VISUALISATIONS AND NARRATIVES IN DIGITAL MEDIA

Methods and current trends

EDITORS

Pere Freixa, Lluís Codina,
Mario Pérez-Montoro & Javier Guallar

DigiDoc
Digital documentation and interactive communication Research group

Profesional de la información

Funding: This work is part of the Project INDOCS "Interactive storytelling and digital visibility in interactive documentary and structured journalism". RTI2018-095714-B-C21 (MICINN/FEDER). Spanish Ministry of Science, Innovation and Universities (Spain).

Contact:

DigiDoc Research Group
Communication Department
Universitat Pompeu Fabra
Roc Boronat, 138, 08018 Barcelona
digidoc@upf.edu
https://www.upf.edu/digidoc

Ediciones Profesionales de la Información SL
Mistral, 36 (Cortijo del Aire)
18220 Albolote (Granada), España
Tels.: +34-639 878 489 / 608 491 521
http://www.profesionaldelainformacion.com

Translations: Iain K. Robinson

1st edition, March 2022.
ISBN: 978-84-120239-9-2
DOI: https://doi.org/10.3145/indocs.2022.0
URL: http://www.profesionaldelainformacion.com/informes_digidoc_epi/visualisations_narratives_digital_media.html

Recommended citation:

Freixa, Pere; Codina, Lluís; Pérez-Montoro, Mario; Guallar, Javier (Ed.) (2022). *Visualisations and narratives in digital media. Methods and current trends*. Barcelona: DigiDoc Research Group & Ediciones Profesionales de la Información SL. https://doi.org/10.3145/indocs.2022.0

CONTENTS

Pere Freixa
Universitat Pompeu Fabra
pere.freixa@upf.edu

Lluís Codina
Universitat Pompeu Fabra
lluis.codina@upf.edu

Mario Pérez-Montoro
Universitat de Barcelona
perez-montoro@ub.edu

Javier Guallar
Universitat de Barcelona
guallar@ub.edu

Visualisations and narratives in digital media: An Introduction

https://doi.org/10.3145/indocs.2022.1

Freixa, Pere; Codina, Lluís; Pérez-Montoro, Mario; Guallar, Javier (2022). Visualisations and narratives in digital media: An introduction. In: Freixa, Pere; Codina, Lluís; Pérez-Montoro, Mario y Guallar, Javier (ed.). *Visualisations and narratives in digital media. Methods and current trends*, (pp. 4-12). Barcelona: DigiDoc-EPI. https://doi.org/10.3145/indocs.2022.1

Abstract

The digital media have undergone an unprecedented transformation in recent years by exploiting the combined communicative potential of interaction and visualisation, generating, in this way, new narrative forms and journalistic stories. But this is not something that can be analysed in isolation. Understanding the digital media requires addressing the study of interactive texts and the platforms of the digital ecosystem from different perspectives.

The digital media today have highly permeable boundaries, the guidelines that define them being subject to constant modification: their texts are dynamic and constantly changing, their systems operate, thanks to artificial intelligence, as just another actor that analyses, collects and manages information. The space separating senders and receivers of messages has become fuzzy and interwoven. For this reason, the study of digital journalism has to assume this unremitting transformation as simply one more element in the debate about the media, and accept it as a characteristic of the digital culture that defines the field of communication.

Keywords

Digital journalism, visualisations, interactive narratives, SEO, journalism, journalistic innovation.

Título

Visualizaciones y narrativas en los medios digitales: presentación.

Resumen

En los últimos años los medios digitales han vivido una transformación sin precedentes al vincular el potencial comunicativo de la interacción y la visualización. Se han generado, con ello, nuevas formas narrativas y relatos periodísticos. Pero no se trata de un aspecto que pueda analizarse de forma aislada. La comprensión del medio digital exige abordar el estudio de los textos interactivos, el medio digital y su ecosistema desde distintas perspectivas.

El medio digital muestra una gran permeabilidad en sus fronteras, modifica constantemente las pautas que lo definen: los textos resultan dinámicos y mutables, los sistemas actúan, por medio de la inteligencia artificial, como un actor más que analiza, recoge y gestiona información. El espacio entre emisores y receptores en muchas ocasiones se entremezcla. Es por ello que la investigación del periodismo digital asume la transformación constante como un elemento más de debate sobre los medios, y la acepta como una característica de la cultura digital que define el ámbito de la comunicación.

Palabras clave

Ciberperiodismo, visualizaciones, narrativas interactivas, SEO, periodismo, innovación periodística.

1. Introduction

The digital media have undergone an unprecedented transformation in recent years by exploiting the combined communicative potential of interaction and visualisation (i.e. infographics and other formats), generating, in this way, new narrative forms and journalistic stories. But this is not something that can be analysed in isolation. As the research carried out over the last three years in the framework of the project *Interactive storytelling and digital visibility in interactive documentary and structured journalism* (funded by the Spanish Ministry of Science, Innovation and Universities – RTI2018-095714-B-C21, MICINN/ERDF) shows, understanding the digital media requires addressing the study of interactive texts and the platforms of the digital ecosystem from different perspectives.

Conducting this project has required a broad vision of the complexity of the media of contemporary communication, in which, as we have often found ourselves verifying, the lines between what have traditionally been separate areas and functions have become significantly blurred. The digital media today have highly permeable boundaries, the guidelines that define them being subject to constant modification: their texts are dynamic and constantly changing, their systems operate, thanks to artificial intelligence, as just another actor that analyses, collects and manages information. The space between senders and receivers of messages has become fuzzy and interwoven. For this reason, the study of digital journalism has to assume this unremitting transformation as simply one more element in the debate about the media, and accept it as a characteristic of the digital culture that defines the field of communication.

The monograph we present here brings together some of the results obtained from the studies carried out as part of this research project. It seeks to shed light on the "changes in communication patterns, social relations, collective and individual behaviour associated with the digitisation of society and culture", the main directive of the research challenge that it specifically addresses (Challenge number 6 – "Social sciences and humanities and science with and for society").

The chapters of this monograph include contributions that help understand how changes in our news consumption habits have led to the transformation of the media and how these changes in the media have, in turn, modified the ways in which users and audiences interact with them and with their content. The observations recorded while undertaking case studies, the development of analytical tools, and the exploration of experimental methodologies allow us to trace a broad panorama of new narratives in the digital media.

The first two chapters design and build analytical tools. Their authors have turned their attention to specific aspects of digital communication – namely interactivity and content curation – and have developed research methodologies for their in-depth analysis.

It is widely accepted that the digital media can be distinguished from the traditional media primarily in terms of the interactivity they offer. The centrality of interaction is such that it is a fundamental determinant of the development of digital products, the provision of access to information and the creation of meaning in reading. Interactivity has established itself as an indispensable concept for understanding a medium characterised by the constant exchange of inputs and outputs between the agents and actors that participate in it. These are exchanges that in themselves serve to define the medium: information that is stored, processed, and interacted with, and procedures that feed into each other to provide a system that allows the user to experience and participate in the information and to be part of it.

Santos, Lopezosa, and Codina design a protocol, in **chapter 2**, for analysing interactivity in the digital press, focusing their study on aspects related to the user experience and their options for interaction with the content that the news media offer. Comparative analyses serve to test and validate a methodological tool that makes it possible to observe the relationship that the medium seeks to forge with its users, the possibilities the audience itself has of generating materials, the community's resources, the relationship established between the users and, finally, the possibilities for personalising content. The definitions provided of each parameter, combined with the choice of examples of good practices, facilitate the adoption of the tool and ensure its good use for future research.

If interactivity is a foundational element of digital journalism, the practice of content curation in the media can be considered an activity still in its implementation phase and one that finds itself in an incipient stage of development. It responds to two joint needs detected by the media: the professional filtering and selection of content, on the one hand, and the personalisation of information, on the other. In a context characterised by the exponential growth of digital content, content curation can be considered the professionalisation of the activity of selecting the most relevant content for a given audience, sharing it with that audience in the best way possible, and endowing it with added value.

Guallar, Pedraza-Jiménez, and Pérez-Montoro present the Curation Analysis System or CAS, a parameterised method for analysing media content curation in **chapter 3**. The system is based around a series of parameters and indicators that allow the two distinct dimensions of content curation to be assessed: those related to curation and those concerning the content itself. For each parameter included in the CAS, the authors offer a definition, a detailed explanation of its meaning, the procedure to be followed for its correct application, and examples of good practices extracted from the newsletters of the digital media in different countries.

The CAS is the first system of its kind to be proposed in the academic literature for the evaluation of content curation in digital media. The system does much more than simply describe this process of curation, as it also enables the analyst to identify the specific components of journalistic curatorial practice, to undertake assessments of quality, and to create derived products, including, for example, rankings and comparative studies.

One of the main challenges facing the news media is getting audiences to access their content. This issue, which might appear trivial from the perspective of analogue journalism, is fundamental in the digital sphere. How can end users be guaranteed access to the most relevant information for them in an environment characterised by the hyper-saturation of information and this in a context in which the virtual immateriality of the digital world makes the visibility of texts and messages even more difficult? Two actors acquire particular relevance here: first, search engines and their ability to select and ascertain the results that best suit the needs of users and, second, SEO positioning strategies, the actions carried out by the media to give visibility to their content and to ensure this content finds its potential user.

In **chapter 4**, Escandell-Poveda, Iglesias-García, and Papi-Gálvez explore the historical evolution of search engines in seeking to understand the dominant role that they play in today's digital ecosystem. Clearly, the development of the digital media, the emergence of the Internet and the subsequent exponential growth of the World Wide Web brought with them the intrinsic need for systems that might facilitate the management, filtering and retrieval of information and content. Search engines quickly became the tools that linked indexes, web pages and multimedia content with user needs, as they established themselves as the main interface and gateway to digital information. As the authors conclude, learning about the origins and gestation of search engines allows us to understand how web visibility and audience attraction have evolved, establishing themselves today as essential qualities for any medium that wishes to stand out in the ocean of information and resources that the web has become over the decades that make up its history.

Indeed, the need for web visibility has meant that the digital news media are obliged to optimise their search engine performance, given that a large part of the traffic received by the online media comes from the searches that users make with these tools, with Google playing a dominant role here. It is for this reason that, for more than ten years, the news media have begun to implement positioning strategies and practices, known as SEO or actions of search engine optimisation, a set of techniques and procedures applied to websites that helps increase their chances of appearing among the main search results and, with it, of having greater visibility and attracting more readers.

However, SEO strategies differ from one news media outlet to another. Not all the premises of SEO are equally useful for all media, in all environments and areas, nor are they common to all types of users. In **chapter 5**, Lopezosa, Trillo-Domínguez, Cabrera, and Codina present an example of the implementation of SEO practices, in this case as carried out by the *Grupo Joly,* a leading press publishing house in Andalusia, on the *Diario de Sevilla*, its flagship newspaper.

The study confirms how SEO has gone from being the specialised task of computer technicians and engineers to occupy a central position in the process of preparing and labelling journalistic content. It also shows how the quantitative concerns – that is, of achieving greater visibility than the competition – have undergone a qualitative reorientation. In this

respect, SEO practices have had a significant impact both on journalists' work guidelines and on the format of the information itself. In short, SEO has become a strategic factor for the transformation of companies, guaranteeing their survival in the digital ecosystem. The good results obtained by the *Diario de Sevilla* allow lessons to be learnt about the different roles that SEO specialists and journalists can adopt to ensure the optimum implementation of SEO guidelines without these being allowed to affect the quality of the newspaper's content.

The processes of audience participation facilitated by social networks and the digital ecosystem as a whole have allowed the development of works of communication and documentation that are characterised by their aim to generate social impact. The ***webdoc*** has found in its exploitation of transmedia strategies the context in which to propose and experiment with new forms of horizontal communication. In these works, presented by Freixa, Soler-Adillon, and Sora in **chapter 6**, the interrelationship between the virtual and the local intersects and feeds back into each other.

Clearly, the online interactive documentary has facilitated the promotion of projects in which participation and interaction between all those involved is essential: authors, community and audience. These are works designed to generate shared spaces for the documentation and exchange of stories which, thanks to the projects, become acts of collective memory for the communities with which they work. The projects serve to activate whole communities via the social networks, the role of which is critical for their development. The three examples analysed in the study each use transmedia strategies in their own unique fashion to organise and establish communication and to channel the contribution of their content, an essential element for the creation and success of the three projects.

Structured journalism has highlighted the importance of the media archive and the need to improve the archiving of information in newsrooms in order to exploit its potential for reuse in the creation of future content. Similarly, media archives and the persistence of citizen stories play an important role in the interactive documentary. The critical reading and the appropriation of the archive become key elements in the constitution of the stories that each individual needs to generate their own life narratives.

The digitisation of traditional media allows us to access once more those analogue audiovisual materials that had become invisible because of their technological obsolescence. Trawling through the archives – both public and private or personal – facilitates critical inquiry. What is shown, what is deemed worth recording, the methods used to do so and the stories themselves are reactivated and enter into dialogue with the gaze and perception of the present day.

In **chapter 7**, Obradors explores different methodologies of *art-practice-as-research* as a source of knowledge. Her work focuses on the family album, which she removes from its private sphere to convert into audiovisual media. As the author stresses – in keeping with the

UNESCO recommendation – amateur audiovisual records and family home films need to be protected and conserved because they form part of our collective cultural heritage. Indeed, it is thanks to them that we can restore part of our history, these records serving as a testimony to a particular age and its customs.

Family albums, be they in the form of home cinema or video recordings, despite an apparent uniformity, can be differentiated according to the expertise of their authors, reflecting the varying degrees of amateurism of the filmmaker. The editing of the film can be considered a fundamental differential, an indication that the author seeks to intervene not only in the recording of events but also in the construction of their narrative. The narrative consciousness of these amateur works means they can be assimilated with other formats of social and historical documents found in public archives. Yet, the editing of these films is also one of the elements that leads us to examine afresh some of the attributes associated with personal archives, that is, their stamp of verisimilitude and authenticity, which are now called into question.

Indeed, the credibility and veracity of multimedia images and journalistic messages has become a major research topic in recent years. Despite being a problem that has been associated with the news media since their very origins, the emergence of the phenomenon of media disinformation and fake news has done much to disrupt the information ecosystem. The infinite possibilities for manipulation afforded by digital technologies oblige those involved in information processes to establish mechanisms that can validate and verify their sources and content. In **chapter 8** of this monograph, Vállez and Boté introduce us to the concept of *deepfakes*, the modification of audiovisual content to generate false content which, to all intents and purposes, appears to be true. *Deepfakes* require the combination of artificial intelligence technologies with the manipulation of image, sound and video in their efforts to mislead, and are typified by the way in which they present content in a decontextualized manner in terms of form, place and time.

In their study, the authors present the characteristics of these synthetic media technologies and their potential for poisoning the work of the news media, and they identify the resources that have been developed to combat them. They also examine the concerns that their existence raises both in relation to the degree of tolerance shown by audiences for these products depending on the uses to which they are put – that is, for entertainment, in film production, or for information purposes – and in relation to other aspects, including respect for the privacy of third parties and their ability to manipulate. As the authors warn, it is becoming increasingly necessary for steps to be taken "to ensure that in all areas, not only in the academic arena and schools, that people are made aware of the need to prevent the development and use of *deepfakes*" and they insist that "digital and media literacy campaigns are now more essential than ever in our rapidly changing digital world".

The research on digital narratives and visualisations undertaken in 2019 within the framework of the research project *Interactive storytelling and digital visibility in interactive documentary and structured journalism* was heavily conditioned by the outbreak of the Covid-19 virus at the end of 2019 and the declaration of the global pandemic in March 2020. Many of the studies initiated had to be adapted to the new scenario, while others had to be put on hold. However, the possibility of monitoring and analysing the impact of the pandemic on the news media appeared as a pressing research opportunity.

The results of this research effort have already been presented at a number of conferences and published in academic journals (Freixa & Redondo; 2022; Freixa, Pérez-Montoro, Guallar, & Codina, 2021; Freixa & Redondo, 2021; Lopezosa et al., 2021; Pérez-Montoro, 2021; Vállez & Pérez-Montoro, 2020). To round off this monograph, in **chapter 9**, we wanted to incorporate the in-depth review that Pérez-Montoro has undertaken of the visualisations that the main news media employed during the early months of the pandemic and which serve as a good demonstration of their enormous communicative potential.

The great uncertainty and confusion generated by the Covid-19 pandemic, as well as the urgent need to provide the population with accurate, readily comprehensible information about the events that were unfolding, led the news media, governments and medical institutions to seek out the best way to transmit complex data to citizens. Once more, data visualisations allowed the creation of effective visual narratives that could be disseminated massively on social networks and which facilitated the establishment of an audiovisual narrative of the pandemic. The combination of image and graphics made the number of deaths and the ratios of infected cases by age and gender groups intelligible, they highlighted the effectiveness of vaccines and permitted transnational comparisons.

We wish to end this preface by thanking all the researchers who have participated in the project *Interactive storytelling and digital visibility in interactive documentary and structured journalism* for ensuring that, despite the outbreak of the Covid-19 pandemic, we were able to meet all our research targets, as the work in this monograph bears witness.

We believe that this monograph is also a way of boosting the social impact of the research carried out, making our results available to a wider audience thanks to its distribution via an open access licence. While some of the results have been communicated in academic publications in the form of journal articles or papers delivered at conferences, we believe that the synthesis presented in this format will make our results much more accessible to audiences that are not only academic, but which also comprise professionals, decision makers and, in general, any citizen interested in the questions we pose.

References

Freixa, P.; Pérez-Montoro, M.; Guallar, J. & Codina, L. (2021). Photography, COVID-19, citizen participation and social networks. Document and preserve the collective witness. In Larrondo, A.; Meso-Ayerdi, K.; Peña-Fernández, S. (ed.). *Disinformation and credibility in the digital ecosystem. XII International Conference on Online Journalism* (pp. 180-197). Bilbao: Servicio Editorial de la Universidad del País Vasco / Euskal Herriko Unibertsitateko Argitalpen Zerbitzua.

Freixa, P. & Redondo-Arolas, M. (2022). Mirando al mar. Tropos visuales de resiliencia y superación en la covid-19. *Arte, Individuo y Sociedad, 34*(2), 777-798. https://dx.doi.org/10.5209/aris.75517

Freixa, P. & Redondo-Arolas, M. (2021). COVID-19, media and photography. A crisis without victims?. In Larrondo, A.; Meso-Ayerdi, K.; Peña-Fernández, S. (ed.). *Disinformation and credibility in the digital ecosystem. XII International Conference on Online Journalism* (pp. 198-216). Bilbao: Servicio Editorial de la Universidad del País Vasco / Euskal Herriko Unibertsitateko Argitalpen Zerbitzua.

Lopezosa, C.; Freixa, P.; Codina, L.; Pedraza, R.; Medina-Bravo, P. & Garin, M. (2021). Protocol and ethical challenges for success in research with minors: the case of cultural and museum studies in the context of COVID-19. Barcelona: Universitat Pompeu Fabra. Departament de Comunicació, (Serie Editorial DigiDoc. DigiDoc Reports). http://hdl.handle.net/10230/48803

Pérez-Montoro, M. (2021). Comunicación visual de una emergencia sanitaria mundial: el caso de la Covid-19. *Anuario ThinkEPI, 15*, 1-15. https://doi.org/10.3145/thinkepi.2021.e15d01

Vállez, M. & Pérez-Montoro, M. (2020). La comunicación periodística en tiempos de pandemia: análisis del tratamiento de la Covid-19 en la prensa europea. *Hipertext. net*, (21), 1-13. https://doi.org/10.31009/hipertext.net.2020.i21.01

Gema Santos-Hermosa
Universitat de Barcelona
gsantos@ub.edu

Carlos Lopezosa
Universitat de Barcelona
lopezosa@ub.edu

Lluís Codina
Universitat Pompeu Fabra
lluis.codina@upf.edu

Analysis of interactivity in digital journalism

https://doi.org/10.3145/indocs.2022.2

Santos-Hermosa, Gema; Lopezosa, Carlos; Codina, Lluís (2022). Analysis of interactivity in digital journalism. In: Freixa, Pere; Codina, Lluís; Pérez-Montoro, Mario y Guallar, Javier (ed.). *Visualisations and narratives in digital media. Methods and current trends*, (pp. 13-25). Barcelona: DigiDoc-EPI. https://doi.org/10.3145/indocs.2022.2

Abstract

We analyse interactivity in the digital press, focusing our study specifically on aspects related to user experience and user options for interaction. In so doing, our objective is the improvement of digital communication, by offering the news media a solid basis from which to evaluate their business model and their readers' behaviour. We exploit a methodology based on the design and testing of a research protocol and undertake a comparative analysis employing an iterative process, thus 'grounding' our final theory in the data gathered. In this way, we obtain a protocol – 'Analysis of interactivity, searchability and web visibility in digital newspapers' or SAIPD in its Spanish acronym – that facilitates the extraction of data from the digital news media. This tool, as its name suggests, is designed around the three characteristic dimensions of the digital media (i.e. interactivity, searchability and web visibility) and employs a series of analytical parameters and indicators. In this chapter we examine four of these parameters related specifically to the dimension of interactivity: the media-user relationship, user-generated content, the user-user relationship and the personalization of content.

Keywords

Digital journalism, digital news media, interactivity, searchability, user's experience, SEO, entrepreneurship in digital communication.

Título

Análisis de la interactividad en el periodismo digital

Resumen

Este trabajo analiza la interactividad en los diarios digitales, situando en el centro del estudio los aspectos relacionados con la experiencia de usuario y sus posibilidades de interacción. El objetivo que se pretende alcanzar es la mejora del emprendimiento en la comunicación digital, proponiendo a las empresas periodísticas bases sólidas de valoración de su modelo de negocio y del comportamiento de sus lectores. Los métodos empleados son el diseño y testeo de un protocolo de investigación y la realización de un análisis comparativo mediante un proceso iterativo desde la teoría a los datos, y a la inversa hasta obtener un protocolo articulado en indicadores operativos que permiten la extracción de datos. El protocolo en el que se enmarca la dimensión de interactividad que presentamos se denomina 'Análisis de la Interactividad, buscabilidad y visibilidad web en periódicos digitales', o SAIPD. Este sistema de análisis ha sido diseñado con base en las tres dimensiones características del medio digital (interactividad, buscabilidad y visibilidad web) y se ha desarrollado a partir de una serie de parámetros e indicadores de análisis. En el presente capítulo nos fijamos en cuatro de estos parámetros; concretamente los vinculados a la dimensión de la interactividad: la relación medio-usuario, los contenidos generados por usuarios, la relación usuario-usuario y la personalización de contenidos.

Palabras clave

Periodismo digital, medios digitales, interactividad, buscabilidad, visibilidad web, SEO, emprendimiento en comunicación digital.

1. Introduction

Journalism has to learn to adapt, whenever the occasion arises, to the particular channel or support that it uses. It had to adapt when it made the move from the print media to radio and again when having to adapt to television. Today, the main challenge journalism faces is adaptation to the digital media. It is for this reason that this study examines one of the most intrinsically important dimensions of the digital medium: namely, interactivity. This is an extension that any digital newspaper must dedicate the utmost attention to because, otherwise, it runs the risk of losing not only its readership, but also something that is considerably more valuable these days, the coveted ***engagement*** of its users.

To be able to analyse and improve this dimension, here we develop a heuristic protocol that can be applied to the digital media and which should prove useful for both researchers and professionals in the field of digital journalism. First, we outline a methodological framework for the overall conceptualization of the study protocol. Then, we focus our attention on the specific dimension of interactivity, presenting in detail each of the analytical parameters and indicators developed. The proposed heuristic places the focus squarely on the consumer of digital media, on their experience as a user and on their options for interaction. The ultimate objective is the improvement of digital communication platforms, by offering the news media a solid basis from which to evaluate their business model and their readers' behaviour.

2. Methodological framework

Our protocol has been christened "Analysis of interactivity, searchability and web visibility in digital newspapers" (or SAIPD in its Spanish acronym) (Santos-Hermosa, Lopezosa & Codina, 2022) and is a heuristic device. On occasions, such protocols are referred to as expert analyses, given that the analytic system so designed is intended for use by experts (and not users). Heuristic or expert analyses have a long history, both in the academic and professional worlds (Nielsen & Molich, 1990; Abdelmaguid et al., 2004; Pedraza-Jiménez et al., 2016), their main characteristic being that they can be applied before, after or in addition to user studies. Another of their characteristics – and one that makes them particularly interesting to us – is that they can be used to carry out comparative analyses and they can support both academic and professional research.

In designing the protocol and its specific dimension of interactivity, we have employed the methods that underpin grounded theory. This involved the initial gathering and analysis of data which served to establish the bases of our theory; from here we returned back to our data sources until we arrived at our final theory, "grounded", that is, in the data. To undertake the initial data analysis, we employed protocols developed and validated in previous

research: PICs (Linares et al., 2015) and PAXBCM (Lopezosa et al., 2020;2021). Indeed, these two protocols have been tested and validated by the Digital Documentation and Interactive Communication (DigiDoc) research group and have been successfully used in conducting funded R&D projects and for writing publications in peer-reviewed journals.

The specific dimension of the protocol that we analyse is, on the one hand, an extension of previous protocols, from which we select and adapt aspects related to user actions in the digital medium, but, on the other, it also incorporates completely new elements with a specific focus on user interaction. All this allows us to obtain more ambitious results that give us greater possibilities of analysis, as well as ensuring greater precision in their use by analysts and professionals of digital media; facilitating, as far as possible, an improvement in the task, by offering news media companies solid bases for evaluating the parameters studied. Finally, the application undertaken here of the protocol itself to specific case studies is completely novel.

3. Dimension of analysis

The dimension of interactivity in digital newspapers consists of four analytic parameters (see Table 1), each of which is made up of different indicators.

Protocol for the analysis of interactivity in digital journalism		
Nº	*Indicator*	*Score*
Parameter 1: Digital medium-user relationship		
1.1	Reader-author written communication	0-1
1.2	Reader-author communication via Twitter	0-1
1.3	Reader-author communication via Facebook	0-1
1.4	Reader-author communication via Instagram	0-1
1.5	Reader-newsroom communication	0-1
1.6	Reader-newsroom communication via Twitter	0-1
1.7	Readers' comments	0-1
1.8	Readers' votes	0-1
1.9	Readers' ratings	0-1
1.10	Comments on the blogs of the digital medium	0-1
1.11	Rectifications	0-1
1.12	Confidentiality channels	0-1
Parameter 2: User-generated content		
2.1	User-generated texts	0-1

2.2	User-generated photographs	0-1
2.3	User-generated videos	0-1
Parameter 3: User-user relationship		
3.1	Contact between registered users	0-1
3.2	Contact between users (readers) and other users	0-1
3.3	Ranking of information according to user activity	0-1
Parameter 4: Personalization of content		
4.1	Offer of newsletters	0-1
4.2	Syndication of content on mobile or via email	0-1
4.3	Specific subscriptions	0-1
4.4	Recommendations based on a user's recent navigation	0-1
TOTAL		N

Table 1. Systematic protocol for the analysis of interactivity in digital journalism

The parameters specifically identify what we seek to measure or evaluate in a digital medium, such as the relationship (that is, the ***engagement***) between the medium and its audience. In turn, these parameters, in order to be examined in depth, need to be specified in operable indicators, that is, characteristics that can be analysed and, where appropriate, measured. The indicators represent the ***how*** (or the way) each parameter is evaluated. Hence, there is always a 1:N relationship between parameters and indicators, whereby each parameter is studied employing two or more indicators. The score associated with each indicator (Table 1) is based on a system that uses binary values (values 0 or 1). This binary system, applied across all the parameters, evaluates whether an indicator is present (score 1) or not-present (score 0). Each of the parameters analysed and their corresponding indicators are presented below. We provide a detailed description of each element and examples to illustrate them.

Parameter 1: Digital medium-user relationship

This parameter comprises indicators of the provision of means of contact between the reader and the medium (indicators 1.1 to 1.6) and indicators related to user feedback (indicators 1.7 to 1.12), key features in the media as highlighted by Paskin (2018). The former capture the procedures offered by the digital medium to enter into contact with it, that is, channels of contact with the journalists or with the newsroom (mail, Twitter, Facebook, Instagram, etc.). The latter capture the options made available by the medium to obtain user feedback: comments on news items, votes, reactions (i.e. 'likes'), participation in surveys, etc. This parameter comprises the following 12 indicators:

1.1 Reader-author written communication:

Users can send messages via a form or by email to the author of the news item to provide information about the story or to expand on it.

1.2 Reader-author communication via Twitter:

Users can send messages via Twitter to the author of the news item to provide information about the story or to expand on it.

1.3 Reader-author communication via Facebook:

Users can send messages via Facebook to the author of the news item to provide information about the story or to expand on it.

1.4 Reader-author communication via Instagram:

Users can send messages via Instagram to the author of the news item to provide information about the story or to expand on it. On occasions, even a link to the social network is available in the author's biography: "Authors" section with information about journalists, photographers, etc. (likes, hobbies, social media accounts). A good example of this is provided by the online newspaper *Uppers* (see Figure 1).

Figure 1. Example of an author profile from *Uppers* (https://www.uppers.es/autores/pepe-serra/)

1.5 Reader-newsroom communication:

Users can contact the newsroom of a news media outlet by email to learn more about the published content or to request further information.

1.6 Reader-newsroom communication via Twitter:

Users can contact the newsroom of a news media outlet via Twitter to learn more about the published content or to request further information. In this case, although it falls outside our protocol, the relationship between digital media websites and external platforms, such as Twitter, could be fruitfully studied. Although almost all media outlets would appear to have Twitter profiles, not all of them are active. A more detailed knowledge of the comments or reactions of Twitter users could improve perceptions of interactivity.

1.7 Readers' comments:

Users can send comments about the news items to which they have access and see the comments of other readers.

1.8 Readers' votes:

Users can participate in surveys created by the digital medium or other proposed voting options.

Surveys are a way of gauging general opinion on a particular topic. However, there are studies that conclude that the survey is not exploited that much as a tool for interaction (Baños-Moreno et al., 2017), yet it can be a good way of engaging readers and finding out what they think on a particular topic.

1.9 Readers' ratings:

Users can cast their vote, or have the sense of expressing a preference or their rejection or a greater or lesser degree of satisfaction (for example, awarding stars) with respect to a news item.

1.10. Comments on the blogs of the digital medium:

Users can comment on the entries published in the blogs of the communication medium.

1.11. Rectifications:

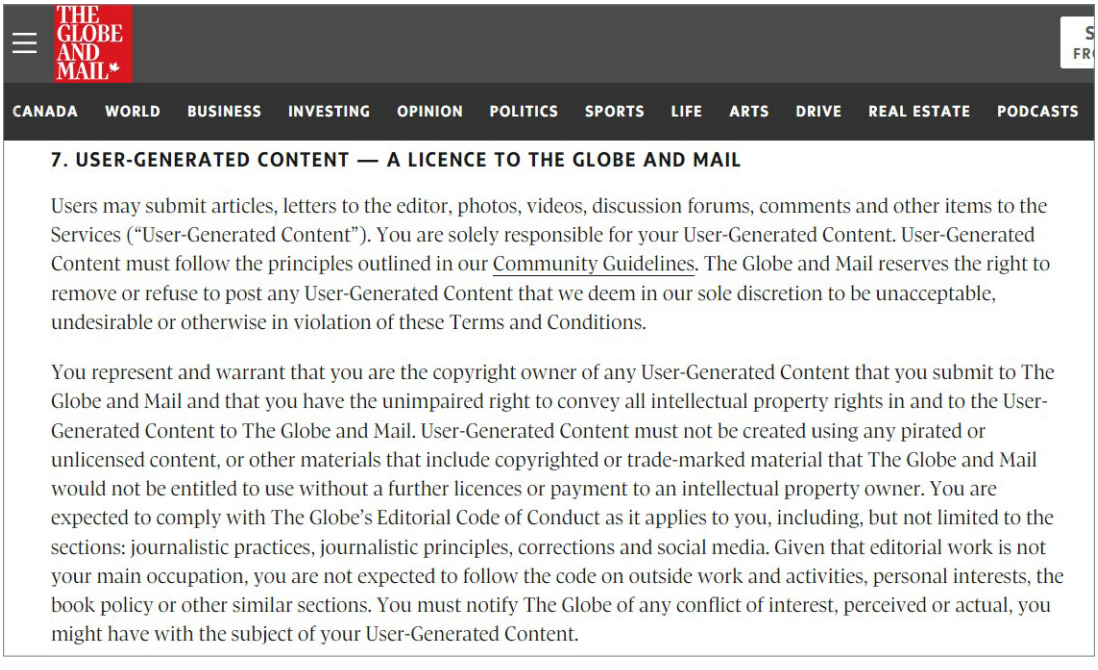
Users of the information published by the media can point out an erroneous fact or piece of information in one of the news items or communicate a technical error or incident so it can be resolved.

1.12. Confidentiality channels:

Users can send complaints or share information with the media anonymously and safely.

Parameter 2: User-generated content

This parameter comprises indicators of the type of content (textual and audiovisual) that the media allow their users to both generate and publish. The development of services for digital consumers (social software, video sharing and Web 2.0 applications) and the increased use of mobile devices and tablets have led to the proliferation of user-generated content (Reyna, Hanham & Meier, 2018). These technological advances have changed the way we communicate and socialize, making it possible for the consumer of information to become an author and co-creator of content for the media. We are witnessing the collective production of content, based on various forms of voluntary participation (Simon, 2016), which entails a change in digital media consumption patterns.


THE GLOBE AND MAIL

CANADA WORLD BUSINESS INVESTING OPINION POLITICS SPORTS LIFE ARTS DRIVE REAL ESTATE PODCASTS

7. USER-GENERATED CONTENT — A LICENCE TO THE GLOBE AND MAIL

Users may submit articles, letters to the editor, photos, videos, discussion forums, comments and other items to the Services ("User-Generated Content"). You are solely responsible for your User-Generated Content. User-Generated Content must follow the principles outlined in our Community Guidelines. The Globe and Mail reserves the right to remove or refuse to post any User-Generated Content that we deem in our sole discretion to be unacceptable, undesirable or otherwise in violation of these Terms and Conditions.

You represent and warrant that you are the copyright owner of any User-Generated Content that you submit to The Globe and Mail and that you have the unimpaired right to convey all intellectual property rights in and to the User-Generated Content to The Globe and Mail. User-Generated Content must not be created using any pirated or unlicensed content, or other materials that include copyrighted or trade-marked material that The Globe and Mail would not be entitled to use without a further licences or payment to an intellectual property owner. You are expected to comply with The Globe's Editorial Code of Conduct as it applies to you, including, but not limited to the sections: journalistic practices, journalistic principles, corrections and social media. Given that editorial work is not your main occupation, you are not expected to follow the code on outside work and activities, personal interests, the book policy or other similar sections. You must notify The Globe of any conflict of interest, perceived or actual, you might have with the subject of your User-Generated Content.

Figure 2. User-generated content conditions of *The Globe and Mail* (https://www.theglobeandmail.com/privacy-terms/terms-and-conditions/)

This parameter comprises the following three indicators:

2.1. User-generated texts:

Users can send texts to the media company for publication in their newspaper (with texts being reviewed by an editorial team before being published).

2.2. User-generated photographs:

Users can send photos to the media company for publication (users must be registered).

2.3. User-generated videos:

Users can send videos they have made themselves to the media company for publication (users are required to provide public and private personal data).

Parameter 3: User-user relationship

This parameter records the interaction between users, those "human" actions that facilitate interpersonal communication (Larsson, 2012), providing a response also to the dimension of socialization. An example could be responding to a comment from another user (Ksiazek et al., 2016).

This parameter comprises the following three indicators:

3.1. Contact between registered users:

Users can communicate with other registered users (community of registered users who can interact with each other).

3.2. Contact between users (readers) and other users:

Users can share information from the digital media with other users via the social networks provided by the same digital news media: Twitter, Instagram, Facebook, etc.

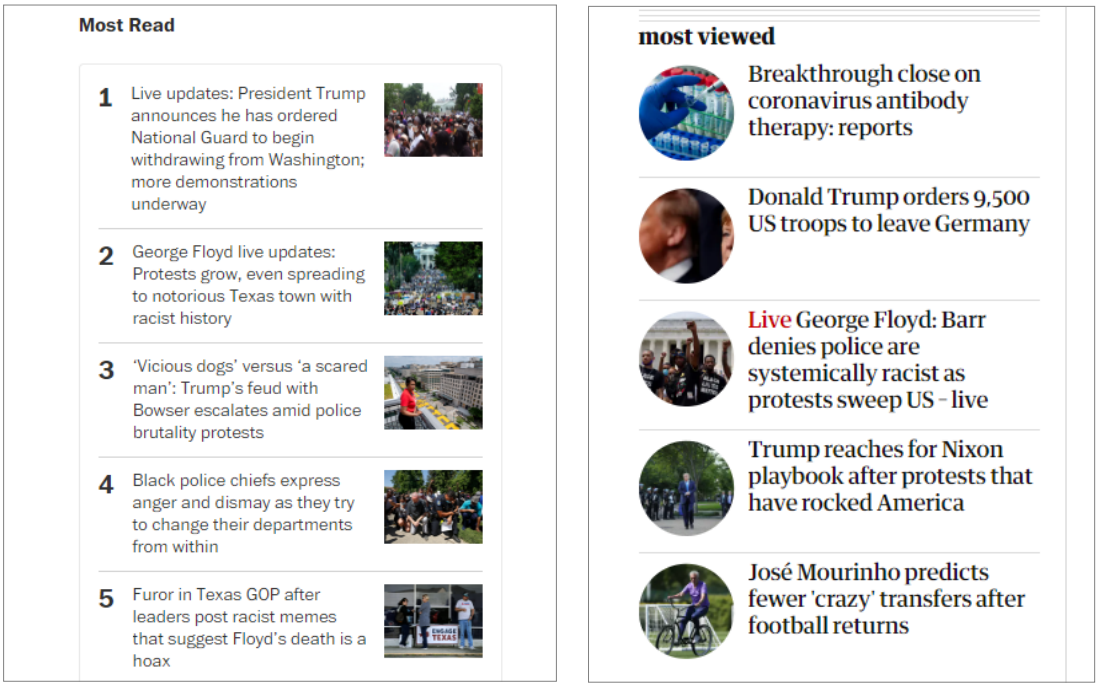


Figure 3. Grouping of the most read (*Washington Post*) and most viewed news items (*The Guardian*) by other users)

3.3. Ranking of information according to user activity:

Users can access information grouped according to the queries of other users (Figure 3).

Parameter 4: Personalization of content

This parameter is based on the possibility – or not – of accessing and adapting the content provided by the digital medium in a more personalized fashion. It is concerned with identifying the type of interactivity with which users can personalize or adapt the website content to their own interests and tastes (Choo et al., 2012; Larsson, 2012; Guallar et al., 2012; Rodríguez-Martínez et al., 2012 and Baños-Moreno et al., 2017).

This parameter comprises the following four indicators:

4.1. Offer of newsletters:

Existence of a newsletter and the possibility that users can subscribe to it to consume specific information that they select according to their interests. According to Baños-Moreno et al. (2017), electronic bulletins are one of the most common services within the media. Guallar et al. (2021) add that newsletters are a very favourable channel for content curation, one of the services considered most relevant in digital journalism in the 21st century.

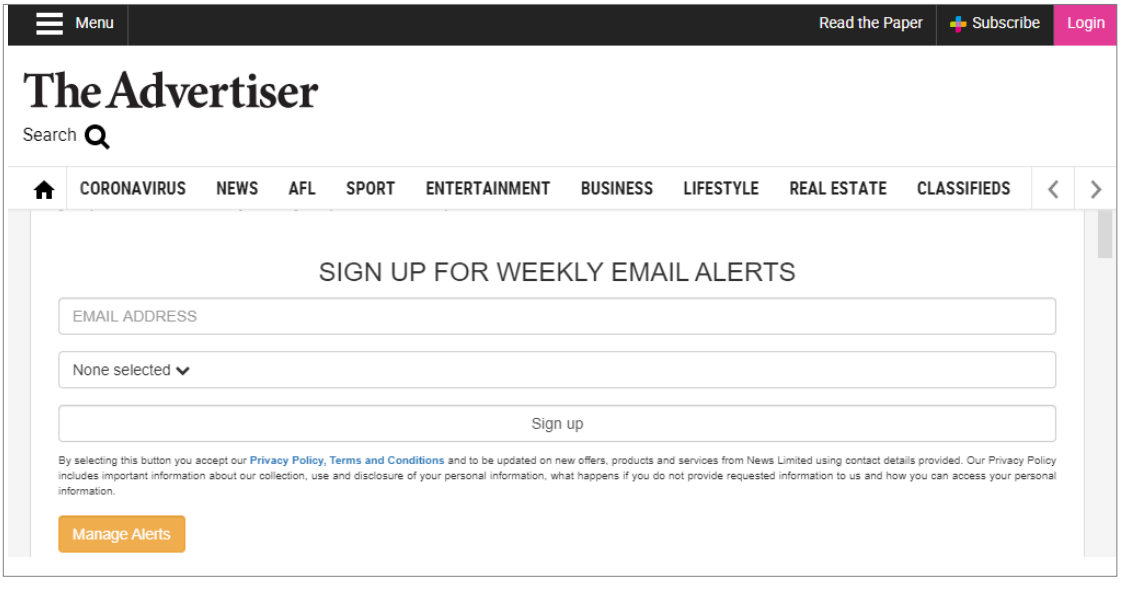


Figure 4. Screenshot of page for signing up for email alerts from *The Advertiser*

4.2. Syndication of content on mobile or via email:

Users can syndicate certain content from the digital medium via their mobile phone or email to consume specific information selected for them according to their interests or

circumstances. For more on this digital experience via applications or adaptations to online devices, see Harvey and Pointon (2019).

4.3. Specific subscriptions:

Users can subscribe to specific content (section headlines, authored articles, topics – dealt with in more depth than in the dedicated section, geographical area, etc.) and/or specific products (Apple Podcast, Google Podcast, Podcast in digital media, blogs, etc.). Although not included in this protocol, subscriptions could also be related to another parallel issue, which would be the presence of digital media on web 2.0 platforms (Figure 5).

4.4. Recommendations based on a user's recent navigation and the most seen/most read rankings:

Users can access content recommendations based on their own recent navigation and on the users' most seen or most read rankings (Figure 6).

Figure 5. *The Guardian* channel on *Spotify*

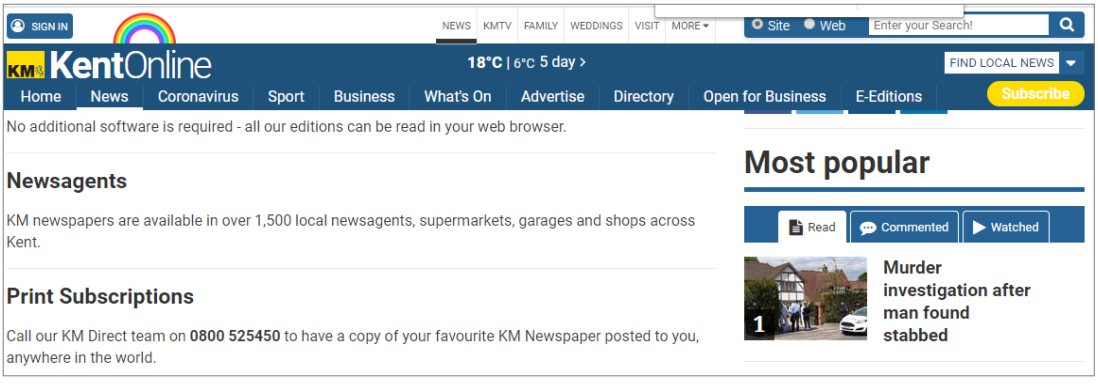


Figure 6. Web page of *KentOnline*, showing the most read, commented and watched news items

4. Conclusions and protocol applications

Interaction has acquired a considerable prominence in many digital industries aimed at improving user experience. For this same reason, the digital news media are aware they must not ignore this feature but rather attempt to integrate it into their products so that they can compete for larger and larger audiences.

The protocol developed is designed in such a way that it can, in the hands of other research teams, be used as it is presented here (first scenario). Equally, other indicators might be incorporated following the same procedure and representing them in a similar fashion to the method employed here (second scenario) or, in the same vein, those indicators that do not fit in with a team's objectives can be eliminated (third scenario). Indeed, research teams might opt to combine scenarios two and three, thus creating a fourth scenario. What we wish to stress is that the ultimate design and application of this protocol is highly flexible and adaptable to the objectives of each study.

Future studies in this line could usefully focus on expanding the comparative analysis to other digital communication media; for example, in a specific geographical area or in a specific sector. Likewise, the application of this protocol to the study of other types of media, such as radio or television portals, streaming platforms, newsletters, etc. could prove insightful.

Finally, in an ongoing study, the authors of this protocol are applying it (publication in process) to a highly select group of digital media, specifically, to a number of recent prize winners in the Online Journalism Awards (ONA) and the World Digital Media Awards (WDMA), as part of their research work in a competitive project.

References

Abdelmaguid, T.; Nassef, A.; Kamal, B. & Hassan, M. (2004). A hybrid GA/heuristic approach to the simultaneous scheduling of machines and automated guided vehicles. *International Journal of Production Research, 42*(2), 267-281. https://doi.org/10.1080/0020754032000123579

Baños-Moreno, M.J.; Pastor-Sánchez, J.A. & Martínez-Béjar, R. (2017). Interactivity features of Online Newspapers: from a facsimile model to a multimedia model one. Interactivity in online journals. *Anales de Documentación, 20*(2). https://doi.org/10.6018/analesdoc.20.2.282401

Choo, J.; Cheong, S.; Lee, Y.L. & Teh, S.H. (2012). I2Navi: An Indoor Interactive NFC Navigation System for Android Smartphones. *International Journal of Electronics and Communication Engineering, 6*(12), 1408 - 1412. https://doi.org/10.5281/zenodo.1330963

Guallar, J.; Abadal, E. & Codina, L.(2012). Hemerotecas de prensa digital. Evolución y tendencias. *Profesional de la Informacion, 21*(6), 595-605. https://doi.org/10.3145/epi.2012.nov.06

Guallar, J.; Anton, L.; Pedraza-Jiménez, R. & Pérez-Montoro, M. (2021). Curación de noticias en el correo electrónico: análisis de newsletters periodísticas españolas. *Revista Latina De Comunicación Social*, (79), 47-64. https://doi.org/10.4185/RLCS-2020-1488

Harvey, M. & Pointon, M. (2019). Understanding in-context interaction: An investigation into on-the-go mobile search. *Information Processing and Management, 56*(6). https://doi.org/10.1016/j.ipm.2019.102089

Ksiazek, T.; Peer, L. & Lessard, K. (2016). User engagement with online news: Conceptualizing interactivity and exploring the relationship between online news videos and user comments. *New Media & Society, 18*(3), 502-520. https://doi.org/10.1177/1461444814545073

Larsson, A.O.(2012). Interactivity on Swedish newspaper websites: What kind, how much and why? *Convergence, 18*(2), 195-213. https://doi.org/10.1177/1354856511430184

Linares, J.; Codina, L. & Freixa, P. (2015). *Tendencias en Cibermedios. Análisis de la interactividad*. Barcelona: Universitat Pompeu Fabra, Departament de Comunicació. http://hdl.handle.net/10230/24636

Lopezosa, C.; Iglesias-García, M.; González-Díaz, C. & Codina, L. (2020). Experiencia de búsqueda en cibermedios: análisis comparativo de diarios nativos digitales. *Revista Española De Documentación Científica*, 43(1), e254. https://doi.org/10.3989/redc.2020.1.1677

Lopezosa, C.; Codina, L. & Freixa, P. (2021). Protocolo de análisis para evaluar la experiencia de búsqueda en medios digitales. *Estudios sobre el mensaje periodístico*, 24(4). https://doi.org/10.5209/esmp.72988

Nielsen, J. & Molich, R. (1990). Heuristic evaluation of user interfaces. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems,* (pp. 249-256). https://doi.org/10.1145/97243.97281

Paskin, D. (2018). News publishing across platforms: Gatekeeping for print, web, Facebook and Twitter. *Newspaper Research Journal, 39*(4), 376-388. https://doi.org/10.1177/0739532918806897

Pedraza-Jiménez, R.; Codina, L. & Guallar, J. (2016*). Calidad en sitios web: método de análisis general, e-commerce, imágenes, hemerotecas y turismo.* Barcelona: Editorial UOC.

Reyna, J.; Hanham, J. & Meier, P. (2018). The internet explosion, digital media principles and implications to communicate effectively in the digital space. *E-Learning and Digital Media, 15*(1), 36–52. https://doi.org/10.1177/2042753018754361

Rodríguez-Martínez, R.; Codina, L. & Pedraza-Jiménez, R.(2012). Indicadores para la evaluación de la calidad en medios digitales: análisis de la interacción y de la adopción de la Web 2.0. *Revista española de Documentación Científica, 35*(1), 61-93. https://doi.org/10.3989/redc.2012.1.858

Santos-Hermosa, G.; Lopezosa, C. & Codina, L. (2022). *Interactividad, buscabilidad y visibilidad web en el periodismo digital: Protocolo de análisis y estudio comparativo*. Departamento de Comunicación. Serie Editorial DigiDoc - Observatorio de Cibermedios. Marzo 2022. RTI04/2021.

Simon, J. P. (2016). User generated content - users, community of users and firms: Toward new sources of co-innovation?. *The Journal of Policy, Regulation and Strategy for Telecommunications, Information and Media, 8*(6), 4-25. https://doi.org/10.1108/INFO-04-2016-0015

Javier Guallar
Universitat de Barcelona
jguallar@ub.edu

Rafael Pedraza-Jiménez
Universitat Pompeu Fabra
rafael.pedraza@upf.edu

Mario Pérez-Montoro
Universitat de Barcelona
perez-montoro@ub.edu

Curation Analysis System (CAS): A method for analysing digital media content curation

https://doi.org/10.3145/indocs.2022.3

Guallar, Javier; Pedraza-Jiménez, Rafael; Pérez-Montoro, Mario (2022). Curation Analysis System (CAS): A method for analysing digital media content curation. In: Freixa, Pere; Codina, Lluís; Pérez-Montoro, Mario; Guallar, Javier (ed.). *Visualisations and narratives in digital media. Methods and current trends*, (pp. 26-46). Barcelona: DigiDoc-EPI. https://doi.org/10.3145/indocs.2022.3

Abstract

We present a method or system for the analysis and evaluation of content curation in the digital media, which we call the Curation Analysis System (CAS). The system comprises two main dimensions – namely, Curation and Content – which, in turn, are made up of a series of parameters and indicators to which we assign a score. Each parameter is described in detail under the following four headings: Definition, Explanation, Procedure, and Examples. Procedure includes a list of indicators, and for each of them an analytical question and its corresponding scoring system. In Examples, a set of good content curation practices for use in the digital media are presented, in this case, related to the same format, that of newsletters.

Keywords

Content curation, news curation, digital media, newsletters, information sources, evaluation systems, indicators.

Título

***Curation Analysis System* (CAS), método para analizar la curación de contenidos en medios digitales**

Resumen

Este trabajo presenta un método o sistema para el análisis y la evaluación de la curación de contenidos en medios digitales, denominado Curation Analysis System (CAS). Este sistema se basa en dos dimensiones (Curación y Contenido), que agrupan una serie de parámetros y de indicadores a los cuales se les ha asignado una puntuación. Para cada parámetro se señalan los siguientes elementos: Definición, Explicación, Procedimiento y Ejemplos. En el Procedimiento se incluye una relación de indicadores, y para cada uno de ellos una pregunta de análisis y una puntuación. En Ejemplos, se presentan un conjunto de buenas prácticas de curación de contenidos por parte de medios digitales, en este caso, en un mismo formato, newsletters.

Palabras clave

Curación de contenido, curación de noticia, medios digitales, boletines, fuentes de información, sistemas de evaluación, indicadores.

1. Introduction

The concept of "content curation" is of recent origin, usually attributed to an article published in 2009 by marketing professor, Rohit Bhargava, entitled "Manifesto for the Content Curator" (Bhargava, 2009). In these times of exponential growth in digital content, the author defends, the need for a new professional activity – "the next big social media job" – specialised in selecting the most relevant content for a given audience, sharing it with that audience in the best way possible, and endowing it with added value.

Since its emergence in the digital marketing sector, the concept has been adopted in other fields that work with digital information, most notably in the Information Sciences, Journalism and Communication, Education, and Engineering and Computer Science. Today, more than a decade after its first appearance, the concept of content curation is well established in the professional and academic literature of various disciplines and specialties.

In the specialist literature, for example, we find works ranging from studies of curation in a broad and transversal sense, including professional monographs (Rosenbaum, 2011; Guallar and Leiva-Aguilera, 2013) and essays (Bhaskar, 2016), to those that address the application of content curation in different sectors, including journalism and digital media, the specific area that concerns us here.

In this field, several studies are worth highlighting, most notably Bradshaw (2013); Guerrini (2013); Cappelletti Júnior and Domínguez Quintas (2014); Diaz Arias (2015); Cui and Liu (2017); Guallar (2017a, b); Bruns (2018); Codina (2018); Guallar and Codina (2018); Guallar et al. (2021a, b); Silva-Rodriguez, 2021; Sixto-García, Escandón Montenegro and Luchessi (2022).

Here, however, we seek to take a different approach to these earlier studies and propose a system for the analysis and evaluation of content curation in digital media, which we baptise, the Curation Analysis System or CAS, and which can be applied in various contexts. An initial version of this method was first described in Guallar et al. (2021b) and a number of reports of its implementation have since been provided (Guallar et al., 2021a; Guallar et al., 2022).

In its development, we have adopted methodologies employed in systems analysis and the analysis of information products as well as in expert observation (Glaser and Strauss, 1967; Creswell, 2009; Denzin and Lincoln, 2011; Pedraza-Jiménez et al., 2016; Ferran-Ferrer et al., 2017; Morales et al., 2020). In this way, we identified a series of elements that enable us to characterise, analyse and evaluate the object of study. Based on this work, we were able to design and build our new system – CAS – for the evaluation of content curation in the online news media.

In an analytical system of this type, the parameters identified refer to general aspects of online journalism products that can be studied. Each parameter, in turn, consists of a set of indicators and it is the analysis of the latter that enables us to make a value judgment about the corresponding parameter. Here, each indicator is assigned a specific scoring system. This might be binary (i.e. 0–1), when assessing the simple presence or absence of a particular feature, or multiple (e.g. 0 –-3), when a more detailed evaluation of the feature in question is required – ranging, for example, from 0 – poor – to 3 – very good.

To obtain the indicators, we conducted an in-depth study of existing proposals in the specialised literature – specifically Deshpande (2013), Barnhurst (2013), Cui and Liu (2017), Guallar (2017a), Orero and Cebrián-Enrique (2019), Guallar et al. (2021b) – which, from a variety of different research angles, have employed a rich diversity of elements in their analyses of content curation. In the following sections, we present the CAS and indicate, for each parameter, the bibliographical references considered and their specific contributions in each instance.

2. Curation Analysis System: general characteristics

The system, as discussed, is based on a set of parameters and their corresponding indicators, which in turn have been grouped into two dimensions: Content and Curation, as shown in Table 1.

A. Content	
A1. Amount of content	A1.1. Quantity
A2. Time range	A2.1. Retrospective or timeless information A2.2. Recent information A2.3. Current information A2.4. Real-time information
A3. Origin	A3.1. Own content A3.2. External content
A4. Source by type of organisation	A4.1 Official sources A4.2 Corporate sources A4.3 News media A4.4 Private citizens

A5. Source by morphology	A5.1 Web sites A5.2 Blogs A5.3 Social networks A5.4 Secondary sources
B. Curation	
B1. Authorship	B1.1. Authorship
B2. Sense making technique	B2.1. Retitling B2.2. Summarising B2.3. Commenting B2.4. Quoting B2.5. Storyboarding B2.6. Parallelising
B3. Function of the hyperlink	B3.1. Without modifying B3.2. Describing B3.3. Contextualising B3.4. Interpreting B3.5. Quoting source B3.6. Citing author B3.7. Calling to action

Table 1. CAS – System for the analysis of the curation of journalism content. Summary of its dimensions, parameters and indicators.

The system is described in detail in section 3 below. For each parameter, we provide the following information:

- ***Definition***: brief definition.
- ***Explanation***: detailed description.
- ***Procedure***: methodology to be employed in the evaluation of its use, including its corresponding list of indicators, and for each of these, the specific analytical question that needs to be asked and the scoring system to be used.

- ***Examples***: examples of the application of the CAS to digital media. In this study, our examples are of the same type of journalism product – that is, newsletters – from different online news media outlets and countries.

The system seeks to be exhaustive, that is, it aims to cover all aspects that might affect the quality of curation in digital media; moreover, in so doing, it uses a scoring method that does not assign greater weight to any one particular quality. Indeed, it is left to the discretion of the individual analyst or evaluator as to whether they opt to make a total or partial use of the system – that is, if they are only interested in some of the parameters – or whether they prefer to modify the scores proposed – that is, if they wish to give greater weight to one or more of the indicators. For example, in the case of the parameter time range (which captures the temporal quality of the curated content), an evaluator might attach greater importance to retrospective content than to real-time content and so opt to score them differently, on the grounds that (from their particular perspective) the better online news media use retrospective information systems.

3. Dimension A: Content

A1. Amount of curated content

Definition

Amount of curated content found in a journalism or information product.

Explanation

Here, when referring to journalism or information products, we understand those units of analysis that constitute the specific object of study – they might include, for example, home page news items, news stories from a given section of the newspaper, newsletters, etc. – and, moreover, the parameter can only be applied in comparative analyses of like products: for example, if we wish to compare the newsletters of different news media outlets or, alternatively, their home page news, etc. In order to identify the curated content in a certain journalism product, we need to determine whether a link exists (be it in the text, in an image or in the form of embedded content) to content that is independent of and external to the content being analysed. If there is, in fact, no link giving access to other original content, even though it might be mentioned, it cannot be considered curated content.

One of the clearest distinctions that can be made between curated journalism products and other types of products is the amount of curated content they offer. In our system of evaluation, a score (from 1 to 3) is assigned on the understanding that the amount of curated content available within a product is in and of itself a criterion of the quality of that product.

Procedure

Only one indicator – that of the amount of curated content – is considered to measure this parameter.

A1.1. Quantity of curated content.

CAS question: How much curated content is there in the analysed product?

Score: 0-none; 1-lower tercile; 2-middle tercile; 3-upper tercile

First, the products analysed need to be of the same type: for example, home pages in digital news media, a given section of the same news medium, a newsletter, etc. Then, a score is assigned corresponding to one of three distinct levels (terciles) of curated content, with the highest number of curated sources being considered the benchmark of good journalistic curatorial practices. This number is then broken down into three parts, with 3 points being assigned to products with an amount of curated content in the upper tercile, 2 points to those in the middle, and 1 point for those in the lower tercile.

Examples

- "Coronavirus Briefing", *The New York Times*, 24/04/2020. A1: 3 points
- "VOX Care", *VOX*, 24/04/2020. A1: 1 point

For an online newsletter, based on the observation of the digital press, a possible score for this indicator might be established as follows: 1 point for amounts of curated content ranging from 1–10 per newsletter, 2 points for amounts ranging from 11–20; and 3 points for amounts above 21.

Applying this system of scoring to two US newsletters of the same type (specialising in providing coronavirus briefings) and published on the same day in 2020: *The New York Times*' "Coronavirus Briefing" included 33 pieces of curated content and so would be assigned 3 points for this indicator, whereas *VOX's* "VOX Care" included 10 pieces and would be assigned 1 point.

A2. Time range of the curated content

Definition

Time range in which the curated content can be delimited based on the date on which it was originally published.

Explanation

Four time frames can be distinguished: retrospective or timeless information (from previous months or years); recent information (from the last few days or weeks), current information (from the last few hours), and real-time information.

In the case of *journalistic curatorial practices*, it is useful, and relatively straightforward, to differentiate information according to these time frames: the first (i.e. retrospective information) is very closely related to the traditional practice of documentation in journalism, while the rest have grown exponentially in recent years as the social media and real-time information have gained in importance in the news habits of audiences.

Procedure

Four indicators are considered, one for each of the time frames established.

A2.1. Retrospective or timeless information.

CAS question: Does the product have retrospective curated content, from previous months or years, or what can be considered timeless content?

Score: 0–No; 1–Yes

A2.2. Recent information.

CAS question: Does the product have recent curated content, from the last few days or weeks?

Score: 0–No; 1–Yes

A2.3. Current information.

CAS question: Does the product have current curated content, from the last few hours?

Score: 0–No; 1–Yes

A2.4. Real-time information.

CAS question: Does the product have any real-time curated content?

Score: 0–No; 1–Yes

The indicators resulting from the application of this parameter are not mutually exclusive, but rather compatible with each other, and, moreover, a measure of quality of a journalism product when it is capable of offering content in all four frames. Thus, the minimum score that a curated product can obtain for this parameter is 1 and the maximum is 4.

Examples

- "Crisis del coronavirus", *ABC*, 24/04/2020. A2: 2 points
- "Coronavirus: lo que debes saber hoy", *Eldiario.es*, 24/04/2020. A2: 4 points

In the Spanish newsletter "Crisis del coronavirus", published by the newspaper *ABC* on 24 April 2020, its curated content can be assigned to two-time frames (current and retrospective) and so is awarded a score of 2 points, whereas "Coronavirus: lo que debes saber hoy", published by *Eldiario.es* presents content in all four time frames and so is awarded the maximum score of 4 points.

A3. Origin of curated content

Definition

The origin of the curated content in terms of the media outlet or organisation that publishes it.

Explanation

The curated content can originate from one of two sources as far as the organisation that produces the analysed product is concerned: it might be either content external to the media outlet itself or its own content. Usually, an outlet's own curated content is very much in the majority in the links of journalism products, although some studies point to the need to qualify this claim in some cases: for example, Guallar et al. (2022) report that pure digital products tend to have more external content than is the case in the legacy media.

Procedure

Two indicators are considered in the assessment of this parameter, one for each of the categories established.

A3.1. Own content.

CAS question: Does the product contain its own curated content?

Score: 0–No; 1–Yes

A3.2. External content.

CAS question: Does the product contain external curated content?

Score: 0–No; 1–Yes

The indicators of this parameter are not mutually exclusive, but rather compatible with each other, and, moreover, in this context, a measure of the quality of a journalism product when it is capable of offering both types of content. Thus, the minimum score that a curated product can obtain for this parameter is 1 and the maximum is 2.

Examples

- "Das 18-Uhr-Update zur Corona-Krise", *Die Welt*, 24/04/2020. A3: 2 points
- "F.A.Z. Newsletter Coronavirus", *Frankfurter Allgemeine Zeitung*, 24/04/2020. A3: 1 point

In the newsletter "Das 18-Uhr-Update zur Corona-Krise", published by the German media outlet *Die Welt*, we find both own and external curated content, and so we assign it 2 points, whereas the "F.A.Z. Newsletter Coronavirus", published by the *Frankfurter Allgemeine Zeitung*, only contains content from the newspaper itself, so it is assigned just 1 point.

A4. Source of the curated content by type of organisation

Definition

The source of the curated content in terms of the media outlet or organisation that produces it.

Explanation

Our proposal considers two distinct parameters in relation to the source of the information: first, according to the type of organisation responsible for producing it and, second, according to its formatting characteristics or morphology. Here, in section A4 – type of organisation – a distinction is drawn between official sources (Government and public agencies), corporate sources (companies or other private organisations), the news media, and private citizens/ online communities (collaborative wikis, hashtags, online forums, etc).

Procedure

For this parameter, four indicators are considered for each of the categories established

A4.1. Content originating from official sources.

CAS question: Does the product contain curated content from official or government sources?

Score: 0–No; 1–Yes

A4.2. Content originating from corporate sources.

CAS question: Does the product contain curated content from corporate sources, companies or other private organisations?

Score: 0–No; 1–Yes

A4.3. Content originating from the news media.

CAS question: Does the product contain curated content originating from the news media?

Score: 0–No; 1–Yes

A4.4. Content originating from private citizens

CAS question: Does the product contain curated content from personal sources or from private citizens?

Score: 0–No; 1–Yes

This parameter, like the previous two, comprises indicators that are not mutually exclusive and where the variety of sources can be considered a criterion of quality. Thus, 1 point indicates a product with a minimum variety of curated sources, while 4 points indicates a product with a great variety of sources.

Examples

- "Politico Nightly. Coronavirus special edition", *Politico*, 0/04/2020. A4: 4 points
- "Coronavirus now", *Boston Globe*, 3/04/2020. A4: 1 point

In these US news media newsletters "Politico Nightly. Coronavirus special edition" published by *Politico* and "Coronavirus now" published by the *Boston Globe*, the former presents four source types by organisation, and so is assigned 4 points, while the latter only uses news media sources and so is assigned 1 point.

A5. Source of the curated content by morphology

Definition

This parameter evaluates the morphology or formatting characteristics of the source of the curated content.

Explanation

A distinction is drawn between web sites, blogs, social networks and secondary sources (including data bases, catalogues, etc.).

It should be stressed that these two parameters – A4 and A5 – of curated content sources based on the type of organisation and their morphology, respectively, are not overlapping parameters, because while they both refer to sources of information, they are classified in accordance with distinct criteria. Thus, for example, the curated content originating from a blog (parameter A5), could originate from an official source, a corporate source, a news media outlet or a private citizen (parameter A4).

Procedure

For the assessment of this parameter, four indicators are considered, one for each of the categories established.

A5.1. Content originating from web sites.

CAS question: Does the product contain curated content originating from web sites?

Score: 0–No; 1–Yes

A5.2. Content originating from blogs.

CAS question: Does the product contain curated content originating from blogs?

Score: 0–No; 1–Yes

A5.3. Content originating from social platforms.

CAS question: Does the product contain curated content originating from social platforms?

Score: 0–No; 1–Yes

A5.4. Content originating from secondary sources.

CAS question: Does the product contain curated content originating from secondary sources, such as data bases, catalogues, etc.?

Score: 0–No; 1–Yes

As in the previous parameter, the variety of source types according to the formatting characteristics of the curated content is considered a criterion of quality and, thus, a score of 1 indicates a minimum and a score of 4 indicates the maximum of variety possible.

Examples

- "Coronavirus Updates", *Washington Post*, 10/04/2020. A5: 3 points

- "Coronavirus Updates", *Washington Post*, 13/11/2020. A5: 1 point

The newsletter of the US newspaper, the *Washington Post*, in the edition published on the 10 April 2020, curates' content from three source types (websites, blogs and social networks), and as such is assigned 3 points, whereas the newsletter corresponding to the 13 November 2020 edition has content from just one type (websites), and so obtains 1 point.

4. Dimension B: Curation

B1. Authorship of curation

Definition

This parameter determines whether the curated product has a signed author.

Explanation

A key element in evaluating the quality of any digital product or service is a clear, unequivocal identification of its authorship.

Procedure

This parameter includes an indicator of the unequivocal identification of the authorship of the journalism product.

B1.1. Authorship of the curated content.

CAS question: Is the authorship of the product analysed visible?

Score: 0–No; 1–Yes

In the case of a product of curated content, the assessment of its authorship has similar implications to those associated with any other kind of digital content. Thus, it is considered an element of quality if its authorship is made clearly and accurately known to the reader.

Examples

- "Coronavirus: l'essentiel à savoir", *Le Parisien*, 27/11/2020. B1: 1 point
- "Coronavirus: l'essentiel à savoir", *Le Parisie*n, 24/04/2020. B1: 0 points

In the newsletter published by the French newspaper, *Le Parisien*, the edition dated 27 November 2020 is signed by Laurence Lefour and Joffrey Vovos, whereas the edition dated 24 April is unsigned – thus the former is assigned 1 point and the latter 0 points.

B2. Sense making techniques

Definition

This parameter assesses the visible use made of techniques of sense making, that is techniques that add value to curated content.

Explanation

The following techniques – drawing above all on Deshpande (2013) and Guallar et al. (2021b) – are taken into consideration:

- retitling: (usually, only applied in the case of using a single source) curating the content with a different title to that used in the original.
- summarising: providing an informative or objective summary of the curated content.
- commenting: providing a personal or subjective summary or a text expressing an opinion.
- quoting: including a textual quote from the curated content.
- storyboarding: combining in one product various pieces of content in different formats (for example, text extracts, photos, tweets, videos) with the text itself.
- parallelising: presenting two or more pieces of content that apparently have no link but for which the curator finds a relevant relation and justifies this link.

These techniques can be used independently of each other or in combination. For example, the use of the last two, which require a considerable degree of elaboration, normally require the use of one or more of the other techniques from the list.

Procedure

For the evaluation of this parameter, six indicators are considered, one for each of the techniques described above.

B2.1. Retitling.

CAS question: Does the analysed product change the title – retitle – of the original curated content?

Score: 0–No; 1–Yes

B2.2. Summarising.

CAS question: Does the analysed product include an informative or objective summary of the curated content?

Score: 0–No; 1–Yes

B2.3. Commenting.

CAS question: Does the analysed product include a personal or subjective summary or a text expressing an opinion?

Score: 0–No; 1–Yes

B2.4. Quoting.

CAS question: Does the analysed product include a textual quote from the curated content?

Score: 0–No; 1–Yes

B2.5. Storyboarding.

CAS question: Does the analysed product combine various pieces of content in different formats?

Score: 0–No; 1–Yes

B2.6. Paralleling.

CAS question: Does the analysed product relate two or more pieces of content that previously presented no apparent link?

Score: 0–No; 1–Yes

This parameter for evaluating the curation quality can cause the score of the evaluated product to vary between 0 points, if the curation is fully automated, and a maximum of 6 points, should all six techniques be combined. However, a high combination of these techniques is very rare.

Examples

- "Coronavirus: lo que debes saber hoy", *Eldiario.es*, 17/04/2020. B3: 4 points
- "Especial Coronavirus", *El Mundo*, 17/04/2020. B3: 0 points

In the Spanish newsletter "Coronavirus: lo que debes saber hoy", published by *Eldiario.es* on 17 April 2020, we find four sense making techniques being used (summarising, commenting, quoting and storyboarding), for which 4 points are assigned; and on the same day in "Coronavirus Special", published by *El Mundo*, no techniques are used given that it is an automated newsletter, and so it is awarded 0 points.

B3. Informative function of the hyperlink

Definition

This parameter assesses the function that the curated content fulfils within the analysed product from a journalistic or informative perspective.

Explanation

Here, we are concerned with the intentionality or the purpose of each piece of curated content within the analysed product. Three of the categories are taken from Cui and Liu (2017) – namely, sourcing curation, contextualising curation and interpreting curation – while the others are based on our direct observations of the digital media. This parameter focuses exclusively on the function of each hyperlink – not on that of the rest of the content written by the curator – and which can be considered as completing the product in question. We include the following categories:

- Without modifying: original content unmodified by the curator. Two types have been identified: hyperlinks in which the text corresponds to the original title of the document, without undergoing any modifications by the curator; and embedded documents originating from social platforms.

- Describing: curated content as description or summary. The hypertexts of the articles reflect, partially or totally, the content of the linked source and without clicking on the original content the reader should be able to form a general idea about it. This function is related with the sense making technique of summarising.

- Contextualising: the curated content is used fundamentally to contextualise information (contextualising curation): while the content might not be directly related to the news story, it can help explain its social, cultural or historical contexts, etc. It exploits the curation of timeless or retrospective information sources.

- Interpreting: the curated content is interpreted by the curator (interpreting curation), that is, the hyperlinks do not have to provide descriptive or contextual information but rather interpretations of the content. This function is directly related with the sense making technique of commenting.

- Quoting source: when the hyperlink text does not refer to the subject matter of the content that is accessed, but rather it identifies the source where it has been published.
- Citing author: as in the case above, but it identifies the author.
- Calling to action: a call to enter a hyperlink, via a text of the type "click here".

Procedure

For the evaluation of this parameter, seven indicators are considered, one for each of the functions described above.

B3.1. Without modifying

CAS question: Does the analysed product provide access to curated content without any modifications in its title or directly as embedded content?

Score: 0–No; 1–Yes

B3.2. Describing.

CAS question: Does the analysed product provide access to curated content with a descriptive or summary type hyperlink or embedded link?

Score: 0–No; 1–Yes

B3.3. Contextualising.

CAS question: Does the analysed product provide access to curated content with a contextualised hyperlink or embedded link?

Score: 0–No; 1–Yes

B3.4. Interpreting.

CAS question: Does the analysed product provide access to curated content with an interpretative or opinion-type hyperlink or embedded link?

Score: 0–No; 1–Yes

B3.5. Quoting source.

CAS question: Does the analysed product provide access to curated content by indicating the source?

Score: 0–No; 1–Yes

B3.6. Citing author.

CAS question: Does the analysed product provide access to curated content by indicating the author?

Score: 0–No; 1–Yes

B3.7. Calling to action.

CAS question: Does the analysed product provide access by making an explicit call to action?

Score: 0–No; 1–Yes

These are non-exclusive functions and they are likely to complement each other, although it is unlikely to find a combination of all seven in one product. The minimum score is therefore 1 point and the maximum is 7.

Examples

- "Themenspezial Coronavirus", *Süddeutsche Zeitung*, 20/11/2020. B3: 4 points
- "F.A.Z. Newsletter Coronavirus", *Frankfurter Allgemeine Zeitun*g, 20/11/2020. B3: 2 points

In the newsletter published by the German newspaper, *Süddeutsche Zeitung*, entitled "Themenspezial Coronavirus", we find four types of function being used in the links (sourcing curation, calling to action, citing author and without modifying), and, as such, it would be awarded 4 points, while in "F.A.Z. Newsletter Coronavirus", published by *Frankfurter Allgemeine Zeitung*, we find two types (sourcing curation and without modifying) and so it is awarded 2 points.

5. Conclusions

This system of curation analysis or CAS is the first to be proposed in the academic literature for the evaluation of content curation in digital media. The system does much more than simply describe this process of curation, as it also enables the analyst to identify the specific components of journalistic curatorial practice, to undertake assessments of quality, and to create derived products, for example, rankings, etc.

The CAS promotes the value attached to the variety of content curation practises, in the belief that a greater variety results in an enhanced quality of that content. For example, parameter A4, Sources of curated content according to the type of organisation, understands that

a journalism product that curates content from different types of source – such as official organisations, private organisations, associations, citizens and news media – will be of a higher quality than a product that only offers content from a single type of information source, e.g. the news media (something that is currently very common). In this case, therefore, the greater the variety of source types, the better the curation is likely to be.

Indeed, more generally, we consider that the greater the variety of use made of each of the elements that combine to make up the curatorial practice, that is, from the source types to the types of sense making techniques, will have an impact in terms of improving the quality of curation and, by extension, the quality of journalism offered to the readership.

As with any analytic system, the CAS needs to be subject to constant testing and fine tuning. Future studies related to its use might profitably involve its application to bounded samples of digital media or journalism products, for example: 1) the digital media of a given country, the main news media outlets of a given geographical area or even globally; 2) given journalistic products, such as newsletters, multimedia reports, social media networks, etc.; or 3) thematic news content, which might include politics, sport, technology, etc.

In short, with this proposal and via successive applications, the aim is to provide a better understanding of the potential of content curation, which, in turn, should result in the improvement of its use by digital media, both in the professional world and in that of academia and research.

References

Barnhurst, KG. (2013). Newspapers experiment online: Story content after a decade on the web. *Journalism, 14*, 3–21. https://doi.org/10.1177/1464884912448898

Bhargava, R. (september 30, 2009). Manifesto For The Content Curator: The Next Big Social Media Job Of The Future?. *Rohit Bhargava.com*. http://www.rohitbhargava.com/2009/09/manifesto-for-the-content-curator-the-next-big-social-media-job-of-the-future-.html

Bhaskhar, M. (2016). *Curation. The power of selection in a world of excess*. [ebook] https://www.curationthebook.com

Bradshaw, P. (september 30, 2013). Journalism *is* curation: tips on curation tools and techniques. *Online journalism Blog*. http://onlinejournalismblog.com/2013/09/30/curation-tools-tips-advice-journalism

Bruns, A. (2018). *Gatewatching and news curation: Journalism, social media, and the public sphere*. Peter Lang.

Cappelletti Júnior, M. & Domínguez Quintas, S. (2014). La curaduría de contenidos y la narrativa colaborativa en el Ciberperiodismo: estudio del caso de Storify en el diario digital elpais.com. *Estudios del mensaje periodístico, 20*(1) https://revistas.ucm.es/index.php/ESMP/article/view/45216

Codina, L. (2018). *Curación de contenidos para periodistas: definición, esquema básico y recursos*. Universitat Pompeu Fabra, Facultad de Comunicación, Área de Formatos Digitales e Interactivos. https://repositori.upf.edu/handle/10230/34369

Creswell, J.W. (2009). *Research Design. Qualitative, Quantitative and Mixed Methods Approach*. Sage.

Cui, X. & Liu, Y. (2017). How does online news curate linked sources? A content analysis of three online news media. *Journalism, 18*(7), 852-870. https://doi.org/10.1177/1464884916663621

Denzin, N. K. & Lincoln, Y. S. (2011). *The Sage handbook of qualitative research*. Sage.

Deshpande, P. (August 13, 2013). Six content curation templates for content annotation. *The Curata blog*. https://www.curata.com/blog/6-content-curation-templates-for-content-annotation

Díaz Arias, R. (2015). Curaduría periodística, una forma de reconstruir el espacio público. *Estudios del mensaje periodístico*, (21). https://revistas.ucm.es/index.php/ESMP/article/view/51129

Ferran-Ferrer, N.; Guallar, J.; Abadal, E. & Server, A. (2017). Research methods and techniques in Spanish library and information science journals (2012-2014). *Information Research, 22*(1). http://eprints.rclis.org/32191/

Glaser, B. & Strauss, A.L. (1967). *The Discovery of Grounded Theory*. Aldine.

Guallar, J. (2017a). Artículos de curación de contenidos. Categorías y ejemplos. *Anuario ThinkEPI, 11*, 210-216. https://doi.org/10.3145/thinkepi.2017.38

Guallar, J. (2017b). Content curation in digital media: Between retrospective and real-time information. In: Campos Freire, F. et al. *Media and Metamedia Management,* (pp. 37-46). Springer International Publishing.

Guallar, J.; Anton, L.; Pedraza-Jiménez, R. & Pérez-Montoro, M. (2021a). News curation by email: Analysis of the Spanish journalistic newsletters. *Revista Latina de Comunicación Social*, (79), 47-64. https://doi.org/10.4185/RLCS-2020-1488

Guallar, J. & Codina, L. (2018). Journalistic content curation and news librarianship: Differential characteristics and necessary convergence. *El profesional de la información, 27*(4), 778-791. https://doi.org/10.3145/epi.2018.jul.07

Guallar, J.; Franch, P; Boté, J.J. & Anton, L. (2022). How do legacy and digital media curate coronavirus content? An assessment of newsletters from the USA and four European countries. *Profesional de la información*. [in press].

Guallar, J. & Leiva-Aguilera, J. (2013). *El content curator. Guía básica para el nuevo profesional de internet*. Profesional de la información & Editorial UOC.

Guallar, J.; Pedraza-Jiménez, R.; Pérez-Montoro, M. & Anton, L. (2021b). Curación de contenidos en periodismo. Indicadores y buenas prácticas. *Revista española de documentación científica, 44*(2). https://doi.org/10.3989/redc.2021.2.1742

Guerrini, F. (2013). *Newsroom curators & independent storytellers: content curation as a new form of journalism*. Reuters Institute for the Study of Journalism, University of Oxford. https://reutersinstitute.politics.ox.ac.uk/our-research/newsroom-curators-and-independent-storytellers-content-curation-new-form-journalism

Karlsson, M.; Clerwall, C. & Örnebring, H. (2015). Hyperlinking practices in Swedish online news 2007-2013: The rise, fall, and stagnation of hyperlinking as a journalistic tool. *Information, communication & society, 18*(7), 847-863. https://doi.org/10.1080/1369118X.2014.984743

Morales-Vargas, A.; Pedraza-Jiménez, R. & Codina, L. (2020). Website quality: An analysis of scientific production. P*rofesional de la información, 29*(5), e290508. https://doi.org/10.3145/epi.2020.sep.08

Orero, P. & Cebrián-Enrique, B. J. (2019). Evolución del uso del hipertexto en la prensa española (2002-2016). *El profesional de la información, 28*(2), e280225. https://doi.org/10.3145/epi.2019.mar.25

Pedraza-Jiménez, R.; Codina, L. & Guallar, J. (2016). *Calidad en sitios web: método de análisis general, e-commerce, imágenes, hemerotecas y turismo*. Editorial UOC.

Rodríguez-Martínez, R.; Codina, L. & Pedraza-Jiménez, R. (2010). Cibermedios y web 2.0: modelo de

análisis y resultado de aplicación. *El profesional de la información, 19*(1), 35-44. https://doi.org/10.3145/epi.2010.ene.05

Rodríguez-Martínez, R.; Codina, L. & Pedraza-Jiménez, R. (2012). Indicadores para la evaluación de la calidad en cibermedios: análisis de la interacción y de la adopción de la Web 2.0. *Revista Española de Documentación Científica, 35*(1), 61-93. https://doi.org/10.3989/redc.2012.1.858

Rosenbaum, S. (2011). *Curation Nation: how to win in a world where consumers are creators*. McGraw-Hil.

Silva-Rodríguez, A. (2021). Emergencia de newsletters especializadas en Covid-19: información curada y actualizada en el email. *Profesional de la información, 30*(4). https://doi.org/10.3145/epi.2021.jul.10

Sixto-García, J.; Escandón Montenegro, P. & Luchessi, L. (2022). Co-creation and Curation of Contents: An Indissoluble Relationship?. In: Vázquez-Herrero, J. ***et al.*** *Total Journalism: Models, Techniques and Challenges*, (pp. 227-238). Springer Nature. https://doi.org/10.1007/978-3-030-88028-6

Raquel Escandell-Poveda
Universitat d'Alacant
r.escandell@ua.es

Mar Iglesias-García
Universitat d'Alacant
mar.iglesias@ua.es

Natalia Papí-Gálvez
Universitat d'Alacant
natalia.p@ua.es

From Memex to Google: The origin and evolution of search engines

https://doi.org/10.3145/indocs.2022.4

Escandell-Poveda, Raquel; Iglesias-García, Mar; Papí-Gálvez, Natalia (2022). From Memex to Google: The origin and evolution of search engines. In: Freixa, Pere; Codina, Lluís; Pérez-Montoro, Mario y Guallar, Javier (ed.). *Visualisations and narratives in digital media. Methods and current trends*, (pp. 47-66). Barcelona: DigiDoc-EPI. https://doi.org/10.3145/indocs.2022.4

Abstract

Search engines have become one of the main channels for accessing information on the Web. Their widespread use means that the media, companies, institutions or any agent whose objective is to be visible to or to attract digital audiences is obliged to ensure its website is ranked at the top of the results pages. In Spain alone, 88% of people use search engines daily, with Google being the most used to conduct these searches, claiming a national market share of 95% and more than 90% worldwide. Given their current importance, an in-depth understanding of these tools becomes essential and it is worth pondering just how we have reached this current situation. What exactly are the origins of search engines and how has Google come to exercise its quasi-monopoly? The objective of this study is to explore the origin of search engines and to describe the main milestones in their evolution. To do so, a bibliographic review has been carried out using the main academic databases in the social sciences comprising a narrative search, complemented by an examination of the grey literature. The result is a journey through the history of search engines from their origins and subsequent technology developments to the creation of the World Wide Web. Likewise, a study is made of the original search engines and their main characteristics, with a particular emphasis on the path taken by Google given its current position of supremacy. Understanding the past has direct implications not only for our understanding of the present reality of web searches and access to information, but it is also essential for managing the continuous digital transformation to which we are all exposed, which has repercussions for all areas of the economy and society and, of course, for communication.

Keywords

Search engines, directories, Google, algorithms, SEO, web positioning, internet history.

Título

De Memex a Google. Origen y evolución de los buscadores

Resumen

Los buscadores se han convertido en una de las principales vías de acceso a la información existente en la Web. Su uso generalizado supone que medios de comunicación, empresas, instituciones o cualquier agente que tenga como objetivo la visibilidad o atracción de audiencias digitales, esté obligado a conseguir las primeras posiciones en los resultados de búsqueda. Solo en España, un 88% de personas utilizan buscadores diariamente, siendo Google el más utilizado, con una cuota en el mercado nacional del 95% y más del 90% mundial. Ante esta tesitura, se hace imprescindible un conocimiento profundo de estas herramientas y cabe preguntarse ¿cómo se ha llegado a esta situación? ¿Cuál es el origen de estos motores de búsqueda y cómo ha llegado Google a ejercer este cuasi monopolio? El objetivo de este estudio es explorar el origen de los buscadores y exponer cuáles han sido los principales hitos en su evolución. Para ello se ha realizado una revisión bibliográfica empleando las bases de datos académicas más destacadas en ciencias sociales así como una búsqueda narrativa, complementada con la consulta de literatura gris. El resultado es un recorrido por la historia de los buscadores desde su origen, el desarrollo de la tecnología y el surgimiento de la World Wide Web. Asimismo, se presentan los buscadores primigenios y sus características más destacadas, con especial énfasis en la trayectoria de Google por su posición de supremacía. Conocer el pasado tiene implicaciones directas no solo en el entendimiento del presente de las búsquedas y el acceso a la información, sino que es imprescindible para hacer frente a la transformación

digital contínua a la que estamos expuestos, la cual tiene repercusiones en todos los ámbitos de la economía, la sociedad y, por supuesto, en la comunicación.

Palabras clave

Buscadores, motores de búsqueda, directorios, Google, algoritmos, SEO, posicionamiento web, historia de Internet.

1. Introduction

The development of digital documentation, the emergence of the internet and the subsequent exponential growth of websites in the era of the World Wide Web have gone hand in hand with an intrinsic need for systems that can facilitate the management, filtering and retrieval of this information (Baeza-Yates and Ribeiro-Neto, 1999). Indeed, it was in response to this need that search tools originated and have since evolved in tandem with an ever-expanding Web. Today, as we enter the third decade of the 21st century, the Google search engine, with a world market share that climbs above 90% (Statista, 2021b), is responsible for 66% of web traffic in the USA (Fishkin, 2019), while, every second, 91,171 searches are conducted worldwide (Internet Live Stats, 2021).

In Spain, according to the *Asociación para la Investigación de Medios de Comunicación* (AIMC, 2020), internet penetration has reached 82.8% of the population and 88% of those use a search engine daily (Fernández, 2021). At the same time, search engines lead investment in digital advertising to the tune of 977.3 million euros, equivalent to roughly a third of the total (Interactive Advertising Bureau [IAB] Spain and PricewaterhouseCoopers, 2021), highlighting the importance of this channel for brands in the digital ecosystem.

The widespread use of search engines puts them in the crosshairs of any entity that needs to gain visibility and attract audience to its website (Escandell-Poveda, 2021). Based on these premises, this study seeks to explore the origins of search tools and learn about their history, by revisiting the many milestones in their evolution down to the present day.

The methodology we adopt in so doing is based on a review of the academic literature dedicated to our object of study, using for this purpose the Scopus and Web of Science databases, complemented by that of Google Scholar. More specifically, the review has been undertaken by conducting a narrative search and by examining the grey literature, while defunct web pages have been explored using the free-access Internet Archive Wayback Machine. The

results are discussed below under different headings, beginning with the origins of hypertext and concluding with a look at current figures for search engine use around the world.

2. Memex, the germ of the concept of hypertext

Humanity has had the need to organise information and to facilitate access to it since the founding of the first libraries. Around 1945, this need became even more pressing, as the military research undertaken during successive world wars boosted the development of scientific knowledge (García-Marco, 1996). Against this backdrop, Vannevar Bush (1945) envisioned a device he called the *memex* (Memory - Index), which would allow the storage and retrieval of information not included in the classic indexes. His system based the classification of documents on their links with each other and these same trails of reference would allow the user to jump from one document to another according to their relationship. Although the device never materialized, Bush's conception of the interrelationship of documents is considered the predecessor of *hypertext*, a term coined by Ted Nelson in the 1960s, and defined as an information management system in which data are stored in a network of nodes connected by links (Smith and Weiss, 1988).

In 1965, Nelson devised Project Xanadu, a global information storage system based on an electronic repository of interlinked documents offering immediate and simultaneous access to millions of people (Rayward, 1994), a futuristic version of the *World Wide Web* that would appear thanks to the efforts of Tim Berners-Lee in the early 90s.

3. The origins: the earliest search engines

Officially, the first *search engine* in history was Archie, created in 1990 by Alan Emtage, Bill Heelan and Peter Deutsch, students at McGill University in Montreal. Archie (a name that derives from the word *archive*) allowed queries to be made in its database comprising the names of files obtained from existing FTP servers. Archie did not index the contents inside the files, which meant searches were performed only on the file names (Schwartz et al., 1992; Seymour et al., 2011) (Image 1).

In 1991, as an alternative to Archie, Mark McCahill and his team created Gopher which permitted category searches to be performed and full texts accessed (Anklesaria et al., 1993). Gopher, considered the predecessor of future directories like Yahoo, displayed information in a structured fashion and allowed navigation between documents using links (Salazar, 2005). The problem was that it was limited to conducting searches in just one server, a drawback that was overcome with the development of Veronica (Very Easy Rodent-Oriented Net-wide Index to Computerized Archives) in 1992. This search engine provided results from the different

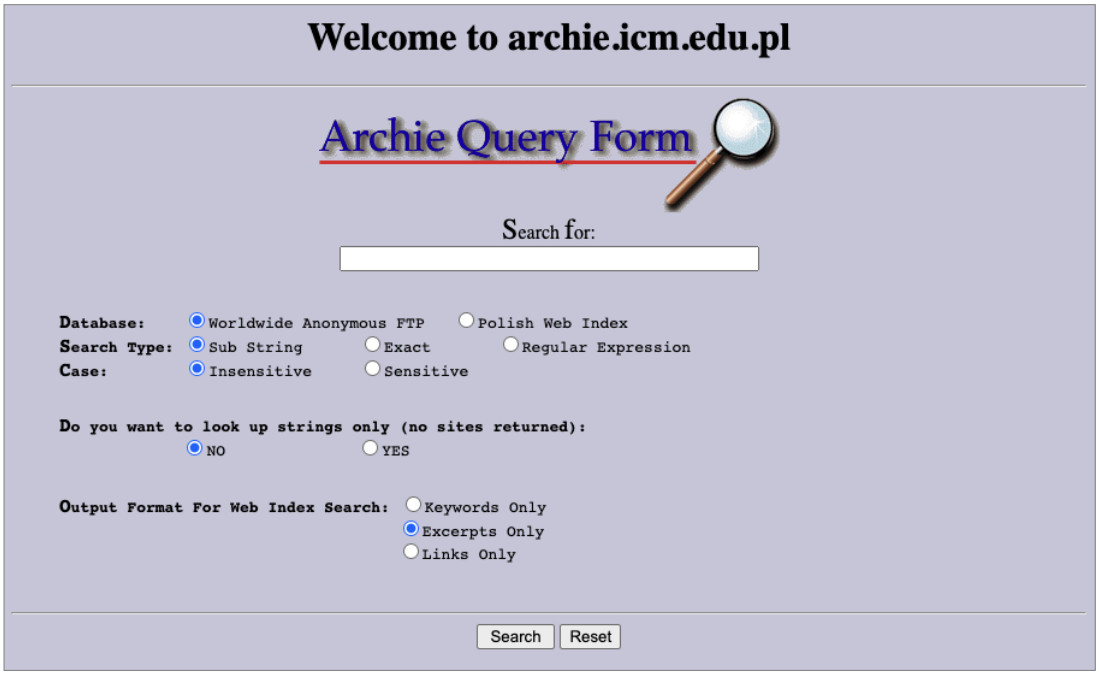


Image 1. *Archie's interface*. Source: http://archie.icm.edu.pl/archie-adv_eng.html.

Gopher servers by conducting searches with Boolean operators and keywords, although only on the menu items, without reaching the full texts. Jughead (Jonzy's Universal Gopher Hierarchy Excavation and Display), created in 1993, was another of the search engines used in "Gopher-space", although it differed from Veronica because it searched a single server at a time (Wall, 2017; Seymour et al., 2011; Smith and Updegrove, 1993; Sonnenreich, 1997; Wagner, 1996).

In parallel with these developments, Tim Berners-Lee, a scientist at CERN (European Council for Nuclear Research) was creating the *World Wide Web*, presented in 1989 as a hypertext-based system (Berners-Lee, 1989). It was created with the aim of facilitating access to information in Berner-Lee's own research, and became popular thanks to being open access, the development of HyperText Markup Language or HTML and the creation, in 1993, of the first graphical web browser called Mosaic (Griffiths, 2002; Tabarés-Gutiérrez, 2012).

With the arrival of the WWW, although lists of websites were available, such as the Index of content on the WWW Virtual Library, searches could not be conducted in them. To overcome this limitation, Oscar Nierstrasz created W3Catalog, the first World Wide Web search tool, which downloaded these resource listings and reformatted them to make them accessible through a searchable catalogue (Nierstrasz, 1996).

In June 1993, at the Massachusetts Institute of Technology (MIT), Matthew Gray developed the World Wide Web Wanderer, the first automated web crawler, whose objective was to chart the growth of the Web and discover new websites (Gray, 1996). The Wanderer robot

only visited the pages and counted them; it did not index them or save them in a database (Sherman and Price, 2001). A later extension allowed the robot to read web addresses or URLs (Uniform Resource Locators) as it went along, which generated Wandex, an index in which searches could be conducted and what is today considered the origin of web search engines as we now know them (Casares, 2008; Duklan and Bahuguna, 2012; Lara, 2014; Sonnenreich, 1997).

In October 1993, Martijn Koster created Aliweb, the Archie equivalent for Hypertext Transfer Protocol or HTTP. The fundamental difference with Wandex was that it did not use any robot crawler to create its index, rather the websites had to be registered manually (Koster, 1994). Its operation solved the problems of overloading the bandwidth that the trackers caused when entering the webs. In fact, it was Koster who created the "Robots Exclusion Protocol", a standard that prevents crawlers from entering web content through the robots.txt file (Jha et al. 2014; Koster, 1994). However, the fact of having to manually upload the websites to the directory limited its growth, since webmasters did not know how to create the necessary files to do so (Sonnenreich, 1997).

In December 1993, three search engines appeared based on automated crawlers or spiders similar to Gray's Wanderer: JumpStation, the first to combine the three essential features in one web browser: crawling, indexing and searching (Seymour et al., 2011), World Wide Web Worm and RBSE (Repository-Based Software Engineering). Like their predecessor, they only read certain fields or web tags and, the first two, presented their results in index date order. In this regard, therefore, RBSE was a pioneer, since it used a ranking based on the relevance of the search term (Eichmann, 1994; Mauldin, 1997; Sonnenreich, 1997).

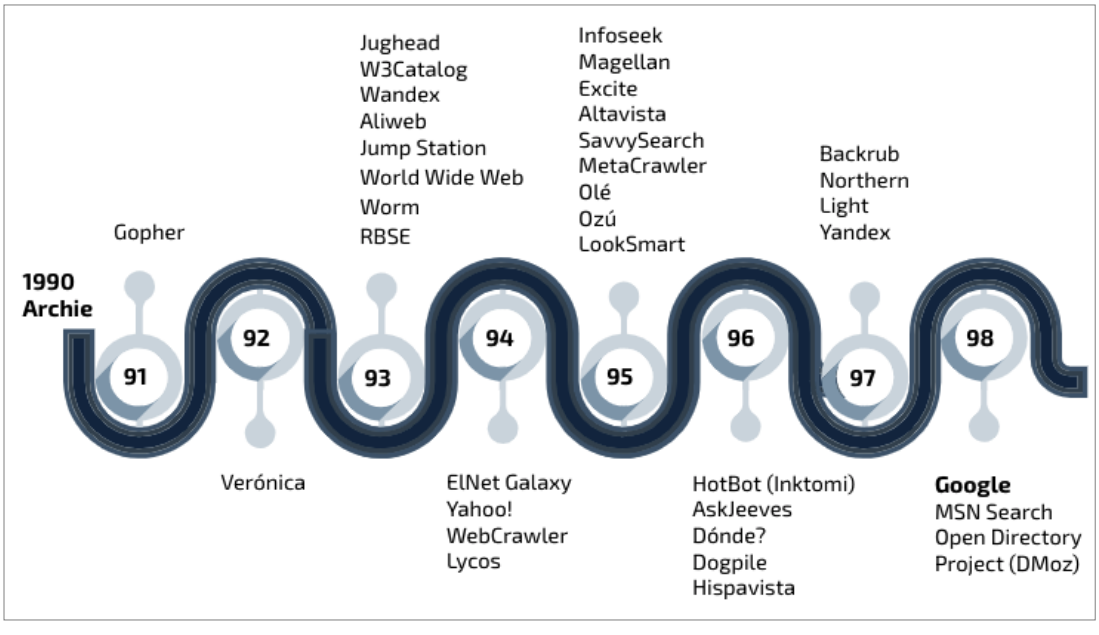


Image 2. Chronology of the earliest search engines.

The problem these early search engines faced was that the query had to use short concise words to ensure successful information retrieval. To overcome this limitation, in 1994 directories such as eiNet/Galaxy emerged, the latter considered the internet's first searchable directory (Galaxy, 2021). Web browsers of this type, rather than exploiting technological solutions, were able to create their database thanks to human editors who manually selected and classified the pages into categories and subcategories, thus allowing users to browse their index and explore the resources included on a specific topic (Sonnenreich, 1997).

Adopting this same philosophy, David Filo and Jerry Yang created Yahoo! in April 1994, a manually catalogued web directory that achieved great success and was one of the leaders in this market in the late 1990s (Trigo-Aranda, 2004). Yahoo! later evolved by incorporating a search engine that automated certain aspects of the collection and categorization of websites, thus blurring the distinction between search engine and directory (Casares 2008; Sonnenreich, 1997).

That same year, Brian Pinkerton created WebCrawler, a search engine that, unlike the then existing crawlers, indexed the entire content of the web (Seymour et al., 2011; Sonnenreich, 1997). In addition, it used a vector-based information retrieval model (Salton, 1989), which improved the results by displaying them according to relevance. This model took into account the frequency and weight of the query terms (Mauldin, 1997; Pinkerton, 1994), something that has since become the standard. Demand for WebCrawler was such that it crashed the servers of the University of Washington, where it had been created, and it was finally bought by the AOL portal (America Online), later being acquired by Excite (1997) and InfoSpace (2001). Lycos was another of the popular search engines launched in 1994. It stood out for its large number of indexed documents, which rose from 54,000 in July 1994 to 60 million in November 1996 (Mauldin, 1997) and its algorithm, which took into account the proximity between words and so improved the quality of its results (Sherman and Price, 2001). Lycos was also the first to display automatic excerpts on the search engine results pages (Sherman and Price, 2001).

In 1995 new competitors emerged in the market, among the most used were Magellan, Excite, Infoseek, Inktomi, Northern Light and AltaVista (Seymour et al., 2011). Infoseek became popular by becoming the default search engine on the Netscape navigator, while Excite introduced the concept of synonyms in its algorithm based on concepts, thanks to which results containing a related but not exact keyword could appear (Casares, 2008; Ballari, 2003).

AltaVista was the first to allow natural language querying, Boolean operators (AND, OR, NOT, ...) for queries, the possibility of searching all websites that linked to a specific page, the manual addition and deletion of websites, the conducting of searches in multimedia files and the inclusion of search tips, among other improvements (Seymour et al., 2011; Sonnenreich, 1997).

In 1995, with the curtailment of US government subsidies, privatization of the Internet began, quickly followed by massive investment in the *dotcom* commercial domains and with it the primary goal of the Web changed from being a place to provide information to serving as a place to attract customers (Ballari, 2003).

Search engines began to increase the size of their indexes, so that a search could produce hundreds of results, many of them irrelevant. Similarly, different search engines would return different results. This problem was addressed with the creation in 1995 of metasearch engines, such as MetaCrawler or SavvySearch, which simultaneously displayed the results of several search robots (Ballari, 2003; Schwartz, 1998; Sonnenreich, 1997). Faced by this new situation, website owners began to manipulate the results so as to appear at the top of the rankings (Ballari, 2003), thus laying the foundations for SEO (Search Engine Optimization).

HotBot, launched in 1996, was noted for its speed and power, being considered the first search engine capable of indexing the millions of websites that then existed, as well as using cookies to return personalized results per user (Casares, 2008; Seymour et al., 2011; Sonnenreich, 1997). That same year saw the launch of Ask Jeeves, which answered questions posed in natural language, and BackRub, the search engine that would be renamed Google in 1997, was started by two students at Stanford University (Brin and Page, 1998). Google began to grow in popularity, becoming a world leader in the early 2000s and has since maintained its dominance over its competitors, acquiring a global market share of between 80 and 90% in the second decade of the 21st century (Johnson, 2021).

The year 1997 saw the launch of Northern Light, a search engine that introduced search suggestions, but which was subsequently developed for private custom only, and Yandex, which became the largest Russian search engine (Casares, 2008; Seymour et al., 2011) until it too was surpassed by Google in 2017 (Liveinternet, 2021).

4. Search engines of the new century

In the late 1990s, search engines also formed part of the investment frenzy attributable to the internet boom. Like a multitude of companies at that time, many saw their finances plummet when the *dotcom* bubble burst between 1999 and 2001 (Seymour et al., 2011). In 1999 it was estimated that the public Web contained more than 800 million pages but that no single search engine was capable of indexing more than 16% of that total (Lawrence and Giles, 1999) and it was predicted that this number would soon reach thirteen billion pages (Sherman, 2000).

In 1999, AlltheWeb appeared, establishing itself as the owner of the largest Internet database, as did GoTo (later becoming Overture), the first to offer pay-per-click advertising (Fain and Pedersen, 2006; Van Couvering, 2008). In 2000, the Teoma search engine was launched, using

a system of link popularity, similar to Google's, based on "ExpertRank" technology. What made it different was it took into account the context of the link, giving it more importance if the topic of the page was related to that of the website receiving it (Casares, 2008; Seymour et al., 2011). In that same year, Baidu, the Chinese and Japanese website, image and audio search engine was also launched, and which in 2021 continues to lead the search engine market in China (StatCounter, 2021a).

Despite the success enjoyed by Google from the early years of that decade – further boosted by the fact that Yahoo! included Google as a search engine on its homepage (Baker, 2021) – new search engines continued to appear: In 2000, Exalead was created in France, displaying thumbnail previews and providing advanced refining of its results; in 2003, Seekport was launched, offering local, regional versions adapted to each country; in 2004, Nutch, an open source web crawler, was created; and, in the same year, Clusty, which grouped results into thematic categories (or clusters), and Yahoo! Search, the search engine evolution of the Yahoo! Directory, were released (Seymour et al., 2011).

In 2005, the Microsoft search engine, MSN, was also released. Its name would later be changed, first, to Windows Live Search in 2006 and then to Bing in 2009.

Other ventures – most notably GoodSearch (2005), Forestle (2008) and Ecosia (2009) – opted for a different business approach, donating part of their profits to charities and non-profit organizations. In 2006, Hakia, a pioneer semantic search engine, was launched. In this respect it differed from all of its competitors who continued to work by keywords, albeit that, since 2003, Google had begun to implement improvements in its semantic searches (Casares, 2008). Natural language processing, aimed at improving the quality of results, was

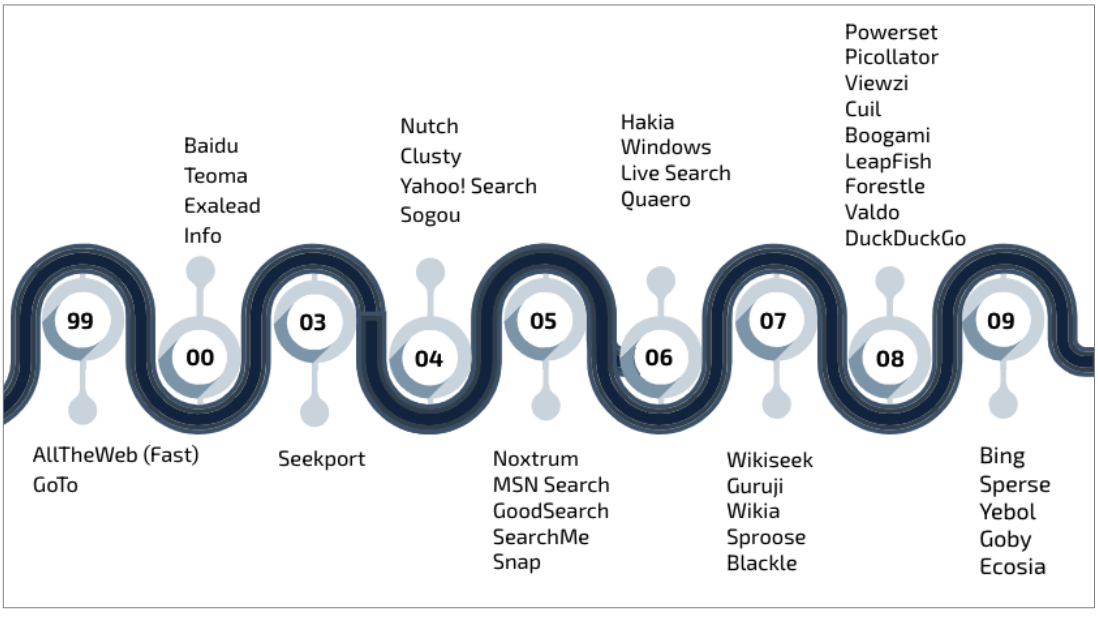


Image 3. Chronology of the main search engines launched between 1999 and 2010.

one of the developments made in those years (Pérez-Agüera, 2008), something that Google would continue to develop in its algorithm with the following updates: Brandy in 2003, Hummingbird in 2013 and Bert in 2019.

5. The arrival of Google

In 1995, on matriculating at Stanford University to pursue his PhD in the Computer Science Department, Larry Page met Sergey Brin and, the following year, both students began tracking and downloading websites and analysing their links (Vise and Malseed, 2006). Page worked on the theory that the more links a website received the more important it was, basing his thinking on the citation system used in the scientific literature, in which having more citations was equivalent to being more popular. However, he understood that not all links had the same weight, but that those originating from pages with more links were more important. He called this link classification system PageRank, a play on words based on his own surname, and used it in the creation in 1997 of a search engine which he called BackRub (Vise and Malseed, 2006; Wills, 2006).

That same year, Page, Brin, and Rajeev Motwani, the Stanford professor who tutored him, worked on the tool and built a prototype search engine that added PageRank to the technology then being used by search engines. All told, they disposed of a repository that stored 24 million pages and which implemented PageRank, building an index of links at the same time as the web crawler systematically browsed the sites (Page et al., 1999).

In the Autumn of 1997, they changed the name of their search tool to Google. On September 15, they registered the Google.com domain, and made it available to the community at Stanford University via the url: www.google.standford.edu (Vise and Malseed, 2006) (Image 4).

Their premise was that existing search robots basically relied on keyword matching, which returned too many low-quality results (Brin and Page, 1998). Given the exponential growth in the number of websites, for Brin and Page, it was critical to obtain more precise results, something they achieved by taking into account the link structure; the link (anchor) text, which worked as a description of the specific page that a link points to; and, the proximity information about the elements of a document (Brin and Page, 1998).

PageRank worked as an objective measure, assigning importance to a website according to the number and weight of each of the citations. To calculate a page's PageRank, they took into account the pages that link to it, the number of outgoing links from that page and the damping factor, which is normally set to 0.85, and which refers to the probability that on any given page a surfer continues to click on the links it contains or stops clicking and moves on to another page. Thus, a page will have a higher PageRank the more links it has or if, even if it

Image 4. Google interface in 1998. Source: https://web.archive.org/web/19981111183552/http://google.stanford.edu/

does not have many links, the few links it receives are from pages with a high PageRank (Brin and Page, 1998).

This combination of factors permitted the development of a system of ranking results by importance and not by a simple connection between the words of the query and those contained on the website, which meant the results were much more accurate (Vise and Malseed, 2006).

In March 1998, they tried to sell the PageRank patent to existing search engines, including AltaVista, Excite and Yahoo but all of them rejected the opportunity, preferring Internet users to stay in their portal rather than have them leave to visit other pages (Vise and Malseed , 2006). In August of that same year, they obtained the necessary financing and a month later they founded Google Inc (Google, n.d.; Wills, 2006).

Google was to revolutionise the world of search engines as it was then known (Seymour et al., 2011) and the company gradually gained market share until establishing itself as the undisputed leader. Indeed, in 2003 it already accounted for more than half of the searches made (Trigo-Aranda, 2004) and today, it has maintained its position, well above the rest of the search engines, with a worldwide use that reaches more than 86% (Johnson, 2021).

6. Evolution of the Google algorithm

Among the various modifications made to the operation of search engines, the launch of Google's PageRank algorithm in the 1990s was the one that revolutionised the industry.

The incorporation of this innovation made the manipulation of search results much more difficult, as they no longer depended on website creators, but rather on third parties who had to link to the sites' pages to increase their popularity and, hence, a page's positioning (Arbildi-Larreina, 2005; Brin and Page, 1998; Morato-Lara et al., 2013).

From the very outset, Google's algorithm has been subject to constant updates. According to Moz (n.d.), in 2018 alone, 3,234 updates were made. In 2003, the first named update – Boston – was released, involving the implementation of monthly structural changes in the algorithm combined with important updates of the index, a process that became known as the *Google Dance* as it introduced sizeable ranking shifts (Baker, 2021).

Among the most important updates (Image 5), we find, in chronological order, that of Florida, in 2003, which penalizes keyword *stuffing* and other abusive techniques carried out to manipulate positioning. Brandy was launched in 2004, which, among other modifications, introduced latent semantic indexing (LSI), which expanded Google's ability to understand synonyms and keyword context and began the rating of the relevance of anchor texts, the text containing a hyperlink (Moz, n.d.).

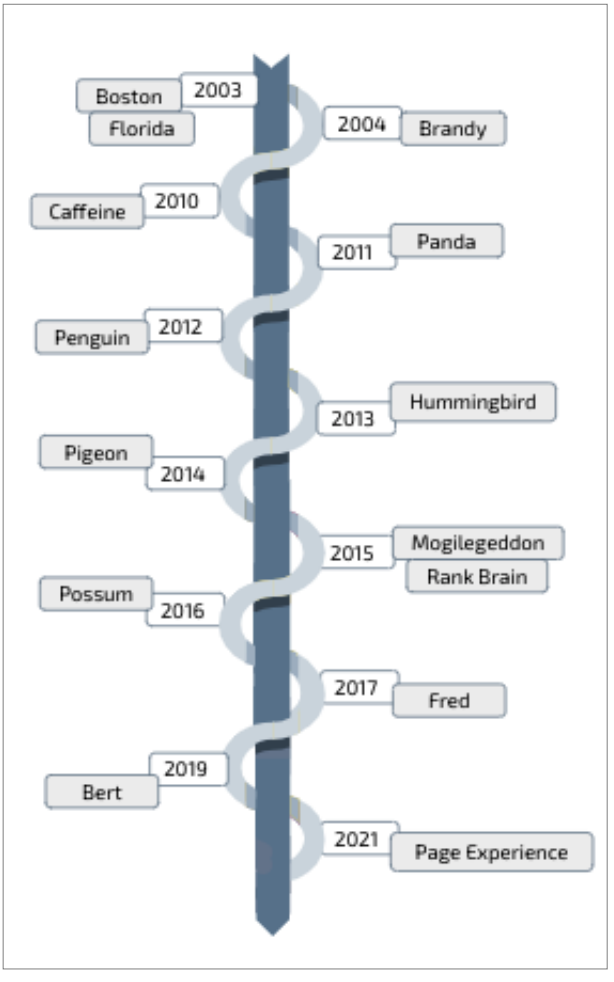


Image 5. Main updtaes to the Google algorithm.

In 2010, the Caffeine update (Bezhovski, 2015; Cutts, 2009; Gonzalo-Penela, 2015; Grimes, 2010; Morato-Lara et al., 2013; Shahzad et al., 2020) provided for a faster indexing of content while search results were now more current than before by up to 50 per cent. And a year later, in 2011, Panda was launched for the first time, one of the most influential updates since it penalizes websites with little or poor quality content, such as the so-called content farms, and those with duplicate and/or copied content (Aswani et al., 2018; Bezhovski, 2015; Egri and Bayrak, 2014; Giomelakis and Veglis, 2016; Gonzalo-Penela, 2015; Goodwin, 2018; Gudivada, 2015;

Mavridis and Symeonidis, 2015; Nanda et al., 2019; Shahzad et al., 2020; Singhal, 2011; Zhang and Cabage, 2017).

In 2012, another of the major updates was made, Penguin. This penalized a series of overoptimization practices considered webspam, such as keyword stuffing – thereby, renewing the 2003 update – and the buying and trading of links (Aswani et al., 2018; Bezhovski, 2015; Egri and Bayrak, 2014; Giomelakis and Veglis, 2016; Gonzalo-Penela, 2015; Gudivada, 2015; Nanda et al., 2019; Shahzad et al., 2020; Zhang and Cabage, 2017). In 2013, Hummingbird represented a complete update of the algorithm to provide more relevant results thanks to a better understanding of natural language and voice queries (Aswani et al., 2018; Bezhovski, 2015; Giomelakis and Veglis, 2016; Gonzalo-Penela, 2015; Goodwin, 2018; Gudivada, 2015; Lopezosa et al., 2018; Mavridis and Symeonidis, 2015; Nanda et al., 2019; Shahzad et al., 2020) and, in 2014, Pigeon, introduced drastic changes in how local searches are interpreted (Aswani et al., 2018; Gudivada, 2015; Nanda et al., 2019).

In 2015, the main updates went by the name of "Mobilegeddon", which involved a transformation in the rankings of mobile search results, giving priority to websites optimised for viewing on these devices (Goodwin, 2018; Shahzad et al., 2020), and Rank Brain, by which Google announced that machine learning now formed part of the algorithm, that is, artificial intelligence that permits the system, by means of automatic learning, to better understand the intentions of search queries (Lopezosa et al., 2018; Rovira et al ., 2018; Nanda et al., 2019; Shahzad et al., 2020). The following year, Rank Brain was confirmed by Google as one of its three most important ranking factors (Schwartz, 2016; Shahzad et al., 2020).

The Possum and Fred updates are assumed to have been released in 2016 and 2017, respectively, although they have never officially been confirmed by Google (Moz, n.d.). Possum gives relevance to local businesses, using a user's geolocation (Hawkins, 2016; Nanda et al., 2019), while Fred penalizes websites with excessive advertising or content created exclusively for monetization, that is, the generation of traffic income from advertisements or affiliate commissions (Gonzalez, 2017; Nanda et al., 2019).

In 2019, the BERT (Bidirectional Encoder Representations from Transformers) update was introduced, aimed at improving understanding of the language used in searches and giving better matches between queries and results (Devlin et al., 2019; Nayak, 2019).

In June of that same year, the Google Page Experience update was rolled out. This incorporates the so-called Core Web Vitals, which measure a website's speed and response factors so as to provide an optimal browsing experience (Google, 2021; Gupta, 2021; Schwartz, 2020).

7. Search engines around the world

The importance of search engines as a means of accessing internet content is clearly evident if we consider their penetration rate around the world. In countries like Belgium, Norway and South Africa they are used by 85% of the population (Figure 1), while in Spain the percentage coincides with that of the world mean, with 74% of Spaniards using search engines frequently (Statista, 2013).

Among the search engine options available worldwide, Google is the most used, with a market share of 91.4% (Statcounter, 2021b) (Figure 2). The degree of quasi monopoly enjoyed by Google varies from country to country. In Russia, its market share is not as pronounced due to

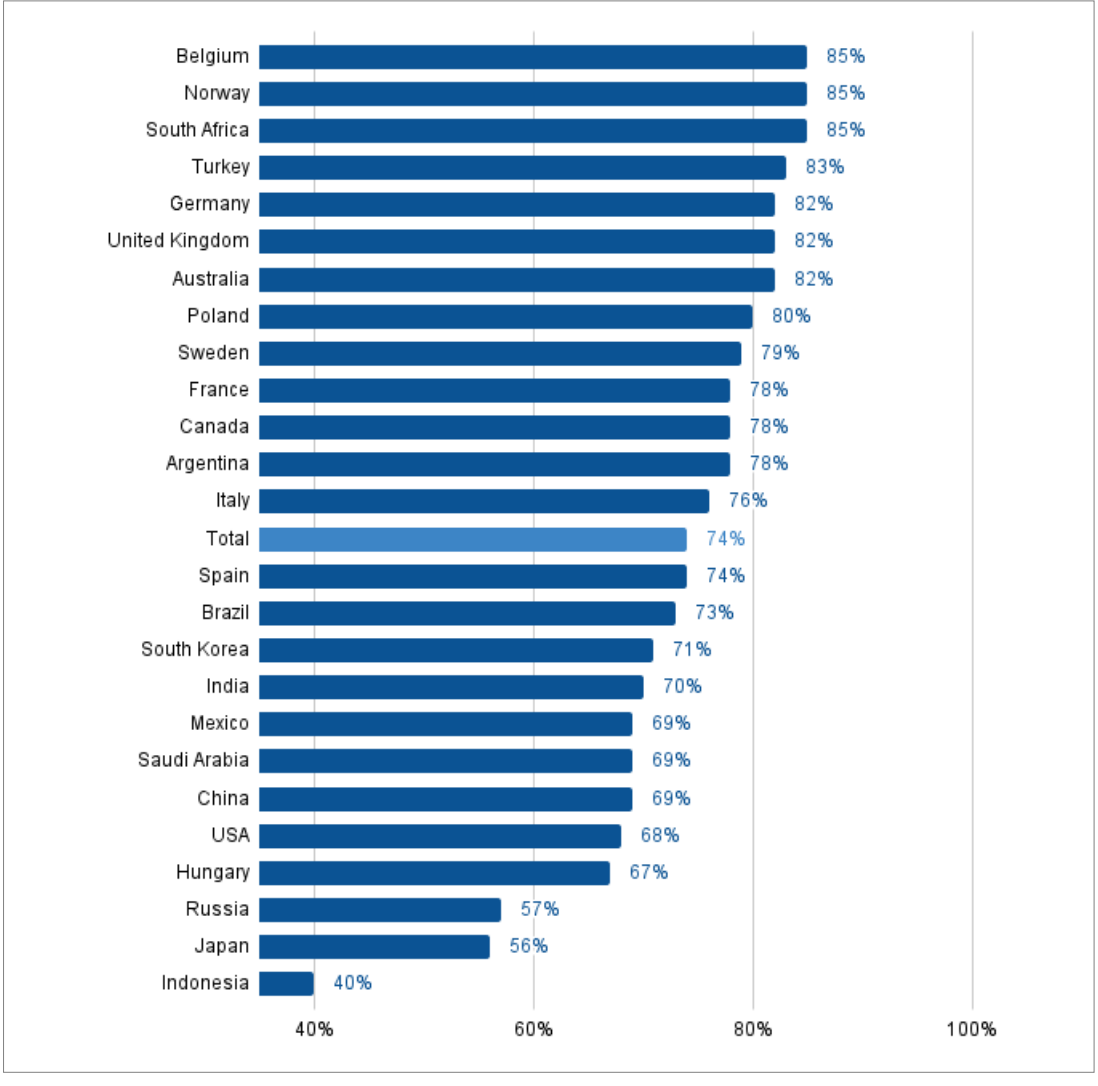


Figure 1. Search engine website usage worldwide as of August 2013, by country. Source: Based on data from Statista, 2013.

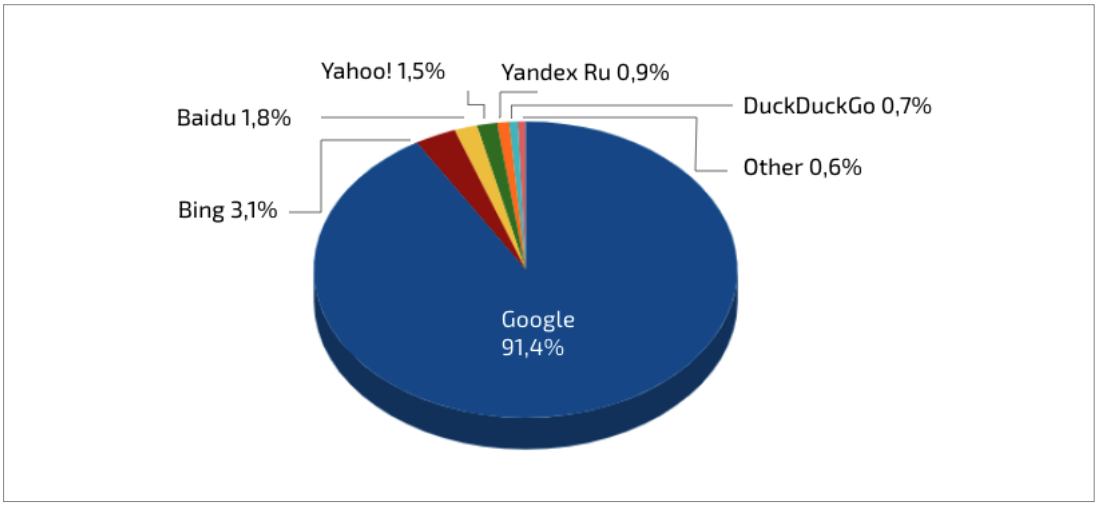


Figure 2. Search Engine Market Share Worldwide. November 2021. Source: Based on data from Statcounter 2021b

the presence of Yandex, although in May 2017 it was ousted by Google as the country's leading search engine and since then they have jousted for leadership. Indeed, in 2021, Google controlled more than 60% of the market, leaving Yandex in second position with 36% (Liveinternet, 2021). In China, the dominant search engine is by far and away Baidu, with an 87% market share in 2021, well above that of Sogou (5%) and Bing (4%), which even leave Google trailing, used by just 2% of the population (Statcounter, 2021a).

In Spain, however, Google is by far the most popular search engine, controlling more than 95% of the market, with Bing lagging a long way behind in second place with just 3% (Statcounter, 2021c).

These figures underline the importance of knowing the origin and evolution of search engines and in the case of Google – the world's market leader – the main updates made to its algorithm. Future research needs to explore further not only their characteristics but also the way they operate, specifically the ranking models used to position their results. Of special interest here is the discipline of SEO – Search Engine Optimisation – and its role in enhancing search result rankings and attracting web traffic, as well as the industry that has grown up around it, which in Spain alone comprises more than a thousand companies offering SEO services (Escandell-Poveda et al., 2021). A better understanding of this discipline will enable webmasters and web developers to understand how to increase visibility and attract audiences, essential qualities for any organization that needs to stand out in the ocean of information and resources that the Web has been converted into the decades that make up its history.

Note: This study forms part of the following doctoral thesis: Escandell-Poveda, R. (2021). *Radiografía del SEO en España: demanda laboral, oferta formativa, sector empresarial y perfil profesional del posicionamiento web.* (Radiography of SEO in Spain: labour demand, training

offer, business sector and professional profile of web positioning. [Doctoral Thesis, University of Alicante].

References

Asociación para la Investigación en Medios de Comunicación, AIMC (2020). *Navegantes en la red – encuesta AIMC a usuarios de internet.* https://www.aimc.es/otros-estudios-trabajos/navegantes-la-red/

Anklesaria, F.; Mccahill, M.; Lindner, P.; Johnson, D.; Torrey, D. & Alberti, B. (1993). The internet gopher protocol (a distributed document search and retrieval protocol). *Network Working Group Request for Comments*: 1436, 1-16. https://dl.acm.org/doi/pdf/10.17487/RFC1436

Arbildi-Larreina, I. (2005). Posicionamiento en buscadores: una metodología práctica de optimización de sitios web. *Profesional de la Información, 14*(2), 108-124. https://doi.org/10.3145/epi.2005.mar.03

Aswani, R.; Kar, A. K.; Ilavarasan, P. V. & Dwivedi, Y. K. (2018). Search engine marketing is not all gold: Insights from twitter and SEOClerks. *International Journal of Information Management, 38*(1), 107-116. https://doi.org/10.1016/j.ijinfomgt.2017.07.005

Baeza-Yates, R. & Ribeiro-Neto, B. (1999). *Modern information retrieval.* ACM press.

Baker, L. (Febrary 27, 2021). 20+ years of SEO: A brief history of search engine optimization. *Search Engine Journal*. https://www.searchenginejournal.com/seo-101/seo-history/

Ballari, T. L. (2003). *Motores de búsqueda en internet. Un enfoque generacional*. https://web.archive.org/web/20040829010656/http://www.unlu.edu.ar/~tyr/tyr/TYR-motor/ballari-motor.pdf

Berners-Lee, T. (1989). Information management: A proposal. *The World Wide Web Consortium* (W3C). https://www.w3.org/History/1989/proposal.html

Bezhovski, Z. (2015). The historical development of search engine optimization. *Information and Knowledge Management, 5*(15) 91-96. https://iiste.org/Journals/index.php/IKM/article/view/27821

Brin, S. & Page, L. (1998). The anatomy of a large-scale hypertextual web search engine. *Computer Networks and ISDN Systems, 30*(1), 107-117. https://doi.org/10.1016/S0169-7552(98)00110-X

Bush, V. (1945). As we may think. *The Atlantic Monthly, 176*(1), 101-108.

Casares, J. (2008). *Guía de referencia SEO*. Ojo Internet, SL.

Cutts, M. (November10, 2009). Expect caffeine after the holidays. *Matt Cutts* [personal blog]. http://www.mattcutts.com/blog/google-caffeine-update/

Devlin, J.; Chang, M.; Lee, K. & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 1*, 4171-4186. https://doi.org/10.18653/v1/N19-1423

Duklan, N. & Bahuguna, H. (2012). Search engine optimization techniques for organic search: A theoretical approach. *Uttaranchal Business Review, 2*(2), 45-55.

Egri, G. & Bayrak, C. (2014). The role of search engine optimization on keeping the user on the site. *Procedia Computer Science,* (36), 335-342. https://doi.org/10.1016/j.procs.2014.09.102

Eichmann, D. (1994). The RBSE spider — balancing effective search against web load. *Computer Networks and ISDN Systems,* 27(2), 308. https://doi.org/10.1016/S0169-7552(94)90151-1

Escandell-Poveda, R. (2021). *Radiografía del SEO en España: demanda laboral, oferta formativa, sector empresarial y perfil profesional del posicionamiento web.* [Tesis Doctoral, Universidad de Alicante].

Escandell-Poveda, R.; Iglesias-García, M. & Papí-Gálvez, N. (2021). Who does SEO in Spain? A cybermetric methodology for the construction of company universes. *Profesional de la Información, 30*(4). https://doi.org/10.3145/epi.2021.jul.19

Fain, D. C. & Pedersen, J. O. (2006). Sponsored search: A brief history. *Bulletin of the American Society for Information Science and Technology, 32*(2), 12-13. https://doi.org/10.1002/bult.1720320206

Fernández, R. (June 3, 2021). Los buscadores online - datos estadísticos. *Statista.* https://es.statista.com/temas/3898/los-buscadores-online

Fishkin, R. (November 13, 2019). Google in 2020: From everyone's search engine to everyone's competitor. *Sparktoro.* https://sparktoro.com/blog/google-in-2020-from-everyones-search-engine-to-everyones-competitor/

Galaxy/eiNet (2021). *Galaxy/eiNet history.* https://www.einet.net/static/history.htm

García-Marco, F. (1996). Vannevar Bush, el hipertexto y el futuro del documento. In Tramullas, J. (Ed.), *Tendencias de investigación* (pp. 185-210). Documentación.

Giomelakis, D. & Veglis, A. (2016). Investigating search engine optimization factors in media websites. *Digital Journalism, 4*(3), 379-400. https://doi.org/10.1080/21670811.2015.1046992

Gonzalez, J. (March 28, 2017). Fred update – what do all losers have in common? *Sistrix.* https://www.sistrix.com/blog/googles-fred-update-what-do-all-losers-have-in-common/

Gonzalo-Penela, C. (2015). *Posicionamiento web y dinámicas de información en motores de búsqueda: propuestas de análisis y estudio comparativo de visibilidad de contenidos digitales en el caso de procesos electorales.* [Tesis Doctoral, Universitat Pompeu Fabra]. Tesis Doctorals en Xarxa. https://hdl.handle.net/10803/292730

Goodwin, D. (June 29, 2018). Why y how to track google algorithm updates. *Search Engine Journal.* https://www.searchenginejournal.com/seo-guide/track-google-algorithm-updates

Google (n/d). *De un garaje al Googleplex.* https://about.google/our-story

Google (2021). Qué es la experiencia de las páginas en los resultados de la búsqueda de google. *Centro de la Búsqueda de Google*. https://developers.google.com/search/docs/guides/page-experience

Gray, M. (1996). Credits and background. *Massachusetts Institute of Technology, MIT*. https://www.mit.edu/people/mkgray/net/background.html

Griffiths, R. T. (2002). Chapter two: From ARPANET to world wide web. *History of the internet, internet for historians (and just about everyone else)*. http://www.universelle-automation.de/1969_Washington.pdf

Grimes, C. (June 8, 2010). Our new search index: Caffeine. *Google Official Blog.* https://googleblog.blogspot.com/2010/06/our-new-search-index-caffeine.html

Gudivada, V. N.; Rao, D. & Paris, J. (2015). Understanding search-engine optimization. *Computer (Long Beach, Calif.), 48*(10), 43-52. https://doi.org/10.1109/MC.2015.297

Gupta, A. (January 4, 2021). Google page experience update is all set to launch in may 2021 - webmasters, hang in there! *Search Engine Watch.* https://www.searchenginewatch.com/2021/01/04/google-page-experience-update-is-all-set-to-launch-in-may-2021-webmasters-hang-in-there/

Hawkins, J. (Septembrer 21, 2016). Everything you need to know about Google's 'Possum' algorithm update. *Search Engine Land.* https://searchengineland.com/everything-need-know-googles-possum-algorithm-update-258900

Internet Advertising Bureau [IAB] Spain y PricewaterhouseCoopers (2021). *Estudio de inversión publicitaria en medios digitales. Resultados 2020.* https://iabspain.es/download/50771/

Internet Live Stats (2021). Google searches in 1 second. *Internet Live Stats*. https://www.internetlivestats.com/one-second/#google-band

Jha, P.; Goyal, S.; Kumari, T. & Gupta, N. (2014). Robots exclusion protocol. *International Journal of Emerging Science and Engineering*, *2*(5). http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.685.2053&rep=rep1&type=pdf

Johnson, J. (2021). Global market share of search engines 2010-2021. *Statista*. https://www.statista.com/statistics/216573/worldwide-market-share-of-search-engines/

Koster, M. (1994). Aliweb - Archie-like indexing in the Web. *Computer Networks and ISDN Systems*, *27*(2), 175-182. https://doi.org/10.1016/0169-7552(94)90131-7

Lara F. J. (2014) Search Engine Optimization and Ethical Leadership Strategies. In I. Gil-Pechuán, D. Palacios-Marqués, M. Peris-Ortiz, E. Vendrell, C. Ferri-Ramírez (Eds.) *Strategies in E-Business*, (pp. 79-90) Springer. https://doi.org/10.1007/978-1-4614-8184-3_7

Lawrence, S. & Giles, C. L. (1999). Accessibility of information on the web. *Nature (London)*, *400*(6740), 107. https://doi.org/10.1038/21987

Liveinternet (2021). *Estadística de uso de buscadores en Rusia.* https://www.liveinternet.ru/stat/ru/searches.html

Lopezosa, C.; Codina, L. & Caldera-Serrano, J. (2018). SEO semántico: Framework ISS para la optimización de sitios intensivos en contenidos. *Cuadernos De Documentación Multimedia*, (29), 97-123. https://doi.org/10.5209/CDMU.60607

Mauldin, M. I. (1997). Lycos: Design choices in an internet search service. *IEEE Expert*, *12*(1), 8-11. https://doi.org/10.1109/64.577466

Mavridis, T. & Symeonidis, A. L. (2015). Identifying valid search engine ranking factors in a web 2.0 and web 3.0 context for building efficient SEO mechanisms. *Engineering Applications of Artificial Intelligence*, (41), 75-91. https://doi.org/10.1016/j.engappai.2015.02.002

Morato-Lara, J.; Sánchez-Cuadrado, S.; Moreno-Pelayo, V. & Moreiro-González, J. A. (2013). Evolución de los factores de posicionamiento web y adaptación de las herramientas de optimización. *Revista Española De Documentación Científica*, *36*(3), e018. https://doi.org/10.3989/redc.2013.3.956

Moz (n/d). *Google algorithm update history.* https://moz.com/google-algorithm-change

Nanda, P.; Patnaik, S. & Patnaik, S. (2019). Towards algorithmic business: A paradigm shift in digital business. In Nanda, P.; Patnaik, S. & Patnaik, S. *Digital business* (pp. 3-22). Springer International Publishing. https://doi.org/10.1007/978-3-319-93940-7_1

Nayak, P. (October 25, 2019). Understanding searches better than ever before. *Google The Keyword.* https://blog.google/products/search/search-language-understanding-bert

Nierstrasz, O. (1996). W3 catalog history. *Software Composition Group*. http://scg.unibe.ch/archive/software/w3catalog/

Page, L.; Brin, S.; Motwani, R. & Winograd, T. (1999). The PageRank citation ranking: Bringing order to the web. *Stanford InfoLab.* http://ilpubs.stanford.edu:8090/422

Pérez-Agüera, J. (2008). Nueva generación de motores de búsqueda basados en procesamiento. *Anuario ThinkEPI*, *2*(1), 39-40. http://recyt.fecyt.es/index.php/ThinkEPI/article/view/32036/17028

Pinkerton, B. (1994). WebCrawler: Finding what people want. *Second International WWW Conference.* http://www.thinkpink.com/bp/WebCrawler/WWW94.html

Rayward, W. B. (1994). Visions of Xanadu: Paul Otlet (1868-1944) and hypertext. *Journal of the American Society for Information Science*, *45*(4), 235-250. https://doi.org/10.1002/(SICI)1097-4571(199405)45:4%3C235::AID-ASI2%3E3.0.CO;2-Y

Rovira, C.; Guerrero-Solé, F. & Codina, L. (2018). Las citas recibidas como principal factor de posicionamiento SEO en la ordenación de resultados de Google Scholar. *Profesional de la Información*, *27*(3), 559-569. https://doi.org/10.3145/epi.2018.may.09

Salazar, I. (2005). *Las profundidades de internet: Accede a la información que los buscadores no encuentran y descubre el futuro inteligente de la red*. Trea.

Salton, G. (1989). *Automatic text processing: The transformation, analysis, and retrieval of information by computer*. Addison-Wesley.

Schwartz, B. (March 24, 2016). Now we know: Here are google's top 3 search ranking factors. *Search Engine Land.* https://searchengineland.com/now-know-googles-top-three-search-ranking-factors-245882

Schwartz, B. (November 10, 2020). Google page experience update to launch may 2021 with new labels in search results. *Search Engine Land.* https://searchengineland.com/google-page-experience-update-to-launch-in-may-2021-343581

Schwartz, C. (1998). Web search engines. *Journal of the American Society for Information Science, 49*(11), 973-982. https://doi.org/10.1002/(SICI)1097-4571(1998)49:11%3C973::AID-ASI3%3E3.0.CO;2-Z

Schwartz, M. F.; Emtage, A.; Kahle, B. & Neuman, C. (1992). A comparison of internet resource discovery approaches. *Computing Systems, 5*(4), 461-493. https://scholar.colorado.edu/csci_techreports/576

Seymour, T.; Frantsvog, D. & Kumar, S. (2011). History of search engines. *International Journal of Management y Information Systems, 15*(4), 47-58. https://doi.org/10.19030/ijmis.v15i4.5799

Shahzad, A.; Jacob, D. W.; Nawi, N. M.; Mahdin, H. & Saputri, M. E. (2020). The new trend for search engine optimization, tools and techniques. *Indonesian Journal of Electrical Engineering and Computer Science, 18*(3), 1568-1583. https://doi.org/10.11591/ijeecs.v18.i3.pp1568-1583

Sherman, C. (2000). The future revisited : What's new with web search. *Online (Weston, Conn.), 24*(3), 27-34. https://search.proquest.com/docview/199926335

Sherman, C. & Price, G. (2001). *The invisible web: Uncovering information sources search engines can't see.* Information Today, Inc.

Singhal, A. (February 24, 2011). Finding more high-quality sites in search. *Google Official Blog.* https://googleblog.blogspot.com/2011/02/finding-more-high-quality-sites-in.html

Smith.J & Weiss, S. (1988). Hypertext. *Communications of the ACM, 31*(7). https://doi.org/10.1145/48511.48512

Sonnenreich, W. (1997). A history of search engines. *Wiley.* https://web.archive.org/web/20040929081754/http://www.wiley.com/legacy/compbooks/sonnenreich/history.html

Statcounter (2021a). Search engine market share China. *Statcounter.com.* https://gs.statcounter.com/search-engine-market-share/all/china/

Statcounter (2021b). *Search engine market share Worldwide. Statcounter.com.* https://gs.statcounter.com/search-engine-market-share

Statcounter (2021c). *Search engine market share Spain. Statcounter.com.* https://gs.statcounter.com/search-engine-market-share/all/spain

Statista (2013). Search engine penetration worldwide by country 2013. *Statista.com.* https://www.statista.com/statistics/282842/share-of-respondents-using-a-serach-engine-regularly-worldwide-by-country/

Statista (2021). Search engine market share worldwide. *Statista.com.* https://www.statista.com/statistics/216573/worldwide-market-share-of-search-engines/

Statista (2013). Search engine penetration worldwide by country 2013. *Statista.com.* https://www.statista.com/statistics/282842/share-of-respondents-using-a-serach-engine-regularly-worldwide-by-country/

Statista (2021). Search engine market share worldwide. *Statista.com.* https://www.statista.com/statistics/216573/worldwide-market-share-of-search-engines/

Trigo-Aranda, V. (2004). Historia y evolución de internet. *Acta, (33)*, 1-12. https://www.acta.es/medios/articulos/comunicacion_e_informacion/033021.pdf

Van Couvering, E. (2008). The history of the internet search engine: Navigational media and the traffic commodity. *Web search* (pp. 177–206). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-75829-7_11

Vise, D. A. & Malseed, D. A. (2006). *La historia de Google: Los secretos del mayor éxito empresarial, mediático y tecnológico de nuestro tiempo*. La esfera de los libros.

Wagner, G. (1996). *Upgrading search strategy skills: Unit three: Developing search strategies for searching some of the indexes of the world wide web: A continuing professional development course for library and information service managers*. Edith Cowan University. https://ro.ecu.edu.au/ecuworks/6757/

Wall, A. (2017). History of Search Engines: From 1945 to Google Today. *Search Engine History.com*. http://www.searchenginehistory.com/

Wills, R. (2006). Google's pagerank. *The Mathematical Intelligencer*, *28*(4), 6–11. https://doi.org/10.1007/BF02984696

Zhang, S. & Cabage, N. (2017). Search engine optimization: Comparison of link building and social sharing. *The Journal of Computer Information Systems*, *57*(2), 148–159. https://doi.org/10.1080/08874417.2016.1183447

Carlos Lopezosa
Universitat de Barcelona
lopezosa@ub.edu

Magdalena Trillo Domínguez
Universidad de Granada
mtrillo@ugr.es

Marga Cabrera Méndez
Universitat Politècnica de València
mcabrera@upv.es

Lluís Codina
Universitat Pompeu Fabra
lluis.codina@upf.edu

Search engine optimisation in online journalism: A case study of the *Grupo Joly* media company

https://doi.org/10.3145/indocs.2022.5

Lopezosa, Carlos; Trillo-Domínguez, Magdalena; Cabrera-Méndez, Marga; Codina, Lluís (2022). Search engine optimisation in online journalism: A case study of the *Grupo Joly* media company. In: Freixa, Pere; Codina, Lluís; Pérez-Montoro, Mario y Guallar, Javier (ed.). *Visualisations and narratives in digital media. Methods and current trends*, (pp. 67-80). Barcelona: DigiDoc-EPI. https://doi.org/10.3145/indocs.2022.5

Abstract

Given the confirmation of the fundamental role of the application of Search Engine Optimization (SEO) in digital news media as part of the media's commitment to its readers, this chapter proposes to analyze the process of implementing SEO to the productive routines of *Diario de Sevilla* and its consequent transfer to the rest of digital news media that make up the *Joly* Group and collect the quantitative results of web visibility resulting from said implementation, specifically from 2018 to 2021. To carry out this research, first of all, the case study was used, specifically of the editorial team of *Diario de Sevilla*. For an adequate triangulation, a participant observation, a battery of semi-structured interviews, and a qualitative analysis of the content of the internal manual on SEO of *Joly* Group, a media group to which *Diario de Sevilla* belongs, have been carried out. All this has been complemented in turn, with a quantitative study, where, thanks to Google Analytics, the web traffic obtained from the group from the year of implementation to the present (2018 - 2021) is collected. The results confirm that in order to comprehensively implement the SEO of a newsroom, it is necessary to clearly separate the work of the journalists/editors and that of the SEO experts or managers at the level of the entire medium. In this sense, the SEO department has a strategic function of analysis and generation of guidelines for editors. While it is up to the newsroom to apply the SEO optimization recommendations without affecting the journalistic quality of the news.

Keywords

Digital news media, SEO, search engineoptimizationonline journalism, website visibility.

Título

El SEO en la empresa periodística: estudio de caso del Grupo Joly

Resumen

Ante la confirmación del papel fundamental de la aplicación del SEO en los cibermedios como una parte del compromiso del medio con sus lectores, este capítulo plantea analizar el proceso de implementación del SEO a las rutinas productivas del Diario de Sevilla *y su consiguiente transferencia al resto de cibermedios que integran el* Grupo Joly *y recoger los resultados cuantitativos de la visibilidad web resultante de dicha implementación, concretamente desde 2018 hasta 2021. Para llevar a cabo esta investigación se ha utilizado, en primer lugar, el estudio de caso, concretamente del equipo de redacción del* Diario de Sevilla*. Para una adecuada triangulación, se ha llevado a cabo una observación participante, una batería de entrevistas semiestructuradas, y un análisis cualitativo del contenido del manual interno sobre SEO del* Grupo Joly*, grupo mediático al que pertenece* Diario de Sevilla*. Todo esto se ha complementado a su vez, con un estudio cuantitativo, en donde se recoge, gracias a Google Analytics, el tráfico web obtenido del grupo desde el año de implementación hasta la actualidad (2018 - 2021). Los resultados confirman que para implantar de manera integral el SEO de una redacción es preciso separar claramente el trabajo de los periodistas/redactores y el de los expertos o responsables SEO a nivel de todo el medio. En este sentido, al departamento de SEO le corresponde una función estratégica de análisis y de generación de orientaciones para los redactores. Mientras que a la redacción le corresponde aplicar las recomendaciones de optimización SEO sin que las mismas afecten a la calidad periodística de las noticias.*

Palabras clave

Cibermedios, SEO, posicionamiento en buscadores, periodismo online, visibilidad web.

1. Introduction

With two decades of accumulated history both in Spain (Salaverría and Martínez-Costa, 2021) and worldwide, the online news media have had to respond to the challenges of a communicative situation that have little in common with those that affected the traditional media. Today's panorama is characterised above all by processes of interactivity (Cebrián-Herreros, 2009; Rodríguez-Martínez et al., 2012; Caminero & Sánchez-García, 2018), the constant renewal of internet technology and, more recently, the Covid-19 pandemic that has, as of 2020, resulted in a heightening of the problems and processes of digital adaptation that have impacted the media as they strive for survival.

These changes mean the news media have few options other than to reorient their content and strategies (Rodríguez-Martínez et al. 2010; Díaz-Noci, 2018; García-Avilés et al., 2018) as swiftly as they can (Suárez-Sucre, 2017) and, clearly, the integration of newsrooms into the online world has become one of the primary initiatives taken by media companies (Muerza-Ferrer, 2017). To this we need to add the growing importance of the role being played by social networks, on the one hand (Trillo-Domínguez & Ollero, 2018), and by search engine positioning, on the other (Giomelakis & Veglis, 2015a), in what is very much a new context of enhanced web visibility and the dynamization of the news.

Indeed, this critical need for web visibility has meant that the news media are obliged to optimise their search engine performance (Carlson, 2007; Norris, 2007; Smyrnaios & Rebillard, 2009; Smyrnaios, 2015), given that part of the traffic the online media receive comes from user searches on platforms such as Google (Norris, 2007; Machill et al., 2008). For this reason, it is essential that newsrooms implement strategies of search engine optimisation (SEO) (Smyrnaios & Sire, 2014;), that is, the techniques or procedures applied to websites that help increase their chances of appearing among the main search results (Gonzalo-Penela et al., 2015; Lopezosa et al., 2018) and, so, of having greater visibility and attracting more readers.

Logically, in terms of their fine detail, SEO techniques and strategies can differ from one medium to another (Dick, 2011; Giomelakis & Veglis, 2015b; García-Carretero et al. 2016), primarily because each medium has at its disposal a series of different resources, be they

of time, finances, technical capabilities or managerial support (Dick, 2011). However, there is little question – and this is a point stressed by more than one study – that the convergence between journalistic and SEO principles has favoured online newspapers (Asser, 2012; Lopezosa et al., 2020).

Given the blanket acceptance of the fundamental role played by the application of SEO in the online news media as part of the newspapers' commitment to its readers (Richmond, 2008; Iglesias-García & Codina, 2016), this chapter has two primary objectives: (1) to analyse the implementation of SEO in the production routines of the *Diario de Sevilla* and its subsequent transfer to the rest of the online papers that make up the Spanish media company, *Grupo Joly* and (2) to study the quantitative outcomes of this implementation in terms of the Group's web visibility between 2018 and 2021.

2. Methodology

In conducting our research, we have drawn, in the first instance, on the case study method described by Yin (2014) and applied here to the work of the newsroom of the *Diario de Sevilla*. To ensure an adequate triangulation of data, we carried out a participant observation, employed a battery of semi-structured interviews, and undertook a qualitative analysis of the content of *Grupo Joly's* in-house SEO manual. All this was, in turn, complemented by a quantitative study, in which, thanks to Google Analytics, we determined the Group's web traffic from the year it implemented its SEO strategies to the present day (that is, the period 2018–2021).

In what follows, we first outline the design of the participant observation, which is based essentially on the methodologies described in Angrosino (2012) and Pons and Monistrol (2017). Second, we describe the configuration of the semi-structured interviews and the steps taken to complete the content analysis of *Grupo Joly's* in-house SEO manual. Third, and finally, we explain how we went about collecting the quantitative data.

The participant observation was carried out during the initial phase of implementation of SEO in the newsroom of the *Diario de Sevilla* (2018). Our objective here was to determine the type of SEO work carried out by journalists within the online news media and the extent of their knowledge and understanding of the techniques of SEO. To obtain these insights, the observer was integrated into the newsroom of the *Diario de Sevilla.* All the paper's journalists were aware they were being observed and were fully informed about the goals of the study.

Additionally, to complement this participant observation, we conducted 10 semi-structured interviews (Coller, 2000; Valles, 2002) with employees of the *Diario de Sevilla* using the NVivo qualitative analysis tool.

The questions put to the interviewees were the following:

- How would you define SEO?
- Do you think it is important to use SEO strategies in the newsrooms of a digital newspaper? Why?
- Do you think journalists have enough information and sufficient guidelines to be able to create news stories that are optimised for SEO?
- What SEO strategies do you apply on a daily basis as part of your productive routines as a journalist?
- Do you think a journalist should be able to implement SEO strategies or, on the contrary, these strategies should be the responsibility of the web positioning department? Why?
- Do you think that having to carry out SEO strategies affects the quality of the news and its broader social role? Why?
- How do you envision the future of journalism and its relationship with SEO?

In selecting the interviewees, we took into account the following so-called inclusion criteria (Valles, 2002), which consider who has the most relevant information, who is most socially and physically accessible, who is most willing to inform and cooperate and who has the greatest communicative fluency to ensure the accuracy of the information reported. Having concluded our participant observation and the interviews, we proceeded to analyse the content of the SEO manual that all the journalists of the *Grupo Joly* employ when working on their news items. To do so, a semantic network analysis was carried out based on the repetition of terms, using the NVivo tool. The *Grupo Joly's* in-house SEO manual explains what SEO is and its importance for news visibility and includes the principal elements that a journalist must optimise and explains how to optimise them.

Finally, we accessed Google Analytics to obtain the quantitative results for *Grupo Joly* between June 2018, the beginning of the implementation of SEO in the Group, and October 2021, the date this study was undertaken.

3. Results

In conducting this study, we are interested in identifying the principal dimensions involved in the implementation of a model of SEO in an online newspaper, the *Diario de Sevilla*. In so doing, we take into consideration the particular point of view of the editor/journalist as well as that of the SEO expert.

The participant observation allowed us to confirm that in order to ensure the full implementation of SEO in a newsroom, a clear distinction has to be drawn between the work of the

journalist (editor) and the work of the SEO expert (as carried out by the SEO department), given that their functions and skills differ greatly.

Consequently, we identified two significant levels at which SEO operates: (1) the strategic level – dependent, in this instance, on the SEO department and with a global impact on the online newspaper, and (2) the tactical level – the direct responsibility of the editors and with an impact on the news and journalistic production of the online newspaper. The outcomes of this study have been transformed into 16 SEO checkpoints that are explained in detail in an earlier study by the authors (Lopezosa et al., 2021).

Each of the semi-structured interviews focused specifically on understanding the processes of strategic assimilation experienced by the SEO department and the processes of tactical application of SEO techniques by the different departments of the *Diario de Sevilla*. Their respective responses confirmed that:

- In general, SEO is perceived as being very important in the newsrooms as it helps attract more readers. The interviewees consider that SEO does not affect the quality of the news.
- The interviewees report having a good understanding of what SEO entails. Their responses point to the emergence of two basic definitions of SEO. The first sees SEO as the set of techniques or practices that ensure news items appear at the top of Google SERPs, while the second is more closely related to the work of the professional whose responsibility it is to position the website in Google.
- They report having received sufficient training in SEO. The interviewees believe that major efforts are being made to train journalists in SEO.
- They consider that when applied correctly, SEO does not affect the quality of the news. In fact, they believe that the future of journalism depends on digital platforms and, specifically, on Google. As such, they feel newsrooms will tend to seek a greater convergence between journalism and visibility in search engines.
- The interviewees recognize that having an SEO department in the newsroom is essential and that, therefore, the newsroom and the SEO department should work shoulder to shoulder.

Additionally, *Grupo Joly's* in-house SEO manual is confirmed as constituting a handbook that addresses the basic strategies that all news items published in its digital edition must comply with.

Finally, after more than three years of the intense digital overhaul of all of *Grupo Joly's* newspapers (2018–2021), the data point to a clear upward trend in the news Group's web visibility. As the following tables show, over the last three years, increases of up to 500% have been recorded in the user traffic of some of its newspapers (Table 1), with two marked peaks,

one coinciding with the first wave of Covid-19 (the May 2020 lockdown) and the other with the second wave of the pandemic in the new year (January 2021). The turning point that the pandemic has represented for digital users is most clearly seen in Table 2.

	June-18	Oct-21	Variation (%)
Grupo Joly	2,835,835	13,984,684	393.14
Diariodesevilla.es	961,332	6,200,538	544.99
Diariodecadiz.es	556,347	2,125,240	282.00
Malagahoy.es	240,408	1,196,173	397.56
Diariodejerez.es	176,329	582,511	230.35
Huelvainformacion.es	189,724	1,143,377	502.65
Eldiadecordoba.es	157,985	557,738	253.03
Europasur.es	160,242	655,325	308.96
Granadahoy.com	228,597	1,030,063	350.60
Diariodealmeria.es	135,333	552,812	308.48

Table 1. Evolution in user traffic between June 2018, date of the introduction of *Grupo Joly's* transformation plan, and October 2021, latest available data. Source: Google Analytics.

	March, 2020	April, 2020	May, 2020	June, 2020	Difference vs same month previous year (%)	Difference 2019 vs 2018 (%)
Grupo Joly	14,349,784	12,825,737	16,631,544	12,076,932	25.59	115.68
Diariodesevilla.es	5,366,841	5,328,324	6,267,981	4,442,040	7.02	131.83
Diariodecadiz.es	2,249,794	2,073,209	2,405,904	1,998,991	18.89	93.52
Malagahoy.es	1,742,258	1,119,655	1,817,575	1,261,998	78.93	164.52
Diariodejerez.es	762,943	715,477	809,482	552,770	57.03	59.75
Huelvainformacion.es	757,134	642,338	754,922	719,780	44.64	96.21
Eldiadecordoba.es	685,523	489,454	649,372	477,032	28.89	105.95
Europasur.es	606,968	547,425	1,258,150	627,706	25.32	123.71
Granadahoy.com	1,443,410	1,309,718	1,877,250	1,326,386	64.53	115.32
Diariodealmeria.es	677,906	612,705	815,033	683,592	35.84	147.04

Table 2. Evolution in user traffic during the months of the pandemic lockdown, when user records were set in all the Group's newspapers. Source: Google Analytics.

As can also be seen in the graph (Figure 1) showing the evolution in the dynamics undergone by each of the different newspaper websites, growth in the number of users has been exponential

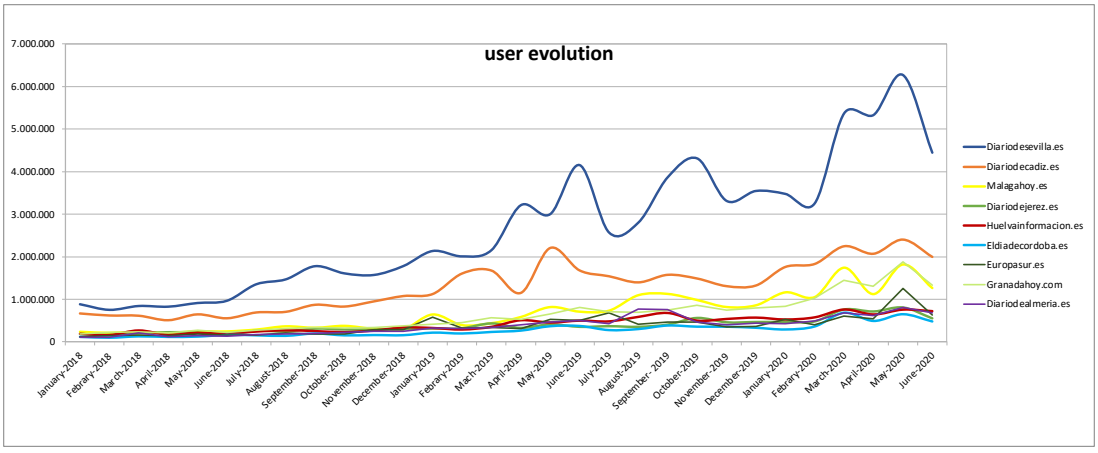


Figure 1. Evolution in number of *Grupo Joly* users, disaggregated by its nine newspapers (January 2018–October 2021). Source: Google Analytics

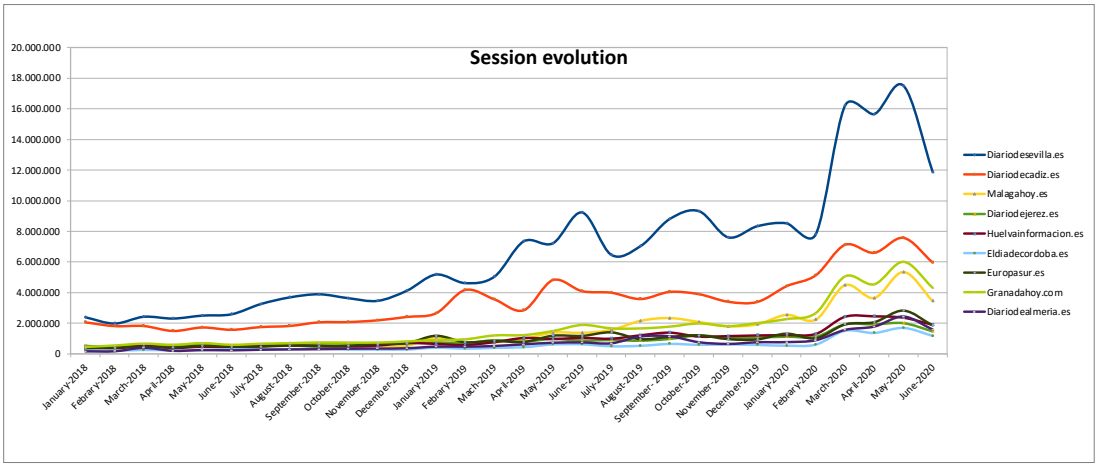


Figure 2. Evolution in number of *Grupo Joly* sessions, disaggregated by its nine newspapers (January 2018–October 2021). Source: Google Analytics.

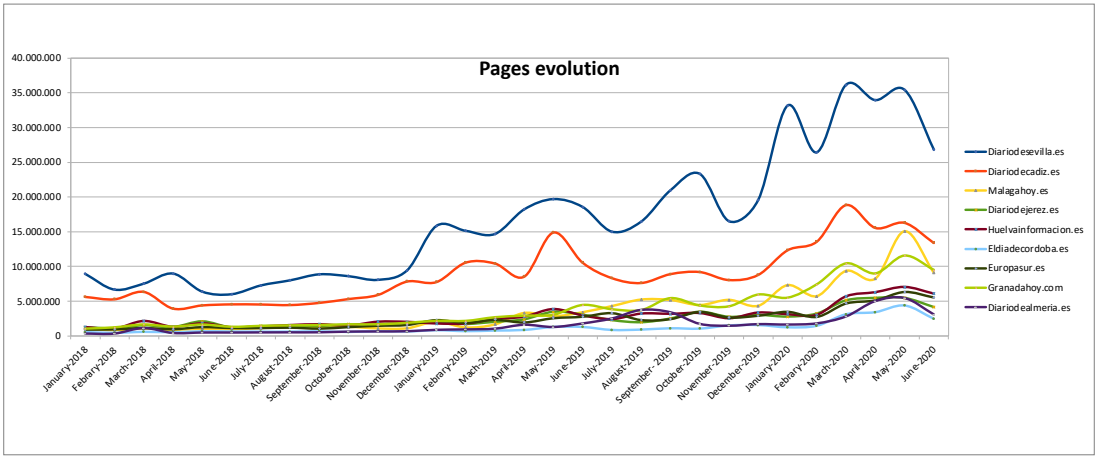


Figure 3. Evolution in number of *Grupo Joly* pageviews, disaggregated by its nine newspapers (January 2018–October 2021). Source: Google Analytics.

in all nine, with the *Diario de Sevilla* leading the way as the Group's leading newspaper and the primary lever of change. This growth is paralleled by the newspapers' metrics both for sessions (Figure 2) and pageviews (Figure 3).

From the perspective of SEO, the following graphs show how the content optimization strategies for search engines (organic traffic) have served to trigger growth in this first phase of

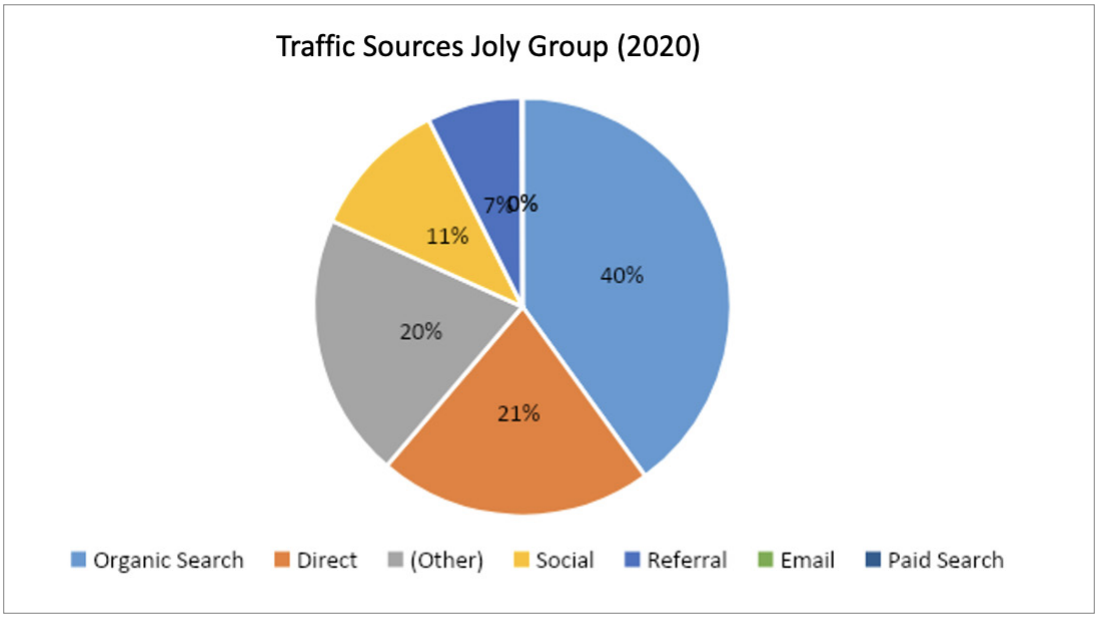


Figure 4. Origin of user traffic for the nine *Grupo Joly* newspapers in 2020, the year recording the greatest rise in the number of users. Source: Google Analytics.

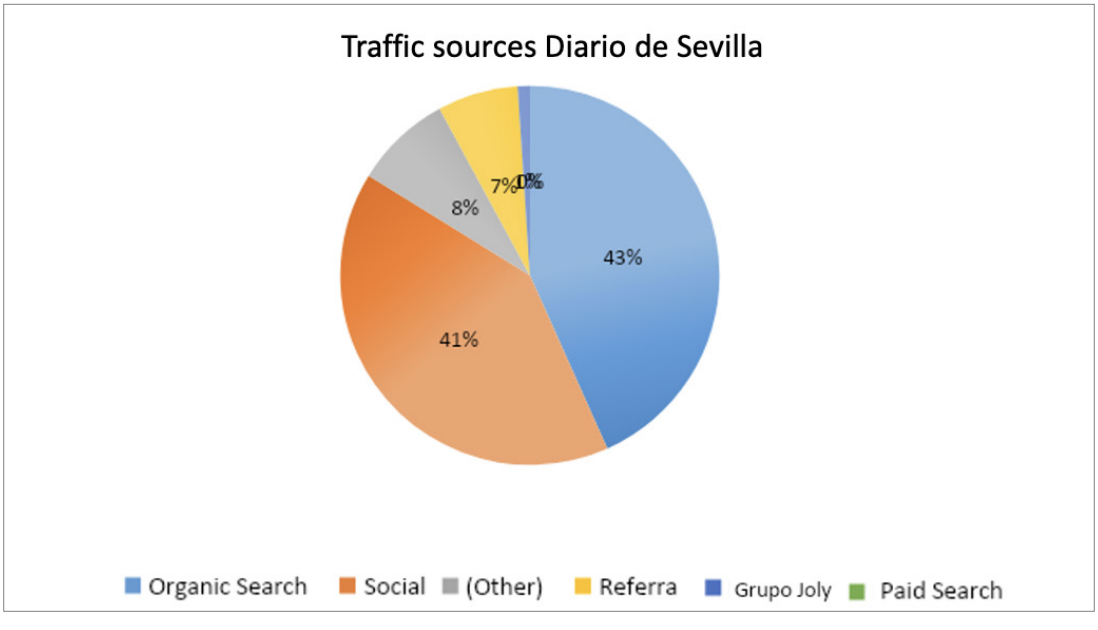


Figure 5. Origin of user traffic for the *Diario de Sevilla* as the Group's leading newspaper and the primary lever of change in 2020. Source: Google Analytics.

the Group's transformation. It is evident that this has become the primary way of accessing content in general in the Group, accounting for 40% of recorded users (Figure 4) and even more (43%) in the case of the *Diario de Sevilla* (Figure 5).

4. Discussion and conclusions

The digital transformation project implemented by the *Grupo Joly*, a leading media group in Andalusia, serves to confirm how SEO has gone from being the specialized task of computer technicians and engineers to occupy a central position in the news production business; how the purely quantitative concerns of SEO – that is, achieving visibility in the digital environment and winning audience as the primary business factor – have undergone a qualitative reorientation with direct repercussions for the way journalists work and for the introduction of new formats and content (Trillo-Domínguez & Alberich-Pascual, 2017); and, also, how the business has gone from incorporating a sole professional specialized in SEO to the dissemination of the SEO culture throughout the newsroom, to the point that SEO is now firmly established as a strategic factor for the digital transformation of the company.

First and foremost, the results confirm that in order to fully integrate the practices of SEO into a newsroom, the work of the journalists/editors and that of the SEO experts or managers need to be clearly differentiated throughout the news media company. Thus, the SEO department fulfils a strategic function of analysis and has a key role in generating guidelines for editors, while it is the job of the newsroom to apply the SEO recommendations without this affecting the journalistic quality of the news.

Another way to understand this distribution of tasks is as follows:

- The SEO department: takes responsibility for the general SEO strategies of the news outlet, analyses the impact on the search engines of applying these strategies, trains and provides support to journalists and editors, and applies the technical SEO required for the maintenance of the portal.
- Editors/Journalists: need to know how search results affect their news items and how they can attract more users, understand the basics of search engines and apply the visibility recommendations prepared by the SEO department to their work.

This study confirms the journalists' obvious interest in the optimization of news items for search engines and the practical application of SEO techniques to the news, in terms of both their planning and production routines. Likewise, the results reported by this study serve to identify the most common SEO strategies employed by journalists, on the one hand, and the SEO department, on the other. Indeed, the implementation of SEO strategies is perfectly evident in both ambits, that is, in the writing of news stories by journalists and in the global

strategic optimization practised by SEO experts. In fact, in the writing and optimization of news items, keywords are used in the headline, the standfirst, the body text, the skyline, the photo captions, and in the words in bold, etc., while in the SEO department the practices have a more strategic role focused on advising the paper's journalists about search engine positioning and identifying trends in news story searches to obtain more readers.

This chapter serves to corroborate that, when it comes to the implementation of SEO in a newspaper with a long tradition in its print edition, the possibilities of being able to do so effectively are high. Proof of this is the coherence demonstrated by the production routines of the main actors involved in the implementation described in this case study: that is, the SEO department, the digital edition of the Diario de Sevilla and the print/digital sections.

Moreover, the methodology employed here for gathering data, based on participant observation, semi-structured interviews and the analysis of Grupo Joly's in-house SEO manual, has proved to be a good way of understanding how an online newsroom operates in a context of the growing importance of web visibility.

In short, these operations are carefully articulated by the SEO department, as the head of search engine positioning strategies; the professionals that work on the digital edition, in their roles as news editors and as points of support for the journalists of the other sections; and the journalists of the different sections that work on the digital/print version, as the main creators of news stories for this news media group.

Likewise, our data serve to verify that the journalists consider SEO strategies essential for the writing of an online newspaper like the *Diario de Sevilla*, principally because many readers search for their information via Google and, so, appearing on the first page of the keyword searches means not only that they can compete for the top ranks in the SERPs but that they can also obtain more web traffic in the form of new users.

In line with previous studies, we are able to confirm the fact that newsrooms today have a good understanding of the SEO mission and its relationship with the visibility of journalistic production. Yet, in contrast with the findings of certain authors (Suárez-Sucre, 2017; López-García, 2017; Muerza-Ferrer, 2017; Trillo-Domínguez & Ollero, 2018), here we have described the case of a newsroom in which journalistic interests can be successfully combined with the objectives of SEO, at least from the important stand point of professionals working in the newspaper sector.

It is apparent that the boundaries between the journalistic, marketing and commercial worlds are increasingly being weakened, in terms of both their content and the specific focus of their respective teams. On the one hand, the most highly specialized professionals of the newsroom have to respond to the needs of the largest projects, whether or not there is a commercial agreement behind it, and on the other, it is the SEO content, fully oriented on the user, that

is doing most to break down the barriers that have traditionally separated the information of opinion from advertising.

As far as the optimisation of search engines is concerned, the experience gained by the *Diario de Sevilla* over the last three years provides a clear example of the importance of combining specialized profiles dedicated entirely to improving positioning (both in terms of technical and content optimization) and of the need for the implementation and dissemination of an SEO culture throughout a newsroom until such knowledge and strategies have been integrated as just one more factor in the journalistic routines associated with agenda setting, writing and specific focus.

However, restructuring online newspapers in this way is a challenge across the board, from those taking their first initial steps to the emerging realities of various leading newspapers in Latin America (the case, for example, of *El Tiempo* in Colombia and *La Nación* in Argentina), newsrooms are being transfigured as they seek to adapt to the communication flows of the digital environment and social networks, to the different challenges brought by this transformation and to the new journalistic constraints of the profession. In this sense, the sector has initiated a debate about the survival or otherwise of the various sections that constituted the central axis around which newspapers have been structured as they shift to a model that can respond, in one of its channels, to the production of SEO content.

In the case study described in this chapter, the good quantitative results reported by Google Analytics for the *Grupo Joly* have even conditioned the decision to reinforce the newsroom with the hiring of new SEO professionals. Moreover, the team that is currently being formed is highly illustrative of the latest trends affecting the labour conditions of newspaper companies: the SEO professionals are hired as freelancers, their renewal in the post depending on their meeting goals measured in terms of increased web traffic.

References

Angrosino, M. (2012). *Etnografía y observación participante en Investigación Cualitativa.* Colección Investigación Cualitativa. Editorial Morata.

Asser, M. (2012). Search Engine Optimisation in BBC News. *BBC*. http://www.bbc.co.uk/blogs/internet/posts/search_engine_optimisation_in

Caminero, L. & Sánchez-García, P. (2018). El perfil y formación del ciberperiodista en redacciones nativas digitales. *Hipertext.net*, (16), 4-15. https://doi.org/10.31009/hipertext.net.2018.i16.04

Carlson, M. (2007). Order versus access: news search engines and the challenge to traditional journalistic roles. *Media, Culture & Society*, *29*(6), 1014-1030. https://doi.org/10.1177/0163443707084346

Cebrián-Herreros, M. (2009). Nuevas formas de comunicación: cibermedios y medios móviles, *Comunicar*, *17*(33). https://doi.org/10.3916/c33-2009-01-001

Coller, X. (2000). *Estudio de caso. Colección Cuadernos Metodológicos.* Centro de Investigaciones Sociológicas.

Díaz-Noci, J. (2018). Hipertexto en periodismo: realidad e investigaciones de futuro.. *Anuario ThinkEPI, 12*, 230-235. https://doi.org/10.3145/thinkepi.2018.34

Dick, Murray (2011). Search Engine Optimization in UK News Production. *Journalism Practice, 5*(4), 462–77. https://doi.org/10.1080/17512786.2010.551020

García-Avilés, J.; Carvajal, M. & Arias, F. (2018). Implantación de la innovación en los cibermedios españoles: análisis de las percepciones de los periodistas. *Revista Latina de Comunicación Social*, (73), 369 - 384. https://doi.org/10.4185/RLCS-2018-1260

Giomelakis, D. & Veglis, A. (2015a). Investigating Search Engine Optimization Factors in Media Websites, The case of Greece. *Digital Journalism, 4* (3), 379-400. https://doi.org/10.1080/21670811.2015.1046992

Giomelakis, D. & Veglis, A. (2015b). Employing Search Engine Optimization techniques in Online News. *Studies in Media and Communication, 3* (1), 22-33. http://dx.doi.org/10.11114/smc.v3i1.683

Gonzalo-Penela, C.; Codina, L. & Rovira, C. (2015). Recuperación de información centrada en el usuario y SEO: categorización y determinación de las intenciones de búsqueda en la Web. *Index comunicación, 5*(3), 19-27. https://journals.sfu.ca/indexcomunicacion/index.php/indexcomunicacion/article/view/197

Iglesias-García, M.& Codina, L. (2016). Los cibermedios y la importancia estratégica del posicionamiento en buscadores (SEO). *Opción, 32* (9), 929-944. http://hdl.handle.net/10045/72320

Lopezosa, C.; Codina, L. & Caldera-Serrano, J. (2018). SEO semántico: Framework ISS para la optimización de sitios intensivos en contenidos. *Cuadernos de Documentación Multimedia, 29*, 97-123. http://dx.doi.org/10.5209/CDMU.60607

Lopezosa, C.; Codina, L.; Díaz-Noci, J. & Ontalba, J. (2020). SEO y cibermedios: De la empresa a las aulas. *Comunicar, 63*, 65-75. https://doi.org/10.3916/C63-2020-06

Lopezosa, C.; Trillo-Domínguez, M.; Codina, L. & Cabrera-Méndez, M. (2021). El SEO en la empresa periodística: percepciones y elementos clave para su adopción en las redacciones. *Revista Latina de Comunicación Social, 79*, 27-45. https://www.doi.org/10.4185/RLCS-2020- 1487

Muerza-Ferrer, A. (2017). Producción de noticias radiofónicas en el contexto de la convergencia periodística: análisis de Radio Marca y sus sinergias con Marca y marca.com. *Estudios sobre el Mensaje Periodístico* 24(1), 758-795. https://doi.org/10.5209/ESMP.59979

Norris, C. (April 13, 2007). SEO For Journalists: Introduction & Keywords. *Moz.com.* https://moz.com/ugc/seo-for-journalists-introduction-keywords-part-1-of-5

Pons, M. & Monistrol, O. (2017). *Técnicas de generación de información en investigación cualitativa II. Curso de Introducción a la Investigación Cualitativa. Máster de Investigación en Atención Primaria*. Sociedad Española de Medicina de Familia y Comunitaria, Universitat Autònoma de Barcelona. Fundació Doctor Robert. https://www.semfyc.es/wp-content/uploads/2018/01/M4_curso_intro_investigacion_cualitativa.pdf

Richmond, S. (2008). How SEO is changing journalism. *British Journalism Review*, (19). 5. doi: https://doi.org/10.1177/0956474808100865

Rodríguez-Martínez, R.; Codina, L.; Pedraza-Jiménez, R.(2010). Cibermedios y web 2.0: modelo de análisis y resultados de aplicación. *El profesional de la información, 19*(1), 35-44. https://doi.org/10.3145/epi.2010.ene.05

Rodríguez-Martínez, R.; Codina, L.; Pedraza-Jiménez, R.(2012). Indicadores para la evaluación de la calidad en cibermedios: análisis de la interacción y de la adopción de la Web 2.0. *Revista española de documentación científica, 35*(1), 61-93. https://doi.org/10.3989/redc.2012.1.858

Salaverría, R. & Martínez-Costa, M.P. (2021). *Medios nativos digitales en España: caracterización y tendencias*. Comunicación Social.

Smyrnaios, N. & Rebillard, F. (2009). L'actualité selon Google. L'emprise du principal moteur de

recherche sur l'information en ligne. *Communication et Langages*, (160), 95–109. https://doi.org/10.4074/S0336150009002087

Smyrnaios, N. (2015). Google and the Algorithmic Infomediation of News. *Media Fields Journal*, (10), 1-10. https://mediafieldsjournal.squarespace.com/google-algorithmic-infomedia/

Suárez-Sucre, E. (2017). Comprendiendo a los cibermedios. Una propuesta estructural para el análisis de los medios de comunicación en la red. *Estudios sobre el Mensaje Periodístico, 23*(1), 273-288. http://dx.doi.org/10.5209/ESMP.55596

Trillo-Domínguez, M. & Alberich-Pascual, J. (2017). Deconstrucción de los géneros periodísticos y nuevos medios: de la pirámide invertida al cubo de Rubik. *Profesional De La Información*, 26(6), 1091–1099. https://doi.org/10.3145/epi.2017.nov.08

Trillo-Domínguez, M. & Ollero, J. (2018). Estrategias de comunicación periodística para publicar en Redes Sociales desde la prensa local: Plan de Acción de Grupo Joly. *Hipertext.net*, (16), 68-77. https://doi.org/10.31009/hipertext.net.2018.i16.11

Valles, M. (2002). *Entrevistas cualitativas. Colección Cuadernos Metodológicos*. Editorial Centro de Investigaciones Sociológicas.

Yin, R. (2014). *Case study research. Design and methods.* SAGE.

Pere Freixa
Universitat Pompeu Fabra
pere.freixa@upf.edu

Carles Sora-Domenjó
Universitat Politècnica de Catalunya
carles.sora@citm.upc.edu

Joan Soler-Adillon
Universitat Oberta de Catalunya
jsoleradillon@uoc.edu

Webdocs: Social interaction and transmedia

https://doi.org/10.3145/indocs.2022.6

Freixa, Pere; Sora-Domenjó, Carles; Soler-Adillon, Joan (2022). Webdocs: Social interaction and transmedia. In: Freixa, Pere; Codina, Lluís; Pérez-Montoro, Mario; Guallar, Javier (ed.). *Visualisations and narratives in digital media. Methods and current trends*, (pp. 81-100). Barcelona: DigiDoc-EPI. https://doi.org/10.3145/indocs.2022.6

Abstract

Webdocs are facilitating the appearance of documentary projects in which the interaction and participation of all the partners implied –authors, community and audience– generate a shared space of documentation which in turn become part of a collective memory. In these projects, the authors articulate the dialog with and among such communities with digital and transmedia communication strategies, in which social networks play a central role. In this paper the results about the participation and transmedia strategies on i-docs are presented. Research is based on a triangulation of methods using study of literature, analysis of 3 case studies and interviews with the i-docs authors and producers. We argue that these projects present the following key characteristics: ability to adapt to each specific context, a main goal is to transcend the digital space and positively affect the participant communities, but also the difficulties to find a role and space within the larger media context.

Keywords

Web documentary, i-docs; transmedia, interactive documentary, participatory culture, webdoc.

Título

Documental Web: interacción social y transmedia

Resumen

El webdoc está facilitando la emergencia de proyectos en los que la participación y la interacción de todas las personas implicadas, ya sean autores, comunidad sobre la que se actúa o audiencia, se articulan para generar un espacio de documentación compartido, de historias que devienen memoria colectiva. En estos proyectos se dinamizan y articulan comunidades ciudadanas por medio de estrategias de comunicación digital y transmedia, en los que las redes sociales juegan un rol determinante. Se presentan resultados de las estrategias de participación transmedia en webdocs. La investigación se realiza a partir de una triangulación metodológica formada por la revisión bibliográfica, el estudio de caso de 3 proyectos y entrevistas con los autores y productores de esos proyectos. En el estudio se constata: la diversidad de procedimientos, la capacidad de adaptación a la realidad de cada escenario, la voluntad de trascender el espacio digital e incidir en las comunidades participantes, así como la dificultad para inscribirse en el escenario mediático.

Palabras clave

Documental web, i-docs, transmedia, documental interactivo, cultura participativa, webdoc.

1. Introduction

Transmedia storytelling, a term first coined by Henry Jenkins in 2006, describes the creation of narrative experiences that are dispersed systematically across multiple channels or platforms of delivery so as to create a unified entertainment experience – a narrative world – incorporating varying degrees of unity and coordination (Jenkins, 2006; 2011). One area in which transmedia storytelling has acquired particular relevance is that of the online interactive documentary or webdoc, a work situated in the documentary tradition but distributed via the web and incorporating multimedia and interactivity (Nash, 2012).

In recent years, many webdoc authors claim to have adopted the postulates of transmedia theory in designing their projects, be it in relation to their methods of production, the construction of their narrative or their means of promoting audience participation. Indeed, this is evident in many of the works compiled by the main directories and specialised web portals (MIT – Docubase, NFB/Interactive, i-Docs, IDFA DocLab, docSHIFT Index, among others). At the same time, preliminary research charts the emergence of transmedia storytelling in the world of the interactive documentary, as well as the diversity and singularity of the projects, which greatly hinders their characterisation (Freixa, Sora, Soler-Adillon & Ribas, 2014; Freixa, 2015; Sora, 2016a; 2016b; Freixa, Pérez-Montoro & Codina, 2017; Soler-Adillon, 2017; Miles, Sora, Fetzner & Aston, 2017; Sora, 2018).

What is becoming increasingly evident is that a number of these projects have the ability to energise and empower citizen groups (whole communities even), and to promote processes of communication based on the creation of collective archives and stories. More often than not they are not large ventures but, rather, projects with strong local roots that exploit transmedia storytelling as a means of interacting with the community in and with which they work. Above and beyond their use of transmedia strategies, these successful productions are characterised by the ties they establish with the territory in which they are made; yet, they cannot, as a matter of course, depend on stable production resources or notable means of distribution, although, in many cases, they can count on the exceptional participation of unidirectional media, such as TV channels, and funding, albeit only ever partial and always insufficient, from public or private grants.

This study focuses its attention on this specific category of transmedia project. More specifically, we present a selection of works that highlight the informative potential and the capacity for citizen empowerment of projects developed and implemented by exploiting transmedia strategies and which concern themselves with local problems capable of generating global communication processes. Each of them can be defined as a digital media product; they each use, exploit and define themselves in terms of social networks and seek to transcend

the digital environment in order to impact social policy and affect the lives of the people that participate in them.

2. Objectives and methodology

This study seeks to identify and define different transmedia strategies by undertaking several case studies of these projects and by interviewing their authors and producers. In so doing, we employ a methodological triangulation that combines the following approaches: case study, bibliographic review and in-depth interview. Unlike transmedia projects in the world of entertainment or fiction, transmedia documentaries seek the incorporation of the audiences and the communities involved in them throughout the different phases of production, that is, from their very ideation to the contribution of their actual content and their ultimate dissemination. Their authors aim to adapt their proposal to the participation of the audience and the flows of communication that the project generates.

Our primary objective here, therefore, is to sketch out an initial characterisation of the transmedia strategies developed by these interactive webdocs. To do so, we present an exploration of the communicational effectiveness and discursive possibilities of these webdocs that use the basic structures of transmedia and interaction as fundamental strategies of their operational dynamics. Specifically, we have selected three examples of what can be considered successful webdocs insofar as they have achieved the results their authors expected from exploiting these novel procedures.

To select our case studies, in an initial exploratory phase, we conducted a review in specialised directories of works meeting the conditions. On this basis, three works were selected on the grounds of their illustrative value, the possibilities they afforded for study (Stake, 1995; Yin, 2013) and the nature of their impact on the community in which and for which they were created. For the interviews, several of their authors were invited to participate in a round table organised within the framework of the *III Conference on Interactive Communication and Cybermedia*, CIC2017, held at the Universitat Pompeu Fabra. In the case of *Proyecto Quipu*, an in-depth interview was conducted individually with the author.

Title	Authors	URL address	Year
Cuentos de viejos (seasons 3 & 4)	Marcelo Dematei, Carlos Smith, Laura Piaggio & Anna Ferrer	http://cuentos-deviejos.com	2016 and 2017
Proyecto Quipu	Rosemarie Lerner & María Court	https://interactive.quipu-project.com	2015
Orgull de Baix	Isabel Fernández	http://www.orgull-debaix.cat/	2016

Table 1. List of webdocs analysed.

3. State of the art

In the course of the last two decades, authors, academics and theorists in the field of communication have debated at length the properties and characteristics of interactive journalism and documentaries in the digital media. The webdoc, i-doc, web documentary or interactive documentary has been studied from a wide range of different approaches, from within many different theoretical frameworks and in many distinct disciplines. As an audiovisual text, deemed complete when accessed by a viewer, authors such as Kate Nash and Richard Walsh have characterised interactive works by their ability to permit the emergence of content as they are explored and read by the user (Walsh, 2011; Nash, 2014). For Aston and Gaudenzi (2012), what makes interactive stories different from other discursive forms is their ability to propose conversational, participatory, experiential and hypertextual solutions of their own. For most authors, interactivity and participation define the specific qualities of the interactive story, where interactivity is construed in relation to the semantic capacity of the markup language that provide the hypertext, later hypermedia, links, as first developed by Nelson and Landow (Nelson, 1983; Landow, 1991; Pavlik, 2001; Aston, 2003). Interactivity, as an element of computational communication, forms part of the technological characteristics that define and condition digital communication, and which have facilitated the definition of tools and resources for dialogue, in the form of platforms and protocols, procedures and processes. The conception of interaction as a set of technological tools has allowed, as a consequence, the study of the limiting and conditioning capacity of communication models (Schultz, 1999; McMillan, 2002).

Another of the fundamental qualities of the interactive documentary is the lengths it goes to involve the participants in it and its efforts to ensure a greater degree of audience engagement than one would expect in a traditional documentary. In a number of projects, these objectives have been pursued using strategies of co-creation (Miller & Allor, 2016) as well as co-design (Green et al., 2017), which means participation is not necessarily limited to content creation, but it can also permeate the whole process of shaping the interactive documentary. Hence, the interface acquires central importance, since it is not merely a means of presentation, but also an essential part of the experience itself. In exploring the documentary, the interface design facilitates the creation of experiences, the different elements of the content forming links with each other in ways that are more complex than in a linear experience, thus simulating "chance encounters" (Stewart, 2019) with the different characters, spaces or objects that make up the documentary.

We are dealing with documentaries that, quite clearly, do much more than simply present their content online: they are medium-specific, native web projects. This makes it possible to create experiences that, according to the creators of i-docs, are remarkable for their playfulness, but also for their technical complexity, which in turn cannot be separated from the

handicaps that this represents in terms of their reaching broad audiences. Yet, such projects call into question the way in which the documentary genre in general is understood, especially with regard to the relationship they establish with their audiences via their very specific form of presentation (Cucinelli et al., 2018) - a form that centres on interaction. However, some authors consider their interactivity to be a limiting factor when compared to the traditional format. They argue that their works sacrifice the narrative force of the linear story for that of free exploration, more akin to navigating in a database (Forceville, 2017).

Moving from analyses of modes of interaction in webdocs to their content, some studies have centred their attention on the discursive qualities of their structures. These new digital narratives – thanks to current web technologies (Sora, 2015) – are nourished by small individual syntactic forms that are distributed and interconnected so as to generate multiple narratives (Miles, 2017). These rhizomatic structures facilitate the configuration of new narrative codes where the contributions of the users not only complement the discourse, but also form an actual part of the narrative construction.

In this sense, the fact that interactive documentaries can be generated from a large number of audience contributions – the case of those studied in this article – means that they can be considered as living products, their shape evolving with the links they forge with the community for which they are created, together with the audiences that become an active part of the dialogue, generating conversations between the witnesses, the audiences and the authors of the projects. These are veritable processes of co-authorship and, often, of co-design too (Rose, 2011), that are generated in longer, more complex periods of production and with greater polyphonic potential (that is, accommodating multiple voices) (Aston & Odorico, 2018), which try to incorporate different participating voices into all their processes. In this regard, ideas concerning models of audience participation in digital documentaries as defined, among others, by Rose (2011), Jenkins & Carpentier (2013) and Nash (2014) are of interest to our discussion here.

Nash differentiates between participation "*in media*" and "*through media*" (2014, p. 5). She uses the former to refer to the classic approach of representing the voices of those bearing testimony within the digital documentary, while the latter is used to emphasise the potential of interactive projects to create spaces for collective participation where testimonies and opinions can be expressed in a more open way, a way that is less dependent on the author of the project. This allows a more extensive dialogue to be obtained, one that can reach more participants not actually present during the production of the project and the greater community surrounding the project. This second category, participation through the digital documentary, is closely related to the technical possibilities afforded by online participation in social networks, or platforms programmed for this use. For Jenkins & Carpentier, being able to contemplate participation through the media "allows us to zoom in on decision-making

processes within media organizations themselves and analyze how equal or unequal the power relations in these settings are" (2013, p. 274).

The three interactive documentaries that we analyse in this article present social dialogues articulated via digital interfaces that respond to the previously mentioned categories of social interaction. They are non-linear projects, involving complex iterations in which personal stories are intertwined in multiple temporal layers (Sora, 2018) offering alternative spaces for communication and social denunciation to the hegemonic circuits of linear documentaries.

4. Case studies

The three projects selected boast remarkable longevity. They are all far-reaching projects, developed using complex participatory procedures and designed to have a permanent active presence on the web, limited only, as we shall see, by cost and maintenance factors. The three interactive documentaries have in common their use of transmedia narrative strategies that seek to give a social value to user participation and, in this way, to create not only projects about a specific local reality, but works that have intrinsic value for the communities involved.

4.1. *Cuentos de Viejos*

Cuentos de Viejos (Old Folks' Tales), a transmedia project that was begun in 2012, continues to have an active presence on the web and is currently preparing its fifth season [http://cuentosdeviejos.com/convocatoria-tu-viejo-en-la-tele]. It defines itself as a "collaborative transmedia documentary [...] that is developed on television and the internet, as well as in homes and schools. A user experience that can start with the animated TV documentary series or via the collaborative online platform, or by participating in a school project" (Dematei, Smith, Piaggio & Ferrer, 2012). The project is a pioneer insofar as it links elements of the collaborative tradition with transmedia documentary and audiovisual production. After five years of online activity, it can be considered quite a rarity given its uniqueness but, perhaps, its most notable quality is its ability for engaging audiences, establishing and maintaining links with the traditional media, and for having created a stable transmedia strategy, in which each part of the project complements each other. In the words of Laura Piaggio, one of the project's co-producers, its success in encouraging audience participation can be attributed to the challenging nature of what the audience is required to do.

> *In our case, I believe the relevance of what we're asking from the audience is also important. I mean, we're asking for a lot. We ask them to go and talk to an elderly person, decide what they want to talk about, and if it seems interesting, to get a camera and record it. But it has to come out well. If not, they have to go back and record it again – then they must edit the video, they have to make cuts, as it can't be longer than three minutes. They have to fill out a form, get a whole bunch of authorisations and, finally, they can upload the video. The whole job takes a*

week, minimum. It's not simply about giving a 'like,' it's not a selfie; we're asking them to do something that is far from trivial. (Piaggio, 2017)

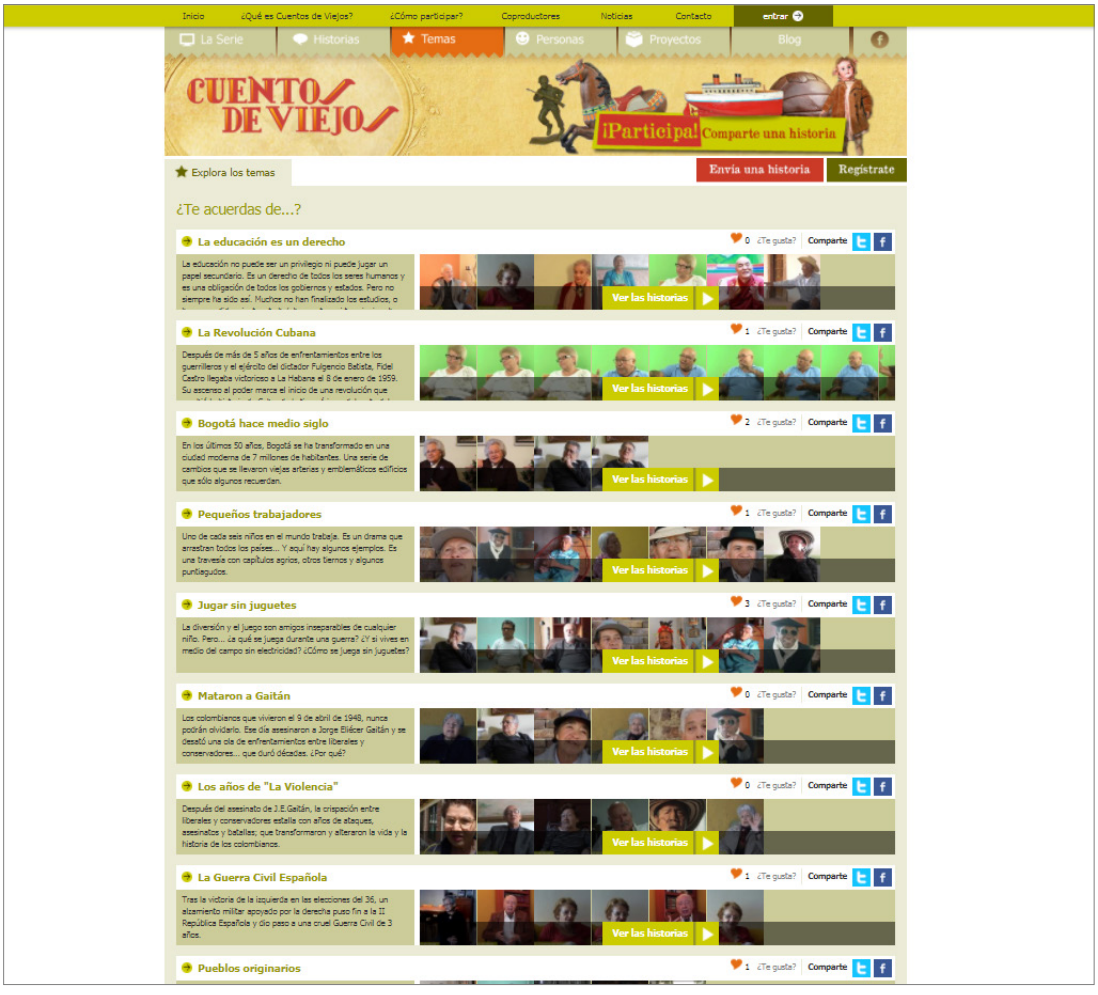


Figure 1. A screen shot of the topics page of the web project *Cuentos de Viejos*. Source: http://cuentosdeviejos.com/explora-los-temas

What the audience is required to do is to record a short video – no more than three minutes long – in which an elderly person recounts a childhood memory, a memory that is of interest both to them and to the person recording it, who could be their son or daughter, grandchild, neighbour or friend. It is, in the words of the authors, a project about both the individual and collective memory of a country, Colombia, in which an entrenched silence has suppressed any talk of the tragedies that have afflicted the country since the middle of the last century. The topics addressed include the forced displacements, war, violence and fear, all told in the first person. The building up of stories in this way has made it possible to create a collective story that had hitherto never been told. Some of the stories – chosen by the public and the producers – have been made into short animated clips and today form part of the free-to-air broadcasts of *Cuentos de Viejos* on the *Señal Colombia* TV channel.

> *I think the important thing, what we've managed to achieve, and here the TV series was fundamental, was that people could see how important their stories were for us. The work we did on their stories, [...] What we ask them to do, the more challenging we make it, the more relevant it has to be and, so, the more it is valued [...] The main thing is that people have to feel that you value what you're asking them to do, and that it makes sense for them to do it. (Piaggio, 2017)*

The transmedia project is made up of three main sections: the animations that are broadcast on television; the webdoc containing the collection of stories uploaded and ordered by location, author and topic; and, finally, agreements and collaborations with libraries and schools to investigate personal stories and to construct micro-stories from the memories and to establish reflective learning practices in relation to collective memory.

The project is interested, above all, in exceptional, unique stories, ostensibly quite small, that refer to, and allow the listener to subtly approach, the bigger stories, those which because

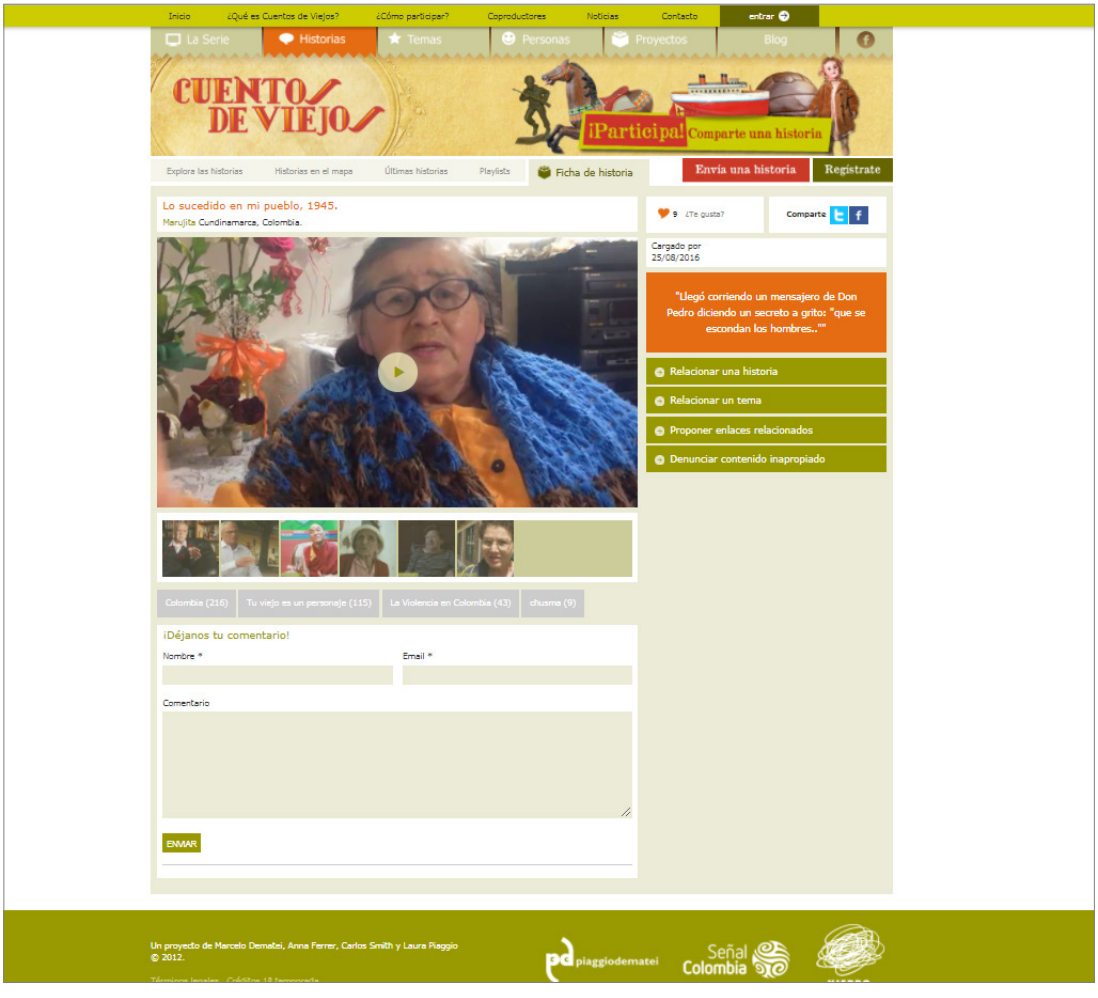


Figure 2. Screen shot of one of the testimonial videos contributed by a user of the web project *Cuentos de Viejos*. http://cuentosdeviejos.com/explora-los-temas

of their content threaten to overwhelm us. All the stories selected form part of the project's online collection, which in the years up to 2016 received more than 200,000 visits from 40,000 unique users. A few, as we have said, are then chosen for animation and are broadcast on free-to-air television, but they must first be edited, scripted, directed and produced, a task involving more than 30 people, including animators, illustrators, musicians and editors (Piaggio, 2017).

> *The project proposed undertaking a very specific exploration of the concepts of oral narratives and portraits from the point of view of animation which, when used as a documentary, proved to be a tool for making the invisible visible: a tool that's allowed us to reconstruct not only the historical dimension of the memories narrated, but more especially, their emotional dimension, highlighting at the same time the mechanisms of memory. (Dematei & Piaggio, 2015)*

The project is one of many recent, fruitful experimental productions based on transmedia narrative to be carried out in Latin America (Alberich-Pascual & Gómez-Pérez, 2016). In common with other works, including *Pregoneros de Medellín* (2015), it shares a commitment to the educational value of the project itself. To this end, it encourages collaboration with schools and libraries to expand the project and so connect with potential end users, inhabitants of the territory with whom it wishes to enter into dialogue and build a community. But probably its greatest potential lies in its commitment to the recovery of the oral memory. Cuentos de Viejos is a tribute to the oral tradition, to the ancestral ability to recount one's memories and stories so as to reconstruct the local microhistory, unique to that place. This recovery of the oral legacy is the most obvious link between the *Cuentos de Viejos* project and the next project we wish to describe, *Proyecto Quipu*.

4.2. *Proyecto Quipu*

Proyecto Quipu was published on the Internet in 2015. However, it was first set in motion in 2011 by Rosemarie Lerner and María Court, when they began their investigation into the National Program of Reproductive Health and Family Planning, established during the 1990s by the then president of Peru, Alberto Fujimori, and the denunciation of thousands of forced sterilisations carried out under the cover of this ministerial program. Proyecto Quipu began life as an experimental project, initiated by the two authors on the completion of their studies in documentary making:

> *I'd been exploring participatory methodologies on my Master's degree that year and had carried out a neighbourhood project in London, in which we invited neighbours to share their memories about a city block that was to be demolished. I'd begun to experiment with ways of opening up the narrative, looking at the neighbours in a way that extended beyond the mere perception provided by an author. It struck us as interesting if this project could do something similar, in some way be interactive between its different areas, as the topic was so huge, making a film seem to fall well short of the mark. (As reported in Freixa, 2019)*

The project involved a huge amount of documentation and gathering of information, but perhaps the most remarkable thing about the project lies in the efforts made by its authors to establish links with the communities affected, in order to activate their participation and to get them to denounce what had happened to them. It is basically about creating a virtual context for and with the collective in which the documentary provides spaces for dialogue about a silenced reality (Court & Lerner, 2015a).

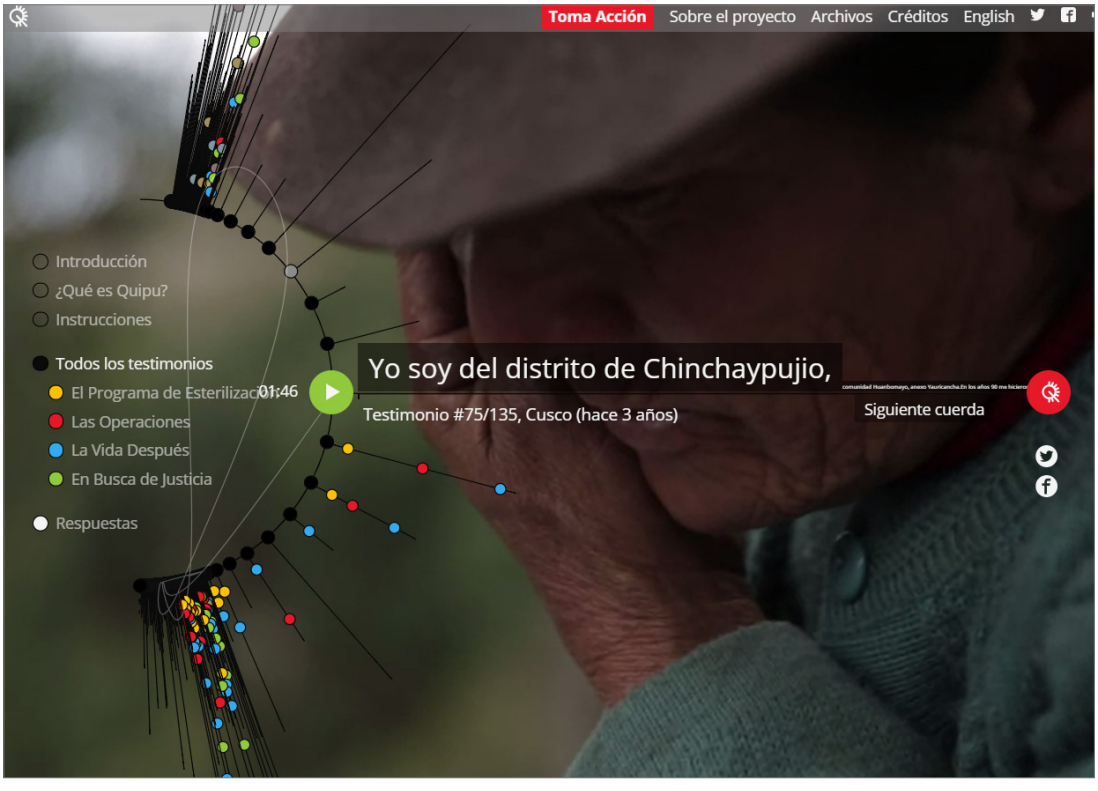


Figure 3. A screen shot of a page from *Proyecto Quipu* showing an interaction with one of the testimonies recorded. https://interactive.quipu-project.com/#/es/quipu/listen/61.

María Court speaks of the difficulties they faced in getting the project up and running:

> *The stage in which we managed to connect on the ground with the women's associations was long and drawn out. It was a year and a half before Cusco gave us the go ahead, that's right, more than a year; things were a bit easier in Huancabamba. Of course, at the outset, when I, a Chilean, turned up, the questions began: Who are you? With Rose, who's Peruvian, it was a little easier. Her father [Salomón Lerner Febres] had been involved in putting together the report of the Truth and Reconciliation Commission in Peru. So, slowly, but very slowly, we began to win their trust. Because obviously, it was a community that had been badly let down, they were frustrated and also distrustful of journalists, in general, and also of anyone who came with a camera to steal their story. It was never easy. (As reported in Freixa, 2019)*

In *Proyecto Quipu*, the authors made interactivity, related above all to participatory processes, their main design feature. Those affected, the protagonists of Quipu, are essentially

inhabitants of the Peruvian Andes. Many of them are speakers of Quechuan and its dialects. Many of them with few notions of Spanish.

The authors of the project designed a detailed transmedia strategy (see Figure 4) so as to establish contact with the inhabitants in their different dialects and to design protocols for obtaining and recording their testimonies. To do so, they established links with local radio stations – spokespersons for the project – and designed a telephone system for the anonymous recording of testimonies. The set of testimonies makes up the collective voice that denounces the genocide perpetrated in Peru. They are presented in the project interface by means of a *quipu*, a set of knotted strings used by the Inca people for collecting data, comprising numerous threads and coloured knots, that facilitates the ordering, classifying and archiving of the collective history of the community. This interactive device allows users to consult and read all the oral records, which are categorised, but not in any pre-established order. The user can activate and listen to them in any order they want and from any corner of the world. The project provides a digital platform, a means by which users can record new testimonies and listen to the archived stories. It proposes the pro-active involvement of the audience in favour of a culture of collective participation and empowerment, in which the authors of the project give control to the collective over future actions and the proposals made via the platform (Aston, Gaudenzi & Rose, 2017; Vázquez-Herrero & Moreno, 2017).

As Rosemarie Lerner and María Court point out, Quipu can be construed as an information system (Freixa, Perez-Montoro & Codina, 2017; Court & Lerner, 2015b): in the initial phases, the basic operating elements related to technology (use of telephone system, servers, file encoding, etc.) were characterised; next, the procedures, operating protocols and objectives of the action plan were defined with the users and agents involved; then, during the implementation stage, which in this case lasted several years, the interface was specified, the actions to be undertaken defined and the discursive features of the work given shape; finally, the videos were edited, and decisions were taken about the visual and sound quality.

As a structured, interactive information system, it was published and is maintained as an active platform on the web and, unless its authors decide to close it down, it has no expiry date. The platform is designed to allow the addition of new testimonies, both from affected parties and viewers who wish to record their responses. It is this participation and contribution of content in the conceptualisation phase that characterises the work as a collaborative transmedia project (Fox, 2017; Ortuño & Villaplana, 2017).

As in Katerina Cizek's webdoc, *Out My Window* (2010), in Proyecto Quipu technology plays a major role not only as support for its content, archiving and processing, but also in the definition of procedures and the development of the documentary process. This choice does not correspond to the fact of having registered to a certain platform, system or technology chosen a priori; each element, the use of radio stations, telephone lines, file storage, data processing

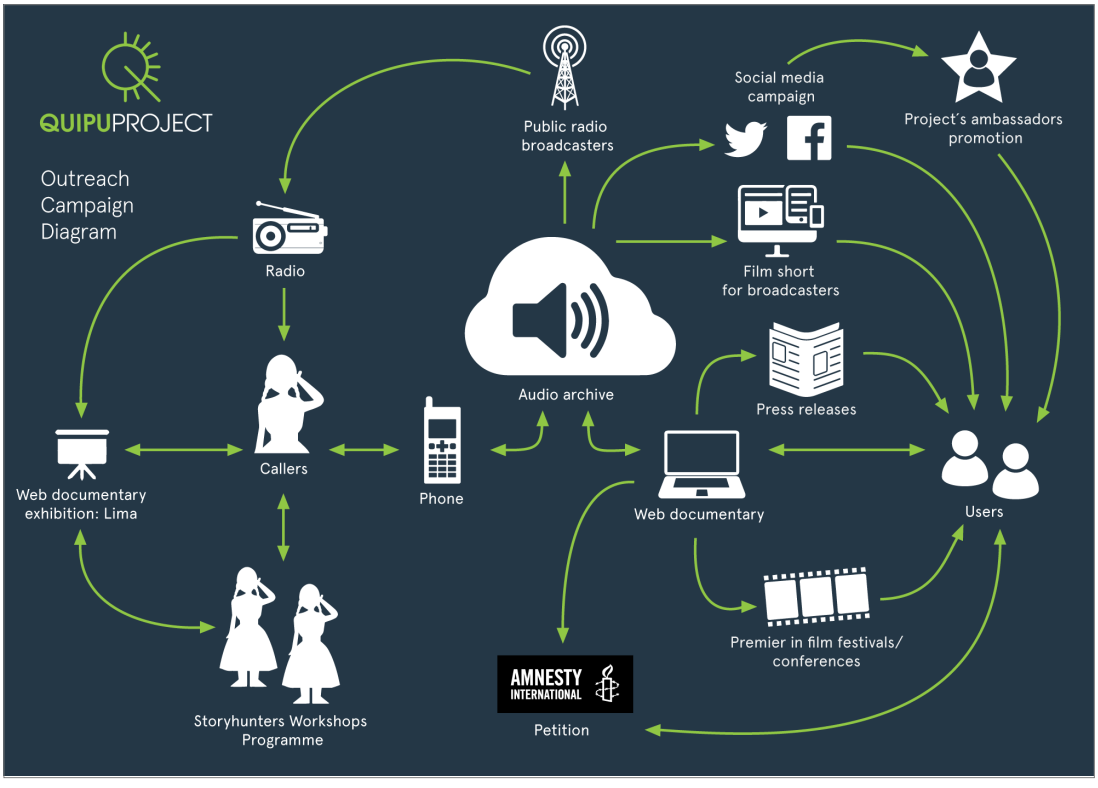


Figure 4. An illustration of the organisation and transmedia strategy deployed by *Proyecto Quipu*. Published at: http://blog.proyectoquipu.com/proyecto/#system.

according to standard languages, etc., responds to the will to adapt to the specific needs of the content (Freixa, 2018).

As Brown and Tucker (2017) are quick to point out, people who connect to the system can listen to the testimonies of all those that have gone before them and send a message back to them. This provides an opportunity to observe the collective dimension that the sterilisation process represented for the country. At the same time, it enables us to understand and to be aware of the potential of collective stories, in which any individual can decide to take part. Indeed, the project maintains the presence of the project's authors hidden – they remain very much in the background – so as to promote at all times the value of the collective and of participation. In this way, the project manages to transcend anonymity and become a collective voice.

Figure 4 highlights the dual global and local dimension of Proyecto Quipu, organised around the use of audio files. The identification and recording of testimonies (left-hand side) is carried out using public radio stations, mobile phones and, above all, as a result of the work of a significant number of volunteers who interact in situ with the affected parties and activate the chain of information essential for the project's success. This approach can be identified

as a form of transmedia activism to use the term proposed by documentary filmmaker Lina Srivastava (in Jenkins, 2016).

Proyecto Quipu has been a project with a fairly lengthy period of gestation, its development being helped by a number of different scholarships and grants and with funding from the British REACT program and the Tribeca Film Institute, among others. The support of the news media has been forthcoming at a somewhat smaller scale; in its final phase, in 2017, The Guardian helped finance the short film associated with the project (Freixa, 2019).

4.3. *Orgull de Baix*

This project took over the baton from a previous initiative, *El Pati del darrere* (The Backyard) a linear documentary, made in collaboration with the Catalan television company, TV3. It presented a portrait of the comarca or county of Baix Llobregat and which allowed, during its production, relationships and links to be established with the majority of actors, groups and entities that live in the territory and who would later go on to participate in the making of *Orgull de Baix*. Isabel Fernández and Anna Carreras joint directors and producers of the project *Orgull de Baix*[1] conclude that the success of the project is due to the role that social networks have played and continue to play in the development and implementation of the project. As they themselves say, "the social networks in Orgull de Baix have not been used as marketing tools, but rather they have allowed the construction of complementary spaces for debate, research, the gathering of information and for forging relationships with the audience" (2017).

Orgull de Baix is presented on the screen by means of a three-dimensional reproduction of the territory, in the form of a map of the whole comarca. A collection of micro-narratives that make up the project have been placed at the surface (Figure 5). As if taking a trip or playing a video game, the viewer can travel around the territory, activate each of the geolocated pieces on the ground and discover the stories, characters and themes it contains.

The project focuses its interest on the survival of agriculture in urban environments under the pressure of the metropolis. It explores the tensions between urban and rural models related to such issues as sustainability, ecological balance, and the determination on the part of the inhabitants to maintain their way of life. Isabel Fernández explains that at the beginning of the project: "We were exploring how to bring the documentary to the web and how best to play with interaction [...] We were very interested in how we could conquer the audiences of the 21st century" (Fernández, 2017). The virtual community created in the development of

1 In Catalan, *Orgull de Baix*, "Baix Pride", is a reference to the Baix Llobregat (Lower Llobregat), a comarca in the metropolitan area of Barcelona, one of the agricultural areas subject to the greatest urban and industrial pressures in the whole of Spain. However, it is also a play on words, with a clear reference to the strong working-class pride in this agricultural county that contrasts starkly with the city of Barcelona, with baix also meaning lowly, or poor – hence, a pride in their social status too.

Figure 5. Main page of the project *Orgull de Baix*. Published at: http://www.orgulldebaix.cat/#es/map.

the project has now taken it over: "They have identified with the *Orgull de Baix* brand and they have promoted it via the #orgulldebaix hashtag, which appears each day in multiple posts related to the agro-ecological and social debate in the comarca [...] on Twitter and Facebook" (Fernandez, 2017).

In common with *Cuentos de Viejos*, a television channel – in this case a public one – Televisión de Catalunya (TV3), has participated in *Orgull de Baix* by disseminating the linear version of the project. Indeed, initially, TV3 explored the narrative possibilities of using a second screen and broadcast the linear version synchronised to the exploration of the online webdoc. In this way, the web used the broadcast as a second screen of the web exploration. The project's transmedia strategy also includes the implementation of training activities so that the people of Baix Llobregat can interact with the project. To this end, a series of workshops entitled "*Orgull de Baix*, memories of the landscape" has been organised, in which some thirty people, ranging from high school students to the over 65-year-olds, have participated. In teams that combined young and old and over a period of two months, the participants produced a two-and-a-half-minute micro-documentary, made and edited on smartphones. These documentaries have been gradually incorporated into the work. For Isabel Fernández, the webdoc, the interactive documentary is one of the best contemporary forms available for exploring citizen empowerment:

Figure 6. Screen shot from one of the microstories that make up the *Orgull de Baix Project*. Fuente: http://www.orgull-debaix.cat/#es/map

What I'm always looking for are doors: "What doors can I open to get people to come in?" Here, what I detected and what really motivated me was that I could really begin to shake things up with this project [...] I was looking for something that would mobilize people, an issue that would show the power civil society has to transform something [...] I thought, if these people, by remaining silent, [...] have managed to preserve a large part of this comarca, there must still be a space from which they can make their voices heard and from where this movement can grow; [...] there is now much greater synergy among these people. Before, although they shared the same space, working for the same cause, they were working from different angles, their paths never crossed and, of course, when they called a meeting or something only ten people showed up. Now, thanks to our work, there is much more communication. And, yes, [...] this is transmedia, but it is no longer digital, it is very much human. [...] You get the idea that when those working with traditional audiovisual speak, we are somehow inhuman, that we are like machines, but no, no, we are more human than anyone. (Fernandez, 2017)

5. Discussion and results

Each of the three projects outlined above has developed its own unique transmedia strategy to organise its various component parts into an interactive participatory project, a project, as we have seen, of transmedia activism. *Quipu* and *Orgull de Baix* prioritised the establishment of a dialogue with the community during the ideation and definition of their projects and with their respective community of participants, the guidelines, the tone and the objectives sought

were all debated and established. The gathering of content was, in all three projects, essential for their existence and success; in each, the content is representative of the high level of commitment to the project. All three, moreover, involved the implementation of the different phases that make up an audiovisual production: the definition of content, its production and direction, the audio and/or video recordings, their editing and publication. *Proyecto Quipu* allows, moreover, the contribution of content from a part of the audience that identifies with its goals, the public not directly affected by the sterilisations but those who want to participate and show their support for the project.

Another important characteristic of these projects is the degree of control that the authors exercise over each part of the process. Thus, they are responsible for the financing and production of the work, as well as for defining the basic elements that make up the information system: the technologies and programs used, the creation and definition of the databases, the coding system and the procedures and functionalities that facilitate their use. They also assume, in all three cases, the maintenance and promotion of the project in the social networks, the engine, according to all of them, of the success in the creation of linked communities, the main objective of all three.

Interaction via the interface takes the form of discovery in *Orgull de Baix* and *Quipu*. Both projects place the viewer in front of an interface that organises the content by means of semantic layers: *Orgull de Baix* uses the metaphor of the map and the territory to geolocate its content; *Quipu* uses the Andean set of knotted strings of the same name to order the stories that make up the project's auditory archive, the voices of the testimonies.

In all three cases, the authors seek to ensure their proposals have an impact beyond the digital space. They aspire to influence media agendas and local and global policy through citizen empowerment. Laura Piaggio, in reflecting on the *Cuentos de Viejos* project, concludes:

> *There is a shroud of silence over the violence that wracked Colombia fifty years ago and which led to forced displacements from rural towns to the big cities, where there were all kinds of violence; sexual violence, physical violence, economic violence, deaths, violence of all kinds. Many people refused to speak about it and just got on with rebuilding their lives. "We came here, we had to come – end of story". And no one asked any questions and everyone kept their mouths shut. But then these stories began to appear. The grandmother telling her story about what she had seen before leaving her village and, in the comments, you get to hear how her family felt about that. And that's where you feel the transformative power lies. Suddenly there is something that, maybe, you didn't think would amount to much when you were planning it, but, no, it's there, and it quickly emerges from the digital medium. (Piaggio, 2017)*

References

Alberich Pascual, J. & Gómez-Pérez, F.-J. (2016). Exploraciones transmedia en la creación cinematográfica colaborativa iberoamericana contemporánea [Transmedia explorations in contemporary Ibero-American collaborative film-making]. *Artnodes*, 18. https://doi.org/10.7238/a.v0i18.3054

Aston, J. (2003). *Interactive multimedia: an investigation into its potential for communicating ideas and arguments* [Doctoral dissemination]. Royal College of Art. https://ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.565936

Aston, J. & Gaudenzi, S. (2012). Interactive documentary : setting the field. *Studies in Documentary Film*, 6(2), 125–139. https://doi.org/10.1386/sdf.6.2.125_1

Aston, J.; Gaudenzi, S. & Rose, M. (2017). *i-Docs: The Evolving Practices of Interactive Documentary*. New York: Columbia University Press.

Brown, M. & Tucker, K. (2017). Unconsented Sterilisation, Participatory Story-Telling, and Digital Counter-Memory in Peru. *Antipode*, 49(5), 1186–1203. https://doi.org/10.1111/anti.12316

Carreras, A. & Fernández, I. (2017). Orgull de Baix: un documental web per explorar i descobrir tot un paisatge i patrimony [Orgull de Baix (Baix Pride): a web documentary dedicated to explore and discober the Baix Llobregat's landscape and heritage]. *Hipertext.net*, 15, 66–73. https://doi.org/10.2436/20.8050.01.47

Court, M. & Lerner, R. (2015). The Project. (Blog Proyecto Quipu). https://interactive.quipu-project.com/#/en/quipu/intro

Court, M. & Lerner, R. (2015). The System. (Blog Proyecto Quipu). https://interactive.quipu-project.com/#/en/quipu/intro

Cucinelli, G.; Réne-Véronneau, É. & Oldford, B. (2018). Interactive documentaries and the connected viewer experience: Conversations with Katerina Cizek, Brett Gaylor, Jeff Soyk, and Florian Thalhofer. *First Monday*, 25(5). https://doi.org/10.5210/fm.v22i5.6182

Dematei, M. & Piaggio, L. (2015). Cuentos de Viejos: Los relatos de la Memoria. [Cuentos de Viejos: The tales of memory]. *Poli[Papers]*, 5, 88–101. https://doi.org/10.4995/caa.2015.3541

Dematei, M.; Smith, C.; Piaggio, L. & Ferrer, A. (2012). ¿Qué es Cuentos de Viejos? [What is Cuentos de Viejos?]. Cuentos de Viejos. https://cuentosdeviejos.com/que-es-cuentos-de-viejos

Deuze, M. (2006). Participation, Remediation, Bricolage: Considering Principal Components of a Digital Culture. *The Information Society*, 22(2), 63–75. https://doi.org/10.1080/01972240600567170

Fernández, I. (2017). *La producción y la difusión de Webdocs. Dentro y fuera de los Medios* [Production and dissemination of Webdocs. In and outside media]. https://youtu.be/q3U_B8S5XUo

Forceville, C. (2017). Interactive documentary and its limited opportunities to persuade. *Discourse, Context & Media*, 20, 218–226. https://doi.org/10.1016/J.DCM.2017.06.004

Fox, B. (2017). *Documentary Media: History, Theory, Practice*. New York: Routledge.

Freixa, P. (2019). Entrevista a María Court [Interview to María Court].

Freixa, P. (2018). Content Access, Storytelling and Interactive Media. In M. Pérez-Montoro (Ed.), *Interaction in Digital News Media. From Principles to Practice* (pp. 9–31). London: Palgrave Macmillan. https://doi.org/10.1007/978-3-319-96253-5_2

Freixa, P. (2015). Reportajes especiales en los cibermedios. Análisis de diez años de Premios Goya y Oscar Awards en Elpais.com (2005-2014) [Special features in online journalism. Ten years of Premios Goya and Oscar Awards in Elpais.com (2005-2014)]. *El Profesional de La Información*, 24(3), 291–300. https://doi.org/10.3145/epi.2015.may.09

Freixa, P.; Pérez-Montoro, M. & Codina, L. (2017). Interacción y visualización de datos en el periodismo estructurado [Interaction and data visualization in structured journalism]. *El Profesional de La Información*, 26(6), 1076–1090. https://doi.org/10.3145/epi.2017.nov.07

Freixa, P.; Soler-Adillon, J.; Sora, C. & Ribas, J. I. (2014). Snow Fall and A Short History of the Highrise: two approaches to interactive communication design by The New York Times. *Textual and Visual Media*, (7), 63-84.

Gaudenzi, S. (2014). Strategies of Participation: The Who, What and When of Collaborative Documentaries. In K. Nash, C. Hight, & C. Summerhayes (Eds.), *New Documentary Ecologies. Emerging Platforms, Practices and Discourses* (pp. 129–148). Houndmills: Palgrave Macmillan.

Green, D. P.; Bowen, S.; Hook, J. & Wright, P. (2017). Enabling Polyvocality in Interactive Documentaries through "Structural Participation." In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (pp. 6317–6329). New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3025453.3025606

Harrigan, P. & Wardrip-Fruin, N. (2009). *Third Person: Authoring and Exploring Vast Narratives*. Cambridge, Massachusetts: MIT Press.

Jenkins, H. (2016). Telling Stories: Lina Srivastava Talks About Transmedia Activism (Part One). *Confesions of an ACA-FAN.* https://henryjenkins.org/blog/2016/01/telling-stories-lina-srivastava-talks-about-transmedia-activism-part-one.html

Jenkins, H. (2011). Transmedia 202: Further Reflections. *Confesions of an ACA-FAN*. https://henryjenkins.org/blog/2011/08/defining_transmedia_further_re.html

Jenkins, H. (2006). *Convergence culture: Where old and new media collide*. New York: New York University press.

Jenkins, H. & Carpentier, N. (2013). Theorizing participatory intensities: A conversation about participation and politics. *Convergence*, 19(3), 265–286. https://doi.org/10.1177/1354856513482090

Landow, G. P. (1991). *HyperText: the convergence of contemporary critical theory and technology*. Baltimore: Johns Hopkins University Press.

Larrondo-Ureta, A. (2011). The Potential of Web-Only Feature Stories. *Journalism Studies*, 12(2), 188–204. https://doi.org/10.1080/1461670X.2010.491300

McMillan, S. J. (2002). A four-part model of cyber-interactivity. *New Media & Society*, 2(271–291).

Miles, A.; Sora, C.; Fetzner, D. & Aston, J. (2017). *The Material Turn and Interactive Documentary: A Panel*. https://zenodo.org/record/1120448#.XixiNRNKi8p

Miller, L. & Allor, M. (2016). Choreographies of collaboration: social engagement in interactive documentaries. *Studies in Documentary Film*, 10(1), 53–70. https://doi.org/10.1080/17503280.2016.1171686

Nash, K. (2014). Clicking on the World: Documentary Representation and Interactivity. In K. Nash, C. Hight, & C. Summerhayes (Eds.), *New Documentary Ecologies* (pp. 50–66). Houndmills: Palgrave Macmillan.

Nash, K. (2014). What is interactivity for? The social dimension of web-documentary participation. *Continuum*, 28:3, 383-395. https://doi.org/10.1080/10304312.2014.893995

Nash, K. (2012). Modes of interactivity: analysing the webdoc. *Media, Culture & Society*, 34(2), 195–210. https://doi.org/10.1177/0163443711430758

Nelson, T. (1982). *Literary Machines*. Sausalito, California: Mindful Press.

Ortuño Mengual, P. & Villaplana-Ruiz, V. (2017). Activismo Transmedia. Narrativas de participación para el cambio social. Entre la comunicación creativa y el media art [Transmedia activism. Participatory narratives for social change. Between creative communication and media art]. *Obra Digital: Revista de Comunicación*, 12, 123–144. https://raco.cat/index.php/ObraDigital/article/view/328402

Palacios, M. & Díaz-Noci, J. (2007). *Online journalism: research methods. A multidisciplinary approach in comparative perspective*. Bilbao: Universidad del País Vasco.

Pavlik, J. (2001). *Journalism and new media*. New York: Columbia University Press.

Piaggio, L. (2017). *La producción y la difusión de Webdocs. Dentro y fuera de los Medios* [Production and dissemination of Webdocs. In and outside media]. https://youtu.be/q3U_B8S5XUo

Rose, M. (2011). Four categories of collaborative documentary. *Collabdocs. Where Documentary Meets Networked Culture*. https://collabdocs.wordpress.com/2011/11/30/four-categories-of-collaborative-documentary

Ryan, M. L. & Thon, J. N. (2014). *Storyworlds across Media: Toward a Media-Conscious Narratology*. Lincoln, USA: University of Nebraska Press.

Schultz, T. (1999). Interactive Options in Online Journalism: a Content Analysis of 100 U.S. Newspapers. *Journal of Computer-Mediated Communication*, 5(1). https://doi.org/10.1111/j.1083-6101.1999.tb00331.x

Soler-adillon, J. (2017). Three Key Concepts in Digital Media. *Hipertext.net*, 15, 4–6. https://raco.cat/index.php/Hipertext/article/view/329766

Sora, C. (2018). Expanding Temporal and Participative Digital Horizons Through Web Documentaries for Social Change. In V. Shafer (Ed.), *Temps et temporalités du Web* (pp. 249–262). Nanterre: Presses universitaires de Paris-Nanterre. https://doi.org/10.4000/books.pupo.6193

Sora, C. (2016a). *Temporalidades digitales. Análisis del tiempo en los new media y las narrativas interactivas* [Digital temporalities. Analysis of time in new media and interactive media]. Barcelona: UOC Press.

Sora, C.s (2016b). *Radiografía del webdoc a Catalunya. Obres i autors* [Webdoc in Catalonia. Works and authors]. Universitat Pompeu Fabra. https://hdl.handle.net/10230/27483

Sora, C. (2015). Etapas, factores de transformación y modelo de análisis del nuevo audiovisual interactivo online [Stages, conversion factors and a model for the analysis of the new interactive online audiovisual products]. *El Profesional de La Información*, 24(4), 424–431. https://doi.org/http://doi.org/10.3145/epi.2015.jul.09

Stewart, M. (2019). Sarcellopolis: virtual cartographies of multicultural living in contemporary i-docs. *Studies in Documentary Film*, 13(1), 53–72. https://doi.org/10.1080/17503280.2018.1538404

Vázquez-Herrero, J. & Moreno, G. (2017). Documental interactivo iberoamericano: proximidad y transformación social [Interactive ibero-american documentary: proximity and social transformation]. *DOC On-Line: Revista Digital de Cinema Documentário*, Special Is, 109–130. http://doc.ubi.pt/ojs/index.php/doc/article/view/312

Walsh, R. (2011). Emergent Narrative in Interactive Media. *Narrative*, 19(1), 72–85. https://doi.org/10.1353/nar.2011.0006

Matilde Obradors
Universitat Pompeu Fabra
matilde.obradors@upf.edu

The expanded intimacy: Home movies transcend the media. *Dad's films*, an *art-practice-as-research* project

https://doi.org/10.3145/indocs.2022.7

Obradors, Matilde (2022). The expanded intimacy: Home movies transcend the media. *Dad's films*, an *art-practice-as-research* project. In: Freixa, Pere; Codina, Lluís; Pérez-Montoro, Mario; Guallar, Javier (ed.). *Visualisations and narratives in digital media. Methods and current trends*, (pp. 101-115). Barcelona: DigiDoc-EPI. https://doi.org/10.3145/indocs.2022.7

Abstract

For years home movies have not been the object of academic study, since they have been seen as a minor cinema. Little could we imagine that family cinema would leave the private space to become part of the mass media. Expanded intimacy. Recovering family films is an increasingly frequent exercise, in order to give it artistic projection and new cinematographic representations. Likewise, Unesco recommends safeguarding and preserving amateur audiovisual records and family films as cultural heritage that is part of our history and testimony of an era and its customs. This chapter deals with the topic of home movies and makes a case study: Dad's films. Places in memory, applying the art practice as research methodology. And finally, it opens the way for further research that addresses the difference between family movies of the 60-80s of the last century and those of today.

Keywords

Home movies, amateur cinema, cinema, new audiovisual formats, Social Media Live Streaming, (SMLS), audiovisual heritage, memory, private-public, performance art, art practice as research.

Título

La intimidad expandida. El cine familiar trasciende a los medios de comunicación. *Dad's films*, un proyecto de *art practice as research*

Resumen

Durante años las películas caseras no han sido objeto de estudio académico, puesto que se han visto como un cine menor. Poco podíamos imaginar que el cine familiar saliera del espacio privado para formar parte de los medios de comunicación de masas. La intimidad expandida. Recuperar películas familiares es un ejercicio cada vez más frecuente, a fin de darles proyección artística y nuevas representaciones cinematográficas. Asimismo, la Unesco recomienda salvaguardar y conservar registros audiovisuales amateurs y películas familiares como patrimonio cultural que forma parte de nuestra historia y es testimonio de una época y sus costumbres. El presente capítulo trata el tema del cine familiar (casero o doméstico), y hace un estudio de caso: Dad's films. Lugares en el recuerdo, *aplicando la metodología* art practice as research*. Y, finalmente, abre el camino para una posterior investigación que aborde la diferencia entre las películas familiares de los años 60-80 del siglo pasado y las actuales.*

Palabras clave

Cine familiar, cine amateur, cine, nuevos formatos audiovisuales, Social Media Live Streaming, (SMLS), patrimonio audiovisual, memoria, público-privado, performance art, art practice as research.

1. Origin and development of home movie formats

In 1909, 35 mm film was finally recognized as the international standard. Years later, substandard formats intended for home use began to appear: In 1922, Pathé Frères launched the 9.5 mm film format; in 1923, Eastman Kodak introduced the 16 mm format as an alternative to 35 mm; and in 1932, Standard-8, with a width of 8 mm, was created to become the support for home movies par excellence. It started with the 16 mm film and 7.5 meters long, in such a way that it was filmed on half of the film and once the roll was finished, it was turned around and the other half was filmed. The roll was then taken to the laboratory and after development the film was cut lengthwise to obtain a coil 8 mm wide and 15 m long. This format allowed laboratories to use the same machinery as they had been using for the 16 mm format, thus cutting costs and making it more competitive than other formats.

In the early 1960s, as an evolution of 8 mm film, the Eastman Kodak company developed the Super-8 format. No longer derived from 16 mm film, the perforations of this stock are smaller and located in the centre of the frame.

In the early 1980s, the first video cameras appeared, a system of recording images that was to change the landscape of home movies. In 1985, Sony introduced Video8 to replace its own Betamax tapes and JVC's VHS tapes. In 1989, the Sony Hi8 (8 mm bandwidth) was released with PCM digital audio stream and an enhanced resolution from the 200 lines of Video8 to 420 lines (560 x 480 pixels in digital terms). Some 120-minute Hi8 video tapes could also be used in Digi-tal8 cameras, but at 60 minutes as the level of quality in digital storage was higher.

Today, although some of these formats continue to be used along with home cameras that offer great image quality, most home movies are recorded on smartphones, as indeed are some films projected in cinemas.

2. Characteristics of home movies

Image 1. Frames from the home movie shot by my father in Super-8 format. *La barraqueta.* 1967

Home movies are a close relative of the family photo album. When the first home cameras appeared, the desire to immortalize the family and to tell its story in first person became a real possibility. They were 16 mm, 9.5 mm or Super-8 films, shot by a member of the family who captured the rest of the family and their lives – births, baptisms, first communions, birthday parties, holidays, etc., key moments in the life of the family, their rites and rituals.

Guerín (1998) likens home movies to a "garden", a paradise, a place free of crises and routines. It shows the joyous, festive side of life, a life of comfort and happiness. When we watch these films, it seems that time stands still; they have the "quality of 'embalming time', the beautiful epithet applied by Bazin."

Home movies are immediately recognisable with their blurred, out-of-focus images; brusque, sweeping camera movements; shaky camera and incorrect framing; sudden, visible cuts; long panoramas and abuse of zoom; changes in the lighting or lack of light; children colliding with the camera and adults crossing in front of it; no speaking, just music. Repetitions are given to emphasize its beauty. It would seem that the fact of tiring the viewer is not important, but that the person who films, gives priority to what seems beautiful to him, showing it over and over again. The roles played by the characters are treated from an entirely affective point of view and this prevails over any concerns related to frame or filming. Objects are used to identify or validate a place, just as tourist postcards do, stereotypes of the place where the action is filmed; the same goes for maps of the world and globes.

We could say that home movies have an aesthetic quality that has little in common with that of the cinema itself. They are made to be watched by the family and discussed in the intimacy of the home. It is a cinema that has undergone little to no editing. They are simply films that were shot, sent to be developed, and then projected at home. They have no coherent narrative structure.

> *Home cinema has no need to produce a narrative structure or a coherent construction, because these already exist in the memory of the participants. All that is required of the film is to rekindle memories, to allow families to relive past events together. (Odin, 2010, p. 45)*

It is worth drawing a distinction here between amateur cinema and home movies. Amateur cinema, although devoid of any commercial interests, seeks to resemble professional cinema, while home movies have no such pretensions. All the traits I have identified above show it to be a cinema full of defects, that does not meet the cinematographic standards and therefore, we would be talking about a "badly made" cinema. However, as Cuevas Álvarez (2018) recognises, "it is these very 'defects', its formal characteristics, that give it that air of authenticity, of non-commercial cinema, that provides that sensation of truth that no other cinema has achieved" (p. 130).

Home movies underwent a transformation in the 1980s with the popularization of video, a medium that overcame the limitations of celluloid, the high costs of development and which

allowed a film to be viewed immediately. This marked the beginning of the boom in homemade images, and tapes that could be reused. However, despite these changes, the specific qualities of home movies were maintained, which according to James M. Moran (2002), allows us to speak of the unique identity of the "domestic mode", which extends beyond the format.

To conclude this section, today home movies constitutes part of our cultural heritage. UNESCO has designated October 27 as World Day for Audiovisual Heritage, which also takes into consideration the work of those who archive and protect this world heritage. Moreover, it proposes investigating and identifying amateur audiovisual recordings and family films, as well as raising awareness among the owners of these films of the need to adopt basic preventive conservation practices. In a similar vein, Keldjian (2015), in the "Cine casero" project, emphasises an additional component to that of conservation, wich is projection, which she states "favours the sharing of these images and their circulation..." (p. 25).

Image 2. Frame from the home movie shot by my father in Super-8 format. *New York*. 1969

3. Uses, recovery, revision, recycling and new representations. Home movies transcend the media

Recovering home movies has become an increasingly frequent exercise, with the objective of giving them artistic projection and new representations in contemporary cinema, both in documentaries and in more experimental spheres. In this appropriation of domestic cinema, several paths have been taken: 1) filmmakers and artists that use home movie techniques when shooting their own works; and, 2) the recovery of home movies material (the so-called *found footage*). The latter is a narrative technique used by artists in their productions and is

frequent in horror movies and in *mockumentaries*. Here, we find two different modalities: a) films in which the home movie is fictitious but is presented as if it were found footage; and, b) films, works of art and video art that use home movies as the basis of the piece. In both cases we are dealing with a genre that draws inspiration from the literary technique of the found manuscript. Within modality b), a distinction should be drawn between filmmakers who recycle other people's domestic material and those who recycle their own family material, so that their work is autobiographical in nature. This is the case of the project I describe in this chapter.

Finding abandoned rolls of film or footage is a kind of fetishism, a desire to discover stories or great hidden truths – a search for our origins. Home movies handle time in a much more intimate fashion, lifting family members out of the stream of time and rescuing them from death. An example of just that is the feature film Ulysses' Gaze (1995) by the Greek director, Theo Angelopoulos. The film, set in the Bosnian war, tells the story of a Greek filmmaker who returns to his hometown on a personal journey in search of three undeveloped reels of film, containing the innocent gaze, the first gaze, of the pioneers of Greek cinema. As such, it is also a journey to the roots of cinema.

If I may draw a parallelism, family movies may also be a journey into oneself. Phototherapy uses family photos to evoke therapeutically relevant memories, feelings and information that are unconsciously contained in the images, which help patients understand the place they occupy in their family. With more reason the movies, due to their quality of moving images, can project scenes that may shed even more light on such questions. "Photo therapy allows subjects to connect with parts of themselves in ways that words alone cannot fully represent or deconstruct" (Weiser, 2010, p. 3). Here we can include memory recovery work in patients with senile dementia or Alzheimer's; in which case, if you have your own family films, these can provide the patient, especially in the initial stages, the possibility of remembering and exercising their biography with the objective of recovering his identity1

A further example of a film that plays with found footage, in this case a fictitious finding, is Daniel Myrick's Blair Witch Project (1999). On October 21, 1994, Heather Donahue, Joshua Leonard and Michael C. Williams entered the Burkittsville woods in Maryland to shoot a documentary about a local legend, the "The Blair Witch". They were never heard of again. A year later, the camera they used for filming was found and the tape revealed the terrifying events that led to their disappearance.

1 "*The Narrator-Narrated: The Participatory Social Documentary and the Expanded Documentary as possible tools for the representation of identity in people affected by memory loss*", is the title of the doctoral thesis by PhD student Virginia Fuentes (PhD Program in Communication, University Pompeu Fabra, 2021–2023) of which I am the director. The study points to the possibility of using patients' family films, where they exist, as material for cognitive stimulation in their search for identity and the improvement of their memory.

Tren de sombras (Train of Shadows) (1997), by José Luis Guerín, is a film of fiction and yet it could be classified as non-fiction, experimental cinema or an *essay film*. It is a tribute to home movies. What Guerín does is reconstruct home footage, invent the characters and record them with the characteristics of family cinema. Yet, according to Cuevas Álvarez (2018), "This invention does not have a strictly narrative purpose, but rather seeks to place domestic cinema as a paradigmatic mode where the most essential features of cinema are brought together, at the moment when it has become centenary" (p.137). In 2003, I interviewed the director of photography Tomás Pladevall for my doctoral thesis. I had dedicated a section to family cinema and I showed him some of my family's films and we also talked about the characteristics of home movies and his cinematography for the film, *Train of Shadows*. Next, I will expose some contents of the Pladevall decalogue (Obradors, 2003) for said feature film according to the precepts of domestic cinema. Pladevall, classifies different sections: Handling, Framing, Panoramic, Actions, Presence of the filmmaker.

As characteristic features we can highlight:

> *"Improvisation even within the same scene. Artistic close-ups or reframers. Often medium shots with plenty of space above. Some ankle shots. Fixed camera and the characters entering or leaving the frame. An empty frame held for some time, waiting for an entry or an action that finally occurs. A hand, arm or something that comes between the camera and the characters. Some blurring in some shots. Panning at erratic speeds. Irregular camera pans following a motion. Panning from characters (coming out of frame) to landscapes. Groups of characters that do not fit in the frame and the camera hesitates. Some 360-degree panoramas. Some camera sweeps. A mother who turns the head of a child towards the camera or places it in the frame. Characters looking at the camera. Reflections in the windows".*

In the line of true found footage, we can highlight the feature film *My mexica bretzel*, by Nuria Giménez Lorang (2019) that recovers Super 8 and 16 mm films from her grandfather and with that material builds a fictional story. Gimenez Lorang, elaborates the intimate diary of a woman, Vivian Barret, who reflects on life and her own existence after the Second World War and illustrates it with the home recordings of her husband León, a rich industrialist in the 40s and 60s of the century. last. A journey through the best of Europe, from the welfare of wealthy families. Lorang enters us, without speech or dialogue, into a world that we do not know if it is real or fictitious.

Another case of domestic cinema that has come to light, thanks to the documentary directed by José Luis López-Linares with a script by Javier Rioyo, and which can be considered the first autobiography in images, is that of Madronita Andreu (daughter of Doctor Andreu). Madronita spent her life filming not only her family, but also her travels (in the United States, Africa, India, Jamaica, Italy, and Switzerland, among others). She started out making home cinema but over time ended up becoming an amateur filmmaker, displaying a natural talent for filming and a sure composition. Indeed, her films constitute a veritable historical legacy of the 20th century. The film, entitled *Un instante en la vida ajena* (A moment in the life of others)

(2003), took the director seven years to complete, between viewing and editing the 150 hours of film shot by Madronita and today housed in Barcelona's *Filmoteca*.

In relation to artists who use home movies for their works, it would be worth high-lighting artists such as María Ruido (1967) who uses archive material (or found foot-age) combining it with images of her own production that have resulted in works on the construction of memory, body and identity. Carlos TMori (1970) works in non-regulated audiovisual, researches on the domestic, and identity and its distortions. Peter Tscherkassky (1958), works exclusively with found footage. His cinema is experimental, he develops 35mm films and intervenes the material to question the conventions of narrative cinema.

The intention here is not to make a complete inventory of these filmic practices– this is not the objective of this chapter– rather I have noted a few examples to illustrate the different uses of home movies in documentary, film and artistic practices[2]

4. *Dad's films. Places in memory*. An *art-practice-as-research* project

Motivated by the interest and fascination for family cinema, in 2018 I plan to develop a project to explore this subject, using the art-practice-as-research methodology[3]. From the hegemonic investigation, I would have studied family films, solely from the analysis of their narrative, whether formal, from semiotics, hermeneutics or other disciplines or methods. In this case, in addition to studying the expression, the message, the transcendence and the purpose, I have developed two artistic actions to approach the subject from experience. I designed a study from the art practice as research methodology that included two actions: a performance and a participatory event. In undertaking a performative act, I was accepting that this was an approach that would enable me seek out other codes, not necessarily those of the word, as a platform for addressing my object of study. It would be like letting my mind be guided by uncoded signals – "Can we put aside the certainties for a while and listen and hear about those things about which we know nothing?" Taylor, (2017, pp. 23-24) – trying to interpret meaning in terms of rites. According to Diana Taylor, performance teaches us to envision knowledge in other ways. Thus, I proposed a study of family cinema through representation.

Years ago, I had transferred some of my father's films to DVD, so I started the project because of my family's cinema.The sample is made up of five films: La Barraqueta, New York, Atlanta, Mexico and Peru, and Llafranc. I visualized the sample, and looked for the variables and meanings common to all of them. The recurring variables that I found are the following: sea, living on the beach, exile, nomadism, travel.

2 In this regard, see Efrén Cuevas Álvarez, (ed.) (2010).
3 On *art-practice-as-research methodology*, see, Matilde Obradors (2020).

Image 3. Workflow event </CODE±NOWNESS>, Livemedia SMLS Fabra i Coats project *Dad's films. Places in memory*. Matilde Obradors.

In 2019, within the framework of the studio/event, workflow set installation, Social Media Live Streaming (SMLS), </CODE±NOWNESS>, carried out by Livemedia at the Fabra i Coats creation factory (Barcelona), I developed a project entitled, Dad's films. Places in memory.

I designed a workflow to examine the theme of my family's home movies that consisted of 1) a performance on the beach, broadcast on social media live streaming (SMLS); and, 2) a presentation and projection of the project in the Fabra i Coats creations factory's multipurpose room, live before a public and retransmitted by Social Media Live Streaming (SMLS)[4]. For both actions, I previously prepared a script.

Image 4. Shot of streamed performance art. Matilde Obradors, 2019.

4 The performance on the beach, as well as the event at Fabra i Coats, were broadcast via Periscope, Twitter live and the Livemedia platform.

1. Performance broadcast via social media live streaming (SMLS).

Performance streamed from one of the places in which my father's films were frequently set, the beach, and a recreation of a campsite, the trip and the elements that have most caught my attention ever since I was a child – food and family photographs. A representation of the past, to talk about the family, emulating my father's home movies. The representation, of the representation of the representation. I become the voice of the places in memory.

I simulated what was a small campsite, evoking a sense of nomadism, with a camping table, on which I laid out all the foods I associate with my grandmother, salt-cured tuna, fish roe, round red *ñora* peppers, etc. Sitting at the table, I replicated my grandmother's life, I ate and talked about her. I understood the bond that united us. I would have liked to have heard her voice. She lived half the year on the beach in a cart. My mother's brother, exiled during the dictatorship in Perpignan, also ended up spending part of the year on the beach at Saint-Cyprien, in a wooden hut built on the sand, with a fence and a garden, the water collected with a hand pump. They called it the *barraqueta*. There, every summer, Spanish exiles in France and Mexico used to meet up to visit my aunts and uncles. At night, they would sit in the sand garden, telling stories about the civil war. All this is captured in the film, *La barraqueta*, filmed by my father. Thus, it is discovered how families, although they separate from the original family nucleus, continue to repeat and reproduce models.

Image 5. Frame from the home movie shot by my father in Super-8 format. *La barraqueta*. 1967.

2. Presentation in the Fabra i Coats creation factory

Presentation in the multipurpose room of the Fabra i Coats creation factory on three giant screens, where images of the beach performance described in point one and fragments of my father's films were projected. This event was open to the public and was also broadcast on Social Media Live Streaming (SMLS).

The public was seated as if in a cinema and I, to the right of the three screens, explained where the idea for the project had come from, the characteristics of family home films in general and those of my Dad's films in particular, and opened the debate on how they were recorded family movies were made in the '60s and '70s, and how they are made today. At that time recording was expensive and you had to wait for it to be developed. My father used to say that he "recorded without haste", he immersed himself in a kind of ritual, with the awareness of perpetuating a historical moment.

I explored this memory, the analogue montage of films my father made with his Moviola. My child's eyes fascinated by family scenes suspended in no time. Dazzled by the films of my father's travels in the Americas. Especially those to New York and Atlanta (1964). I also discussed with the public the meaning of the live streaming of this intimacy, of these places in memory that I had decided to visit. And we spoke of just how little home filmmakers could have imagined that their films would be of interest to people outside the family circle.

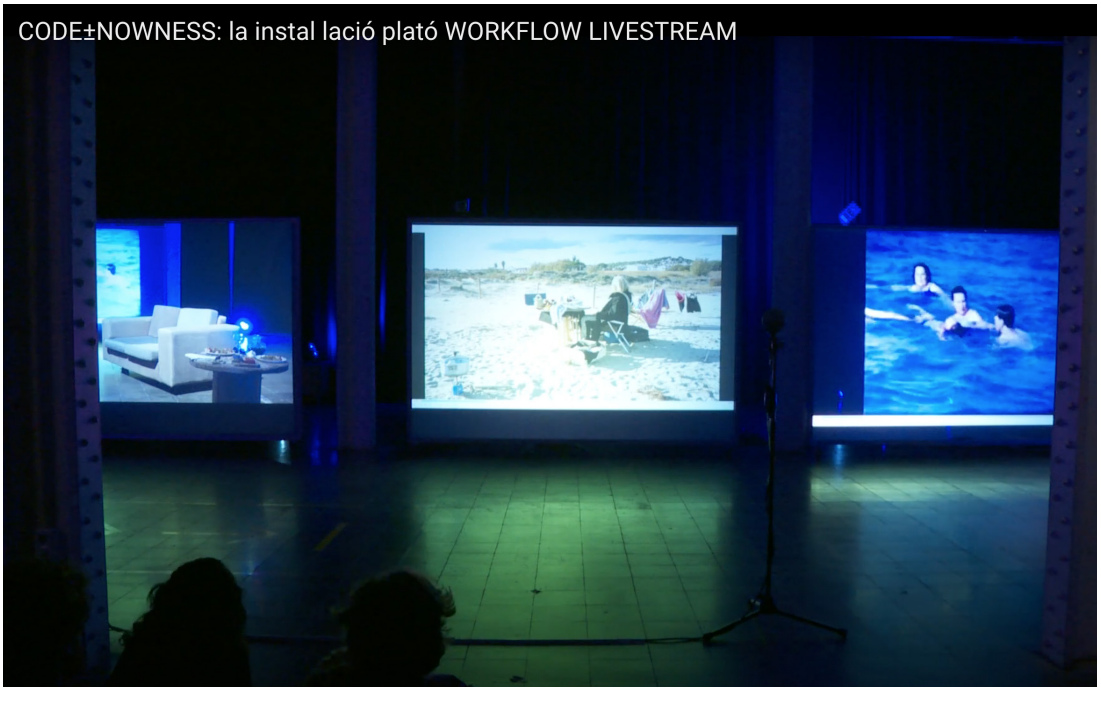


Image 6. Projection in the multipurpose room of the Fabra i Coats creation factory.

Image 7. Presentation and colloquium in the multipurpose room of the Fabra i Coats creation factory.

5. Conclusions. Expanded Intimacy

Family cinema (shot between the 1950s and the 1980s), because of its domestic condition, prioritized the affective over the professional. It is a cinema that has very specific characteristics, traits that it continued to preserve even with the appearance of video in the eighties. The fact that all home movies tend to resemble each other means that, starting from the variables of a movie camera and a dearth of cinematographic knowledge, the outcome is always the same. What should be stressed however is that, by dint of practice, some of these home movie filmmakers learned how to make cinema and their films improved.

Among the characteristics of home movies described in this chapter, variations have been detected that are a function of the degree of amateurism shown by the filmmaker. In the case of the films analysed in the sample for this study, it is evident that the filmmaker (my father), does edit his films on his Moviola, a clear demonstration of an amateur's trait, that is, that he wishes to make good cinema and is no mere *aficionado*. What is more, he tries to make sure not to bore his viewers, so he seems to be seeking to reach an audience wider than that of his immediate family. He is aware that he is making a social historical document and assumes the responsibility that goes with that task. In other words, we are not only talking about a playful, eminently home-made gaze – films made for the family's own consumption – but rather there is a desire to achieve social and collective permanence. Consequently, we can speak of

degrees of amateurism within home cinema, which will vary its condition as "badly made" cinema.

One of the conclusions endorsed by specialists that is repeated when qualifying family cinema and that has been verified in the present study, is that the sensation of truthfulness is greater than in professional cinema, probably due to that air of imperfection that makes one think of spontaneity and therefore authenticity.

However, this is something of a contradiction since it is a cinema that seeks to show only the family's good side, the "garden", where everything is contentment and no one has any worries. It does not present the family's true private life. This characteristic seems to be the unspoken agreement that is repeated in all family home films of that time. It is curious, then, that family cinema gives this sensation of veracity, when it is far from being completely honest and does not even seek to be authentic.

What no one could have imagined is that "badly made" family cinema would one day break through the barriers of the private to become part of the mass media. This fascination with family cinema, as a starting point for other audiovisual works, is closely related to feelings of curiosity, comparable in that respect to the feelings generated by a found manuscript or treasure map. As well as being documents that appear to be authentic and true, they are testimonies that could reveal a new, previously hidden, truth. They are also documents that play with time, by both containing it and retaining it. They stimulate feelings of nostalgia, take people out of the current of time and defy the death of our kith and kin. And finally, they serve as testimonies of an era; they are a window on past customs, lifestyles, fashions, and more.

I am able to confirm that performance is a useful tool for reflecting on a subject as it means a complete immersion in that subject via representation, and it enabled me to contemplate other points of view and to formulate new questions.

My work has been a reflection on family cinema, one that allowed me to reflect also about myself, my family and the place I occupy in it. I have found answers but, at the same time, it has found many new questions. I have drawn conclusions that, due to their personal nature, I have not thought fit to capture in this chapter. I have verified that watch own home movies helps cognitively reintegrate insights and can improve personal aspects. Therefore, it is highlighted that family cinema not only has recycling uses for films, documentaries or artistic expression, but also that it is a great tool for therapy. And it shows how convenient it is to combine psychology and communication to advance self-awareness. Going through the images of the past and returning to the origins was Freud's maxim, -based on the history of religions and cosmogony-, to achieve healing.

A follow-up of the investigation is proposed regarding how family films were recorded from the 50s to the 80s of the last century and how they are recorded now. From the outset as a

starting point, we could analyze a difference: in those years, the best face of the family was presented, "the garden" to which Guerín refers, in which there are no worries or conflicts. In current family cinema, even when the happy part of families is also reflected, there is no fear in showing the most conflictive part, ridiculing family members or even showing arguments. Under the influence of programs like Big Brother or the reality shows in which anything goes, crude ways of showing intimate and private spaces emerge. And even, we could say more, the conflict is provoked and promoted to attract the viewer's attention. All these modes of voyeurism have permeated the domestic practices that circulate through the networks. Above all, taking into account that videos are currently being made showing one's own house, House Tour, or showing the products in a box that has been received, YouTube unboxing, or videos about what do I carry in my bag? Intimacy comes out of its private environment to be shown without shame.

A paradox emerges here that is worth reflecting on: we can see that fashion is what we all follow to differentiate ourselves from the rest. Moving images take on and perpetrate invisible collective forms that help us feel unique. On these invisible and common forms, our differentiation is based, the search for our own identity. As I have already observed in previous research, the norms of the mainstream, its stereotyped reiterations, sneak into all the practices of users in the digital age, like invisible canons. In future studies, it will be possible to delve deeper into this aspect.

Finally, there is one point about what we agree on studying and verifying different researchers, and that is that many of the audiovisual practices that exist today and wich are attributed to the digital age and the social networks, have their bases in practices invented and rooted in previous centuries. The study of audiovisual hybridizations (Obradors, 2021), reveals that these practices date from the 1920s and with the appearance of video and artistic manifestations in the 1960s, they are the precursors of audiovisual transdisciplinarity. In the same way, it is worth mentioning the work of Freixa and Redondo-Arolas, (2021) in which it has been possible to verify that,

> *Many of the practices that are now recognized as belonging to hyperconnected photography are based on learning initiated during the last decades of the 19th century through postcards, visiting letters, and the social circulation that they provoked. (p.135)*

References

Angelopoulos, T. (Director). (1995). *La mirada de Ulises* [película]. Paradis Films. Basic Cinematografica. Greek Film Centre. Theo Angelopoulos Film Production.

Code±Nowness, la instalación plató Workflow Livestream. (s.f.). Livemedia https://www.codenowness.online/

Cuevas-Álvarez, E. (2018). Sombras reflejadas. Apropiaciones del cine doméstico español en el cine documental. En: Vicente, Pedro; Gómez-Isla, José (ed.). *Álbum de familia y prácticas artísticas. Relecturas*

sobre autobiografía, intimidad y archivo (pp. 129-137). Huesca: Diputación General de Huesca.

Cuevas-Álvarez, E. (ed.) (2010). *La casa abierta. El cine doméstico y sus reciclajes contemporáneos.* Madrid: Ocho y Medio.

Fabra i Coats (s.f.) *Fabra i Coats web*. https://www.barcelona.cat/fabraicoats/

Freixa, P. & Redondo-Arolas, M. (2021). Fotografiar, compartir, circular: el tramado invisible de la imagen fotográfica. En: Renó, Denis; Lopes Cardoso, Fátima; Freixa, Pere. *Poéticas fotográficas*. Aveiro: Ria editorial, pp. 118-41.

Giménez-Lorang, N. (Directora). (2019). *My Mexica Bretzel* [película]. Bretzel & Tequila Film Productions, Avalon P.C.

Manrique-Sabogal, W. (January 1th, 1998). Guerin estrena un filme mudo donde el espectador debe imaginar la historia. Homenaje del director catalán al cine. *El País.com.* https://elpais.com/diario/1998/01/16/cultura/884905212_850215.html

Guerín, J. L. (Director). (1997). *Tren de sombras* [Película]. Institut del Cinema Català.

Keldjian, J. (2015). Las proyecciones de Cine Casero desde la perspectiva semio-pragmática. *Dixit*, julio-diciembre, (23), 16-25. https://revistas.ucu.edu.uy/index.php/revistadixit/issue/view/128/256

López-Linares, J. L. & Rioyo, J. (Directores). (2003) *Un instante en la vida ajena* [Película]. Cero en conducta S.L, Madrid Film S.A, Telson.

Moran, J. M. (2002). *There's No Place Like Home Video*. Minneapolis: University of Minnesota Press.

Moya-Jorge, T. (July 3th, 2013). Tren de sombras desde la perspectiva de Andrè Bazin. *Entretanto Magazine.* https://entretantomagazine.com/2013/07/03/tren-de-sombras-desde-la-perspectiva-de-andre-bazin/

Myrick, D. & Sánchez, E. (Director). (1999). *Blair witch Project* [Película]. Haxan Films.

Obradors, M. (2021). Hybridizations and overflows between disciplines and sectors in art and communication. Tracing the transdisciplinary creative potential for teaching. *Icono 14*, 19(2), 212-234. https://icono14.net/ojs/index.php/icono14/article/view/1708/1880

Obradors, M. (2020). *Practice as research: la investigación a través de la práctica en las artes creativas: enfoques y métodos.* Barcelona: Universitat Pompeu Fabra. Departament de Comunicació. https://hdl.handle.net/10230/44877

Obradors, M. (2019). Dad's Films. Lugares en el recuerdo. *Code±Nowness*. http://www.codenowness.online/llocs.html

Obradors, M. (2007). *Creatividad y generación de ideas. Estudio de la práctica creativa en el cine.* Barcelona: Aldea Global.

Obradors, M. (2003). *Creatividad y generación de ideas. Estudio de la práctica creativa en el contexto cinematográfico* [Tesis doctoral, Universitat Pompeu Fabra].

Odin, R. (2010). El cine doméstico en la institución familiar. In: Cuevas-Álvarez, Efrén. *La casa abierta. El cine doméstico y sus reciclajes contemporáneos* (pp. 39-60). Madrid: Ocho y Medio.

Onieva, C. (November 6, 2014). Madronita Andreu, La primera española que manejó una cámara de cine. *Cincuentopía.* https://cincuentopia.com/madronita-andreu-primera-espanola-con-una-camara-de-cine/

Taylor, D. (2017). ¡Presente! La política de la presencia. *Investigación Teatral*, 8(12), 11-34. https://investigacionteatral.uv.mx/index.php/investigacionteatral/article/view/2550/4432

Weiser, J. (2010). *PhotoTherapy Techniques: Exploring the Secrets of Personal Snapshots and Family Albums Paperback*. Vancouver: PhotoTherapy Centre Press. https://www.lianalowenstein.com/Weiser_PhotoTherapy_article.pdf

Juan-José Boté-Vericad
Universitat de Barcelona
juanjo.botev@ub.edu

Mari Vállez
Universitat de Barcelona
marivallez@ub.edu

Image and video manipulation: The generation of *deepfakes*

https://doi.org/10.3145/indocs.2022.8

Boté-Vericad, Juan-José; Vállez, Mari (2022). Image and video manipulation: The generation of *deepfakes*. In: Freixa, Pere; Codina, Lluís; Pérez-Montoro, Mario; Guallar, Javier (ed.). *Visualisations and narratives in digital media. Methods and current trends*, (pp. 116-127). Barcelona: DigiDoc-EPI. https://doi.org/10.3145/indocs.2022.8

Abstract

The growth of fake news is nowadays a growing reality, leading to users feeling insecure when consuming information. Text-based news is perhaps the most manipulated, but the video is rapidly gaining ground. Technology allows content to be easily adulterated, leading users to become misinformed. In this study, we explain the different image, video, and audio manipulation techniques that are carried out with software and advanced audiovisual techniques. The most prominent methods are morphing and warping, together with machine learning techniques. Also, artificial intelligence is a critical element in the growth of fake video generation. This brings with it the need to identify the fake videos to verify the facts being told. It is also essential to understand social tolerance, especially in humour programmes, as some media outlets use these mechanisms. Moreover, it is analysed the importance of privacy policies that affect how the users' personal information is collected and their implications in their fundamental rights. Finally, it is concluded with the need for media and digital literacy campaigns in the education system to minimise this problem.

Keywords

Deepfake, disinformation, video manipulation, apps, mobile applications.

Título

La manipulación de vídeos e imágenes y la generación de *deepfakes*

Resumen

El crecimiento de las noticias falsas es hoy en día una realidad creciente, que lleva a los usuarios a sentirse inseguros cuando consumen información. Las noticias basadas en texto son quizás las más manipuladas, pero el vídeo está ganando terreno rápidamente. La tecnología permite adulterar fácilmente los contenidos, lo que lleva a los usuarios a estar mal informados. En este estudio se explican las diferentes técnicas de manipulación de imagen, vídeo y audio que se llevan a cabo con software y técnicas audiovisuales avanzadas. Los métodos más destacados son el morphing *y el* warping*, junto con las técnicas de aprendizaje automático. Además, la inteligencia artificial es un elemento crítico en el crecimiento de la generación de vídeos falsos. Esto conlleva la necesidad de identificar los vídeos falsos para verificar los hechos que se cuentan. También es esencial entender la tolerancia social, especialmente en los programas de humor, ya que algunos medios de comunicación utilizan estos mecanismos. Además, se analiza la importancia de las políticas de privacidad que afectan a la recogida de información personal de los usuarios y sus implicaciones en sus derechos fundamentales. Finalmente, se concluye con la necesidad de realizar campañas de alfabetización mediática y digital en el sistema educativo para minimizar este problema.*

Palabras clave

Deepfake, desinformación, manipulación de vídeos, apps, aplicaciones móviles.

1. Introduction

The generation of fake news, in text, image and video formats, seems to increase daily, reaching such levels that citizens can, at times, face difficulties in being correctly informed about what is going on in the world. Relevant examples include the transmission via instant messaging applications of hoax information related to the appearance of COVID-19 (Atehortua and Patino, 2021) and the generation of disinformation from Macedonia about the 2016 US presidential elections (Hughes and Waismel-Manor, 2021).

Fake news is mostly textual; however, images or videos can also be manipulated. These *deepfakes*, as they have come to be known, can be defined as the generation of deliberately misleading or false audiovisual content by manipulating images, sound and video. They are further characterised by the fact that this content is decontextualized in terms of time, form and place.

Both fake news and deepfakes are created with a similar goal in mind: the deception of the user. Yet, because it is clearly much easier for the user to process visual information than other types of information, manipulated images have the potential to be much more viral. Indeed, the impact that a manipulated and decontextualized image or video might have is likely to be much greater than that of a text. For example, they can cause a journalist to commit an error when reporting a news story, which means reporters must be constantly aware of this danger and always corroborate carefully their sources of information (Joseph, 2019).

Deepfakes integrate different types of content (image, video, and audio) which when woven together allow fake videos to be created. These deepfakes aim to deceive the user by manipulating, for example, a character or by using techniques of animation. In this way, customisations can be created by exploiting audiovisual technology at a range of different levels, including, for example, the face, the voice, the lip movement and the head and body posture. Although there are mobile applications that allow deepfakes to be generated very cheaply, the creation of professional videos – those that really deceive the public – is costly since it requires the participation of experts from different fields, including linguists, video editors, and animators, among others. For this reason, the production of deepfakes tends to be associated with much more elaborate objectives than those of simple satire or jokes. In an extreme case, a video created using these techniques could lead to a candidate standing for and winning an election.

In the news media, there are mechanisms that can help identify manipulated information. In the case of fake news, verification filters are used for fact-checking. In the case of deepfakes, biometric analysis or blockchain technology is used (Hasan and Salah, 2019); however, it should be noted that their application is costly in the extreme.

For technical reasons, it is much less difficult to modify a photograph than a video. The image is static and lacks elements that form part of a person's physiognomy or anatomy, such as their voice or movement. In the case of the manipulation of video, additional difficulties have to be overcome, including those of resolution, digital format and its runtime. In the case of image manipulation, different techniques have traditionally been used, perhaps the best known being the *morphing* technique that emerged in the 70s as applied in aeronautical applications. This involves modifying an image by means of metamorphosis via which image A can be transformed into image B and vice versa (Ivakhiv, 2016). This technique allows, for example, one face to be changed for another, integrating the face of one person onto another, or for caricatures to be created by exaggerating facial features.

In the video world, artificial intelligence has originated more advanced technology so that it is possible not only to change a person's physiognomy but also to integrate other characteristics, including the voice and lip movement. However, this technique still runs into a number of technical barriers as is evident in the video created by *Future Advocacy and UK Artist Bill Posters* in which Boris Johnson and Jeremy Corbyn endorse each other's candidacies to be UK Prime Minister (BBC News, 2019), a video that serves as an example of how such a vital factor as democracy itself can be undermined. The use of such videos and manipulated photographs in an informative context, such as the news, could mean that they are not perceived as fake precisely because of the prestige of the news media outlet.

If we start from the basic fact that the objective of fake news – which today continues to be primarily textual – is to deceive the user, then the objective of deepfakes cannot be said to differ greatly. Indeed, the decontextualization of a video or an image can also be classified as fake news, although in this article our focus is more specifically on the artificial manipulation of video.

A video piece can incorporate images, video or audio. This set of elements when woven together allows a fake video to be created; however, even if treated separately they can constitute part of a deepfake video. Therefore, as regards their morphology, deepfakes can include false footage, images, audio as well as video. Conceptually, a deepfake seeks to deceive the user through character customisation or by means even of animation. The customisation that is achieved with audiovisual technology involves manipulation of certain characteristics of a person, including their face, voice, lip movement and even their cloning (Khodabakhsh et al., 2018). Satire, which exists in a textual format, can also be represented in an audiovisual object such as video. However, in the production of deepfakes, there is one factor whose influence is especially marked: the economic cost. Although there are apps that allow deepfakes to be generated cheaply, creating videos that genuinely deceive is expensive.

From the point of view of the end user, the generation of fake images and even fake videos is becoming easier, since it is possible to find both desktop computer programs and apps that allow a person's physiognomy to be modified, especially the face by morphing one into

another. This reduces the obstacles to the generation of false images. Likewise, computerized and more complex versions can be found that use machine-learning algorithms to create computer-generated faces (BBC News, 2018).

In the following sections, different aspects that directly or indirectly impact the implications of deepfakes are analysed. The specific domains examined are techniques of image and video manipulation, mechanisms for detecting manipulated videos, social tolerance of video manipulation, and the importance of privacy policies.

2. Techniques of image, audio and video manipulation

The principal technique used for manipulating images or videos is, as we have seen, *morphing*, consisting of the identification of patterns between two photographs and the dynamic transformation of one image into another with movement as it shifts from point A to point B (Ivakhiv, 2016; Scherhag et al., 2017). The technique has had a wide range of applications, one use being in major cinema productions, including *Indiana Jones and the Last Crusade*, which was one of the first films to use it (Puerto, 2018). In the field of psychology, the *morphing* technique is used to help in the perception of identity and different types of expression (Kramer et al., 2017). As the technology has evolved and as its results have improved, the technique has incorporated elements of 3D, for example, in facial expressions (Tang and Ni, 2019).

Another of the current techniques in use is *warping*, which allows the shape of some part of the image to be digitally modified for creative purposes, correcting possible dysfunctions or even creating distortions (Prathap et al., 2016). This technique has different applications including the generation of caricatures by way of the exaggeration of personal features. The technique is also used in the field of healthcare in radiotherapy (Veiga et al., 2015) and in photography to correct panoramic images in sports cameras (Li et al., 2015). Finally, *warping* is used for the post-production of images, and is commonly used for aesthetic enhancements of photos on social media (Islam et al., 2017; Krylov et al., 2014).

In the case of audio manipulation, different techniques are also being exploited and include, for example, voice exchange based on text-to-speech technology, which allows the audio in a recording to be changed and a new text to be rewritten (Cole, 2019; Somers et al., 2006; Wijethunga et al., 2020).

When audio and video manipulation are combined, another technique is used, that of lip-syncing, which involves modifying the movements of the speaker's mouth to match the fake words. This technique is used when a person is speaking and the camera shot focuses solely on them, in what is known as a *talking-head* video (Kietzmann et al., 2020). The technique is also used, for example, in the modeling of 3D video games (Ali et al., 2018).

Therefore, different techniques are available that facilitate total or partial changes to be made to a person's physiognomy, making it easier for deception to occur.

3. Detection of manipulated videos

Just as there are tools to detect fake news texts, tools have been developed to detect fake videos. These tools are high-tech in their specifications and have been developed at high costs. Joseph (2019) points out that the term deep fake, a portmanteau of "deep learning" and "fake", involves the application of machine learning and artificial intelligence techniques aimed at generating videos showing people saying or doing things that they have never said or done. Various mechanisms exist to detect fake videos without having to use a specific technological tool. For example, conducting an internet search to see if someone else is reporting on the content of a video, taking screen grabs and conducting reverse image searches to try to find the original or using verified trusted sources of information.

To use technology to identify fake videos, several methods are available. The use of recurrent neural networks where inconsistencies between video frames are used to detect whether or not a video has been manipulated (Güera and Delp, 2018). Big data analysis made it possible to detect that many fake videos employed faces that did not blink, and so fake video detection software was developed based on this fact (Li, Chang, and Lyu, 2018). There is also the option of performing a multimedia forensic analysis as a way to ensure the authenticity of a video and its origin (Rossler et al., 2019).

Although there is no definitive solution for detecting a fake video manipulating a person, the key is to develop tools that can go some way in helping to identify possible deepfakes (Nguyen et al., 2019). Analyses using methods based on automatic-learning and machine-learning algorithms can also provide solutions, such as image frequency analysis that helps to recognise different behaviours within a video. In short, the use of cutting-edge technologies can assist in the detection of fake videos.

4. Social tolerance of video manipulation

There are certain social domains in which modifications of a person's anatomy using different graphic techniques are tolerated and even accepted – examples include simulations of the outcomes of cosmetic surgery, as well as for the purposes of humour and satire – and where fake videos are professionally generated.

In the case of cosmetic surgery or other applications in the field of healthcare, there are various justifications for creating video simulations. Both female and male facial features can

have a bearing on physical attraction and this fact has generated considerable demand for such simulations (Foo et al., 2017). Thus, in plastic surgery, a prospective patient's expectations regarding the outcomes can be increased by the exploitation of fake before-and-after testimonies from patients that have undergone operations (Crystal et al., 2020).

In the case of videos created for the purposes of parody or satire, it should be borne in mind that in some countries they are subject to specific legislation. Copyright laws also determine whether or not these videos can be made. One well-known case was the release of a video parodying former UK Prime Minister David Cameron, showing him 'rapping' a speech to the rhythms of hip-hop at a Conservative Party Conference. Thanks to legislative changes in copyright in the UK, the video could be published on YouTube, where it attracted millions of views (Baker, 2014). It should also be noted that satire and fake news have a different narrative. While in fake news or deepfakes, the objective is to deceive, in satire, the objective is to criticise someone from a position where the full informational context is available (Das & Clark, 2019).

In the audiovisual sector, there seems to be a degree of social tolerance for fake videos used for humour. An example of this is provided by the programme "*El Intermedio*", broadcast on the Spanish television network La Sexta (La Sexta, 2019), in which the manipulation of videos is very much part and parcel of the programme. Videos are shown in which previous interventions of politicians, actors and sportsmen and women are manipulated, putting fake words into their mouths with a humorous slant on some topical news story or other.

Radio stations, too, transmit programmes of humour in which the voice of a person is imitated or changed. Applying voice synthesis technology facilitates this type of programme, but it also allows fake news to be created (Stark, 2016). However, while humour is accepted in the media, especially the radio, the protection of the basic rights of individuals is an issue that needs to be carefully reflected on (Bendel, 2019).

Other cases have also come to light in which, for example, the manipulation of videos in the tourism sector have been used to help promote a tourist destination and influence an area's economy (Kwok and Koh, 2021), although the authors of this particular study acknowledge that it is an aspect that has yet to be explored in much depth.

Thus, we identify situations in which the creation of manipulated videos has become democratized and socially accepted, despite the fact their effects can be both positive and negative. However, we should not forget that reality is being distorted with all the potential dangers that this might entail.

5. The importance of privacy policies

As discussed above, smartphone applications have been developed to generate deepfakes. Such apps are covered by privacy policies, however, that do not always fulfil their function, namely to inform customers about how the application handles and processes the personal information they provide it with. Indeed, many apps share information with third parties and users should be informed of this fact. This occurs because applications tend to use third-party software, in many cases libraries, regarding which little is known as to what personal data they collect since the libraries fail to indicate what information is saved (Balebako et al., 2014).

Moreover, these policies are not always clear and, in some cases, can put the user's individual reputation at risk and even, in some instances, impact their health (Parker et al., 2019). The authors of this last study carried out an analysis of 61 mental health apps, and found that 41% of the apps analysed did not have privacy policies to inform users about how and when their personal data were collected and shared with third parties.

Similarly, in a study conducted in India, also in the field of health, a total of 70 apps were analysed, comparing the complexity involved in the reading of privacy policies for apps linked to mental health and diabetes (Powell et al., 2018). The study concluded that the privacy policies were written for users with a university level of education and that this complexity in the interpretation could be a barrier for decision making. In another study carried out on 369 mental health apps, it was determined that the information in their privacy policies was not transparent enough for users, that they were too generic and required a level of university literacy to be understood (Robillard et al., 2019).

Another of the problems associated with apps is their incorrect documentation, which can generate a considerable lack of confidence among users. Yu et al. (2016) conducted a study of the trustworthiness of the privacy policies of mobile applications. To do this, they adopted a systematic approach employing a system they named *PPChecker* which, by means of natural language processing techniques, allowed them to dissect the apps' privacy policies, focusing on three types of problem: a) incomplete privacy policies, b) incorrect privacy policies and c) inconsistent privacy policies. In their study, they analysed 1,197 apps and found that 282 (23.6%) presented at least one of these three problems, and that 222 (18.5%) had incomplete privacy policies.

To address this problem, Yu et al. (2017) developed novel software, named AutoPPG, to automatically construct and write correct and readable descriptions and so facilitate the generation of privacy policies for smartphone applications. To do this, they compared 20 privacy policies randomly selected from the sample of 7,781 applications used in their study. They found that AutoPPG tended to reveal more operations related to users' personal data than actually appeared in the existing privacy policies.

Zimmerle and Wall (2019) have drawn up a set of guidelines for the evaluation of privacy policies in mobile applications for children, as well as for children's websites. The authors define the key elements that should be included in privacy policies: a description of all personal data collected; details of the use that third parties can make of the information collected; specification of the parental control options that parents/guardians can implement; and, finally, a highly visible link to the privacy policies.

6. Conclusions

There are many issues that need to be addressed in relation to mobile applications, and not only those that allow the manipulation of images to generate deepfakes, but also in relation to all other types of application.

In the case of manipulation techniques, such as morphing, it is possible that, with the development of new algorithms and more innovative technological processes, it will be possible to improve the process of changing a person's physiognomy, which is the key to producing deepfakes. Future advances in artificial intelligence and machine learning look set to facilitate these changes.

Likewise, the detection of manipulated videos will become increasingly complex as well as more efficient, given that the same advances permitting the modification of a video will also be used in the detection of falsifications. However, there will always be a certain time lag in this regard. And this may be enough to make other types of audiovisual manipulations.

The application of privacy policies is a complex issue and not solely because their global distribution leads to problems of legislative jurisdiction. There are also inconsistencies in the information offered to users about the data collected and shared by the application. Perhaps the creation of a standard could make it easier for privacy policies to be more consistent with the operations that apps actually perform.

To conclude, it is critical that steps be taken to ensure that in all areas, not only in the academic arena and schools, that people are made aware of the need to prevent the development and use of deepfakes. Digital and media literacy campaigns are now more essential than ever in our rapidly changing digital world.

References

Ali, I. R., Kolivand, H., y Alkawaz, M. H. (2018). Lip syncing method for realistic expressive 3D face model. *Multimedia Tools and Applications*, 77(5), 5323-5366. https://doi.org/10.1007/s11042-017-4437-z

Atehortua, N. A., y Patino, S. (2021). COVID-19, a tale of two pandemics: Novel coronavirus and fake

news messaging. *Health Promotion International, 36*(2), 524–534. https://doi.org/10.1093/heapro/daaa140

Baker, V. (2014). Humour on record: Why parody videos need to be protected. *Index on Censorship, 43*(4), 134-136. https://doi.org/10.1177/0306422014560520

Balebako, R.; Marsh, A.; Lin, J., Hong, J. & Cranor, L. (2014). The Privacy and Security Behaviors of Smartphone App Developers. *Figshare. Journal contribution*. https://doi.org/10.1184/R1/6470528.v1

BBC News (March 1, 2018). Deepfakes: The Face-Swapping Software. *BBC News*. https://www.bbc.com/news/av/technology-43118477/deepfakes-the-face-swapping-software-explained

BBC News (November 12, 2019). Are You Fooled by This Johnson-Corbyn Video? *BBC News*. https://www.bbc.com/news/av/technology-50381728/the-fake-video-where-johnson-and-corbyn-endorse-each-other

Bendel, O. (2019). The synthetization of human voices. *AI & Society, 34*(1), 83–89. https://doi.org/10.1007/s00146-017-0748-x

Cole, S. (April 26, 2019). A Site Faking Jordan Peterson's Voice Shuts Down After Peterson Decries Deepfakes. *Vice.com*. https://www.vice.com/en/article/43kwgb/not-jordan-peterson-voice-generator-shut-down-deepfakes

Crystal, D. T.; Cuccolo, N. G.; Ibrahim, A. M. S.; Furnas, H. & Lin, S. J. (2020). Photographic and video deepfakes have arrived: How machine learning may influence plastic surgery. *Plastic and Reconstructive Surgery, 145*(4), 1079–1086. https://doi.org/10.1097/PRS.0000000000006697

Das, D. & Clark, A. J. (2019). Satire vs fake news: You can tell by the way they say it. 22-26. *In First International Conference on Transdisciplinary AI, TransAI,* (pp. 22-26). https://doi.org/10.1109/TransAI46475.2019.00012

Foo, Y. Z.; Simmons, L. W. & Rhodes, G. (2017). Predictors of facial attractiveness and health in humans. *Scientific Reports, 7*(1), 39731. https://doi.org/10.1038/srep39731

Güera, D. & Delp, E. J. (2018). Deepfake video detection using recurrent neural networks. In *15th IEEE International Conference on Advanced Video and Signal Based Surveillance, AVSS*, (pp. 1–6). https://doi.org/10.1109/AVSS.2018.8639163

Hasan, H. R. & Salah, K. (2019). Combating deepfake videos using blockchain and smart contracts. *IEEE Access*, (7), 41596–41606. https://doi.org/10.1109/ACCESS.2019.2905689

Hughes, H. C. & Waismel-Manor, I. (2021). The macedonian fake news industry and the 2016 US election. PS: *Political Science & Politics, 54*(1), 19–23. https://doi.org/10.1017/S1049096520000992

Islam, M. B.; Lai-Kuan, W. & Chee-Onn, W. (2017). A survey of aesthetics-driven image recomposition. *Multimedia Tools and Applications, 76*(7), 9517–9542. https://doi.org/10.1007/s11042-016-3561-5

Ivakhiv, A. (2016). The Art of Morphogenesis: Cinema in and beyond the Capitalocene. In: Shane Denson, J. L. (Hg.): *Post-Cinema. Theorizing 21st-Century Film*, (pp. 724-749). REFRAME Books. https://doi.org/10.25969/mediarep/13475

Joseph, R. (2019). Fakebusters strike back: How to spot deep fakes, the manipulated videos that are the newest form of "fake news" to hit the internet. *Index on Censorship, 48*(1), 76-79. https://doi.org/10.1177/0306422019841326

Khodabakhsh, A.; Busch, C. & Ramachandra, R. (2018). A taxonomy of audiovisual fake multimedia content creation technology. In I*EEE Conference on Multimedia Information Processing and Retrieval, MIPR*, (pp. 372–377). https://doi.org/10.1109/MIPR.2018.00082

Kietzmann, J.; Lee, L. W. ,McCarthy, I. P. & Kietzmann, T. C. (2020). Deepfakes: Trick or treat? *Business Horizons, 63*(2), 135-146. https://doi.org/10.1016/j.bushor.2019.11.006

Kramer, R. S. S.; Jenkins, R. & Burton, A. M. (2017). InterFace: A software package for face image warping, averaging, and principal components analysis. *Behavior Research Methods, 49*(6), 2002–2011. https://doi.org/10.3758/s13428-016-0837-7

Krylov, A.; Nasonova, A. & Nasonov, A. (2014). Image warping as an image enhancement post-processing tool. In Paulus, D.; Fuchs, C. & Droege, D. (ed.) *Proceedings of the 9th Open German-Russian Workshop on Pattern Recognition and Image Understanding*, (pp. 132-135). Universität Koblenz-Landau. https://kola.opus.hbz-nrw.de/frontdoor/index/index/docId/915

Kwok, A. O. J. & Koh, S. G. M. (2021). Deepfake: A social construction of technology perspective. Current Issues in Tourism, 24(13), 1798–1802. https://doi.org/10.1080/13683500.2020.1738357

La Sexta (2019). Videos Manipulados - *El Intermedio*. https://www.lasexta.com/programas/el-intermedio/videos-manipulados/

Li, D.; He, K.S, un, J. & Zhou, K. (2015). A geodesic-preserving method for image warping. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, CVPR*, (pp. 213–221). https://doi.org/10.1109/CVPR.2015.7298617

Li, Y.; Chang, M.-C. & Lyu, S. (2018). In ictu oculi: Exposing ai created fake videos by detecting eye blinking. In *IEEE International Workshop on Information Forensics and Security, WIFS*, (pp. 1–7). https://doi.org/10.1109/WIFS.2018.8630787

Nguyen, H. H.; Yamagishi, J. & Echizen, I. (2019). Use of a capsule network to detect fake images and videos. *ArXiv:1910.12467 [Cs]*. http://arxiv.org/abs/1910.12467

Parker, L.; Halter, V.; Karliychuk, T.& Grundy, Q. (2019). How private is your mental health app data? An empirical study of mental health app privacy policies and practices. *International Journal of Law and Psychiatry, 64*, 198–204. https://doi.org/10.1016/j.ijlp.2019.04.002

Powell, A.; Singh, P. & Torous, J. (2018). The complexity of mental health app privacy policies: A potential barrier to privacy. *JMIR MHealth and UHealth, 6*(7), e158. https://doi.org/10.2196/mhealth.9871

Prathap, K. S. V.; Jilani, S. A. K. & Reddy, P. R. (2016). A critical review on Image Mosaicing. In *International Conference on Computer Communication and Informatics, ICCCI*, (pp. 1–8). https://doi.org/10.1109/ICCCI.2016.7480028

Puerto, S. (2018). Técnicas de animación e interrelación de imágenes bidimensionales. *Mosaic*, 165. https://doi.org/10.7238/m.n165.1842

Robillard, J. M.; Feng, T. L.; Sporn, A. B.; Lai, J.-A.; Lo, C.; Ta, M. & Nadler, R. (2019). Availability, readability, and content of privacy policies and terms of agreements of mental health apps. *Internet Interventions*, 17, 100243. https://doi.org/10.1016/j.invent.2019.100243

Rossler, A.; Cozzolino, D.; Verdoliva, L.; Riess, C.; Thies, J. & Niessner, M. (2019). Faceforensics++: Learning to detect manipulated facial images. In *Proceedings of the IEEE/CVF International Conference on Computer Vision, ICCV*, (pp. 1–11). https://doi.org/10.1109/ICCV.2019.00009

Scherhag, U.; Nautsch, A.; Rathgeb, C.; Gomez-Barrero, M.; Veldhuis, R. N. J.; Spreeuwers, L.; Schils, M.; Maltoni, D.; Grother, P.; Marcel, S.; Breithaupt, R.; Ramachandra, R. & Busch, C. (2017). Biometric systems under morphing attacks: Assessment of morphing techniques and vulnerability reporting. In *International Conference of the Biometrics Special Interest Group, BIOSIG*, (pp. 1–7). https://doi.org/10.23919/BIOSIG.2017.8053499

Somers, H.; Evans, G. & Mohamed, Z. (2006). Developing speech synthesis for under-resourced languages by «faking it»: An experiment with Somali. In *Proceedings of the Fifth International Conference on Language Resources and Evaluation, (pp.* 2578-2581). http://www.lrec-conf.org/proceedings/lrec2006/pdf/483_pdf.pdf

Stark, J. (September 11, 2016). Adobe stellt Sprach-Software Voco vor!-Das Computer-Magazin. *Com! Professional*. https://www.com-magazin.de/news/adobe-systems/adobe-stellt-sprach-software-voco-1146967.html

Tang, J. & Ni, B. (2019). Progressive face dynamic morphing. In *International Conference on Intelligent Computing, Automation and Systems, ICICAS*, (pp. 48–53). https://doi.org/10.1109/ICICAS48597.2019.00019

Veiga, C.; Lourenço, A. M.; Mouinuddin, S.; van Herk, M.; Modat, M.; Ourselin, S.; Royle, G. & McCle-

lland, J. R. (2015). Toward adaptive radiotherapy for head and neck patients: Uncertainties in dose warping due to the choice of deformable registration algorithm: Dose warping uncertainties due to registration algorithm. *Medical Physics*, *42*(2), 760–769. https://doi.org/10.1118/1.4905050

Wijethunga, R. L. M. A. P. C.; Matheesha, D. M. K.; Noman, A. A.; De Silva, K. H. V. T. A.; Tissera, M. & Rupasinghe, L. (2020). Deepfake audio detection: A deep learning based solution for group conversations. In *2nd International Conference on Advancements in Computing, ICAC*, (pp. 192-197) https://doi.org/10.1109/ICAC51239.2020.9357161

Yu, L.; Luo, X.; Liu, X. & Zhang, T. (2016). Can We Trust the Privacy Policies of Android Apps? In *46th Annual IEEE/IFIP International Conference on Dependable Systems and Networks, DSN*, (pp. 538-549). https://doi.org/10.1109/DSN.2016.55

Yu, L.; Zhang, T.; Luo, X.; Xue, L. & Chang, H. (2017). Toward Automatically Generating Privacy Policy for Android Apps. *IEEE Transactions on Information Forensics and Security*, *12*(4), pp. 865-880. https://doi.org/10.1109/TIFS.2016.2639339

Zimmerle, J. C. & Wall, A. S. (2019). What's in a policy? Evaluating the privacy policies of children's apps and websites. *Computers in the Schools*, *36*(1), 38–47. https://doi.org/10.1080/07380569.2019.1565628

Mario Pérez-Montoro
Universitat de Barcelona
perez-montoro@ub.edu

Visual communication in the management of a global health emergency

https://doi.org/10.3145/indocs.2022.9

Pérez-Montoro, Mario (2022). Visual communication in the management of a global health emergency. In: Freixa, Pere; Codina, Lluís; Pérez-Montoro, Mario y Guallar, Javier (ed.). *Visualisations and narratives in digital media. Methods and current trends*, (pp. 128-147). Barcelona: DigiDoc-EPI. https://doi.org/10.3145/indocs.2022.9

Abstract

As the Covid-19 pandemic took hold, governments and the media suddenly found themselves facing the difficult task of communicating details of the evolving health emergency to the public and of making technical concepts and massive volumes of data intelligible to the world's citizens. The challenge was to find a way of making this new, unfamiliar content accessible to all, by exploiting known, easily grasped formulas. Against this backdrop, it was quite clear what the strategy had to be for optimising the efficiency of this communication process: the transmission of this vital information by means of data visualisation. In this study, I analyse the problems associated with the data employed in these visualisations and critically review both good and bad practices that emerged to underpin these communication strategies.

Keywords

Information visualisation, coronavirus, pandemic, Covid-19, SARS-CoV-2, data visualisation, interactive visualisation, infographics, multimedia visualisation.

Título

Comunicación visual en la gestión de una emergencia sanitaria global

Resumen

Frente a la pandemia generada por el coronavirus, los gobiernos y los medios se encontraron de repente con el difícil reto de comunicar sobre este tema con la población y trasladar al gran público conceptos técnicos y grandes volúmenes de datos. Había que buscar una manera de hacer accesibles esos contenidos no familiares para todos, a través de fórmulas conocidas y asimilables. En ese contexto, la estrategia elegida para hacer más eficiente esa comunicación estaba clara: transmitir toda esa información vital a través de la representación visual de datos. En este trabajo analizamos los problemas asociados con los datos involucrados en esas visualizaciones y revisamos una selección de buenas y malas prácticas utilizadas en esas estrategias comunicativas.

Palabras clave

Visualización de la información, coronavirus, pandemia, COVID-19, SARS-CoV-2, visualización de datos, visualización interactiva, infografía, visualizaciones multimedia.

Note

This work is the result of reviewing, updating and editing a previous work (Pérez-Montoro, 2021).

1. Introduction

We woke at the start of 2020 to learn of the unsettling news that the Chinese were to build a hospital in just a week to care for those suffering a strange respiratory disease. Following the inevitable flood of memes in which this meteoric feat of construction was contrasted with that of the slowest ever (140 years and counting) of the Sagrada Familia temple in Barcelona, we were soon able to verify how an entire city (Wuhan) of more than 11 million inhabitants had been locked down in their homes, in almost military fashion, to stem the spread of this disease. The situation was obviously no longer a laughing matter.

We all know how the story was to pan out and, as I write, we continue to suffer its consequences: fear, uncertainty and, in the worst cases, grief at the loss of loved ones. Our social habits were revolutionised as the clouds of a major economic crisis formed over our heads and threaten to remain there for some time to come.

In the midst of this bleak, heart-breaking panorama – as if we had suddenly found ourselves characters in a dystopian science fiction novel – we began to employ a *neo-language* made up of foreign-sounding words hitherto unheard: *coronavirus, PCR, antigen test, IgG-IgM, lock-down, infection rate, de-escalation, cumulative incidence, hydroxychloroquine, FFP2, Covid-19, mortality rate*, and *SARS-CoV-2*, to mention just a few. A profusion of new concepts, guidelines, and directives, all extremely difficult for the lay population to digest.

2. Flattening the curve: information visualisation as a communication strategy

Against the backdrop of this global health emergency, the world's governments suddenly found themselves facing the exacting task of informing the public about this rapidly evolving situation. To stem the pandemic, they would have to find a way of making these technical concepts and massive volumes of associated data intelligible to their citizens. The challenge was to find a way of making this new, unfamiliar content accessible to all, by exploiting known, easily grasped formulas. It was quite clear what the strategy had to be for optimising the efficiency of this communication process: the transmission of this *vital* information by means of data visualisation.[1]

Such was the commitment to data visualisation that the principal motto underpinning the actions to be taken was "Let's flatten the curve!". Never before had a line plot been placed at

1 A number of authors have reviewed the attempts made in response to this challenge centred on the exploitation of data visualisation. In one such initiative, led by Paul Kahn (2021), some of these proposals have been classified according to their source type, country of origin, language used, the publisher, subject, technique, and date of production.

the epicentre of a communications campaign. The strategy underlying the motto was clear: the urgent need to reduce the number of Covid-19 infections so that the sick did not swamp the world's health systems. Without a vaccine, the goal was to ensure the system did not collapse, leaving the sick untreated and, so, avoid a heavy death toll.

One of the first news media outlets to disseminate this slogan was *The New York Times*. In an article entitled "Flattening the Coronavirus Curve", published at the end of March 2020, the editor headed the day's top story with a graph made up of two epi curves (Figure 1). One of the curves predicted the high number of Covid-19 cases that would be seen if no protective measures were taken, the other the number of cases with the protective measures in place. Cutting across the chart was a line of dashes indicating the capacity of the healthcare system, the maximum number of cases that hospitals could treat. The first curve rose up and clearly peaked at a point well above this dashed line, highlighting the massive number of citizens that would be denied care in a scenario without any protective measures. The message was clear: we would have to flatten the curve of Covid-19 cases by implementing the necessary measures if we wanted to avoid these terrible consequences.

Many other news media companies were to adopt this same strategy and a countless number of information visualisations were published that included some version or other of this graph. Arguably, however, one of the most effective did not appear in the news media but rather on a social networking service. The Japanese digital artist, Kenta, in a tweet posted on 3 April 2020, uploaded a video with a three-dimensional animated recreation of this same

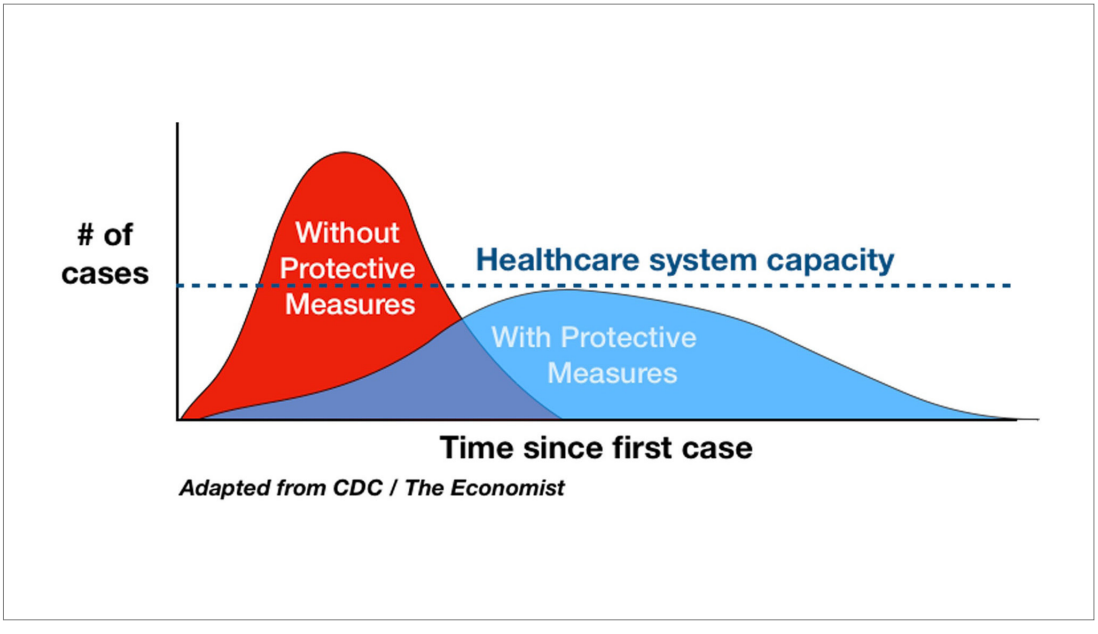


Figure 1. Line plot included in the header to the article "Flattening the Coronavirus Curve", *The New York Times*, March 2020. https://www.nytimes.com/article/flatten-curve-coronavirus.html

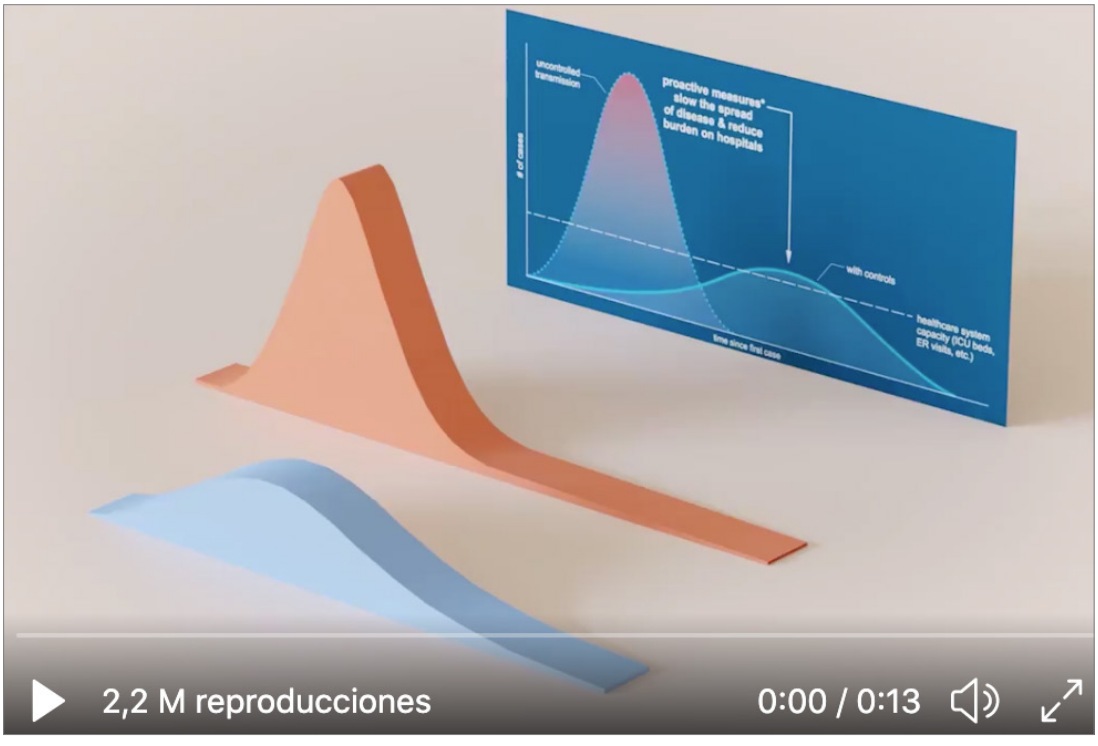


Figure 2. 3D animated representation of the line plot explaining the slogan "Flattening the Coronavirus Curve". https://twitter.com/kntktnk/status/1246011725293318151

graph conveying the same message, but emphasising much more persuasively, if that were possible, the need to stay home.

Turning 3D animation to good effect, the artist cleverly recreates the volume represented by the two epi curves. The two curves are seen moving along in parallel like two trains headed towards two tunnels that represent the capacity of the healthcare system. The curve of infections without any preventive measures in place is too high to pass through the tunnel of medical services and loses a sizeable chunk of its volume, whereas the other curve passes through without any problems (Figure 2).

3. Pandemic data: a source of problems

Despite the efforts made to ensure the danger of overstretching hospital services reached the public, the visual communication of the health crisis faced major problems from the outset. A large part of these problems centred on the data used in the visualisation, that is, on the actual content of these visual strategies. If the quality of the data was not good or the data were incomplete, the efficiency of the visual representations behind the communication strategies ran the risk of being drastically undermined.

From a conceptual perspective, a number of authors argue that there is no such thing as "raw data", that is, data that has yet to be processed. The expression "raw data", it is claimed, is a veritable oxymoron (Bowker, 2008). The simple fact of their being considered *data* means they are no longer neutral, that they have been endowed with a certain theoretical and conceptual meaning. And, to a greater or lesser extent, this controversy was never far from the surface when working with the pandemic data.

Without losing sight of this particular problem, the first difficulty I wish to highlight was that of data collection. During many long months, governments faced the challenging task of gathering data about the pandemic in order that they might implement measures on the basis of those data. In the first few months, in the absence of PCR and antigen tests, it was difficult to establish the real number of infected and asymptomatic cases and to be able, therefore, to calculate Covid-19 incidence and propagation rates. Unfortunately, the only data that could be considered reliable were that of the number of deaths. And even here some of those who died could not be tested and were not included on these lists. These problems of data collection made it impossible to predict, among other things, the real evolution of the pandemic or the effect of the measures that were being taken to alleviate it.

Leaving data collection to one side for a moment, the second notable problem concerned the management of the collected data. A joint national or international database or repository where these data could be managed, recorded and shared by the relevant experts has yet to be created. Each government managed their data on an individual basis and used their own criteria for recording them, thus often preventing any meaningful comparative analyses. To make matters worse, no steps were taken to enhance the interoperability of these data repositories, which could have improved the sharing and generation of aggregate data with respect to any political or geographical unit (region, country, continent, etc.) that experts might have wanted to analyse.

The third problem was that of updating the data. Clear criteria as to when and how to update data were never established. Some countries or governments opted to do so on a daily basis. Others weekly. Some, because of the time lag in obtaining the results of the antigen tests, collected the number of daily infections and then corrected these figures weeks later as the test results became known. This problem did little to help provide a clearer picture of reality, one, that is, on which decisions could be taken about what measures to maintain or implement.

The last of the data-related problems – and one that might at first glance appear unrelated to the specific topic we are addressing here – was a political one. Unfortunately, some governments (local, regional and national) saw an opportunity in their role of collecting, managing and updating pandemic data to reap political benefit and to present themselves to the electorate as exemplary managers of the health emergency. It lies outside the scope of the present study to analyse the degree of intentionality of these institutions, but among these distorting

actions (wilful or otherwise), it is worth highlighting, for example, a possible reduction in the number of tests reported so as to lower the infection rate, the *a posteriori* updating (even months later) of the pandemic data assigned to a specific day to correct a particular trend, and the delay (or blocking, on some occasions) in assigning a code to those infected so as not to trigger a tracking mobile application as proposed by a higher government body.

However, in the midst of these problems, a number of proposals stand out for being especially creative initiatives that helped ensure the responsible collection and management of data. A good example is provided by the solution devised by *Public Health England* and the UK Government unit, *NHSX* (Figure 3). On its open website, the UK Government offers us a data dashboard of clean, simple graphics describing pandemic metrics for the country that can be visualised in a user-friendly way and which, moreover, are subject to constant updates (the date and time of the last updating being clearly indicated). Its home page provides graphs and charts of the number of people testing positive each day, the number of patients admitted to hospital, the number of deaths and the number of virus tests processed. The page also includes a search box so that users can access these data for a specific area of the country as well as an interactive map for conducting more detailed searches. The site's navigation system allows the user to access different visualisations and tables based on data sets built around the same home-page topics (i.e. cases, deaths, healthcare and testing). At this second level, each of the visualisations can be further explored, the metrics being disaggregated for each of the nations that make up the UK as well as by daily and cumulative data. Users, moreover, can access the (downloadable) data and obtain further information about that data and the visualisation that they are viewing.

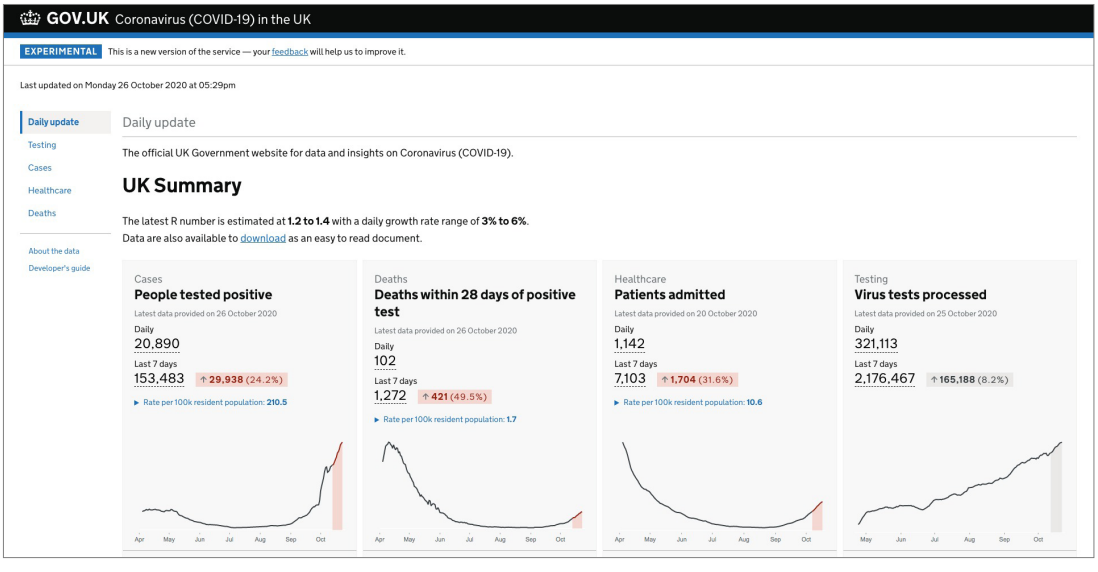


Figure 3. UK Government web resource offering a dashboard for visualising the country's pandemic data. https://coronavirus-staging.data.gov.uk

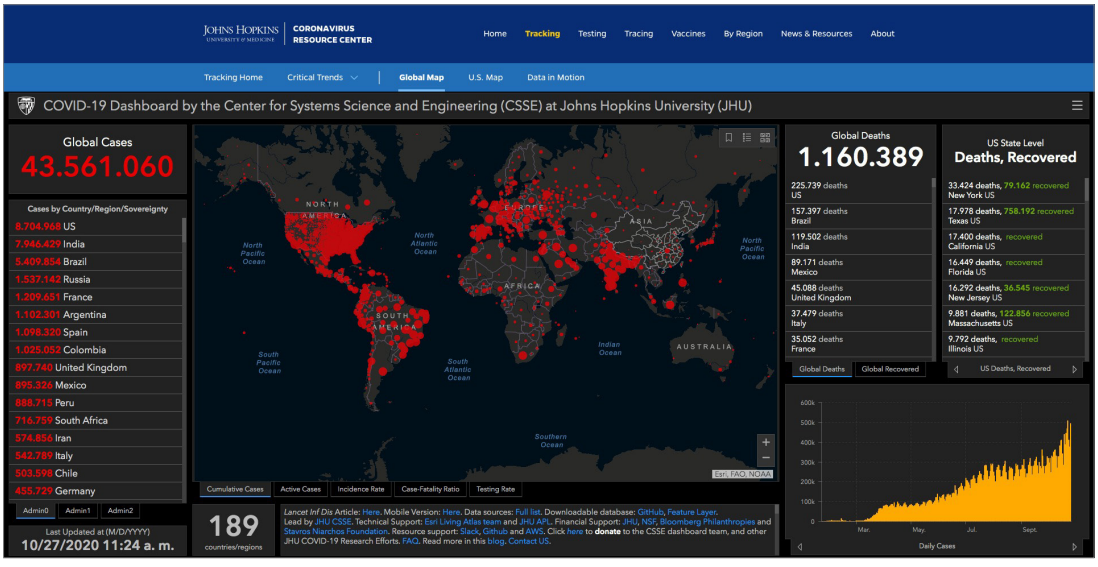


Figure 4. *Johns Hopkins University* web resource which collects and visualises global pandemic data. https://coronavirus.jhu.edu/map.html

A second solution that stands out for its value added is the resource generated by the *Center for Systems Science and Engineering* at the *Johns Hopkins University*. At a time when the epidemic was starting to acquire the characteristics of a pandemic and it was unclear where we should turn to obtain data on the health crisis, this research centre began to collect and visualise global epidemiological data, establishing itself as an obvious point of reference. The result is a data dashboard that allows the user to navigate a global map of symbols (red circles) representing the number of Covid-19 cases in each of the countries (the larger the circle, the higher the number of cases). It should perhaps be mentioned that some researchers have complained about the map's visual saturation, the result of the fact that the data for the United States are disaggregated by county and not included as a single metric for the whole country. This can create the mistaken impression that the U.S. presents many more cases of Covid-19 – given that each circle represents a separate county – while in Europe, for example, each circle represents a separate country.

To the left of the map, a table ranks the number of positive cases by country, state, province and city; while, to the right, there are two tables, one reporting the number of global deaths and recoveries from infection, and another with this same information, plus the number of tests administered for each of the U.S. states.

4. Good (and bad) communicative proposals

Leaving to one side the problematic nature of the pandemic data, their visualisation emerged as one of the most efficient communication strategies for maintaining the public informed and for controlling the health emergency. The list of communicative proposals here is

endless: some offering innovative solutions, others characterised by quite severe limitations. In this section, I review a series of good practices that focus on the nature of the virus and its process of transmission, the spread of the disease and associated issues of mobility, as well as a number of examples of what might be considered bad practices, or inefficient visualisations, in this particular context. I conclude by presenting some of the most efficient communicative proposals related to the pandemic, which stand out for their effectiveness and communicative impact.

From the beginning of the pandemic, one of the first things the population needed to understand was the nature of the virus they were facing and how it was transmitted. Meeting this need was not an easy task as the information had to reach and be understood by the majority of the public, independent of their scientific background and knowledge.

One of the most interesting proposals was published on 11 May 2020 in the Spanish newspaper, *El País*, in an article entitled "*ccu cgg cgg gca. Las doce letras que cambiaron el mundo*" ["ccu cgg cgg gca. The twelve letters that changed the world"]. This article, using a series of attractive infographics and the narrative technique of parallax scrollytelling (interactive viewing of the piece by means of the scroll and overlapping layers on the screen), describes in an engaging, intelligible fashion the virus' genetic material, its physical structure, and the way it invades

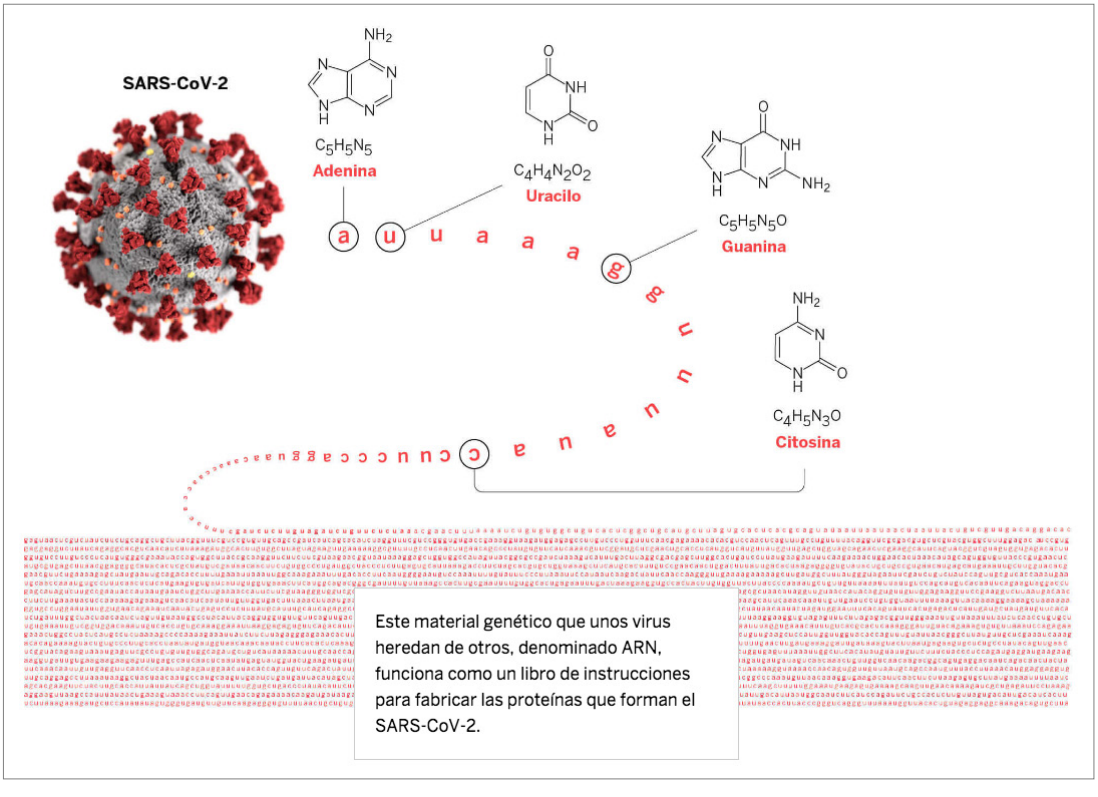


Figure 5. *El País* article offering an in-depth look at the nature of the coronavirus. https://elpais.com/elpais/2020/05/09/ciencia/1589059080_203445.html

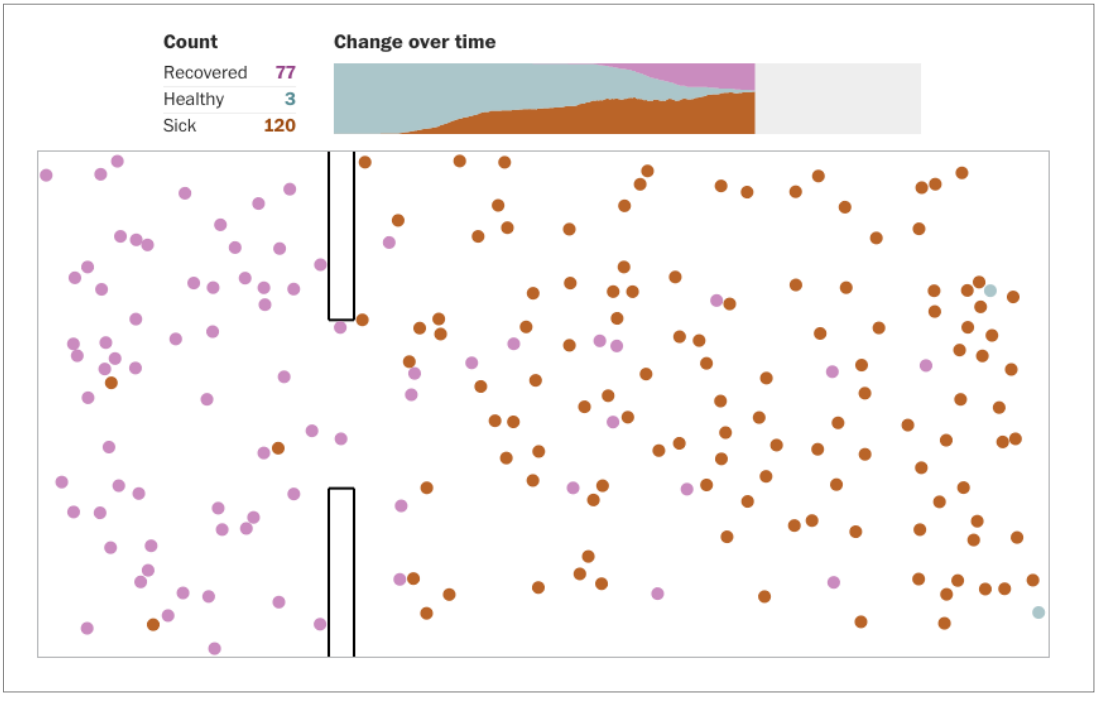


Figure 6. *The Washington Post* article in which the speed of contagion by the coronavirus is explained visually. https://www.washingtonpost.com/graphics/2020/world/corona-simulator/

and reproduces in the human body and also compares it with other viruses such as the flu (Figure 5).

Communicating how the virus was transmitted and how people could be infected was another challenge. On 14 March 2020, *The Washington Post* published an article entitled "Why outbreaks like coronavirus spread exponentially, and how to 'flatten the curve'", explaining visually and by means of simulations the speed of contagion and the importance of reducing social interaction (Figure 6).

Along similar lines, we find the excellent work carried out by the *Reuters* agency and the Spanish newspaper, *El País*. The graphics section of Reuters, in an article entitled *The Korean clusters*, dated 20 March 2020, offers us a journalistic product which explains visually how an outbreak of coronavirus spread in a church and a hospital in South Korea (Figure 7). Thanks to the information obtained from the intense tracking of those involved in the outbreak (through the use of mobile technologies), the article exhaustively describes the sequence of infections (including the family relationships and social interactions) from the first confirmed case of a patient who had flown in from Wuhan in China. The article concludes by analysing other outbreaks of contagion in the same geographical area.

Likewise, *El País*, in a piece entitled "*Un salón, un bar y una clase: así contagia el coronavirus en el aire*" ["A sitting room, a bar and a classroom: this is how the coronavirus spreads through the air"], offers an exhaustive visual explanation of the role played by aerosols in the spread

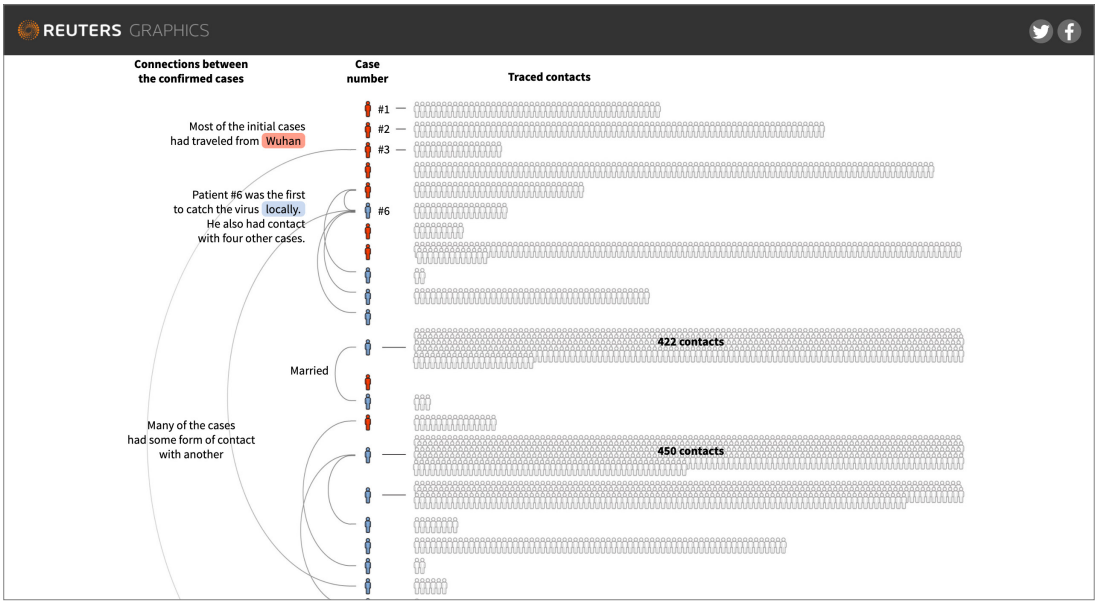


Figure 7. Analysis of the sequence of contagions in South Korea as described by *Reuters*. https://graphics.reuters.com/CHINA-HEALTH-SOUTHKOREA-CLUSTERS/0100B5G33SB/index.html

Figure 8. Explanation of the role of aerosols in the propagation of coronavirus in enclosed spaces published in the newspaper, *El País*. https://elpais.com/ciencia/2020-10-24/un-salon-un-bar-y-una-clase-asi-contagia-el-coronavirus-en-el-aire.html

of the coronavirus. Once again, using a series of attractive infographics and the narrative technique of parallax scrollytelling, the piece illustrates the variation in the rate of transmission of the virus by aerosols in enclosed places (a sitting room, a bar and a classroom) taking into account three variables: exposure time, the quality of ventilation and the use of a mask (Figure 8). Such has been the impact of the item that it has become the most visited digital content in the history of the newspaper's website.

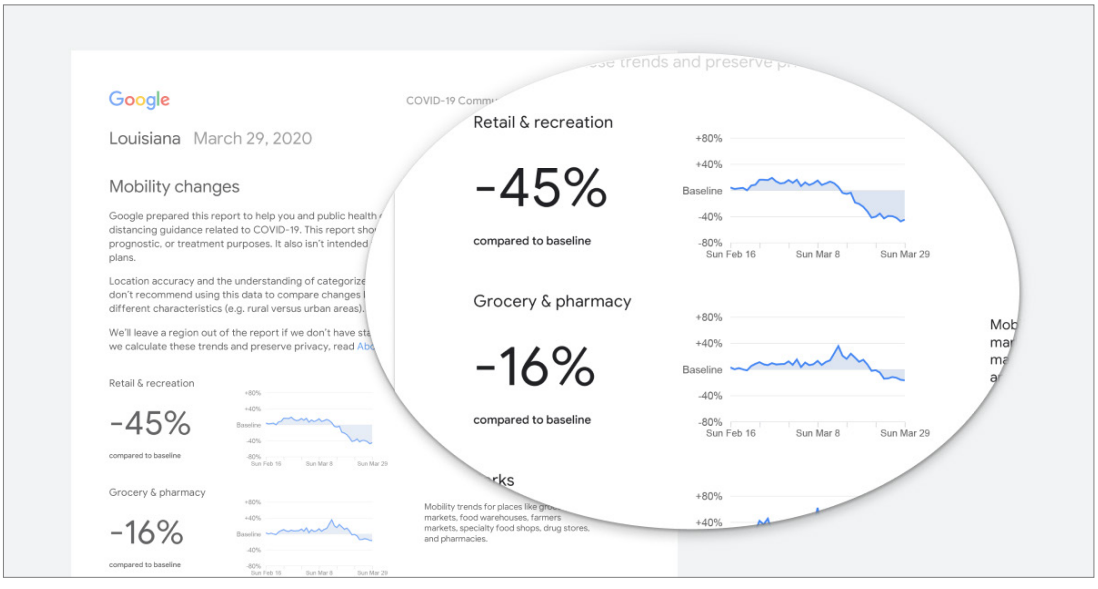


Figure 9. Google mobility report during the pandemic for the State of Louisiana. https://www.google.com/covid19/mobility/

Another area of interest explored by the media is the specific subject of mobility. In response to this information need, two visual proposals – one from Google and one from Apple – are worth highlighting. Both companies collect aggregate and anonymised (or so they claim) mobility data associated with mobile phones that use their respective operating systems. Google, using its *Google Maps* application, provides us with data and a mobility report describing trends for the geographic region of our choice. In this report, by using graphs, it shows us how visits to retail stores and places of recreation, groceries and pharmacies, parks, transit stations, workplaces and residential areas have varied during the pandemic compared to a benchmark for the previous year (Figure 9).

Likewise, Apple, using its *Maps* application, also provides us with data and a mobility report for any given region. In this report, a graph shows how mobility – that is, journeys on public transport, by car and on foot – has varied during the pandemic (Figure 10).

Continuing with good practices, I wish to highlight some pieces that are characterised by the use of unusual visual resources. In this group, the application of a very little used graphic stands out: the stream graph, a type of stacked area chart that can be employed to represent

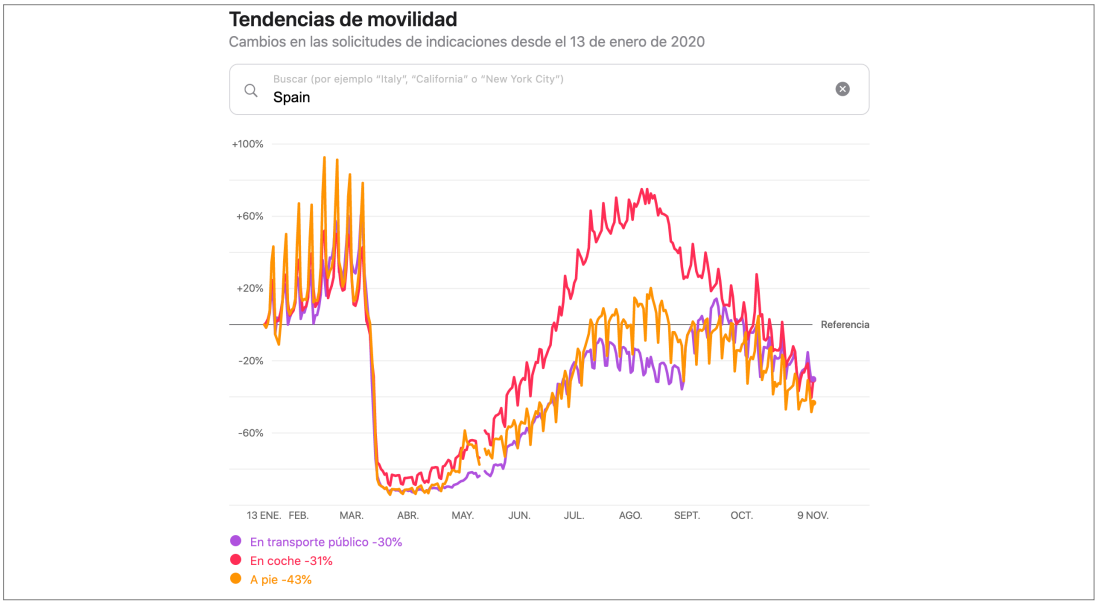


Figure 10. Apple mobility report during the pandemic for Spain. https://www.apple.com/covid19/mobility

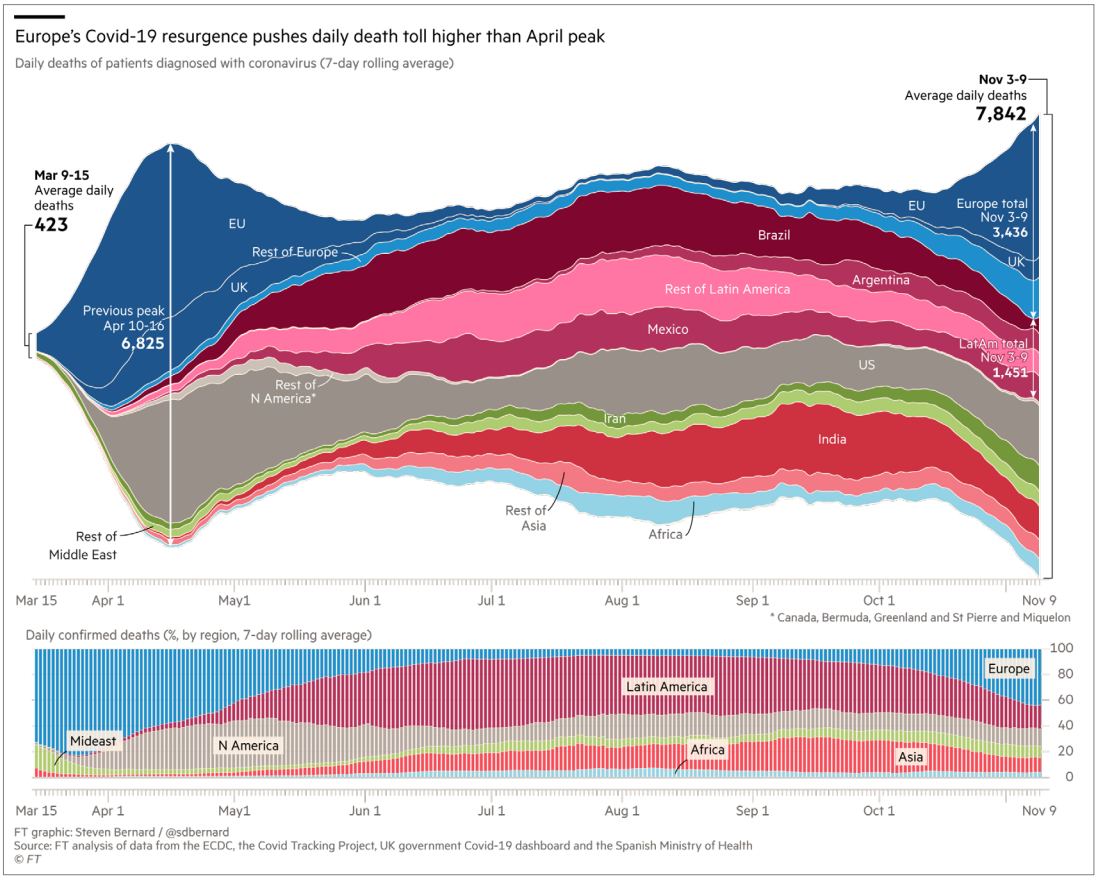


Figure 11. Evolution of daily deaths attributed to Covid-19 by country during the pandemic according to the *Financial Times*. https://www.ft.com/content/a2901ce8-5eb7-4633-b89c-cbdf5b386938

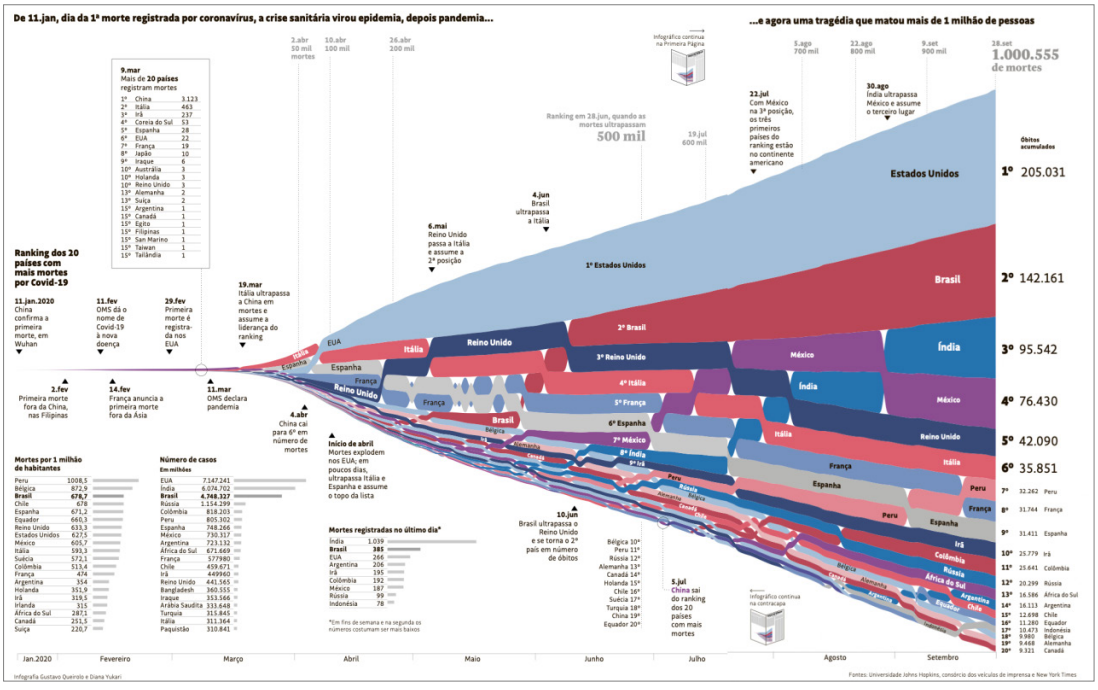


Figure 12. Evolution of daily deaths attributed to Covid-19 by country during the pandemic according to the *Folha de Sao Paulo*. https://64.media.tumblr.com/d25c0dfd70fc1b37d18dd21b6bad8c37/57290016272081f1-c3/s1280x1920/7e0a171159a02785ca7f0d3226874f3b65d199aa.png

several time series in the form of lines making up aggregate surfaces. Their defining feature is that the total stack (not just one category, but the sum of all categories at that time point) is displaced around a central point on the X-axis (located at a central point on the Y-axis). In this way they generate a visual effect for each category similar to that of the flow of a river. Two good examples of stream graphs can be found on the *Financial Times* website (Figure 11) and on the cover of the 29 September 2020 issue of the Brazilian newspaper *Folha de Sao Paulo*, where both media outlets show us the evolution of daily COVID-19 deaths by country during the pandemic (Figure 12).

But not all proposals visualising the pandemic are illustrative of good practices; indeed, there are some that lose their communicative efficiency due to visual design issues. The list is extensive, but here I have selected just a few.

The first, we have already seen. As pointed out in the previous section, the resource generated by the *Center for Systems Science and Engineering at Johns Hopkins University* has become a point of reference for obtaining and understanding data about the ongoing health crisis. Without ignoring its obvious value in this respect, the visualisation suffers from a certain visual saturation caused by the fact that U.S. data are disaggregated by county and not presented as a single value or, at least, on a state by state basis. As discussed, this can lead to the erroneous belief that in the U.S. there are many more cases of contagion (recall, each circle here represents a county) compared to Europe, for example, where each circle corresponds to a single country.

Figure 13. Evolution of daily deaths attributed to Covid-19 by country during the pandemic according to the Folha de Sao Paulo. https://64.media.tumblr.com/d25c0dfd70fc1b37d18dd21b6bad8c37/57290016272081f1-c3/s1280x1920/7e0a171159a02785ca7f0d3226874f3b65d199aa.png

The second item is offered by the Spanish Government's National Security Department. In a report on the prevailing situation (published on 7 March 2020), this body includes a map of graduated symbols representing the number of cases and deaths from coronavirus in the various Spanish regions (Figure 13). The problem with this representation is the symbol used for the quantitative coding: the circle. Leaving aside the inherent difficulty of comparing areas, if we look at the map it is immediately apparent that something is amiss. Circles of the same size and colour (such as those corresponding to Madrid and the Valencian Community) are used to represent very different quantitative values (174 cases and 4 deaths in Madrid vs 30 cases and one death in the Valencian Community).

The third visualisation I wish to describe was published on the *RTVE* website (Spain's television broadcasting corporation) on 13 April 2020 (Figure 14). Made using Flourish visualisation software, this graph provides a visual comparison of the evolution of pandemic data in Catalonia, on the one hand, and the Community of Madrid, on the other. The graph in itself presents no problems. The problem emerges when it is viewed on a mobile device. The fact of not having a *responsive* design version that adapts to the dimensions of the screen means that

when the content is viewed on such a device the numeric data overlap preventing the user from reading the visualisation clearly.

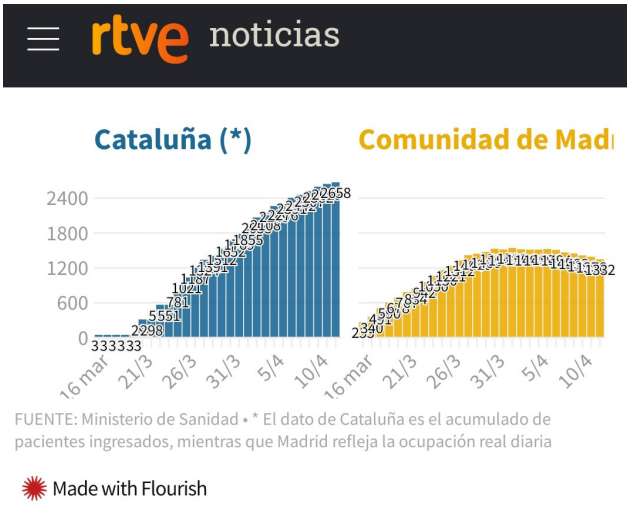


Figure 14. A comparison of the evolution of the pandemic in Catalonia and the Community of Madrid. https://www.rtve.es/noticias

The last of the visualisations that does not adequately fulfil its communicative function is the work of Danny Dorling, Professor of Geography at the University of Oxford. With the help of the illustrator Kirsten McClure, he offers us an unusual visualisation on his personal website (Figure 15). The graph shows the average number of deaths per day and also the rate of change in that number for seven countries: France, the USA, Spain, Italy, China, Germany and the United Kingdom as a whole. When a line curves to the right, deaths per day are increasing. When it curves to the left, they are decreasing. Loops indicate a change in trend.

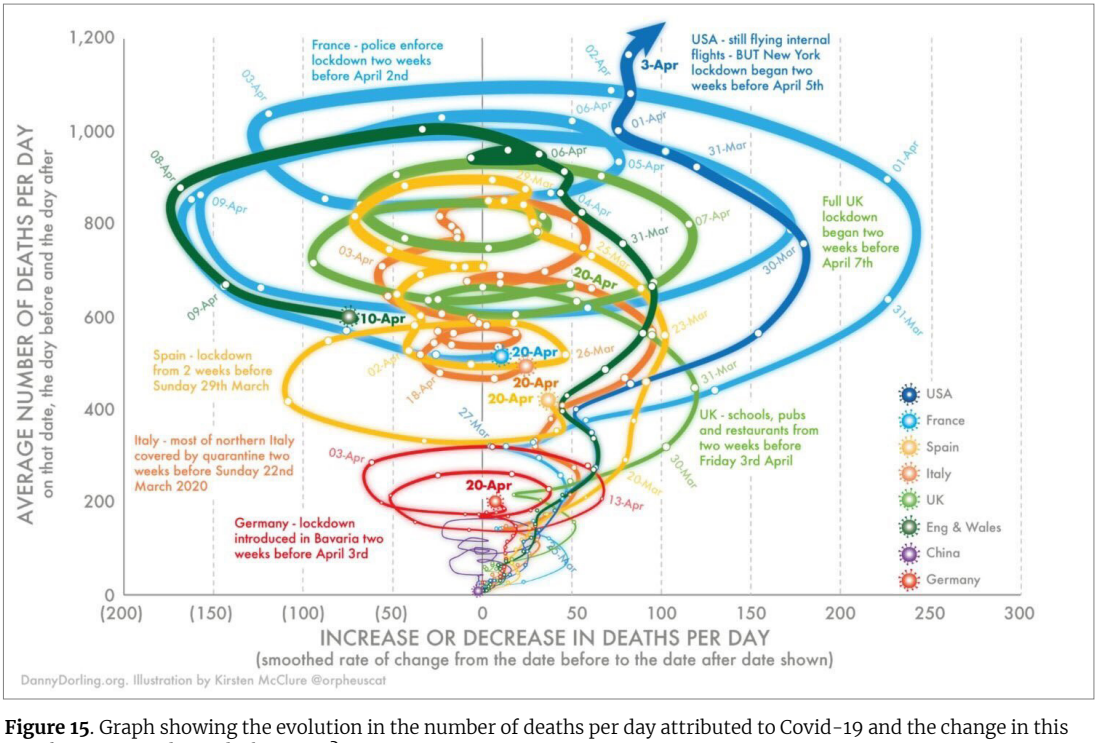


Figure 15. Graph showing the evolution in the number of deaths per day attributed to Covid-19 and the change in this rate. http://www.dannydorling.org/?p=7758


"All the News That's Fit to Print"

The New York Times

Late Edition

VOL. CLXIX ... No. 58,703

NEW YORK, SUNDAY, MAY 24, 2020

$6.00

U.S. DEATHS NEAR 100,000, AN INCALCULABLE LOSS

They Were Not Simply Names on a List. They Were Us.

Figure 16. Frontpage of *The New York Times* on 24 May 2020.

The original idea is not bad. However, the overlapping of lines, the overpopulation of the chart with additional text, and the poor choice of colour palette that does not facilitate visual discrimination means that the end result fails to communicate effectively the data analysed.

I would not want to finish this review of information visualisations without mentioning one that arguably does more than most to capture the reader's attention thanks to its effectiveness and communicative impact. It is not a visualisation that exploits an innovative idea or one that combines advances in visualisation; rather, the piece employs a classic strategy of visual representation.

Often, when wanting to commemorate or pay tribute to those that have died for some common cause, the decision is taken to read out their names in a roll call or to inscribe them on a monument. Putting names and surnames to the anonymous dead is a way of making these people visible, of escaping the coldness that comes from simply citing the number of victims, a way of moving from aseptic statistics to "convey the individual arcs of life".

And that's exactly what the American newspaper *The New York Times* did on its frontpage of 24 May 2020 (Figure 16). From the obituaries published in newspapers across the country, they selected the names of a thousand people to represent all those killed by coronavirus in the United States. The result is quite compelling. And it is even more so if we look at the interactive version launched on its website, where the deaths are shown in a more gradual fashion, as they occurred over time, and small biographical reviews are presented of some of the dead (Figure 17).

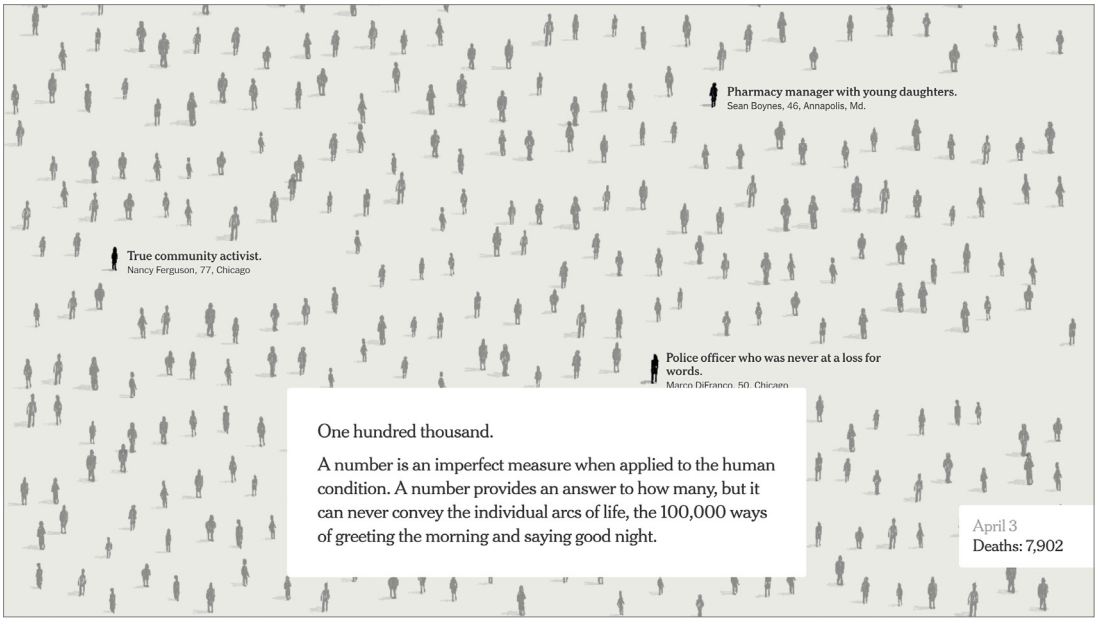
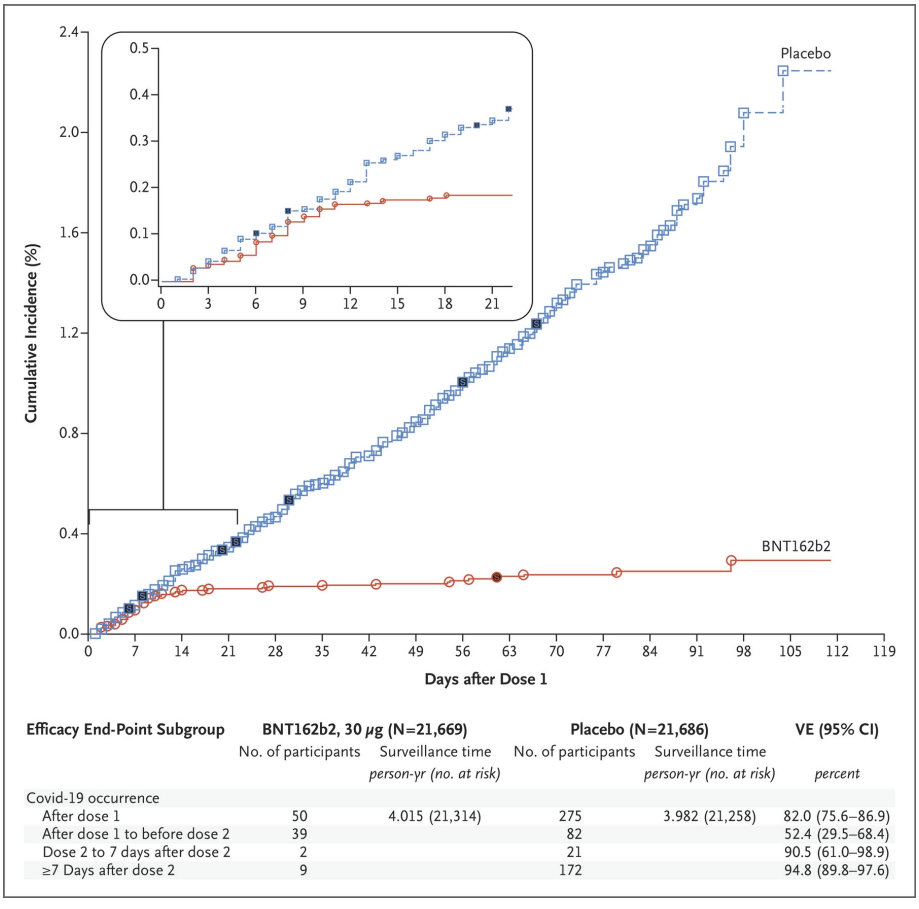


Figure 17. Piece entitled "*An Incalculable Loss*" published by *The New York Times* on 24 May 2020. https://www.nytimes.com/interactive/2020/05/24/us/us-coronavirus-deaths-100000.html



Efficacy End-Point Subgroup	BNT162b2, 30 μg (N=21,669)		Placebo (N=21,686)		VE (95% CI)
	No. of participants	Surveillance time *person-yr (no. at risk)*	No. of participants	Surveillance time *person-yr (no. at risk)*	*percent*
Covid-19 occurrence					
After dose 1	50	4.015 (21,314)	275	3.982 (21,258)	82.0 (75.6–86.9)
After dose 1 to before dose 2	39		82		52.4 (29.5–68.4)
Dose 2 to 7 days after dose 2	2		21		90.5 (61.0–98.9)
≥7 Days after dose 2	9		172		94.8 (89.8–97.6)

Figure 18. Graph appearing in an article published by *The New England Journal of Medicine* showing the effectiveness of the Pfizer/BioNTech vaccine. https://www.nejm.org/doi/full/10.1056/NEJMoa2034577

I would like to finish this review with one more graphic: what I would call "the graph of hope". A very simple, rudimentary graph that is included in an article in the prestigious journal *The New England Journal of Medicine* and which uses data to show that the Pfizer/BioNTech vaccine is 95% effective (Figure 18). In this work, and together with other scientists collaborating in the research, Polack (2020) shows us by means of just two lines how the two groups of subjects studied are being infected by COVID-19 during their trial: on the one hand, those who received a placebo – where cases grow linearly, and, on the other, the group of vaccinated – where incidences of the virus are barely recorded.

References

Apple (2020). Informe de movilidad en España durante la pandemia. *Apple.com*. https://www.apple.com/covid19/mobility

Bowker, G. (2008). *Memory Practices in the Sciences*. The MIT Press.

Dorling, D. (2020). Three charts that show where the coronavirus death rate is heading. *Danny Dorling.*

org. https://www.dannydorling.org/?p=7758

Departamento de Seguridad del Gobierno de España (March 7, 2020). *Coronavirus (COVID-19) - 07 de marzo 2020*. https://www.dsn.gob.es/es/actualidad/sala-prensa/coronavirus-covid-19-07-marzo-2020

El País (March 9, 2020). ccu cgg cgg gca. Las doce letras que cambiaron el mundo. *Elpais.com*. https://elpais.com/elpais/2020/05/09/ciencia/1589059080_203445.html

El País (October 24, 2020). Un salón, un bar y una clase: así contagia el coronavirus en el aire. *Elpais.com.* https://elpais.com/ciencia/2020-10-24/un-salon-un-bar-y-una-clase-asi-contagia-el-coronavirus-en-el-aire.html

Financial Times Visual & Data Journalism team (December 20, 2020). Coronavirus tracker: the latest figures as countries fight Covid-19 resurgence. *Financial Times.com* https://www.ft.com/content/a2901ce8-5eb7-4633-b89c-cbdf5b386938

Google (2020). Informe de movilidad durante la pandemia del Estado de Luisiana. *Google.com*. https://www.google.com/covid19/mobility

GOV.UK (2020). Coronavirus (COVID-19) in the UK. *GOV.UK.com*. https://coronavirus-staging.data.gov.uk

Hernandez, M.; Scarr, S. & Sharma, M. (2020). The Korean clusters. *Reuter Graphics*. https://graphics.reuters.com/CHINA-HEALTH-SOUTHKOREA-CLUSTERS/0100B5G33SB/index.html

Johns Hopkins University & Medicine Coronavirus Resource Center (2020). COVID-19 Dashboard by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU). *Johns Hopkins University & Medicine*. https://coronavirus.jhu.edu/map.html

Kahn, P. (2021). COVIC Visualizer. *Covic*. http://covic-archive.org/index.html

Kenta (April 3, 2020). For those who feel that "I haven't come to this area and don't have to worry about it?" [personal Twitter count]. https://twitter.com/kntktnk/status/1246011725293318151

Pérez-Montoro, M. (2021). Comunicación visual de una emergencia sanitaria mundial: el caso de la COVID-19. *Anuario ThinkEPI, 15*. https://doi.org/10.3145/thinkepi.2021.e15d01

Polack, F. P. & et al. (2020). Safety and Efficacy of the BNT162b2 mRNA Covid-19 Vaccine. *The New England Journal of Medicine*. https://doi.org/10.1056/NEJMoa2034577

Queirolo, G. & Yukari, D. (September 29, 2020). De 11.jan, dia da 1a morte registrada por coronavirus, a crise sanitária virou epidemia, depois pandemia... e agora uma tragédia que matou mais de milhâo de pessoas. *Folha de Sao Paulo*. https://64.media.tumblr.com/d25c0dfd70fc1b37d18dd21b6bad8c37/57290016272081f1-c3/s1280x1920/7e0a171159a02785ca7f0d3226874f3b65d199aa.png

RTVE (2020). Noticias. *Rtve.com*. https://www.rtve.es/noticias

The New York Times (May 24, 2020). An Incalculable Loss. *nyties.com.* https://www.nytimes.com/interactive/2020/05/24/us/us-coronavirus-deaths-100000.html

The New York Times (2020). Flattening the Coronavirus Curve. *nytimes.com*. https://www.nytimes.com/article/flatten-curve-coronavirus.html

The New York Times (May 24, 2020). US deaths near 100.000, an incalculable loss. *The New York Times*, p.1.

Stevens, H. (March 14, 2020). Why outbreaks like coronavirus spread exponentially, and how to "flatten the curve". *The Washington Post*. https://www.washingtonpost.com/graphics/2020/world/corona-simulator/

AUTHOR'S LIST

Juan-José Boté-Vericad

Universitat de Barcelona | juanjo.botev@ub.edu | https://orcid.org/0000-0001-9815-6190

Juan-José Boté-Vericad is a lecturer in the Department of Library Science, Documentation and Audiovisual Communication in the Faculty of Information and Audiovisual Media at the University of Barcelona (Spain) and is a member of the *Centre de Recerca en Informació, Comunicació i Cultura*. He teaches on the bachelor's degree in Digital Information and Documentation Management and on the Master's Degree in Digital Content Management. His academic and research career has focused on the verification and long-term preservation of information, information and media literacy, digital content management systems, as well as the management of audiovisual content in social networking sites. Speaker at numerous conferences, courses and seminars on information verification, types of audiovisual formats in higher education.

Marga Cabrera Méndez

Universitat politècnica de Valencia | mcabrera@upv.es | https://orcid.org/0000-0001-7164-1815

Marga Cabrera has a Ph.D. in Audiovisual Communication from the Polytechnic University of Valencia and is an associate professor of subjects related to digital communication since 2003, both in undergraduate and graduate degrees. She is the director of the master's degree in social media, master CALSI. She is the founder and codirector of the Comunica2 International congress on social networks. She is the coordinator of the Fundéu BBVA book "Writing for the Internet: a guide for new media and social networks". She is the author of the UOC book "Entrepreneurship in journalism, tools for the new online professional". Editor of the book "Digital Tools for Academic Branding and Self-Promotion". She is currently immersed in her project "Observatory of new media in Spanish: www.nuevosmedios.es" funded by the BBVA Foundation, now Leonardo scholarships. Since 2016, she has been a member of the Board of Directors of SembraMedia, a non-profit organization based in Los Angeles (USA) dedicated to promoting entrepreneurship in digital native Spanish-language media.

Lluís Codina
Universitat Pompeu Fabra & Barcelona School of Management | lluis.codina@upf.edu
https://orcid.org/0000-0003-2779-9740

Lluís Codina is a professor at the Faculty of Communication, in the Degrees of Journalism and Audiovisual Communication. He is coordinator of the University Master's Degree in Social Communication (MUCS) of the Communication Department and coordinator of the Journalism and Digital Documentation Research Unit (UPF). He also teaches at the Barcelona School of Management – UPF. He is co-founder and member of the editorial team of Hipertext.net Academic Magazine. In research, one of his latest projects has been the start-up of the Cybermedia Observatory, a production of the DigiDoc Group, dedicated to the study and monitoring of trends in communication and digital media. He is currently the second IP of the project "Interactive storytelling and digital visibility in interactive documentary and structured journalism". RTI2018-095714-B-C21 (MICINN/FEDER). Spanish Ministry of Science, Innovation and Universities (Spain).

Raquel Escandell Poveda
Universitat d'Alacant | r.escandell@ua.es | https://orcid.org/0000-0002-8398-1873

Raquel Escandell-Poveda is an Associate Lecturer at the University of Alicante since 2008, in the Department of Communication and Social Psychology. She graduated with special award in Advertising and Public Relations from the University of Alicante and obtained a diploma in the Advanced Course in Online Marketing at ESIC, Barcelona. In 2021, she obtained her PhD from the University of Alicante, in the Business, Economics and Society programme, Communication, Behaviour and Society line. She teaches in the Degree of Advertising and Public Relations mainly in: Media Research and Planning and Information Management in Communication. In the Official Master's Degree in Digital Communication, she coordinates and teaches Search Engine Positioning and Advertising and has taught in the University's Own Degrees such as the Master's Degree in Online Communication or the Specialist Course in Web Design. She has supervised several degree and master's degree final projects. Her lines of research focus on search engine optimisation (SEO), professional skills and profiles, communication and new technologies. She combines teaching and research with her professional career since 2004 in various communication and advertising agencies and companies. She currently works as a senior consultant in communication and digital marketing.

Pere Freixa
Universitat Pompeu Fabra | pere.freixa@upf.edu | https://orcid.org/0000-0002-9199-1270

Pere Freixa is professor and researcher of the Department of Communication at Pompeu Fabra University, UPF, specialized in photography, digital image, audiovisual media, interactive design and interactive communication. He has participated, coordinated and conducted several courses and postgraduate courses on digital image, photography and interactive communication. He is a member of DigiDoc Research Group and he is Principal Investigator (PI) of competitive project *Interactive storytelling and digital visibility in interactive documentary and structured journalism*, (MICINN/FEDER), and IP of UPF team in *Mironins, redesign of the cultural experience of visiting a museum through transmedia narrative, gamification and new technologies* research project. He is editor in chief of *Hipertext.net* academic journal.

Javier Guallar
Universitat de Barcelona | jguallar@ub.edu | https://orcid.org/0000-0002-8601-3990

Javier Guallar. PhD in Information and Documentation from the University of Barcelona and in Communication from Pompeu Fabra University. He is a professor at the Faculty of Information and Audiovisual Media of the University of Barcelona and academic secretary of this Faculty. He is a member of the CRICC Information, Communication and Culture Research Center at the University of Barcelona. He is deputy director of the journal Profesional de la información, and director of the book collections Profesional de la información and EPI Scholar at Editorial UOC. He has a long professional career as a news librarian in the media. H Index in Google Scholar 25. Content curator's newsletter editor https://www.getrevue.co/profile/jguallar

Mar Iglesias García
Universitat d'Alacant | mar.iglesias@ua.es | https://orcid.org/0000-0001-7926-5746

Mar Iglesias-García is a journalist and professor in the Communication and Social Psychology Department at the Universidad de Alicante. She teaches the Advertising and Public Relations degree and the Tourism degree. With regard to research projects, she has participated in different public R & D & I projects of public competitiveness, such as, for example, "Digital Convergence in the Media" (2006-2009), "The evolution of cyber media in the framework of digital convergence Analysis of the message" (2010-2012), "Active audiences and journalism. Interactivity, web integration and search Information" (2012-2015) and collaborates in the project "Creation and interactive content in the communication of audiovisual information: hearings, design, systems and formats" (2016-2019). She was director of *RàdioUA*, and director of the Secretariat of Cultural and Linguistic Promotion of the University of Alicante (Assistant of Vicerrector of Culture, Sports and Languages). Her primary line of research is online journalism, but she also works on ICT applied to education, and advertising. She was director of ComunicandoUA, an online journal produced by Advertising undergraduates from this university. Also, she has four teaching quinquennials granted by University of Alicante; and one recognized six-year research period granted by CNEAI. Currently she is president of the Valencian Media Corporation (CVMC).

Carlos Lopezosa
Universitat de Barcelona | lopezosa@ub.edu | https://orcid.org/0000-0003-2779-9740

Carlos Lopezosa has a Ph.D. in Information Sciences. He is a visiting researcher at the University of Barcelona (Margarita Salas postdoctoral fellowship). His doctoral thesis focused on the study of the positioning factors of content-intensive sites, and especially of online media, as well as the evaluation of SEO analysis tools. He is a specialist in search engine optimization and monetization systems based on quality content strategies. He has been an associate professor at the Universitat Pompeu Fabra, teaching at the Faculty of Communication, in the Degrees of Journalism, Audiovisual Communication and Advertising and Public Relations, collaborating professor at the Open University of Catalonia teaching the subjects of audiences and consumption of the Degree of Journalism and audiences and consumption of the Master in Journalism, and coordinator and professor of the Online Master in Search Engines and SEO/SEM at the Barcelona School of Management (UPF).

Matilde Obradors
Universitat Pompeu Fabra | matilde.obradors@upf.edu | https://orcid.org/0000-0003-2779-9740

She's a professor of Creativity and Audiovisual Ideation in the Department of Communication of the Pompeu Fabra University (UPF). She has been Vice Dean of the UPF Faculty of Communication, director of the postgraduate course, Creativity and Innovation. Interdisciplinary strategies and applications. She is coordinator of the Master in Television Quality and Innovation of TV3 and Co-director of the Laboratory of Audiovisual and Cultural Production "Live streaming" (Livemedia / UPF), in Arts Santa Mónica. She is the author of the book: "Creativity and generation of ideas. Study of creative practice in the cinema". She is a specialist in creative processes and analysis of new narratives and audiovisual formats. Visual artist, she works on artistic practice as research. She is a researcher in the research project of the National R + D + I Plan "Interactive narration and digital visibility in interactive documentary and structured journalism", within which this deliverable is inscribed.

Natalia Papí Gálvez
Universitat d'Alacant | natalia.p@ua.es | https://orcid.org/0000-0002-4871-1691

Natalia Papí-Gálvez is an Associate University Professor (Senior Lecturer) since 2011 in the Depart. of Communication and Social Psychology-University of Alicante (UA). She graduated in Sociology (concentration in Marketing and Advertising) and PhD from the University of Alicante (Social Welfare and Inequalities Program) with special undergraduate and doctoral awards. She teaches in the Degree in Advertising and Public Relations, mainly in Media Research and Planning, in Official Master's Degree in Communication and contributes to the PhD programme Business, Economy, and Society , on strand 2. Communication, Behaviour and Society. Her research has two lines of specialization: 1. "Gender, policies and effectiveness of advertising", centring on work and family life conciliation; and 2. "Research, strategy and planning of advertising media", focusing mainly on new media and digital communication; with three recognized six-year research periods. She has held the position of Vice Dean of the Degree in Advertising and Public Relations (for 9 years) and the coordination of the Master's Degree in Digital Communication (since 2020).

Rafael Pedraza-Jiménez
Universitat Pompeu Fabra | rafael.pedraza@upf.edu | https://orcid.org/0000-0002-6918-6910

Associate Professor and Serra Húnter Fellow in Information Science and Communication Studies at Universitat Pompeu Fabra (UPF). He develops his teaching and research activity in Communication Studies, applying techniques from Information Science and ICTs to this field. He has coordinated and participated in multiple research projects, financed by public and private institutions. He has also published dozens of works, many of them in high-impact journals. In addition, he has been Coordinator of the DigiDoc Research Group at UPF. He has been a lecturer in various undergraduate, master and doctorate degrees, although mainly in Journalism Studies. He has led and participated in teaching innovation projects, and he has received the Universitat Pompeu Fabra Board of Trustees Award for teaching quality, category: teaching activities in the Journalism Studies programme. Likewise, he has developed academic administration tasks at UPF, among which the following stand out: Member and Secretary of the Works Council at Universitat Pompeu Fabra; Secretary of the Faculty of Communication.

Mario Pérez-Montoro
Universitat de Barcelona | perez-montoro@ub.edu | https://orcid.org/0000-0003-2426-8119

Dr. Mario Pérez-Montoro is a Professor in the Faculty of Information and Audiovisual Media at University of Barcelona (Spain), where he researches and lectures in the areas of interaction design and information visualization. He is currently the Director of the PhD Program in Information and Communication at this University. He studied at the Istituto di Discipline della Comunicazione at the University of Bologna (Italy), he has been a professor at different universities in Spain (Complutense University of Madrid, Open University of Catalonia, and Autonomous University of Barcelona), and he has been visiting scholar at the Center for the Study of Language and Information (CSLI) at Stanford University (California, USA) and at the School of Information at UC Berkeley (California, USA).

Gema Santos-Hermosa
Universitat de Barcelona | gsantos@ub.edu | https://orcid.org/0000-0002-2776-871X

Gema Santos-Hermosa holds a PhD in Information and Communication Sciences from the University of Barcelona (UB). She is a full member of the Faculty of Information and Audiovisual Media (FIMA) at the UB, where she teaches on the undergraduate degree course in Information and Documentation and co-directs the Advanced University Course in 'Open Science: Promotion, Support and Evaluation' (UB). Previously, she has been an associate lecturer at the same university and in the Faculty of Communication at the Pompeu Fabra University (UPF), teaching on the undergraduate degree in Journalism and Audiovisual Communication. She has also been on the teaching staff of the Open University of Catalonia (UOC), lecturing on their PhD program. She is a researcher at the Information, Communication and Culture Research Centre (CRICC) and the Digital Documentation and Interactive Communication Research Group (DigiDoc) at the UPF. Over the last ten years she has participated in a number of national and European funded research projects, including "Interactive storytelling and digital visibility in interactive documentary and structured journalism" (RTI2018-095714-B-C21, MICINN/FEDER), the framework for the current study.

Joan Soler-Adillon
Universitat Oberta de Catalunya | jsoleradillon@uoc.edu | https://orcid.org/0000-0002-7959-4192

Joan Soler-Adillon is an artist and Associate Professor at the Faculty of Computer Science, Multimedia and Telecommunications at the Universitat Oberta de Catalunya (UOC). He has previously held academic positions at Royal Holloway, University of London, and Universitat Pompeu Fabra. His research and practice revolve around interactive digital media and its manifestation in digital art - especially in interactive installation, interactive narrative and documentary, and Virtual Reality. From a full-body interaction game on an inflatable slide to experimental RV documentary, he has participated in several projects focused on system behavior design and interactivity, as well as the collaboration and participation of the audience in these experiences. In recent years, his work has been awarded and selected at major European XR festivals, and exhibited at cities such as London, Paris, Berlin, Zurich, Barcelona, New York and Beijing.

Carles Sora-Domenjó

Universitat Politècnica de Catalunya | carles.sora@citm.upc.edu | https://orcid.org/0000-0003-2761-2384

Carles Sora is an academic researcher whose critical and creative work explores the intersections of technology, audiovisual, culture and the arts. He is currently director and professor at CITM (Centre de la Imatge i la Tecnologia Multimèdia) at the Universitat Politècnica de Catalunya (UPC), in Terrassa. He holds a PhD in Digital Communication from the Universitat Pompeu Fabra (UPF), where he was a lecturer in digital culture and interactive narratives. From 2017 to 2019 he was a Fulbright Fellow at the MIT Open Doc Lab, Massachusetts Institute of Technology, Boston. His PhD thesis developed a theoretical model on digital time and interactive narratives, published by UOC Press (Digital Temporalities, 2016). He holds a master's degree in Cognitive Systems and Interactive Media (UPF, 2009), and a bachelor's degree in Multimedia (UPC, 2004). He has worked on more than 25 interactive projects for visual arts, performance and scientific dissemination. His artistic and academic work has been presented in Quebec City (Méduse Center), Paris (ISCC), Buenos Aires (VE), Mexico DF (ExTeresa), Bristol (i-Docs), Lisbone, Lugano, (ECREA). He has published in international journals and publishers and has worked as reviewer for European research agencies, the European Commission (HE) and ANEP. His current lines of research are critical media studies, media effects of immersive media, interactive documentaries, emerging forms of automated narratives and their potential to transform society.

Magdalena Trillo Domínguez

Universidad de Granada | mtrillo@ugr.es | https://orcid.org/0000-0001-9505-1203

Magdalena Trillo-Domínguez is director of *Granada Hoy*, Deputy Director of *Grupo Joly*. Since the 2010 academic year, she has been working as a professor of 'Journalistic Communication' and 'Media System' at the Faculty of Granada and, for four years, she has been collaborating in the analysis and debate spaces on Canal Sur Andalucía in the morning radio program and on television. She has a degree in journalism from the University of Malaga and a Ph.D. in communication from the University of Granada, she has specialized in digital journalism and cultural management from the UOC and is part of the Nar-Trans research group on narratives and transmedia communication and collaborates habitual in the spaces of analysis and debate of actuality of the RTVA. She is also a current political and social analyst for Televisión Española1 (TVE). She was recognized in 2017 by the Junta de Andalucía with the Meridiana Award for her defense of women's equal rights.

Mari Vállez

Universitat de Barcelona | marivallez@ub.edu | https://orcid.org/0000-0002-3284-2590

Mari Vállez is a Lecturer in the Department of Library, Documentation and Audiovisual Communication at the Faculty of Information and Audiovisual Media of the University of Barcelona (Spain), where she lectures in the bachelor's degree Digital Information and Documentation Management and the master's degree Digital Content Management. She is the double bachelor's degree coordinator in Digital Information and Documentation Management + Audiovisual Communication. Her academic and research career has focused on information access from different perspectives: information retrieval, natural language processing, indexing, SEO, although always centred on the information science and communication disciplines, a sector in which she has professional experience of 27 years.

You are encouraged to print it only when absolutely necessary

March 2022

VISUALISATIONS AND NARRATIVES IN DIGITAL MEDIA

Methods and current trends

The digital media have undergone an unprecedented transformation in recent years by exploiting the combined communicative potential of interaction and visualisation, generating, in this way, new narrative forms and journalistic stories. But this is not something that can be analysed in isolation. Understanding the digital media requires addressing the study of interactive texts and the platforms of the digital ecosystem from different perspectives.

The digital media today have highly permeable boundaries, the guidelines that define them being subject to constant modification: their texts are dynamic and constantly changing, their systems operate, thanks to artificial intelligence, as just another actor that analyses, collects and manages information. The space separating senders and receivers of messages has become fuzzy and interwoven. For this reason, the study of digital journalism has to assume this unremitting transformation as simply one more element in the debate about the media, and accept it as a characteristic of the digital culture that defines the field of communication.

DigiDoc
Digital documentation and interactive communication Research group

Profesional de la información