de comunicación en función de los enfoques que privilegian y los que omiten.

Más allá de las definiciones, y tal como indica Ardèvol-Abreu (2015: 434-435), para detectar los encuadres noticiosos existen dos vías: la deductiva y la inductiva. La vía deductiva parte de encuadres predefinidos por otros autores que se aplican al análisis concreto que se quiere realizar. Esta opción tiene una mayor replicabilidad y es más sencilla de realizar, pero el principal inconveniente es que hay una pérdida de información. Por otro lado, la vía inductiva requiere una aproximación abierta a la muestra para detectar los posibles encuadres que están presentes en ella, ya que no parte de unos encuadres predefinidos. Se trata de un proceso más complejo, pero que permite detectar con mayor precisión y detalle los *frames* del objeto de estudio. En esta investigación, hay una mezcla de ambas vías. Con ello, además, se pretende dar respuesta a uno de los objetivos secundarios: determinar si la información sobre estos hechos prioriza un enfoque de suceso criminal, es decir, de atentado violento contra la propiedad privada, o de problema social, esto es, centrado en la falta de recursos habitacionales de las personas que ocupan. Se entiende que la cobertura que favorece un enfoque de suceso criminal se corresponde con los marcos de conflicto, interés humano y moralidad, aquellos con un discurso más superficial, ligado al espectáculo y que no dan pie a llegar a la raíz de los problemas sociales. Por otra parte, el tratamiento que podría privilegiar un enfoque de problema social se vincularía con los *frames* de consecuencias económicas y atribución de responsabilidades ya que, tratados en profundidad, permiten poner sobre la mesa cuestiones de calado y, por tanto, hacer una lectura de los problemas de fondo.

5. Resultados

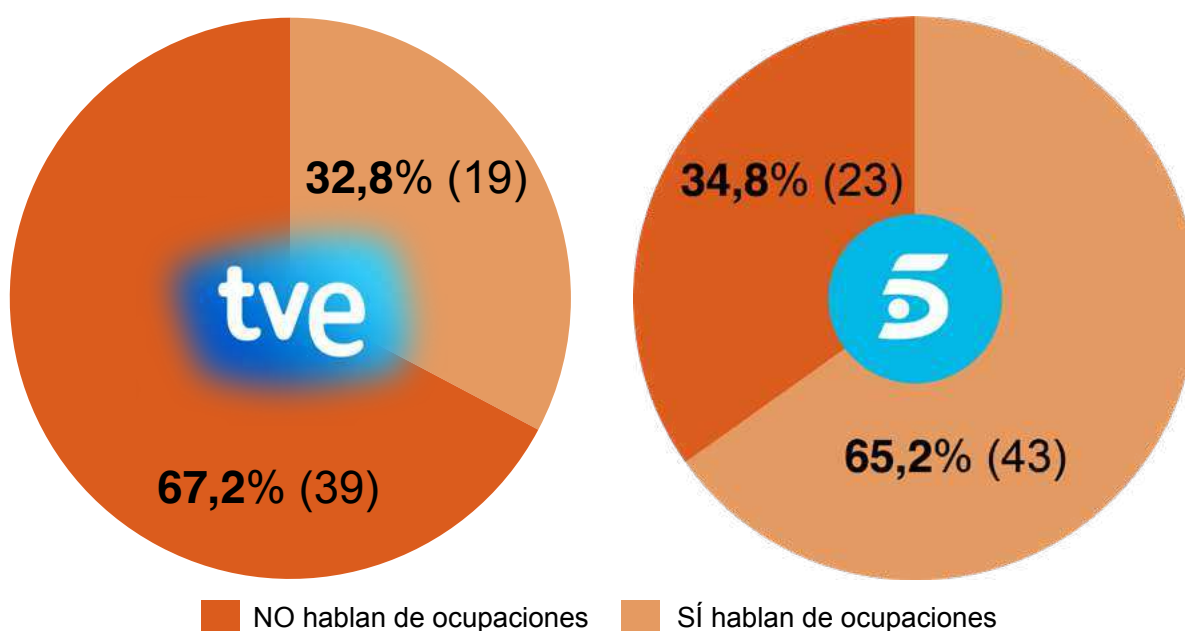
5.1 Presencia de las informaciones sobre ocupación de viviendas en los magazines analizados

En el programa La Mañana, de TVE, de los 58 programas emitidos durante los meses de junio, julio y agosto de 2020 se han detectado 19 magazines, un 32,7% del total, en los que se habla sobre algún caso de ocupación ilegal de viviendas¹. De ellos, se

1. En la web de RTVE no aparecen los programas de los días 2, 3, 5, 10, 11, 18 y 19 de junio, y no se han podido visionar. Consultado el servicio de Documentación, estos programas se retiraron de la web por la protección de datos y no pueden facilitar su visionado en ningún caso.

han escogido los 10 que más tiempo dedican a esta cuestión que se sitúan entre los 11 y los 35 minutos y se ha calculado el porcentaje de minutos que dedican a este tema respecto del total de minutos que dura cada uno de los programas. De media, La Mañana dedica el 9,93% del total del tiempo de los programas analizados a hablar sobre ocupaciones de viviendas. Por lo que se refiere al Programa de Ana Rosa, de Telecinco, de los 66 magacines emitidos durante los meses de junio, julio y agosto de 2020 se han detectado 43 programas en los que se habla sobre algún caso de ocupación ilegal de viviendas, esto representa un 65,1%. De ellos, se han escogido los 10 con mayor duración, que se sitúan entre los 20 y los 44 minutos. Así, dedica una media del 13,17% del total de los programas analizados a hablar sobre ocupaciones de viviendas. Cabe apuntar que estas informaciones se ubican junto a otras noticias de sucesos, hacia la mitad y final del programa, e intercaladas, en la mayoría de ocasiones, entre noticias que se enmarcan en la inseguridad (narcotráfico, peleas, asesinatos...).

Gráfico 1. Presencia de noticias sobre ocupaciones de viviendas en TVE y Telecinco



5.2 Enmarcado de las informaciones sobre ocupación de viviendas

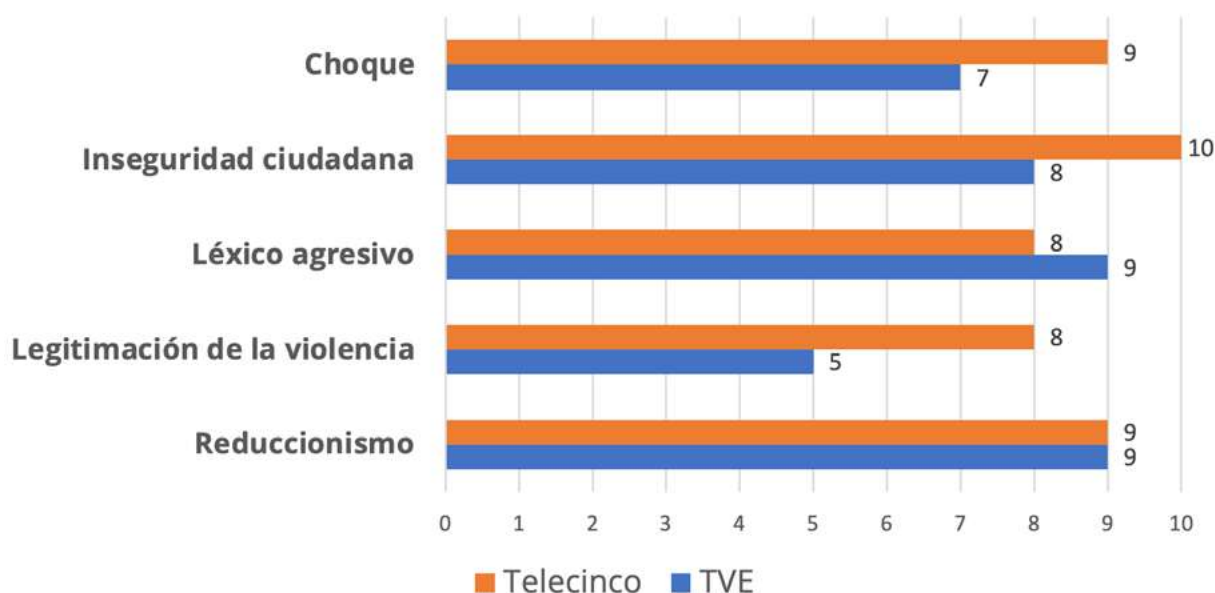
La cobertura desde el marco de conflicto

El *frame* preferente en la cobertura mediática de las ocupaciones de viviendas es el marco de conflicto. En La Mañana, todos los rasgos de este marco están presentes simultáneamente en cinco de los diez magacines analizados; mientras que en El Programa de Ana Rosa, en seis de los diez programas. Sin embargo, algunos de estos rasgos están presentes de manera especialmente destacada. Se trata, principalmente, de una cobertura en la que aparecen explicaciones reduccionistas que dan respuestas rápidas, cortas y, a priori, sencillas al problema de la ocupación y que, en ocasiones, están descontextualizadas, detectado en nueve de los diez programas de ambas cadenas. Estos argumentos giran en torno a afirmaciones falsas como, por ejemplo, que las viviendas que suelen ser ocupadas son viviendas principales, utilizando un discurso que se ha extendido entre la opinión pública de que *sales de tu casa y al volver*

la han ocupado, o que no pertenecen, principalmente, a entidades bancarias. En TVE se puede observar en el programa del 31 de agosto cuando la tertuliana, Anna Grau, señala: “Normalmente la excusa suele ser que hay muchos pisos vacíos de los bancos... pero fíjate que esta gente nunca okupa pisos de los bancos porque los bancos sí se saben defender, se preocupan. Pisos de gente normal que es la que está, perdón por la expresión, jodida porque realmente se ve atada de pies y manos con la ley” (1:34:36-1:34:57).

En la cadena privada, se pueden apreciar otras explicaciones reduccionistas que insisten en el argumento, desacreditado por las asociaciones judiciales², de que “no se puede hacer nada”. En el del 3 de agosto hay afirmaciones de tertulianos como esta: “Si alguien se mete dentro de tu piso después de que tú hayas salido a trabajar y vuelvas por la tarde, estás muerto. No puedes hacer nada” (2:03:05-2:03:12). En Telecinco también se puede observar reduccionismo al plantear esta cuestión compleja y con aristas de manera simplista y, en ocasiones, aportando soluciones *mágicas* que abogan por *tirar* a las personas ocupadoras independientemente de las situaciones de vulnerabilidad que puedan vivir, o sin discernir si es propiedad de una persona física o una entidad bancaria. Con estos argumentos, se elimina el aspecto garantista del sistema judicial español que, a diferencia de otros, parte de la presunción de inocencia. Otro de los argumentos más repetidos referente a este rasgo es que existe un límite de 48 horas en los que la policía puede actuar y, pasado ese tiempo, ya no se puede hacer nada para que la persona ocupadora abandone la vivienda, una afirmación que expertos y juristas han desmentido³.

Gráfico 2. Presencia de los rasgos del Marco de conflicto en TVE y Telecinco



2. Según El Periódico (2020), las tres principales asociaciones judiciales, Jueces para la Democracia, la Asociación Profesional de la Magistratura, y la Asociación Judicial Francisco de Vitoria, coinciden en que las penas y la legislación existente es suficiente para hacer frente a esta cuestión, y solicitan más medios de la administración pública.

3. Según recoge El País (2020) o Nius Diario (2020), expertos y juristas han desmentido que exista ningún límite de 48 horas a partir del cual ya no se pueda actuar en un caso de ocupación de viviendas, señalan que se trata de un bulo, un mito y una leyenda urbana porque ninguna ley contempla este plazo.

Un léxico ligado a la agresión y a la violencia es otro de los rasgos del marco de conflicto presente en este análisis, detectado en nueve de los diez programas de TVE. Así como una cobertura basada en la representación de la ocupación como atentado o ataque, es decir, como una noticia que pone el foco en la inseguridad ciudadana, una característica que se observa en todos los programas analizados de Telecinco, como el del 3 de agosto, donde se alude al “trapicheo de drogas”, las peleas, o el “efecto llamada”: “Los vecinos tienen miedo por el trapicheo de drogas y las peleas, pero también por el efecto llamada”; “son agresivos y nada se les resiste”; “tienen atemorizados a los vecinos, pero es que además, según nos han contado ellos, los vecinos, por el efecto llamada han okupado ahora una sucursal bancaria”.

La cobertura desde el marco de interés humano

Los *frames* de interés humano, moralidad y atribución de responsabilidades también están presentes de manera destacada, aunque en menor medida que el marco de conflicto. En los veinte programas analizados se incide, en algún momento determinado, en el drama de quien ha sufrido una ocupación o de quien ha ocupado una vivienda, un rasgo que pertenece al marco de interés humano. Cuando se trata de la desdicha que sufren las personas a quienes les han ocupado una vivienda, se traslada a través del discurso de presentadores, tertulianos, reporteros o la voz en off que narra los hechos. Pero el drama de las personas que han tenido que ocupar una vivienda proviene, en la mayoría de las ocasiones, de las propias personas ocupadoras que narran su situación. En Telecinco las cuatro ocasiones en la que se hace alusión a las dificultades de las personas que han tenido que ocupar una vivienda, lo hacen a través de los propios ocupadores; mientras que en TVE únicamente hacen alusión al drama que están viviendo las personas ocupadoras en tres ocasiones y, en dos de ellas, la narración la hacen los propios ocupadores.

Otro rasgo presente en ambas cadenas se basa en utilizar un caso concreto para ilustrar un problema que se asegura que es mucho más general. Esto lo hacen a través de argumentos vagos y generalizadores como que existe un vacío legal, que se tarda años en resolver las ocupaciones de viviendas, que se trata de mafias, que las personas que ocupan una vivienda ya no lo hacen por necesidad porque tienen artículos de lujo, o que podría pasarse a cualquiera y haber un efecto llamada. Por otra parte, el rasgo que menos se puede apreciar es el de una cobertura que utiliza un caso concreto para explicar un problema mucho más amplio y complejo. En los programas analizados aparecen problemas más complejos y profundos, pero no se hace hincapié en ellos, como la construcción de un parque público de vivienda, las ayudas sociales a las personas en situación de vulnerabilidad, la especulación inmobiliaria, la falta de recursos de la administración pública o los procesos de degradación o gentrificación. Estos aparecen nombrados en algún momento de la cobertura informativa a raíz de algún caso en concreto, pero de manera extremadamente fugaz. Otros temas, como el precio de los alquileres o los problemas de acceso a la vivienda de las personas jóvenes, ni siquiera aparecen.

La cobertura desde el marco de moralidad

Por lo que respecta al marco de moralidad, se puede detectar un rasgo que predomina respecto del resto y que está presente en todos los programas de la cadena privada y en nueve de los diez de la cadena pública: expresarse en términos morales dicotómicos (bien/mal, trabajador/vago), un rasgo ligado al de una cobertura que reparte papeles de víctimas y villanos que también está notablemente presente, en siete de los diez programas analizados en ambos casos. Estos términos morales giran, en gran medida, alrededor del concepto de meritocracia, ya que se hace distinción de manera reiterada entre aquellas personas merecedoras de una vivienda y que han trabajado y se han esforzado para serlo, y aquellas que no son merecedoras porque no se han esforzado lo suficiente. Esto se logra a través de afirmaciones que dan a entender que las personas ocupadoras de viviendas no trabajan, que escogen viviendas de lujo o que adoptan el papel de víctimas, cuando realmente tienen recursos suficientes. Este discurso se puede apreciar en esta intervención del tertuliano Julio César Herrero en el programa del 31 de agosto de TVE: “Acabamos de ver las imágenes de alguien que yo no sé si solamente, y esto entrecomillas, tiene 400 y pico euros. Desde luego conozco a gente que tiene 400 y pico euros y que no viola la ley y no entra en sitios que no son suyos [...] A lo mejor lo que toca hacer por los hijos es buscar ayudas sociales, buscar un trabajo y, sobre todo, enseñarles que violando la ley no se va a ningún sitio” (1:31:50-1:32:54). En Telecinco se detectan estos mismos rasgos, por ejemplo, en el programa del 20 de agosto con estas intervenciones de dos tertulianos mientras entrevistaban a dos personas que habían ocupado una vivienda vacía y que ya habían sido desalojadas (Tabla 2 e Imagen 1).

Tabla 2: Conversación del 20/08/20 en Telecinco

José Miguel Gaona (tertuliano): A ver si os buscáis un trabajito también, como todos los curritos que hay en este país y os pagáis un alquiler. Y si no, con la edad que tenéis, ya que sois tan niñas, compartid un piso como hemos hecho todos cuando estábamos estudiando.
[...]
Ocupadora: ¿Me lo pagas tú?, ¿Vienes tú conmigo y das la paga y señal?, ¿Vienes tú y me ayudas?
Javier Negre (tertuliano): ¿Pero que paga y señal? Pero no te has dado cuenta que vas de mártir y realmente eres cómplice de un delito.
Ocupadora: ¿Vienes tú y me das trabajo?, ¿Vienes tú y me das un trabajo? Con toda la situación que hay en España y encima os lo tomáis con los okupas.
Negre: Ponte a currar.
Ocupadora: Que nos estamos muriendo de hambre.
Negre: En vez de participar con una estructura mafiosa ponte a trabajar, tardarías menos y viviríais en un piso decentillo. (2:55:51-2:58:27)

Imagen 1: Programa del 20/08/20 en Telecinco



La cobertura desde el marco de atribución de responsabilidades

A pesar de que en ningún programa se pueden detectar todos los rasgos del marco de atribución de responsabilidades al mismo tiempo, uno de ellos es protagonista indiscutible: ambas cadenas reproducen en prácticamente todos los programas —en todos los de Telecinco y en nueve de los diez programas de TVE— el argumento de que para resolver la cuestión de la ocupación ilegal de viviendas es necesario cambiar la ley, bien sea porque consideran que es insuficiente, que está mal orientada o que protege a la persona ocupadora en lugar de al propietario. Este rasgo es destacable porque, además, va en contra del criterio de prácticamente todas las asociaciones judiciales españolas, que consideran suficientes las herramientas existentes. Algunos de los ejemplos más evidentes de La Mañana se pueden observar en esta conversación del 30 de junio (Tabla 3 e Imagen 2).

En el El Programa de Ana Rosa también se pueden encontrar numerosos ejemplos en los que sentencian que “la ley no está legislando bien”, que se trata de “la ley al revés”, o que no se cuentan con “ninguna herramienta legal”. Incluso, la voz en off de uno de los reportajes, emitido el 12 de agosto, termina lanzando esta pregunta a la audiencia: “Los okupas se hacen con cualquier vivienda indistintamente. ¿Se creen impunes a la ley o es que realmente lo son?” (2:27:55-2:28:03).

En el marco de atribución de responsabilidades, se puede apreciar a quiénes señalan como culpables y a quiénes como posibles solucionadores, tanto de los casos concretos como del fenómeno de la ocupación en general. En todas las noticias analizadas, los culpables concretos son los ocupadores de viviendas. Pero, más allá de esta apreciación, el discurso mediático señala otros culpables: la justicia, a la que reprochan lentitud y estar mal orientada; o los partidos políticos, a quienes les achaca falta de acuerdo y

de acción. Por lo que respecta a quiénes podrían solucionar las ocupaciones, ambos canales apuntan a las Fuerzas y Cuerpos de Seguridad del Estado, aunque advierten que no tienen herramientas para hacerlo y proponen un cambio de ley que les permita actuar de inmediato, eliminando cualquier atisbo garantista de la norma, a pesar de que en varios de los casos que aparecen logran resolver la situación. Además, aluden a los responsables políticos, a quienes reprochan una falta de acuerdo y de acción, así como a los ayuntamientos en lo relacionado con servicios sociales y seguridad ciudadana. En contraposición a la inacción o incapacidad de estos agentes, existen dos actores que, según el discurso mediático observado, sí resuelven de manera efectiva estos escenarios: la presión vecinal agrupada en las conocidas como *patrullas vecinales* y las empresas de desocupación como Desokupa o Desocupación Legal, quienes incluso realizan intervenciones en directo.

Tabla 3: Conversación del 30/06/20 en TVE

Luis Fernando Durán (tertuliano): Es que tenemos un problema en este país: la inviolabilidad de la vivienda. O sea, es un derecho constitucional que ampara a los okupas porque nadie puede entrar en tu casa aunque no sea tuya.

Euprepio Padula (tertuliano): y ahí está el fallo porque ampara a los okupas pero no al propietario de la vivienda. [...] Por eso hay que cambiar la ley.

Durán: Efectivamente. Hay que cambiarla para que haya desahucio exprés no solamente cuando hay impago de una vivienda, sino cuando te la okupan de esta forma.

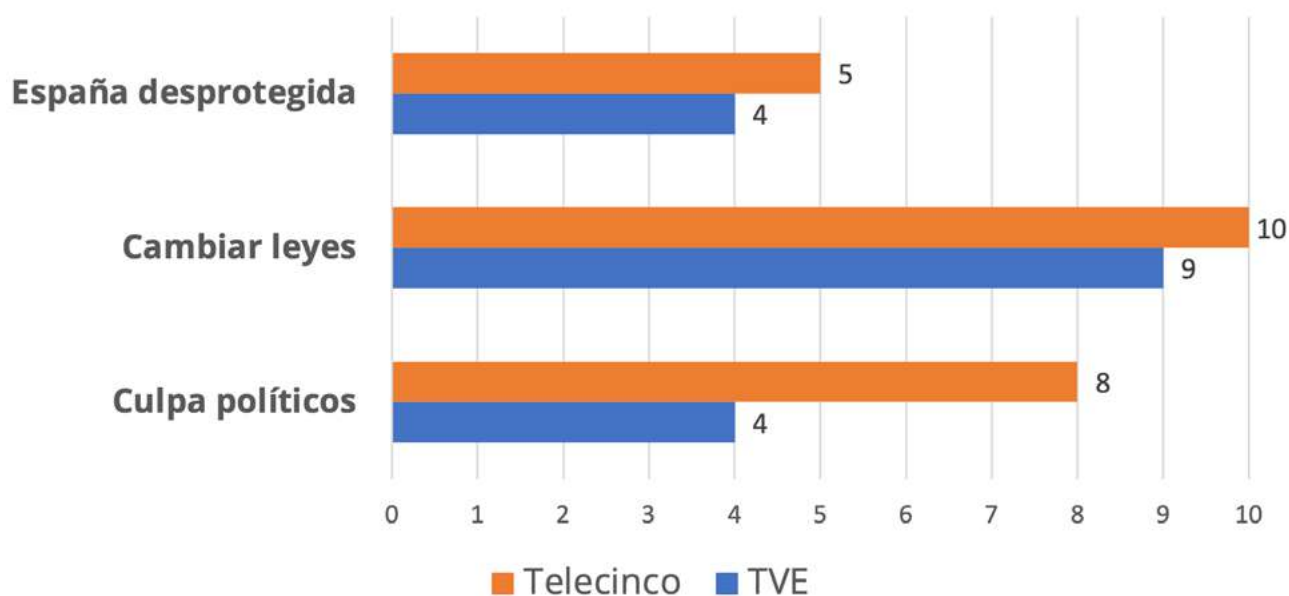
Padula: Exacto! Como en otros países, como en todos los países de nuestro entorno.

Durán: Y que el derecho a la propiedad privada esté por encima del derecho de los okupas. (1:14:38-1:15:13)

Imagen 2: Programa del 30/06/20 en TVE



Gráfico 3. Presencia de los rasgos del Marco de atribución de responsabilidades en TVE y Telecinco

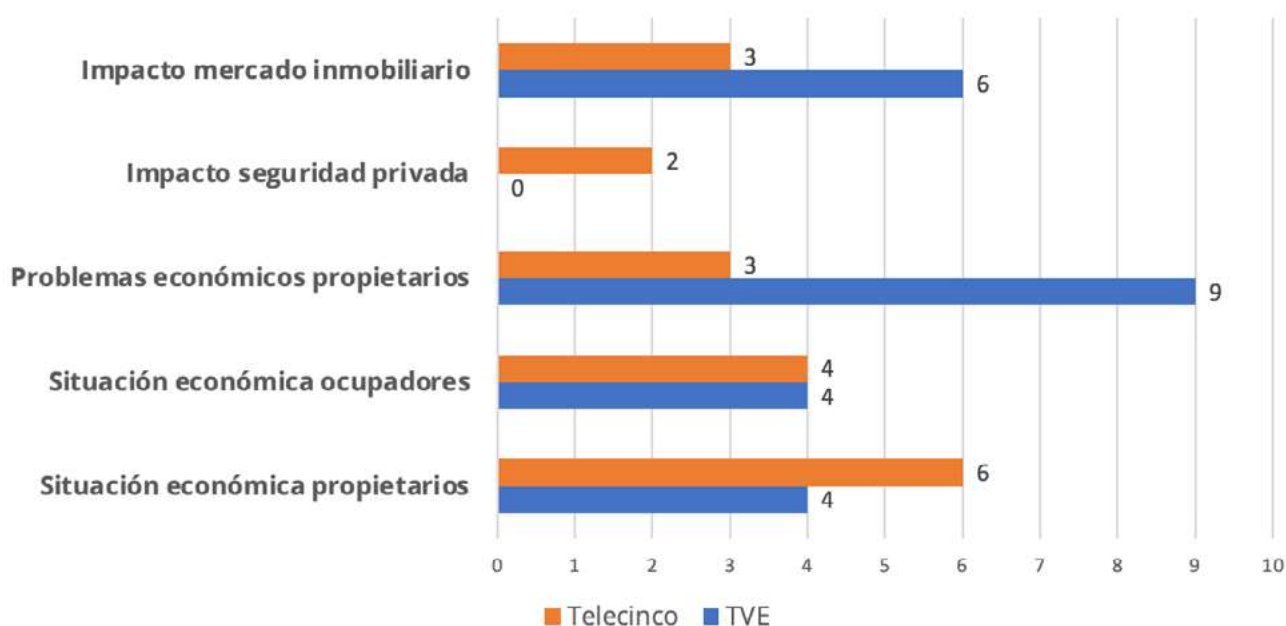


La cobertura desde el marco de consecuencias económicas

Por el contrario, el marco con menor presencia en los programas analizados es el que gira en torno a las consecuencias económicas. En ninguno de los programas analizados de TVE se cumplen simultáneamente todos los rasgos que forman parte de este *frame*; mientras que, en Telecinco, solo un programa incluye todos los rasgos que pueden atribuirse al marco de consecuencias económicas, y otros dos no presentan ningún rasgo de este *frame*. De los rasgos de este marco, el que está menos presente es la alusión al impacto que las ocupaciones de vivienda pueden tener en el mercado de la seguridad privada. De hecho, en la cadena pública no se hace referencia en ningún momento a esta cuestión y en la cadena privada sólo se hace en dos ocasiones. Por lo que respecta a la situación económica en la que se encuentran las personas a quienes les han ocupado la vivienda o las personas que han tenido que ocupar una vivienda está presente en cuatro de los diez programas de La Mañana, y en seis de los diez de El Programa de Ana Rosa. En las ocasiones en las que se hace referencia a la situación económica de las personas que han ocupado una vivienda —en cuatro ocasiones en cada cadena— es, o bien para explicar que son familias vulnerables, en cuyo caso el discurso no se realiza por parte de los presentadores, los tertulianos, los reporteros o los vecinos, sino únicamente a través del discurso de los propios ocupadores; o bien para desacreditar estos argumentos por parte de presentadores, tertulianos, reporteros, vecinos o propietarios poniendo en cuestión que realmente sean personas vulnerables económicamente. Esto se puede detectar en el programa del 17 de junio de la cadena pública, cuando una propietaria, a quien le han ocupado su segunda vivienda, asegura que no se trata de personas sin recursos que necesitan una vivienda: “Alguien que necesita un techo para vivir no va en un BMW” (1:09:47-1:09:52). Y en la cadena privada se puede apreciar en el programa del 27 de agosto, cuando la presentadora sentencia: “Me temo, en fin, que no es un tema de necesidad. Es un tema de que son chalets

de lujo, con piscinas, con parcelas grandes, que les conviene y les interesa instalarse y usurpar una de ellas”; apoyada por la narración del reportero: “Ya no es solo el hecho de que vengan con coches, es que vienen también con ropa de marca, el otro día cuando hablamos con ellos venían vestidos con ropa de marca, con zapatillas que, por experiencia, sé que cuestan bastante dinero. Es lo que venimos contando, no es gente necesitada, son mafias que se dedican a esto” (2:58:16-2:58:52). En cambio, en el discurso sobre la situación económica de las personas a quienes les han ocupado una vivienda la narración queda refrendada, no solo por vecinos y propietarios, sino también por presentadores, tertulianos o reporteros.

Gráfico 4. Presencia de los rasgos del Marco de consecuencias económicas en TVE y Telecinco



5.3 Diferencias entre los *frames* que se priorizan en la cadena privada y en la pública

La respuesta a esta tercera pregunta de investigación se complementa con todo lo respondido en la segunda pregunta, pero entrando más en el detalle de las diferencias que se pueden observar en la cobertura de determinados rasgos. Cabe reiterar que en ambas cadenas el marco preferente para cubrir estas noticias es el de conflicto, pero existe un rasgo en el que se aprecian diferencias considerables: la legitimación de la violencia en un sentido amplio, es decir, no solo la física, sino también la verbal o la intimidación. En la televisión pública es el menos frecuente de todos los rasgos de conflicto y solo aparece en la mitad de los programas, mientras que en la televisión privada aparece en ocho de los diez programas. Esto se refleja en el programa del 19 de junio de Telecinco cuando la presentadora y las tertulianas defienden y legitiman la acción de los vecinos de tomarse la justicia por su mano: “La solidaridad vecinal está demostrando ser un la única alternativa contra los okupas” (2:45:59-2:46:04); “revueltas antiokupas”. O en el programa del 3 de agosto cuando, no solo muestran una actuación en directo de la empresa Desokupa, sino que la presentadora blanquea y avala la contratación de este tipo de empresas: “El ciudadano tiene que recurrir a otras

alternativas, o cogerse la justicia por su mano y sacar a los señores que han usurpado su vivienda con un palo directamente, o recurrir a una empresa de este tipo" (2:03:30-2:03:41). Así, una de las principales diferencias en la cobertura mediática del marco de conflicto entre TVE y Telecinco es que, a pesar de que ambas utilizan argumentos reduccionistas, léxico ligado a la violencia o un enfoque de inseguridad ciudadana, en la cadena privada hay una legitimación mucho más explícita y frecuente de la violencia que en la pública.

En el caso del marco de interés humano, se pueden apreciar diferencias en el discurso mediático al referirse a la situación de vulnerabilidad de las personas a quienes les han ocupado la vivienda o de quien ha tenido que ocupar una vivienda. En TVE, de las cuatro veces en las que se habla sobre la situación de vulnerabilidad que padecen las personas que han tenido que ocupar una vivienda, a parte de la narración de los propios ocupadores, y de algunas réplicas por parte de tertulianos que ponen en cuestión esta condición, también se puede observar cómo, de manera puntual, algunos tertulianos, el presentador e incluso una de las propietarias a quien le han ocupado la vivienda, trasladan en su mensaje la preocupación por la situación de vulnerabilidad de las personas ocupadoras. En cambio, en Telecinco, de las cuatro veces en las que se realiza la situación de vulnerabilidad en la que se encuentran las personas que han tenido que ocupar una vivienda, ninguna se encuentra en el propio discurso mediático. Solo una de ellas proviene del testimonio de los vecinos, pero no desde una perspectiva empática, sino más bien criminalizadora; y las otras tres provienen del testimonio de las personas ocupadoras.

En el marco de atribución de responsabilidades, se pueden apreciar diferencias en el rasgo que privilegia una cobertura en la que se insinúa o se dice de manera explícita que el problema, y por ende la solución, tiene que ver con los partidos o dirigentes políticos. Esta disparidad reside en que, a diferencia del resto de rasgos donde hay un equilibrio y semejanza en el tratamiento de ambos magacines, en esta ocasión la cadena privada alude a esta cuestión en el doble de ocasiones que la cadena pública: este rasgo está presente en ocho de los diez programas de Telecinco, mientras que solo lo está en cuatro de los diez de TVE. En ambas cadenas, no solo hay referencias generales a los políticos, sino que se pueden observar menciones explícitas a Podemos, y a su posible connivencia con las ocupaciones de vivienda, así como a las propuestas de partidos como Vox, PP o Ciudadanos. Pero, en el magacín de la cadena privada, este rasgo es mucho más frecuente con afirmaciones como esta de la presentadora: "El Parlamento tendrá que legislar y los jueces tendrán que actuar de forma rápida". U otras como: "Creo que no hay voluntad de cambiar la legislación, y que se puede y se debe hacer".

O en los programas de los días 14 y 20 de agosto, en los cuales el tertuliano Javier Negre hace declaraciones en las que, de manera implícita, da a entender que existen partidos que utilizan la ocupación de viviendas como un "caladero de votos" en referencia a Podemos, mientras que otros como Vox o el Partido Popular son quienes están reivindicando una ley de desahucio exprés para *okupas*: "¿Por qué los gobernantes no gobiernan contra la ocupación? Porque el PP lo está pidiendo, Vox lo está pidiendo,

Ciudadanos lo está pidiendo. ¿Por qué? Porque considerará el PSOE y compañía y Podemos que son sus votantes, porque, si no, no me cabe otra explicación” (2:41:15-2:41:26). Finalmente, cabe destacar el programa del 12 de agosto donde se relatan las propuestas de los partidos PP, Vox y Ciudadanos respecto a esta cuestión y la periodista Isa Costa sentencia: “El PSOE tendrá que elegir entre contentar a sus socios de gobierno, Unidas Podemos, que defienden a los que ocupan viviendas de manera ilegal con un objetivo social, o a Ciudadanos, y esta propuesta que acabamos de escuchar, de la Ley Antiokupas” (1:46:54-1:47:07).

Respecto al marco de moralidad, uno de los rasgos estudiados hace referencia a situaciones en las cuales los presentadores, tertulianos o reporteros empujan a otros actores de las noticias, como propietarios o vecinos afectados, a que se expresen en términos morales, o les preguntan incluyendo en la cuestión, de manera más o menos explícita, la que esperan que sea la respuesta. Esto es algo más frecuente en Telecinco, ya que está presente en ocho de los diez programas analizados, y algo menos en TVE, presente en la mitad de los programas. Los ejemplos son tan diversos como situaciones se dan en los distintos programas. Un ejemplo es el programa del 3 de agosto en el que una propietaria acaba de recuperar su vivienda tras la intervención en directo de Desokupa y graban la entrada a su vivienda. Entonces, la reportera y la presentadora le preguntan constantemente incorporando la respuesta en la pregunta y empujándola a que se exprese en términos morales: “Porque esto para ti es una noticia impresionante, ¿no?”; “¿No ha visto el estado en que han dejado estos okupas su vivienda?”; “¿Cómo la encuentras? ¿Está un poco sucia?”; “¿Sientes que faltan cosas? ¿Te han podido tirar tus pertenencias?”; “Me imagino que algo habrás notado, ¿no?”. En este mismo programa, pero en otra noticia, el reportero accede al lugar donde viven los ocupadores y, además de enfocar objetos como latas de cerveza o el desorden de la casa, les pregunta si las peleas habían sido causadas por el alcohol, relacionando ocupación con consumo de alcohol y haciendo que responda en términos morales y tenga que justificarse: “En alguna ocasión habéis peleado, ¿pero porque habíais bebido a lo mejor o...? [...] ¿Traficáis con droga y demás?”.

Finalmente, en el *frame* menos frecuente, el de consecuencias económicas, también se pueden apreciar diferencias en el discurso sobre los problemas económicos que conlleva la ocupación de viviendas para los propietarios y sobre el impacto que generan las ocupaciones de viviendas en el mercado inmobiliario. La primera de estas características es predominante en TVE, ya que aparece en nueve de los diez programas analizados, mientras que en el caso de Telecinco tiene un protagonismo más reducido y, únicamente, se puede apreciar en tres de los diez programas estudiados. Ahora bien, en ambos casos los ejemplos son bastante similares y consisten en evidenciar los costes económicos, habitualmente en concepto de pago de suministros, que deben afrontar los propietarios de viviendas que han sido ocupadas. Y, en menor grado, aparecen problemas como la devaluación de los precios de los pisos, robos de objetos o deudas. La segunda de las características donde se pueden observar diferencias entre las dos cadenas es en el impacto que generan las ocupaciones de viviendas en el mercado inmobiliario. En este caso, la diferencia reside en la frecuencia en la que aparece este

rasgo: en más de la mitad de los programas analizados de TVE -seis de los diez-, y solo en tres de los diez de Telecinco. Esto se puede observar en el programa del 8 de junio de La Mañana en el que una de las tertulianas, Teresa Viejo, señala: "Tienen que convivir con una situación insalubre que además deprecia el valor de sus viviendas, que hace incómoda la convivencia... a mí me parece que es absolutamente injusto" (1:15:14-1:15:23).

6. Discusión y conclusiones

Tras esta investigación, y a la luz de los resultados obtenidos, se puede concluir, en primer lugar, que la cadena de titularidad privada le ha prestado una mayor atención a las noticias sobre ocupaciones de viviendas que la pública, tanto en el número de magacines, como en la cantidad de tiempo que le dedican en ellos.

En segundo lugar, que el *frame* preferente es el de conflicto, seguido de los marcos de interés humano, moralidad y atribución de responsabilidades, mientras que el menos frecuente es el de consecuencias económicas. Además, los rasgos más frecuentes del *frame* de conflicto en ambas cadenas son las explicaciones reduccionistas que dan respuestas rápidas, cortas y, a priori, sencillas al problema de la ocupación y que, en ocasiones, están descontextualizadas. En el caso de TVE, también existe una cobertura que privilegia un léxico ligado a la agresión; y en el caso de Telecinco, se muestra la ocupación como atentado o un ataque poniendo el foco en la inseguridad ciudadana.

En los 20 programas analizados está presente el drama de quien ha sufrido una ocupación o de quien ha ocupado una vivienda, un rasgo que corresponde al marco de interés humano pero, dependiendo de si es una situación u otra, el discurso proviene de voces distintas. Cuando se trata de propietarios o vecinos, el drama lo expresan y refrendan presentadores, tertulianos o reporteros. En cambio, si se trata de las dificultades de las personas ocupadoras, no solo aparece en menor medida, sino que proviene de su propia narración. Otra de las conclusiones es la utilización frecuente de un caso concreto de ocupación para ilustrar un problema que se asegura que es mucho más general, aunque los datos sobre el alcance del fenómeno no lo sustenten. En cambio, es menos habitual la utilización de un caso concreto para explicar problemas más amplios y estructurales, como la especulación inmobiliaria, el incremento de desahucios, las dificultades de los jóvenes para acceder a una vivienda o los elevados precios de alquiler, de las que pasan *de puntillas*.

Respecto al marco de moralidad, predomina un discurso que se expresa en términos morales, especialmente relacionado con el concepto de meritocracia. Esto se consigue a través de afirmaciones que dan a entender que las personas ocupadoras de viviendas son personas que no trabajan, que escogen viviendas de lujo, que van de víctimas cuando realmente tienen recursos suficientes, e incluso plantean si las personas ocupadoras deberían tener, o no, derecho a un abogado.

En el marco de atribución de responsabilidades se aboga casi de manera unánime por un cambio de la legislación. Esto coexiste con un perfil del culpable compartido por ambas cadenas: las personas ocupadoras de viviendas, la justicia y los partidos

políticos. Los dos programas también coinciden en quiénes podrían solucionar esta cuestión: las Fuerzas y Cuerpos de Seguridad del Estado, sin armas para hacerlo; y los políticos, sin voluntad para ello. Pero en el discurso de ambas cadenas se identifica, asimismo, dos actores que resuelven de manera efectiva estas situaciones: las *patrullas vecinales* y las empresas de desocupación

Respecto al marco de consecuencias económicas, en la mitad de los programas analizados de cada cadena hay alusiones a la situación económica de las personas que han tenido que ocupar una vivienda o de quienes les han ocupado la vivienda. Eso sí, con diferencias. Cuando se trata de la situación de las personas ocupadoras, o bien explican que son familias vulnerables —en cuyo caso únicamente se hace a través de su propia voz—, o bien desacreditan la versión de los ocupadores. Sin embargo, la situación económica de propietarios o vecinos está presente en su propio discurso y en el de presentadores, tertulianos o reporteros, quienes en ningún caso les desacreditan o ponen en cuestión.

En tercer lugar, que existen diferencias entre ambas cadenas en la presencia de algunos de los rasgos: en Telecinco es mucho más frecuente un discurso que legitima la violencia. También existen diferencias en el modo de reflejar la situación de vulnerabilidad de las personas que han ocupado una vivienda: en La Mañana, puntualmente, presentadores, tertulianos o propietarios se hacen eco de ello; y en El Programa de Ana Rosa en ningún caso este relato viene avalado o se realiza por parte de las presentadoras, tertulianos, reporteros, propietarios o vecinos. La atribución de responsabilidad que pesa sobre los partidos políticos es más habitual en la cadena privada; y la presencia en el discurso mediático de los problemas económicos que conlleva la ocupación de viviendas para los propietarios es predominante en TVE.

Estas tres conclusiones, y especialmente las dos últimas, contribuyen a alcanzar el principal objetivo: comprobar que las noticias analizadas se corresponden con una cobertura generadora de pánicos morales. Esta afirmación queda demostrada por la posición privilegiada dentro del discurso mediático del marco de conflicto, especialmente con la existencia de ciertos rasgos que contribuyen a la generación de pánico moral, como son las explicaciones reduccionistas y descontextualizadas, un léxico ligado a la violencia y el enfoque de la información como noticias de inseguridad ciudadana. También redunda en ello la presencia de otros *frames* notablemente frecuentes que contribuyen a la generación de esta sensación colectiva, como el de interés humano o el de moralidad. Con este análisis, también se da respuesta a uno de los objetivos secundarios: determinar si la información sobre estos hechos prioriza un enfoque de suceso criminal, o de problema social.

Así, se concluye que prima un enfoque de suceso criminal representando la ocupación como un atentado violento contra la propiedad privada, ligado a los *frames* más predominantes, como son los de conflicto, de interés humano y de moralidad, y al hecho de que no aprovechan las noticias sobre ocupación para abrir un debate que aborde las dificultades de acceso a la vivienda en España como un problema social de carácter estructural. Estos resultados coinciden con la descripción del pánico moral

realizada por Cohen (2017) y, además, se observa cómo la construcción de la figura del ocupador de viviendas que se proyecta en los programas estudiados se corresponde con la creación de un enemigo adecuado, alguien con poco poder, escaso acceso a los medios de comunicación, fácilmente denunciado y sin legitimidad. Así, parten desde una posición pasiva en la que deben demostrar que no son culpables, cualquier inclusión de su voz en el discurso de los medios de comunicación se enfoca desde el descrédito y se les otorga el papel de responsables de todo lo ocurrido. Esto provoca, a su vez, un efecto en las relaciones de poder convirtiendo a personas vulnerables en los responsables de su propia suerte, en lugar de ahondar en las causas y problemáticas mucho más complejas que entaña, no solo el mercado de la vivienda en un contexto de hipermercantilización de la misma, sino toda una política del bienestar que debería evitar que las personas se vieran obligadas a situarse al margen de la legalidad para tener un bien tan básico como la vivienda.

También se presenta a una víctima correcta, alguien con quien uno pueda identificarse fácilmente: los propietarios o los vecinos. Sin embargo, en contadas ocasiones se trata de una vivienda habitual y, cuando esto ocurre, la situación se resuelve inmediatamente. Por el contrario, los casos más frecuentes son pisos que pertenecen a bancos o personas que tienen más de una vivienda, un matiz que no se subraya de manera especialmente destacada en el discurso mediático, simplemente se menciona y se acompaña de un rostro humano, como el propietario o los vecinos afectados. Conviene señalar que, según el informe *Radiografía del mercado de la vivienda 2018-2019* (2019) realizado por Fotocasa, el 61% de las personas mayores de edad tiene alguna propiedad y, de estos, un 78% tiene una única propiedad, mientras que solo el 18% tiene dos, y únicamente el 4% tiene tres o más. Por tanto, se ofrece a la audiencia una imagen de las personas a quienes les han ocupado una vivienda mucho más corriente de lo que realmente es, y un perfil de las personas que han ocupado una vivienda mucho más extraordinario e inusual de lo que realmente es. Esto propicia que la audiencia se vea identificada con los propietarios o vecinos, más que con las personas ocupadoras. Este sentimiento de identificación también responde a un discurso en el que señalan que “nos puede pasar a cualquiera”, cuando lo cierto es que lo más habitual es que le ocurra a entidades bancarias o multipropietarios, que no son *cualquiera*, sino una parte muy reducida de la sociedad.

Esta cuestión enlaza con el tercer ingrediente que Cohen (2017) considera necesario para lograr un caso exitoso de pánico moral: generar un consenso acerca de que no se trata de un problema aislado sino de una problemática que compete a toda la sociedad. Así, cabría preguntarse si las personas a las que los medios dirigen una representación de la ocupación como un peligro generalizado que puede afectar a cualquier ciudadano/a, ¿no están más cerca de encontrarse en una situación de vulnerabilidad, que no de disponer de más de una vivienda y que esta sea ocupada?

Por último, conviene resaltar las limitaciones que presenta este artículo, restringido a un período y una muestra muy concreta de dos cadenas de televisión. Por ello, se podría continuar investigando este fenómeno ampliándolo, no solo a otros programas, cadenas de televisión o periodos, sino también a otros medios de comunicación.

Por ejemplo, se podría comprobar si este enfoque de suceso criminal responde a la propia lógica de la televisión que prioriza la espectacularización, el interés humano o el sensacionalismo, y si existe otro tipo de tratamiento informativo en la radio y la prensa. Además, se podrían realizar otros trabajos basados en la teoría de la *agenda setting* que analicen los posibles efectos que esta cobertura de pánico moral ha tenido en la agenda pública a través de estudios de audiencia o recepción, así como la presencia de esta cuestión en los discursos políticos y analizar sus características.

REFERENCIAS

- Alonso, N. (2021): "¿Cuánto se destina del salario al alquiler de vivienda en España?". Disponible en: <https://www.newtral.es/cuanto-se-destina-salario-sueldo-alquiler-vivienda-casa-espana-70-40-30/20210405/>
- Amadeo, B. y Aruguete, N. (2012): "Encuadrando el delito: pánico moral en los periódicos argentinos", en *América Latina Hoy*, nº 62, pp. 177-196. <https://doi.org/10.14201/alh.9350>
- Amadeo, B. y Aruguete, N. (2013): "Medios y miedos. La cobertura de la inseguridad en la Argentina", en *Revista de Investigación del Departamento de Humanidades y Ciencias Sociales*, vol. 1, nº 3, p. 14-31. <https://rihumso.unlam.edu.ar/index.php/humanidades/article/view/27/80>
- Ardèvol-Abreu, A. (2015): "Framing o teoría del encuadre en comunicación. Orígenes, desarrollo y panorama actual en España", en *Revista Latina de Comunicación Social*, nº 70, pp. 423-450. <http://www.revistalatinacs.org/070/paper/1053/23es.html>
- Bauman, Z. (2007): *Miedo líquido. La sociedad contemporánea y sus temores*. Barcelona: Paidós.
- Beck, U. (2006): *La sociedad del riesgo: hacia una nueva modernidad*. Barcelona: Paidós.
- Cohen, S. (2017): *Demonios Populares y "pánicos morales"*. Barcelona: Gedisa.
- Danieli, N. E. y Paz García, A. P. (2020): "Opinión pública de la inseguridad: dimensiones de análisis psicosocial en torno al miedo al delito y su tratamiento mediático", en *Comunicación y Sociedad*, nº 17, pp. 1-23. <https://doi.org/10.32870/cys.v2020.7586>
- D'Adamo, O. & García Beaudoux, B. (2007). "Medios de comunicación de masas y percepción social de la inseguridad", en *Boletín de Psicología*, nº 90, pp. 19-32. <https://www.uv.es/seoane/boletin/previos/N90.htm>
- Dee, E. T. C. (2016). "The Production of Squatters as Folk Devils: Analysis of a Moral Panic that Facilitated the Criminalization of Squatting in the Netherlands", en *Deviant Behavior*, vol. 37, nº 7, pp. 784-794. <https://doi.org/10.1080/01639625.2016.1145019>
- Fernández, J. J. (2020): "Los jueces no ven necesario endurecer las penas contra los ocupas". Disponible en: <https://www.elperiodico.com/es/sociedad/20200824/ocupas-jueces-endurecer-penas-808554>
- Fotocasa (2019): "Radiografía del mercado de la vivienda 2018-2019". Disponible en <https://prensa.fotocasa.es/wp-content/uploads/2019/05/Radiograf%C3%ADa-del-mercado-de-la-vivienda-en-2018-2019.pdf>
- Francés, M. (2020): "Joaquim Bosch, magistrado: El plazo de 48 horas para desalojar una vivienda okupada es una leyenda urbana". Disponible en: <https://www.niusdiario.com>

[es/nacional/tribunales/joaquim-bosch-magistrado-anos-no-conocido-okupacion-vivienda-okupa_18_3002220183.html](https://www.boe.es/nacional/tribunales/joaquim-bosch-magistrado-anos-no-conocido-okupacion-vivienda-okupa_18_3002220183.html)

Gil Calvo, E. (2003). *El miedo es el mensaje. Riesgo incertidumbre y medios de comunicación*. Madrid: Alianza Editorial.

Institut Cerdà. (2017): "La ocupación ilegal: realidad social, urbana y económica... un problema que necesita solución". Disponible en: <https://www.icerda.org/es/producto/la-okupacion-ilegal-realidad-social-urbana-y-economica#un-problema-que-necesita-solucion-pd>

Ley Orgánica 10/1995, de 23 de noviembre de 1995, del Código Penal. Boletín Oficial del Estado, nº 281, de 24 de noviembre de 1995. Art. 202, p. 34011. <https://www.boe.es/boe/dias/1995/11/24/pdfs/A33987-34058.pdf>

Ley Orgánica 10/1995, de 23 de noviembre de 1995, del Código Penal. Boletín Oficial del Estado, nº 281, de 24 de noviembre de 1995. Art. 245, p. 34015. <https://www.boe.es/boe/dias/1995/11/24/pdfs/A33987-34058.pdf>

Madden, D. y Marcuse, P. (2016): *En defensa de la vivienda*. Madrid: Capitan Swing

Morán San Juan, L. (2018): "El relato del miedo en las noticias de televisión: análisis de los informativos de TVE 1 y Telecinco". [Tesis doctoral, Universidad Complutense de Madrid], en *E-Prints Complutense*. <https://eprints.ucm.es/id/eprint/47289/>

Obra Social Barcelona. (2018): "Informe sobre la ocupación de vivienda vacía en Catalunya". Disponible en: https://obrasocialbcn.net/wp-content/uploads/2018/10/Informe_castellano_portada.pdf

Palomera, J. y Virgós, D. (2019): "Okupas: el gran montaje de la derecha populista". Disponible en: <https://ctxt.es/es/20190522/Firmas/26344/Jaime-Palomera-Diana-Virgos-casa-okupas-03-desahucios-especulacion-derechas.htm>

Semetko, H. y Valkenburg, P. (2000): "Framing European politics: a content analysis of press and television news", en *Journal of Communication*, nº 50, pp. 93-109. <https://doi.org/10.1111/j.1460-2466.2000.tb02843.x>

Valera Ordaz, L. (2016). "El sesgo mediocéntrico del framing en España: una revisión crítica de la aplicación de la teoría del encuadre en los estudios de comunicación", en *ZER: Revista de Estudios de Comunicación*, vol. 21, nº 41, pp. 13-31. <https://doi.org/10.1387/zer.17259>


Villena, M. (2020): "El inexistente plazo de 48 horas: la policía no tiene límite de tiempo para intervenir en una okupación". Disponible en: https://verne.elpais.com/verne/2020/09/01/articulo/1598938999_179935.html

Walsh, J. P. (2017): "Moral panics by design: The case of terrorism", en *Current Sociology*, vol. 65, nº 5, 643-662. <https://doi.org/10.1177/0011392116633257>



The Story Behind the Story: Reporting Conditions in China During the Coronavirus Crisis as Told by the Correspondents of El País and ABC

La historia detrás de la historia: Las condiciones de trabajo periodístico en China durante la crisis del coronavirus contada por los corresponsales de El País y ABC

 **Adrià Calatayud-Vaello**
adria.calatayud01@estudiant.upf.edu
Universitat Pompeu Fabra

ABSTRACT This study seeks to investigate the reporting conditions in the early coverage of the coronavirus crisis by the China correspondents of two Spanish newspapers: El País and ABC. After conducting in-depth, semi-structured interviews with the correspondents, this study finds that they only started to focus on the viral outbreak once Wuhan city had been locked down. The interviewed correspondents faced logistical, political, professional and personal challenges while covering coronavirus in China due to the impossibility of entering Wuhan, censorship, difficulties in finding people that would speak to them and the general uncertainty of the situation, respectively. They also showed diverging perceptions with regards to the implications of their role for intercultural dialogue. The present study helps to contextualize foreign correspondents' coverage of Covid-19 in China by shedding light on the factors that constrained their professional practice.

KEYWORDS Coronavirus, Covid-19, foreign correspondents, reporting conditions, China, censorship

RESUMEN Este estudio busca investigar las condiciones de trabajo durante la cobertura inicial de la crisis del coronavirus por parte de los corresponsales en China de dos diarios españoles: El País y ABC. Después de realizar entrevistas en profundidad semiestructuradas con los corresponsales, este estudio muestra que los corresponsales solo empezaron a centrarse en el brote de coronavirus cuando la ciudad de Wuhan fue confinada. Los corresponsales entrevistados se encontraron con barreras de tipo logístico, político, profesional y personal en su cobertura de la crisis sanitaria, por la imposibilidad de entrar en Wuhan, la censura, la dificultad para encontrar fuentes y la incertidumbre general de la situación, respectivamente. Asimismo, mostraron percepciones diferentes en lo que respecta a las implicaciones de su trabajo para el diálogo intercultural. El presente estudio ayuda a contextualizar la cobertura de la Covid-19 en China por parte de los corresponsales internacionales al arrojar luz sobre los factores que constriñeron su práctica profesional.

PALABRAS CLAVE

Coronavirus, covid-19, corresponsales internacionales, condiciones laborales, China, censura

Edita: Lorena Cano-Orón

Recibido: 05/12/2021 | Aceptado: 27/03/2022

Cómo citar este artículo:

Calatayud-Vaello, A. (2022). The Story Behind the Story: Reporting Conditions in China During the Coronavirus Crisis as Told by the Correspondents of El País and ABC. *Dígitos. Revista de Comunicación Digital*, 8: 273-293. DOI: 10.7203/rd.v1i8.231



The Story Behind the Story: Reporting Conditions in China During the Coronavirus Crisis as Told by the Correspondents of El País and ABC

1. Introduction

Journalists tell stories, but there is always a story behind the story. With regards to the Covid-19 pandemic, researchers have studied what specific stories journalists told, when and where; who was featured in them; how reporters told those stories, and why they might have done it that way. This study aims to go back to the origins of the pandemic and tell the story behind that story. It seeks to investigate the reporting conditions in the early coverage of the coronavirus crisis in China as experienced by the correspondents there of two Spanish newspapers, El País and ABC.

This study queries the correspondents of El País and ABC about their experiences, challenges and regrets in the early days of their reporting on Covid-19 in China. This is achieved through in-depth semi-structured interviews with Macarena Vidal Liy, China correspondent for El País, and Pablo M. Díez, ABC's correspondent in Asia.

From a broader perspective, this study seeks to link foreign correspondence with intercultural communication. Do foreign correspondents act as intercultural communicators or as constructors of foreign realities? Thus, another key element to be uncovered through the interviews is correspondents' approach to intercultural communication, focusing on their identity, skills and attitude toward interculturality.

In order to better understand the findings of this research, however, it is necessary first to summarize what research on media coverage of Covid-19 has found so far, focusing on Spanish media and on media coverage of the pandemic in China, and to offer an overview of the reporting conditions in China. Complementing this literature review, some brief theoretical notes portray foreign correspondence in the framework of intercultural communication. The goals of this study and the methodology used are outlined next, before delving deeper into the findings. Finally, a conclusion encapsulates the main takeaways from this research and compares it to previous studies.

2. Literature review on media coverage of Covid-19

Media coverage of Covid-19 has drawn attention from the scholarly community and has been studied from many angles. There have been studies looking at the health consequences of media exposure, linking the use of stigmatizing language in coronavirus media coverage in the U.S. to discrimination and bias toward Asian Americans, examining the social implications of labeling choices in news headlines or analyzing the use of metaphors—a classic topic in the academic literature on media coverage of diseases—

to refer to health workers (Garfin, Silver, and Holman, 2020; Darling-Hammond et al., 2020; Prieto-Ramos, Pei, and Cheng, 2020; Mohammed, Peter, Killackey, and Maciver, 2021).

With regards to media framing of Covid-19, Gui (2021) observed the deployment of the metaphorical war fame dominated news reports of Chinese mainstream media, although its meaning evolved over time and went from preventive war in the early days to people's war and total war at the height of the crisis. The evolving use of these metaphors entailed diverging expectations from the public's role in containing the spread of the virus and might have contribute to make people more willing to accept sacrifices in order to control the outbreak (Gui, 2021). Fernández-Pedemonte, Casillo and Jorge-Artigau (2020) noted that metaphors of war, care and time were predominant in the framings of the pandemic by digital media in Argentina. They found that each metaphor was promoted by a particular group of enunciators —war by politicians and economic experts, care by health specialists and the media, and time by health institutions and corporations— and that their use was linked to a point of view on the phenomenon that defined the political and social relationships activated by the pandemic (Fernández-Pedemonte et al., 2020).

In the case coverage of Covid-19 by Spanish media —the media system to which the objects of study of this research belong—, there is a considerable body of literature as well. Argiñano and Goikoetxea Bilbao (2020) analyzed the headlines and pictures in the front pages of five dailies from January 2020 to April 2020 and found that coronavirus was treated as a foreign matter in the first two months of their sample and flooded newspapers' front pages once it became a Spanish issue in March. Picazo Sánchez, de Frutos Torres and Gutiérrez Martín (2020) concluded in their analysis of the framings of alarm and reassurance used in the front pages of three daily newspapers that their framing choices were based on each outlet's ideological orientation toward the Spanish government rather than on the news events they reported or health reasons.

Another study observed that civil society and Spain's government were the subjects that attracted the most headlines in the coverage of the outbreak and initial expansion of Covid-19 by Spanish digital media (Abuín Vences, Sierra Sánchez, Mañas-Viniegra, and Núñez Gómez, 2020). Moreover, Iranzo-Cabrera and López-García (2021) studied how mass media used leading scientific figures' statements as sources of information during the first wave of Covid-19 in Spain to provide certainty in an uncertain context, but also paid attention to their aesthetics and communication style.

Casero-Ripollés's research showed that audiences reconnected to legacy media, especially television, during the height of the Covid-19 pandemic, reducing inequalities in media consumption and revitalizing the public sphere (Casero-Ripollés, 2020). This suggests that, at least in the context of a major health crisis, the public turned away from social media and toward established media outlets with a long trajectory (Casero-Ripollés, 2020). Still, Sanjuán-Pérez et al. found that content on the coronavirus epidemic in the 24 hours of free-to-air Spanish television was slow to pay attention to Covid-19, as references to the disease were not significant until March, when some people infected with the virus had already died in Spain and death counts were spiking

in Italy (Sanjuán-Pérez et al., 2020). This evidences that Spanish generalist broadcasters failed to anticipate the effects of the pandemic until they were already visible (Sanjuán-Pérez et al., 2020).

As for media coverage of coronavirus in China—which constitutes the object of study of this research—, there have been some studies too. China's response to Covid-19 was depicted by Spanish mainstream newspapers as slow and passive initially and as draconian and bold following the lockdown of Wuhan, with criticisms of censorship and data manipulation and allegations of a cover-up of the first cases, according to Huo and Sautman (Huo and Sautman, 2021). However, these depictions did not make Spanish citizens' perception of China significantly more negative and newspapers' coverage of Covid-19 in China did not alter the image of China in Spain in a meaningful way (Huo and Sautman, 2021).

Zhang and Shaw looked at the coverage of China's handling of the coronavirus outbreak by three U.K. outlets (Zhang and Shaw, 2020). Taking articles published by the BBC, The Guardian and the Daily Mail over a 20-week period from January to May, Zhang and Shaw found that the three outlets covered similar topics and coincided in reporting coronavirus as a threat, although some topics as well as the sentiment they conveyed were different: the BBC tended to be neutral, with fewer direct criticism of China and absent praise; the Daily Mail packaged stories in a more sensational way and paid attention to social topics; while The Guardian was somewhat in the middle in terms of emotion and put human rights at the center of its coverage of China's handling of the crisis (Zhang and Shaw, 2020).

Another study undertook a critical discourse analysis of China's official media discourse during the Covid-19 epidemic by focusing on their tilting toward nationalist and globalist stances (Yang and Chen, 2020). While both globalism and nationalism were present, China's official media was seen as polarized in a nationalistic way. As Yang and Chen put it, "[...] China's media representation of the COVID-19 outbreak is characterized by a deliberate construction of a positive self-representation and a negative other-representation based on binary or antagonistic opposites of us versus them" (Yang and Chen, 2020). Yang and Chen's study showed that Chinese official media discourse sought to legitimize the Chinese Communist Party and defend the international reputation of China, pointing to a political influence on news reports that was also noted by Gui regarding the use of war metaphors (Yang and Chen, 2020; Gui, 2021).

Thomas, Wilson, Tonkin, Miller and Ward analyzed how Australian media attributed responsibility for the Covid-19 pandemic and observed that explicit blame was sparse, indirect and infrequent during the period they studied, ending in March 2020 (Thomas et al., 2020). While the virus was frequently described as originating from China, it was only at the end of their sample that direct criticism toward the response from the Chinese government was deployed, these researchers found (Thomas et al., 2020). In another analysis of Australian media coverage of China's efforts to contain Covid-19, Sun observed that there was a convergence across the ownership—public and commercial media—and ideological—liberal and conservative—spectrum around unfavorable reporting about China in relation to Covid-19 (Sun, 2021).

Hubner's analysis of the reporting on Covid-19 by major U.S. newspapers revealed that their early news coverage focused on the spread of the disease and subsequent detrimental effect on society (Hubner, 2021). This focus on the societal impact rather than on the health risks might have established Covid-19 as a threat to one's way of life rather than a serious threat to one's health, Hubner (2021) wrote. Meanwhile, Jia and Lu (2021) also observed an alignment between coverage of China's handling of Covid-19 by U.S. media and the strategy of the Trump administration toward Beijing.

Nevertheless, the most relevant piece of research for the purposes of this study is a critical discourse analysis of the coverage of the coronavirus crisis in China by the correspondents of *El País* and *ABC* previously carried out by the author (Calatayud Vaello, 2021). Since the present study follows on from an explicit recommendation outlined in that article—to interview correspondents who covered coronavirus for those newspapers to learn about their experiences—and builds on its findings, it is important to recall the main conclusions from that paper were that Chinese authorities had an overarching presence in the coverage by correspondents of both newspapers, as social actors represented in their stories, as sources of information and as agents of the story lines that gave context to their coverage, such as: lack of transparency, lack of freedoms, government control, and Xi's accumulation of power (Calatayud Vaello, 2021). By contrast, Covid-19 patients attracted comparatively little attention, with the exception of Li Wenliang, an ophthalmologist from Wuhan who died from Covid-19 and who was portrayed as a symbol of the health crisis because he warned his colleagues about the outbreak and was reprimanded by police for it. That study found that correspondents' reporting focused on the Chinese government's response to the crisis rather than on the disease or the health situation, and that Chinese authorities were the main social actors found in their coverage and were represented in negative terms (Calatayud Vaello, 2021). As a result, coronavirus patients had a minor role in these correspondents' coverage and were usually represented as speechless statistics and, thus, othered. Moreover, it found a marked presence of references to Chinese culture. "Even when dealing with a public health issue, there is a notable cultural dimension in the articles of the China correspondents of *El País* and *ABC*. After all, they may be covering an epidemic, but they are still covering China" (Calatayud Vaello, 2021: 75).

3. Reporting conditions in China

In an analysis of the voices shaping the image of China portrayed by five weekly newspapers from the U.K., the U.S., Germany and France, Mokry found that access to sources, or lack thereof, due to China's reporting conditions was crucial (Mokry, 2017: 662):

"Journalists seem to be aware of these effects and try to mitigate them by drawing on diverse strategies: when access to sources within the government proves impossible, they quote Party-state media. When Chinese scholars refuse to talk to them, they turn to foreign experts. In cases where state institutions intimidate individual sources, foreign journalists draw on different measures of source protection, for example quoting sources anonymously by changing their name, location and gender."

Sun (2015) also put under the spotlight the often tense relations between China's authorities and foreign correspondents based there. Sun showed that China's public diplomacy efforts have not been fruitful in courting foreign correspondents due to, among other factors, the Chinese government's inability to understand how foreign correspondents work (Sun, 2015).

Zeng (2018) looked into the institutional constraints under which foreign correspondents work in China and concluded that state coercion is a major source of external pressure on the field of foreign correspondence. Zeng noted that blocking access to certain areas and information channels was the most common obstacle correspondents had to clear to do their job, but there were instances of surveillance, harassment and intimidation as well (Zeng, 2018). Although correspondents are aware of these challenges and try to anticipate and negotiate them, none of their tactics is enough to address the structural coercion (Zeng, 2018).

In this regard, the annual reports elaborated by the Foreign Correspondents' Club of China, or FCCC, can provide insights on the reporting conditions in the country too. The FCCC report from 2021 denounced that at least 20 foreign journalists had been expelled from the country the prior year, that Chinese nationals working for foreign media faced increasing pressure and that the coronavirus pandemic —and associated public-health measures— became another way to control and limit foreign journalists, as part of a broader decline in media freedom in China (FCCC, 2021). According to the report, none of the correspondents surveyed by the club responded that working conditions in China had improved during the previous year (FCCC, 2021).

4. Foreign correspondents within intercultural communication

This study aspires to shed some light on foreign correspondents' role within intercultural communication. The author has outlined elsewhere a more comprehensive approach to a theorization of foreign correspondents within intercultural communication (Calatayud Vaello, 2019). At the very least, the links between the field of foreign correspondence and intercultural communication can be placed at two levels: at the interpersonal level, correspondents engage in intercultural communication processes; at the mediated level, their work may contribute to shaping intercultural expectations in their audiences.

A previous paper by the author cited above already explored the intercultural dimension of the discourses produced by the correspondents contemplated in this study about the coronavirus in China (Calatayud Vaello, 2021). By reviewing correspondents' role as participants in intercultural communication processes, this study can help understand the approach to intercultural communication adopted by the correspondents of *El País* and *ABC*.

Two concepts that have been developed by intercultural communication studies seem relevant for understanding foreign correspondents' condition as participants in processes of intercultural communication: identity and competence. While the former enables an examination on the socialization correspondents undergo once they are dispatched to their foreign postings, as newcomers, the latter looks at the skills participants in intercultural communication acquire.

With regards to identity, Kim's cross-cultural adaptation theory and Casmir's third-culture building theory can be used to gauge the degree to which correspondents' identity has changed as a result of their intercultural socialization (Kim, 2009; Casmir, 1997). Kim describes the intercultural transformations experienced by individuals that have completed their primary socialization in a culture and then have prolonged first-hand contact with a new culture (Kim, 2009). In contrast, Casmir's theory is purposive as it aims to be used by intercultural communication practitioners in real-life situations, not just for theoretical purposes in academic environments (Casmir, 1997). When it comes to competence, Chen's intercultural communication competence theory can aid in understanding the skills developed by correspondents (Chen, 2009).

5. Objectives and methodology

This study aims to understand the reporting conditions foreign correspondents encountered in China during their coverage of the Covid-19 outbreak specifically and, from a broader perspective, their approach to intercultural communication as correspondents. It focuses on the China correspondents of *El País* and *ABC*, as these two correspondents stayed in mainland China during the first quarter of 2020 —the peak of the health crisis in that country— and work for two generalist outlets that were selected due to their large circulation within Spain.

In order to achieve that general goal, this paper aspires to answer the following research questions:

- RQ1: How did the studied correspondents approach their coverage of the coronavirus crisis?
- RQ2: What obstacles did correspondents find during their coverage of the coronavirus crisis?
- RQ3: What skills have foreign correspondents developed to engage in intercultural communication processes?
- RQ4: Do foreign correspondents try to foster an intercultural dialogue through their work?
- RQ5: Are reporting conditions reflected somehow in correspondents' coverage of the crisis?

In order to answer those research questions, this study used a qualitative method: semi-structured interviews. This method is appropriate for the purposes of this study, as it is useful to gain an understanding of key practitioners within a given field (Ahlin, 2019). In journalism studies, semi-structured interviews have previously been deployed to research exchanges between media practitioners, news production practices, perceptions on reporting approaches, foreign correspondents' daily work routines or their interactions with sources, to name a few (Archetti, 2011; Cheng and Lee, 2015; Cooper, 2018; Madrid-Morales and Wasserman, 2018; Ouariachi and Peralta, 2021; van Antwerpen, Turnbull, and Searston, 2022; Zhang and Zhang, 2018).

The preparation of the interview questionnaire for this study followed the phases outlined by Kallio et al. (2016). Table 1 shows the questionnaire used as a starting point in the interviews:

Table 1: Questions for interviews with China correspondents who covered coronavirus

1 st Block: Coverage of the coronavirus crisis	
When did you first hear about the coronavirus?	
How did you plan your coverage? Was it more a matter of following events as they developed, or did you think about a series of topics you wanted to address?	
What difficulties did you find in your coverage?	
Did you have any problems finding sources?	
Would you attribute any of these difficulties to cultural elements?	
Did people around you affect your coverage in any way?	
In hindsight, is there anything you would have done differently or written differently?	
Did your approach to the coverage change at the time? Has it changed since then?	
2 nd Block: Intercultural communication	
How would you define your cultural identity?	
Being a foreign correspondent, you have to navigate between different cultures. Does that change the way you perceive them? How?	
Would you say that you have acquired intercultural competence by being a foreign correspondent?	
To obtain information while working as a foreign correspondent, do you use intercultural competence?	
In the stories you write, do you try to favor an intercultural dialogue?	
In your opinion, which of the two options below defines better what a foreign correspondent is? Why?	- An intercultural communicator
	- A constructor of foreign realities

Source: Own elaboration

Once the questionnaire had been prepared and tested, the next move was to reach out to correspondents to request the interviews. Given that the correspondents were in China and it was not possible to go there to conduct in-person interviews, the interviews were conducted using WeChat's videocall function. Both interviews took place in January 2021 and were recorded. The interview with ABC's correspondent lasted 1 hour and 18 minutes and the interview with El País's correspondent, 1 hour and 23 minutes.

During each interview, the questionnaire was followed with only small variations. Occasionally, follow-up questions were asked. Having done the interviews, the last step before analyzing the contents obtained from correspondents' responses was the transcription.

6. Findings

As for the findings, the contents generated from the interviews have been presented under two sections: their experiences on the coverage of coronavirus and their views on intercultural communication. Wherever there is a coincidence or pattern in correspondents' responses, that is singled out. The same happens where there is disagreement between them. Those observations are supplemented and supported with relevant quotes that illustrate correspondents' thinking. In order to shorten references to the interviewees, El País's correspondent will be referred to as EPC and ABC's correspondent, as ABCC.

6.1. Coverage of coronavirus crisis in China

6.1.1. "It caught everyone off guard": Covid-19 breaks out

Both EPC and ABCC were outside of China when they first heard about the new coronavirus. For both, the initial news was when the Chinese government informed the World Health Organization about a viral outbreak affecting Wuhan on December 31, 2019. "I think it caught everyone off guard because of the date, just on New Year's Eve, and with most correspondents who cover China on vacation abroad," ABCC said.

On their return to Asia after the Christmas break, both correspondents went straight to Taiwan to cover the presidential election which was held on January 11, 2020. After Taiwan's vote, EPC went to Vietnam on a reporting trip and only returned to Beijing at the end of January, while ABCC headed to Shanghai. The correspondents coincided in highlighting the difficulty of making calls about the seriousness of the disease in the early days and pointed to another factor: recent precedents of infectious outbreaks that had ended up being lighter than they initially seemed. At the risk of sounding false alarms, the correspondents thought, news of the new virus had to be viewed with a grain of salt.

They had historical precedents in mind. EPC said she thought it would be a something like the MERS outbreak South Korea suffered in 2015. ABCC explained that he regarded a situation like SARS as the worst-case scenario that, hopefully, would not materialize:

"At the beginning —at least I talk from my own experience— you have that fear that it could be something like SARS. In any case, the worst you can think of is that this is going to be like SARS. That is what I thought at first. But since during the first weeks China doesn't say that there are infections among humans, that it doesn't seem likely or that, if there are, they are very limited, I believe at the beginning we didn't give it the importance it later had, because we didn't see it with that information."

Indeed, things would soon take a turn for the worse. That meant that new decisions over how to proceed with their coverage of the unfolding crisis awaited foreign correspondents.

6.1.2. "The biggest regret of my career": Wuhan calling

The first important decision in the coverage of the coronavirus crisis was whether to go to Wuhan or not. EPC did not go to Wuhan, but a colleague did. On the day EPC came back

to Beijing from the reporting trip to Vietnam, a colleague had been told to go to Wuhan, just before the city was locked down. Nevertheless, EPC's colleague was evacuated from Wuhan days later, which meant that EPC had to assume the responsibility of covering Wuhan's lockdown from Beijing.

In ABCC's case, he finally made up his mind to go to Wuhan, but it was too late. The lockdown of the city had already been announced, and ABCC had missed the chance to get in. For ABCC, not being able to enter Wuhan represented a "failure." "That will be, I think, the biggest regret of my career," he said.

ABCC and EPC agree on pointing to January 23, 2020 the day Wuhan was locked down, as the turning point. Up until that day, the dimension of the crisis had not been clear for EPC. Out of China for weeks, EPC had been reassured by the information given by the Chinese government, so that she did not think that the situation could get out of hand the way it did.

ABCC, by contrast, acknowledged that he had been in a wait-and-see mode until January 23. From that point, ABCC focused on following the news flow out of Wuhan, and on reporting the situation in other parts of China. The coverage took a heavy toll on ABCC. There were days when he had to file several stories until late at night and wake up after a few hours' sleep the following only to check out the latest figures of infections and deaths.

6.1.3. "We were all very afraid": Uncertainty and stress

One issue that came up in the interviews was the emotional toll that the coronavirus crisis—that is, living in it, on top of reporting on it—left on the foreign correspondents. Uncertainty, stress, worries about their health and that of relatives, and physical fatigue were all part of the coverage as well.

Although Wuhan suffered the biggest coronavirus outbreak in China, EPC explained there so there were restrictions in Beijing too, as well as concerns about whether things could complicate further. "In my case, I spent 59 days in a row working without being able to take even one day off. I only had three half-days off that I requested, to be able to go buy groceries," EPC said. Adding to the fatigue, EPC was left alone for months, as their relatives left China following advice from their home government. ABCC admitted that he suffered from fear and uncertainty. At that time, it was not known what the disease was like, how deadly it could be, and how people could get infected. "The truth is we were all very afraid, and that affected the approach to the news. It wasn't only about telling what happened, it was also about what happened to you and how your life had been blown up at that time," ABCC said. Having reported on several natural disasters and revolutions in Asia, ABCC admitted that what worried them most about coronavirus was the fact that it could reach their own country, where ABCC feared for their family. "Of all the stories I have covered here, this is the one I like the least because it has been the most anxious one, the one that has scared me the most, the one I have had the worst time with," ABCC said.

6.1.4. “There wasn’t time to think”: How correspondents approached the crisis

In terms of their approach to coverage of the crisis, there was a key difference between the two correspondents, at least at the beginning: EPC had a colleague in Wuhan, while ABCC was on his own in Shanghai.

For EPC, the last week of January was all about “putting fires out and patching things up,” trying to understand the dimension of the crisis. That involved reading the local press—with a degree of skepticism—and talking to people, EPC explained. Moreover, EPC had to coordinate from Beijing with a colleague who was in Wuhan. A division of labor between the two was implemented: EPC would focus on the more newsworthy aspects—like the building of makeshift hospitals in Wuhan, figures of infections and deaths, shortages of equipment, cases of discrimination against Hubei residents in other cities or censorship—, while the colleague explained the situation on the ground. According to EPC, planning anything on the coverage was challenging:

“There wasn’t much time to think or to plan anything because the accumulation of events was so big that we could barely report them. Doing that was enough. Planning big stories... was just not possible. A lot of things happened those days, and it all culminated with the death of physician Li Wenliang.”

Once ABCC was fully focused on covering the coronavirus crisis and realized that he would not be able to get into Wuhan, he tried to warn readers at home of what was about to come. To explain how China had adapted to the new circumstances brought about by the epidemic, as an advance of what awaited other parts of the world, was an objective for ABCC. “For me, reporting on coronavirus was a personal crusade,” ABCC said. That effort eventually had rewards for ABCC:

“I have been reached out by people to thank me because, while governments downplayed coronavirus, thanks to the reports we filed from China, they could protect themselves. That is the greatest satisfaction that your work can give you, especially in such a tough coverage.”

However, both EPC and ABCC highlighted that they had to overcome many obstacles, in addition to the uncertainty and stress mentioned above, in their coverage of the coronavirus crisis in China.

6.1.5. “People were quite afraid to talk”: Difficulties in the coverage

Being barred from entering Wuhan, little transparency from the Chinese government, censorship of Chinese media, fake information, tensions with authorities and Chinese people’s reluctance to speak to foreign media were some of the top difficulties singled out by the correspondents of El País and ABCC in the interviews.

For EPC, the main difficulty was not being able to access Wuhan. EPC’s colleague managed to get into the city before the lockdown was introduced, but was evacuated a week later, leaving her in charge of covering the crisis from Beijing. “Even if you talked to people [in Wuhan] and they told you, you couldn’t verify at 100%. You could get an idea, but...” EPC recalled.

Moreover, EPC highlighted that, after a few days of relative freedom for Chinese media, authorities changed tack in the wake of Li Wenliang's death. According to EPC, the public anger caused by that event triggered a fierce wave of censorship:

"During the first days, Chinese media were able to report with relative freedom, but from then on that was cut. Everything moved to a much more benevolent narrative of what was happening, and it became more difficult to get people to talk to you, they started to be afraid to talk... As the situation improved and there were beds for the sick people, families that until then had been desperate and talked to anybody to see if anyone was able to help them get a bed was less willing to talk..."

The fact that the lockdown there became harsher from mid-February was another factor. "What people knew was little, only what they could see from their windows or what their friends told them via social media. It was difficult to access direct sources," EPC said. EPC also referred to a spike in fake information, and to the difficulties of verifying what circulated on social media. "In all that maze, trying to figure out what really happened was more difficult than it is under normal circumstances," she said.

EPC also described tensions with authorities. "The moment you start to ask you are being negative, not helping, and 'this isn't about blaming anyone, but about rowing all together in the same direction to solve it,'" EPC said.

ABCC coincided in pointing to the lack of transparency of the Chinese government as a key obstacle. The government did provide official information about the balance of figures. Nevertheless, ABCC never believed that the numbers of coronavirus infections of deaths from Wuhan given by authorities were accurate, and was frustrated by authorities' refusal to answer questions about that topic. Beyond that, ABCC found difficulties to dig deeper. "It is very difficult that unofficial sources from the Chinese government talk to a foreign journalist," ABCC said. Therefore, ABCC resorted to scientific studies to try to get a more comprehensive understanding of what the coronavirus was.

ABCC attributed people's reluctance to talk to propaganda disseminated by the Chinese government:

"This is a political issue. I have seen in the 16 years I have been here that it is increasingly hard to work for journalists. Chinese people have more reluctance, fear, and apprehension to Westerners and, obviously, to journalists because the regime's propaganda has labeled us as enemies of China, even if what we are is critics of a government, not enemies of any country. We are journalists and our mission is to report what happens, denounce things that aren't right so they can be fixed, and that is what we do in our own country as well. But that isn't understood here, because the media's mission is propaganda."

EPC coincided with ABCC in pointing out that Chinese propaganda against foreign media did not help them in their coverage of coronavirus. To overcome that reluctance to speak to foreigners, having a Chinese fixer was crucial. "If the assistant goes on his own, it's easier that people open up," she insisted.

While correspondents faced diverse difficulties during their coverage of the coronavirus in China, they also had some regrets about what they did and how their stories were

received back at home.

6.1.6. "If I had known what I know now...": Correspondents' regrets

EPC wished she had known about the potential danger of the disease to warn it was more serious than SARS. "We all thought that it would be a matter of a few months and it would stay limited to a geographic area more or less large but limited there, and that by the summer this would be over. But we didn't know," she said. On a more personal level, EPC said their stories would have been different if she had had at the time the information she later learned.

For ABCC, the biggest regret is a different one: he wished he had booked a flight ticket to Wuhan earlier to be able to cover the city's lockdown from inside. However, both ABCC and EPC showed a certain feeling of being ignored by their editors and audiences. EPC explained how once Wuhan and Hubei were put under lockdown, she realized that the virus was likely to spread all over the world, but her worries were met with a prejudiced "cultural self-indulgence":

"I talked to senior editors and they were like: 'Can you imagine that happening here?' 'No, that would be controlled quickly here.' Maybe we all fell for a certain degree of cultural self-indulgence. Since this was in China... It was like: 'In China, this can happen, because people are Chinese, hospitals are worse, but we are the first world and we are going to control it well.'"

ABCC emphasized how he tried to sound the alarm in their stories about the severity of the disease and its potential danger, and how some readers eventually expressed their gratitude for that. At the same time, ABCC believes not all the warnings were listened to. "At the beginning, we were accused of being alarmists for explaining what happened here," ABCC lamented. ABCC insisted that he tried to relay a sense of urgency and to give the impression that the outbreak could reach his home country as well. "You realize that, actually, people think catastrophes won't occur to them, but they can happen. This is what the coronavirus demonstrated," ABCC said.

6.2. Correspondents' views on intercultural communication

6.2.1. "Who am I?": Identity

To induce the foreign correspondents to do some introspection about culture and interculturality, the block of the interviews focused on intercultural communication began with a personal question about their cultural identity.

"My cultural background is so mixed," EPC said. And EPC elaborated: she was born and raised in Spain but left more than half of their life ago, so EPC has spent more than half of their life abroad in different places. The cultural mix does not end there. "My mother is Chinese-Cuban, we have also that part of Cuban culture and Chinese culture. I have that in my DNA, I've had it breast-fed. I married to a Briton, we've lived most of our marriage in the U.S., now in China," EPC added. That made EPC conclude: "Where I'm from or where I feel from, I wouldn't know what to say. I don't know." After being asked whether a citizen of the world would be an answer, EPC hesitated somewhat before going for it: "I'm not very [from the country where EPC was born], but here certainly I'm

not Chinese. In the U.S., I'm an alien.... Who am I? I don't know. Citizen of the world, I guess, a bit from everywhere."

Meanwhile, ABCC gave multiple answers, which indicate a multi-layered identity or a feeling of multiple belongings to different cultural groups. But ABCC added an important caveat: "After 16 years [in Asia], and before I had traveled a lot around the world, I want to have a global vision of society and of the world we live in. Of course, you are always marked by your culture and your education."

ABCC explained that, in China, he felt like someone different. "Here, I know I am a Westerner in China. For being a Westerner in China, only by getting on the street, especially here in China that people are very curious, a Westerner is someone different, right?" ABCC said.

6.2.2. "Things are neither black nor white": Adaptation to the host country

An important aspect to grasp correspondents from the perspective of intercultural communication is to understand how they adapted to their host country, which in this case was China.

EPC explained that having a Chinese background made her adaptation to China easier. "Part of my family is Chinese, so there are cultural issues that are built in my DNA, so they don't shock me," EPC said. Moreover, working as a foreign correspondent allowed EPC to be exposed to different cultures, changing her view of the world in the process. "For sure, that helped me to be open to the fact that things can be done in a lot of ways. There is no single solution or one single response to things. Most often, things are neither black nor white, they have many nuances," EPC said. "Trying to explain those nuances is one of my obsessions as a journalist," she added.

ABCC explained that his condition of being culturally different can create communication problems, but has positive aspects as well. "Being seen as someone different also generates curiosity and can generate a certain empathy with people that approaches you and wants to talk to you," he said. ABCC defended that, by being a foreign correspondent, one gains humanity, a capability to understand perspectives that are different from their own, and empathy. However, ABCC argued that those are abilities all journalists need, regardless of where they work. "If you want people to talk to you, you have to be a forthcoming person, with empathy, able to win people's trust so that they tell you everything," ABCC said. Working in a foreign country, according to ABCC, boosted those abilities: "You have a translator to say whatever you want, but first, come your face and your eyes, and I believe that is a very important part." ABCC's reflection on cultural adaption has brought up language barriers, which are addressed fully next.

6.2.3. "You always lose some freshness": Language

The two correspondents had diverging views with regards to the importance of speaking the local language for their job, and their engagement in intercultural communication processes with local people.

ABCC downplayed the importance of language differences. "The linguistic differences

or differences that may exist can be easily solved with translators. That isn't a problem," ABCC said. According to ABCC, what matters most is to be able to find interviewees and people who tell good stories. Then, it is important to be able to engage with them and establish a trust relationship, which in ABCC's opinion, is more often achieved through non-verbal means. "Those things are a matter of good luck, certain karma, and also of having that kindness or empathy that here in Asia, [...] at least initially, are fundamental weapons to relate to people. It is part of that intercultural, non-verbal communication," ABCC said.

EPC mentioned "the problem of language" as one of the key limitations for working as a foreign correspondent. Contrary to ABCC, EPC was not particularly enthusiastic about resorting to translators. "Simply due to vocabulary problems, just to interview a doctor about the subtleties of coronavirus symptoms you need a translator. You always lose some freshness by needing a third person, however good they are," EPC said.

6.2.4. "You try to avoid going native": Intercultural communication competence

In addition to speaking language or finding a bridge language, correspondents need other skills to successfully participate in intercultural communication processes. Having an open mind, being culturally self-aware, empathy, treating different kinds of people differently, curiosity about other perspectives, learning new languages were the specific skills that came up in the interviews.

ABCC believes working as a journalist made him a more social being. Humanity, an open mind to understand different perspectives, empathy, forthcomingness are qualities ABCC claims to have acquired since he entered journalism in his home country and were boosted after moving abroad. "Being in contact with very different people makes you see, makes you understand, how each one of us handles their own human condition," ABCC said. ABCC claimed that the intercultural skillset needed to be a foreign correspondent had been integrated into his personality. "It's not that I use skills, I think I'm like that. My job has turned me this way," ABCC said.

EPC explained her multicultural upbringing meant intercultural competence was there almost from the crib. "I was raised in an environment of tolerance toward difference. In any case, being an international correspondent has fostered that feature that I had from the beginning. Of course, I've always been very keen on languages and I love to learn them," EPC expanded. However, EPC believes correspondents need to be cautious in their assimilation or they risk going native:

"Traveling around makes you keener on learning languages, and also wanting to learn everything about the culture of reception in different ways: go to eat local food, reading as much as possible about that place, try to adopt even hand gestures... [...] But then comes the correspondent's perspective to try to avoid what is called 'to go native'—to assimilate yourself so much that you lose objectivity and you see everything from the local point of view, without paying attention to the problems of the nuances that that culture may have."

6.2.5. Attitudes toward intercultural dialogue

When asked about whether they tried to favor an intercultural dialogue with their stories, the foreign correspondents interviewed were divided. Some of them held views that could not be more contradictory.

"I don't think it's a journalist's mission. I think a journalist's mission is to tell what happens, to give as much information as possible, as many perspectives as possible, and let the reader act accordingly," ABCC said. However, EPC replied to the same question by saying: "That's one of the pillars and the mission of a correspondent."

ABCC added that intercultural dialogue is not correspondents' job, as there is a risk that for fear of reporting something that might not be taken well, they miss important parts of the news, which would amount to self-censorship. "Almost all stories that we journalists publish are negative, which affect the image people have about a country. They affect the image people may have of China obviously, or of Asia... That's not something I can control," ABCC said.

Even though ABCC personally favors an intercultural dialogue, he does not believe that a correspondent's job is to foster it:

"What we have to do is to tell what there is as best as we can, and then people can dialogue if that's what they want. Obviously, I'm personally in favor of that intercultural dialogue, and of knowing other countries and knowing other societies and accepting the differences. I'm for that and try that my work contributes to show something else. I hope that's the case, but can't take for granted that that's the case either."

EPC disagreed with that. "We have a responsibility and a role in creating a narrative or an image, or allowing to understand the country we cover," EPC said. And she elaborated:

"I think a correspondent's role is to build bridges and to facilitate that the readers' culture—that of the country of the medium you work for— can better understand the culture of the country or the continent you are covering. I think we play a paramount role as the front line. When you are curious about a country, the first thing you do is check the news about that country."

Thus, the correspondents' views on the issue of intercultural dialogue are not just different, they are contradictory. It is important, though, to look closely at the arguments that support those confronted views.

ABCC was against the notion of foreign correspondents as facilitators of intercultural dialogue and argued that news takes priority, and that news should aspire to reflect reality as best as a reporter can. Other considerations—that is, considerations other than news such as preoccupation with fostering an intercultural dialogue— entail a risk of missing parts of the news, thus misrepresenting reality.

By contrast, EPC, who favors the view of foreign correspondents as facilitators of intercultural dialogue, defends that the way stories are framed leaves scope to try to build bridges between cultures. Since all news stories are written from a given perspective, let that perspective be one that allows readers to understand the country that is being covered.

7. Conclusion

This study set out to respond to five research questions. Namely, it aimed to understand how the China correspondents of El País and ABC approached their coverage of the coronavirus crisis (RQ1), what obstacles they faced (RQ2), what intercultural skills they developed (RQ3), whether they sought to foster an intercultural dialogue through their work (RQ4), and whether reporting conditions were somehow reflected in their coverage of the crisis (RQ5).

As for RQ1, the findings of this research show the correspondents interviewed for the study adopted a reactive attitude in the early days of the health crisis, as the potential impact of the outbreak was unknown and the authorities claimed the situation was under control. That wait-and-see mode was abandoned when Wuhan was locked down. From then on, correspondents became more proactive in their coverage of the crisis, although they encountered some constraints.

In response to RQ2, this study found that the China correspondents of El País and ABC encountered logistical, political, professional and personal difficulties while covering the coronavirus crisis. They could not enter Wuhan, they had to deal with the Chinese government's lack of transparency and media censorship, they were not able to talk to people who could tell them their personal stories about the crisis and the correspondents themselves suffered stress and fear because of the situation they found themselves in.

The respondents claimed to have gained cultural self-awareness, empathy, curiosity and interaction management skills in their career as correspondents, which addresses RQ3. That is, their responses point to the cognitive, affective and behavioral skills that Chen and Starosta (1996) identify as the basis for intercultural communication competence. It must be noted, though, that the correspondents disagreed over the importance of speaking the local language. Moreover, correspondents' perceptions of their cultural identity are somewhat aligned with what Casmir calls "third culture", but not with Kim's cross-cultural adaptation as that would entail going native and, perhaps, losing the perspective of their audience back at home (Kim, 2009; Casmir, 1997).

With regards to RQ4, the interviewed correspondents had opposed viewpoints when asked about whether they should foster an intercultural dialogue or not. El País's correspondent thought that was an essential part of a correspondent's job, but ABC's correspondent disagreed and insisted that journalists must reflect reality. In other words, El País's correspondent sees herself as an intercultural communicator that makes sense of other cultures to allow her audience to understand them, whereas ABC's correspondent understands that his task is to report on reality, without worrying about other considerations.

To aid in answering RQ5, it is necessary to review the main findings obtained in a critical discourse analysis of the coverage of the coronavirus crisis in China by the correspondents of El País and ABC carried out by the author elsewhere (Calatayud Vaello, 2021). As explained in the literature review, the analysis found that correspondents' coverage predominantly focused on Chinese authorities and that it was largely critical of the

Chinese government's handling of the health crisis. At the same time, the discourse analysis noted a significant presence of references to Chinese culture.

Thus, one initial takeaway from a comparison between the results of the analysis of correspondents' coverage and the findings from the interviews conducted for the present study may be that the reporting conditions faced by the correspondents explain the story behind the story. Without being able to access the center of the coronavirus crisis in China and of finding people to interview that could tell the human side of the crisis, the correspondents focused on what they could report: the Chinese government's response.

Indeed, there is a direct link between the reporting conditions the correspondents experienced and the content of their coverage. Chinese authorities' lack of transparency was the main story line found in the coverage of *El País* and ABC's coverage about the coronavirus crisis in China. Covid-19 patients were nearly absent from these correspondents' coverage because they were unable to interview people who had been affected by the disease or their relatives. And correspondents' critical representation of authorities should be read against the background of tensions between the Chinese government and foreign journalists.

As for the abundance of references to Chinese culture also observed in the analysis of correspondents' coverage, that may be seen as a way of enlarging—or updating—their audiences' knowledge of a foreign culture. In the interviews, the correspondents showed opposed attitudes toward whether they should foster intercultural communication, with *El País's* correspondent tilting toward a more interventionist role that seeks a greater understanding between cultures and ABC's correspondent more inclined toward a detached position. This might suggest that reporters' individual preferences end up somewhat blurred by the professional practices to which they adhere.

That being said, the fact that this study is based on only two interviews represents a severe limitation when it comes to the extrapolation of the findings. Zeng's study of China correspondents classified nearly half of the correspondents surveyed within the "populist watchdog" role—characterized by a willingness to play a watchdog role for Chinese elites while trying to maintain a strong audience orientation—, more than one third were considered "detached disseminators"—with emphasis on detachment and neutral reporting—and 13.9% were defined as "facilitative change agent"—showing the strongest interventionist emphasis to keep the Chinese government in check, while willing to cooperate with their home government—(Zeng, 2018). Meanwhile, Humanes and Roses (2018) found that the professional practices in the Spanish national press are mainly defined by an interventionist profile, a high presence of the watchdog role, and a conceptualization of the audience as citizens, but that a homogeneous model of journalism does not exist. If anything, the findings from this qualitative study of the reporting conditions for Spanish correspondents in China align with those observations.

References

- Abuín Vences, N., Sierra Sánchez, J., Mañas-Viniegra, L., and Núñez Gómez, P. (2020). Tratamiento informativo de la pandemia del coronavirus en los medios digitales españoles / Coronavirus pandemic news coverage in spanish digital media. *Hipertext. Net*, (21), 15–26. <https://doi.org/10.31009/hipertext.net.2020.i21.02>
- Ahlin, E. (2019). *Semi-Structured Interviews With Expert Practitioners: Their Validity and Significant Contribution to Translational Research*. <https://doi.org/10.4135/9781526466037>
- Archetti, C. (2011). Reporting the Nation: Understanding the Role of Foreign Correspondents in 21 st Century Public Diplomacy. *Paper Prepared for the International Studies Association (ISA) Annual Convention Montreal, 16-19 March 2011*, (March), 16–19.
- Argiñano, J.-L., and Goikoetxea Bilbao, U. (2020). Análisis de los titulares y las fotografías de portada en España en el contexto de la crisis del coronavirus: protagonistas, frames y lenguaje bélico. *Revista de Comunicación y Salud*, 10(2), 1–23. [https://doi.org/10.35669/rcys.2020.10\(2\).1-23](https://doi.org/10.35669/rcys.2020.10(2).1-23)
- Calatayud Vaello, A. (2019). Foreign Correspondents as Intercultural Mediators: Toward a Theoretical Approach to Foreign Correspondence Based on Intercultural Communication. *II Congrés de La Societat Catalana de Comunicació*. Barcelona: Unpublished.
- Calatayud Vaello, A. (2021). Othering in the Times of Coronavirus: Critical Discourse Analysis of the Early Coverage of a Major Health Crisis by the China Correspondents of El País and ABC. *Comunicació: Revista de Recerca i d'anàlisi*, 38(1), 53–79. <https://doi.org/10.2436/20.3008.01.202>
- Casmir, F. L. (1997). *Ethics in intercultural and international communication*. New York: Routledge.
- Chen, G.-M. (2009). Intercultural Communication Competence Theory. In S. W. Littlejohn and K. Foss (Eds.), *Encyclopedia of Communication Theory* (pp. 529–532). Thousand Oaks: Sage.
- Chen, G.-M., and Starosta, W. J. (1996). Intercultural Communication Competence: A Synthesis. *Annals of the International Communication Association*, 19(1), 353–383. <https://doi.org/10.1080/23808985.1996.11678935>
- Cheng, D. X., and Lee, F. L. F. (2015). Journalist–Source Relations: The case of Taiwan correspondents in China. *Journalism Studies*, 16(6), 850–867. <https://doi.org/10.1080/1461670X.2014.947146>
- Cooper, G. (2018). “Our Relationship? It’s the Odd Mucky Weekend, Not a One Night Stand”: Journalists and aid agencies in the UK, and the current challenges to sourcing in humanitarian disasters. *Journalism Practice*, 12(8), 954–965. <https://doi.org/10.1080/17512786.2018.1513813>
- Darling-Hammond, S., Michaels, E. K., Allen, A. M., Chae, D. H., Thomas, M. D., Nguyen, T. T., ... Johnson, R. C. (2020). After “The China Virus” Went Viral: Racially Charged Coronavirus Coverage and Trends in Bias Against Asian Americans. *Health Education and Behavior*, 47(6), 870–879. <https://doi.org/10.1177/1090198120957949>

- FCCC. (2021). *Track, Trace, Expel: Reporting on China amid a Pandemic*. Beijing.
- Fernández-Pedemonte, D., Casillo, F., and Jorge-Artigau, A. I. (2020). Communicating COVID-19: Metaphors we “Survive” by. *Tripodos*, 2(47), 145–160.
- Garfin, D. R., Silver, R. C., and Holman, E. A. (2020). The novel coronavirus (COVID-2019) outbreak: Amplification of public health consequences by media exposure. *Health Psychology : Official Journal of the Division of Health Psychology, American Psychological Association*, 39(5), 355–357. <https://doi.org/10.1037/hea0000875>
- Humanes, M. L., and Roses, S. (2018). Journalistic Role Performance in the Spanish National Press. *International Journal of Communication*, 12, 1032–1053. Retrieved from <http://ijoc.org>.
- Iranzo-Cabrera, M., and López-García, G. (2021). Scientific power in the Spanish press during the pandemic: A portrait of new leaders while explaining its risk. *Communication and Society*, 34(2), 263–280. <https://doi.org/10.15581/003.34.2.263-280>
- Kallio, H., Pietilä, A.-M., Johnson, M., and Kangasniemi, M. (2016). Systematic methodological review: developing a framework for a qualitative semi-structured interview guide. *Journal of Advanced Nursing*, 72(12), 2954–2965. <https://doi.org/https://doi.org/10.1111/jan.13031>
- Kim, Y. Y. (2009). Cross-Cultural Adaptation Theory. In S. W. Littlejohn and K. Foss (Eds.), *Encyclopedia of Communication Theory* (pp. 243–247). Thousand Oaks: Sage.
- Madrid-Morales, D., and Wasserman, H. (2018). Chinese Media Engagement in South Africa: What is its impact on local journalism? *Journalism Studies*, 19(8), 1218–1235. <https://doi.org/10.1080/1461670X.2016.1266280>
- Mohammed, S., Peter, E., Killackey, T., and Maciver, J. (2021). International Journal of Nursing Studies The “ nurse as hero ” discourse in the COVID-19 pandemic : A poststructural discourse analysis. *International Journal of Nursing Studies*, 117, 103887. <https://doi.org/10.1016/j.ijnurstu.2021.103887>
- Mokry, S. (2017). Whose voices shape China’s global image? Links between reporting conditions and quoted sources in news about China. *Journal of Contemporary China*, 26(107), 650–663. <https://doi.org/10.1080/10670564.2017.1305480>
- Ouariachi, T., and Peralta, L. (2021). Reality on the Ground: Exploring News Production Practices by Syrian Journalists in Times of Conflict. *Journalism Practice*, 0(0), 1–17. <https://doi.org/10.1080/17512786.2021.1916403>
- Picazo Sánchez, L., de Frutos Torres, B., and Gutiérrez Martín, A. (2020). La COVID-19 en la prensa española. Encuadres de alarma y tranquilidad en las portadas de El País, El Mundo y La Vanguardia. *Revista de Comunicación y Salud*, 10(2), 355–384. [https://doi.org/10.35669/rcys.2020.10\(2\).355-384](https://doi.org/10.35669/rcys.2020.10(2).355-384)
- Prieto-Ramos, F., Pei, J., and Cheng, L. (2020). Institutional and news media denominations of COVID-19 and its causative virus: Between naming policies and naming politics. *Discourse and Communication*, 14(6), 635–652. <https://doi.org/10.1177/1750481320938467>
- Sun, W. (2015). Configuring the foreign correspondent: New questions about China’s

public diplomacy. *Place Branding and Public Diplomacy*, 11(2), 125–138. <https://doi.org/10.1057/pb.2014.20>

Thomas, T., Wilson, A., Tonkin, E., Miller, E. R., and Ward, P. R. (2020). How the Media Places Responsibility for the COVID-19 Pandemic—An Australian Media Analysis. *Frontiers in Public Health*, 8(August), 1–14. <https://doi.org/10.3389/fpubh.2020.00483>

van Antwerpen, N., Turnbull, D., and Searston, R. A. (2022). What's Positive in a Pandemic? Journalism Professionals' Perspectives on Constructive Approaches to COVID-19 News Reporting. *Journalism Studies*, 23(4), 506–524. <https://doi.org/10.1080/1461670X.2022.2032804>

Yang, Y., and Chen, X. (2020). Globalism or Nationalism? The Paradox of Chinese Official Discourse in the Context of the COVID-19 Outbreak. *Journal of Chinese Political Science*. <https://doi.org/10.1007/s11366-020-09697-1>

Zeng, Y. (2018). *Reporting China on the Rise: A Study on the Field of Foreign Correspondence in Contemporary China (Doctoral dissertation)*. City University of Hong Kong.

Zhang, S. I., and Zhang, X. (2018). Foreign Correspondents: A case study of China in the digital and globalization age. *Journalism Studies*, 19(12), 1804–1824.

Zhang, X., and Shaw, G. (2020). The UK media coverage of China's handling of the coronavirus outbreak. *Global Media and China*. <https://doi.org/10.1177/2059436420980072>