

Reconstructing memory narratives on Facebook with Digital Methods

Òscar Coromina¹ and Adrián Padilla Molina²

¹ Universitat Autònoma de Barcelona. C/ de la Vinya S/N Facultat de Ciències de la Comunicació, Campus de la UAB Bellaterra - 018193
e-mail: oscar.coromina@uab.cat

ORCID iD: <http://orcid.org/0000-0002-6306-4154>

² Universitat Autònoma de Barcelona. C/ de la Vinya S/N Facultat de Ciències de la Comunicació, Campus de la UAB Bellaterra - 018193
e-mail: adrian.padilla@uab.cat

ORCID iD: <http://orcid.org/0000-0001-7668-1322>

Submitted: 9 December 2017. Accepted: 9 April 2018

ABSTRACT: Social platforms are playing an increasingly more important role in different aspects of our daily lives, which can range from the most mundane to such crucial processes as the transmission of meanings and values in our society. Much of the interest lies in the fact that storytelling on social networks is a collective process in which users participate by creating, sharing and commenting on content. All these actions generate digital traces that are easily accessible, in a non-intrusive and automated manner, which represents an unprecedented opportunity to investigate the mediation of social and cultural phenomena.

Digital Methods is an epistemological proposal that, aligned with Studies in Science, Technology and Society (STS), assumes the existence of a technological mediation of social and cultural practices and uses computational techniques, not only to extract the digital traces left by the users of these social networks, but also to analyze and display their content. This article takes as a case study the representation of the Spanish Civil War on Facebook to exemplify the affordances of such a methodological approach to investigate the processes of generation, diffusion and representation of historical knowledge.

KEYWORDS: Digital methods; Social media; Facebook; Spanish civil war

Citation / Cómo citar este artículo: Coromina, Òscar and Padilla Molina, Adrián (2018) "Reconstructing memory narratives on Facebook with Digital Methods". *Culture & History Digital Journal*, 7 (2): e014. <https://doi.org/10.3989/chdj.2018.14>

RESUMEN: *Reconstruyendo las narrativas de la memoria en Facebook con métodos digitales.*- Las plataformas sociales juegan un papel cada vez más importante en distintos aspectos de nuestra vida cotidiana, que pueden ir de lo más banal a cuestiones tan cruciales como el proceso de transmisión de significados y valores en nuestra sociedad. Buena parte de su interés radica en el hecho de que la construcción del relato en las redes sociales es un proceso colectivo en el que los usuarios de dichas plataformas participan creando, compartiendo y comentando contenido. Todas estas acciones generan trazas digitales accesibles de manera fácil, no intrusiva y automatizada que suponen una oportunidad sin precedentes para investigar la mediación de fenómenos sociales y culturales. Los Métodos Digitales son una propuesta epistemológica que, alineada con los Estudios en Ciencia, Tecnología y Sociedad (CTS), asume la existencia de una mediación tecnológica de las prácticas sociales y culturales y se sirve de técnicas computacionales, tanto para extraer la información relativa a las trazas digitales generadas por los usuarios en las redes sociales, como para analizar y visualizar su contenido. Este artículo toma como caso de estudio la representación de la Guerra Civil española y el Franquismo en Facebook para ejemplificar las potencialidades de dicha aproximación metodológica para investigar sobre los procesos de generación, difusión y representación del conocimiento histórico.

PALABRAS CLAVE: Métodos digitales; Plataformas sociales; Facebook; Guerra Civil Española

Copyright: © 2018 CSIC. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International (CC BY 4.0) License.

INTRODUCTION

Social platforms like Facebook are playing an increasingly more central role in our daily lives by mediating and shaping numerous social activities. We use them to communicate and manage our relationships with family and friends (Ellison *et al.*, 2014; Lenhart, 2009), follow and comment on the news (Rader and Gray, 2015; Hille and Bakker, 2014), keep track of sporting events (Fernández *et al.*, 2017), organize political protests (Poell *et al.*, 2015; Coromina, 2017), for entertainment, etc. We use the term social platforms to refer to a series of Internet-based software, services and applications that are grounded upon the technological principles of the Web 2.0 in which users are able to create contents and share them with other users (Kaplan and Haenlein, 2010). Under the umbrella of social platforms, we can find blogs, wikis, social networking sites, virtual worlds and other applications that today serve as the most representative exponents of so-called social media.

Among all social platforms, Facebook stands out as the most popular with more than 2 billion users worldwide (Facebook Inc., 2017). It is defined as a social networking site since it fulfills the premise of being an application that allows individuals to build a public or semi-public profile within a closed system, to create a list of other users with whom they share a connection and to view and cross their list of contacts with others created by users of the same system (Ellison and Boyd, 2013). Beyond the basic characteristics that define this and other social networks, Facebook includes the interpersonal communication features of other Internet services such as chat, instant messaging and e-mail and allows its users to publish different types of content (videos, photos, texts, links, etc.) with which other users can interact through reactions and comments or even help to spread said content by sharing (Musial and Kazienko, 2012; Rains and Brunner, 2015; Thelwall, 2009; Kim *et al.*, 2010). These contents are published on a user's Timeline, a vertical space in the user profile in which the newest contents displace the older ones below. The timeline is organized as a biographical narrative "that smartly disciplines its user into combining self-expression (in this case memory and emotion) with self-promotion in a uniform format" (Van Dijck, 2013). The timeline and its functionalities are a characteristic shared by Fan Pages and Groups (semi-public spaces designed to promote the creation of communities around companies, brands, public figures and interests). Although both types, pages and groups, are managed by a limited number of users, the members of these communities can also participate in the creation and shaping of the narrative that is published in the Timeline. Such participation is not confined solely to publishing content. Certain actions performed by users also play a significant role in the mechanisms of visibility and dissemination of the platform. Thus, reactions, comments and, more obviously, shares help to make the contents more visible. Likewise, comments fulfill different functions such as extending, completing, debating or questioning the published contents.

The active role of users in the construction of narratives shows the transition from a traditional mass communication model, which was characterized by a one-way relationship between the highly professionalized producers of contents and a mass audience, to a new model in which the relationship between producers and the audience (users) is two-way and the division of roles is much less compartmentalized. In this regard, the concept of 'Participatory Culture' has been used to define the consequences stemming from this new model, in which "audiences, empowered by these new technologies, occupying a space intersection between old and new media, are demanding the right to participate within the culture" (Jenkins, 2006:24). The active role of users in social media contributes to the fact that the narratives disseminated on these platforms are projects constructed by a variety of different actors by reorganizing communicative flows, sharing news and framing issues. It can be described as a collective effort, negotiation or even struggle and is characterized by the fact that professionals and the general public play a crucial role in the story that emerges from this collective process.

The building of narratives implies choosing some aspects of reality and stressing them in the message to promote a given definition, interpretation, assessment and way of acting in relation to the matter described. Social media are gaining ground as platforms that construct the narrative of the facts in an effort to collectively choose a certain agenda and interpretation of facts. These collective processes involve not only the elites but also everyday citizens. Social media offer different mechanisms that facilitate collaboration among individuals, in what we could describe as a process in which the narratives are constructed by many citizens independently, yet in a connected fashion through networked structures (Meraz and Papacharissi, 2013). We can also refer to network structures to describe what is known as gatekeeping, i.e. the process of selecting which noteworthy events are allowed through the gates of a medium and into the news stream (McQuail, 1994). While in the mass media, gatekeeping is an organic function that is performed by a limited number of professionals, on social media this process is totally distributed. Thus, the items to be included in the narratives are also chosen, filtered and disseminated through the collective filtering process through which citizens and elites determine the importance of information when interacting with the mainstream and alternative media and different kinds of audiences (Papacharissi, 2015).

Despite the apparent richness of new media and its expressive resources, it is important not to lose sight of the fact that software and, therefore, social platforms channel communication, publication and interaction through action grammars integrated into the software, means by which actions might be compounded (Agre, 1994). This has two important consequences: a) the participation of users is structured and subject to a limited number of operations and b) precisely because of this formalization, user interactions are integrated into the database in a logical and organized manner. Thus, a social

platform like Facebook is organized from instances or predefined items (profiles, pages, posts, groups, URLs, likes, comments, shares, etc.) and connections between them. It is through these items that the expression and exchange of information is formalized and channeled,

SOCIAL PLATFORMS AS A SOURCE OF DATA AND METHOD

While users create, share and comment on social platforms, a large flow of data is created, which is not solely limited to what we see onscreen on different user devices. The flow of data runs in two different layers: the front end, where users interact, and the back end, where only platform administrators and, to a lesser extent, other computer applications, interact. The dynamics of each of these layers do not follow the same rules, as illustrated by the fact that while the democratization of access to the communicative space is clear in the front end, in the back end access is much more restricted. Thus, the data flowing in the invisible layer offers valuable information on the user, which can be used to articulate the entire social apparatus of the platforms (likes, shares, comments, etc.), while it can also generate financial value (Gerlitz and Helmond, 2013; Langlois and Elmer, 2013; Puschmann and Burgess, 2013). The accumulation of a user's digital traces has a clear impact on the entertainment industry, marketing, advertising, security forces and, as the Snowden case has spotlighted, international espionage, among others. However, at the same time it means that these data can be examined by the social sciences to inquire into the forces that shape our imagination (Latour, 2007).

The volume of information generated through these means is very high since it includes a large number of users and a high level of activity. Facebook, for example, has more than 2 billion active users per month (Facebook Inc., 2017) who generate a vast amount of data by creating content, liking and commenting. Beyond the size and dimensions hinted at by these indicators, it is even more important to stress the granular nature of the information contained in social media, which translates into the existence of a wide variety of objects and items. For example, we can distinguish between the contents published by users (photos, texts, URLs, etc.) and the data generated by the interactions within the different platforms (reactions, shares, comments, etc.).

The large critical mass of users and the granular nature of the data produces a large amount of information. This explains why we often refer to it as 'big data', an expression that has recently risen to prominence and is now trending in both digital marketing and academic circles. Traditionally, 'big data' has been used to refer to datasets that are so large that they require special computers and programs to be processed. Today, making this distinction based on the resources that are mobilized is somewhat unclear because standard computers and programs can be used to analyze large amounts of data (Boyd and Crawford, 2011).

However, the interest for the social sciences does not lie in the volume of information contained but in its capacity to represent social and cultural processes and feelings that millions of people around the world express on social media through comments, photos, videos, articles in blogs, maps, recording lists, etc. While research of culture and the social sciences in the 20th century was based on surface data obtained using quantitative methodologies and on deep data from qualitative studies, big data offers the chance to blur this boundary or at least not to have to choose between the size and depth of analysis and instead to use vast amounts of data to examine issues that are closer to qualitative studies (Manovich, 2011). In any case, the limitations of Facebook and any other social platform should be taken into account in terms of their social and cultural representativeness since access to technology and the knowledge necessary to participate in social networks is not yet universal. Likewise, there are people who despite having the necessary resources decide not to participate in social networks. If both groups played a more active role, the narratives on a certain issue might be substantively different.

This is an opportunity to reconsider the antagonistic relationship between those who follow the research methods of the natural sciences and those who lean towards interpretative or hermeneutic disciplines. Traditionally, the latter have criticized the former because their methods erase particularities, details and subtleties. Meanwhile, the former criticize the latter because their observations are not generalizable. The digital traces left by human beings on social media are objects that we can use to quantify elements that until now would have been placed in the qualitative realm and vice versa (Latour, 2010). This methodological dichotomy is also overcome when questioning the two levels of analysis of quantitative approaches (the aggregate of elements) and qualitative approaches (the individual) since by situating ourselves at the individual level as a point of departure, we can trace and display a social phenomenon without resorting to the aggregate level (Latour *et al.*, 2012).

Aside from these approaches, the task of inquiring into social factors using large volumes of data, either at the individual or the aggregate level, requires us to use methods from computer science to gather and analyze data that are extensive not only in volume but also in the breadth and depth of the phenomena they represent. Thus, when we are faced with the challenge of researching in a digital environment like social media, we must use techniques, tools and procedures from computer science in at least two parts of the process.

The first is obtaining vast amounts of data in an automated manner, which ostensibly facilitates the process of gathering data and building the corpus of analysis. The second is analysis of these data, a process that entails exploring, organizing, cleaning and providing the data obtained with the right format in the extraction process, as well as inferring knowledge from these data. This need leads to the emergence of a computational social science in which teams made up of social scientists and IT ex-

perts work together to further our understanding of our lives and society (Lazer *et al.*, 2009). This approach adds a series of instruments to the repertoire of tools available to researchers that go beyond the translation of statistical methods of quantitative studies, among other reasons because the data are not organized into a matrix of variables and cases but instead into a structure that is more similar to a relational database in which the connections between objects define the models of analysis (Giglietto *et al.*, 2012).

This kind of research also has other consequences that cannot be ignored: the epistemological change that it entails, the false appearance of the objectivity and accuracy of the type of quantitative studies that can be performed, the fact that a larger amount of data is not equivalent to higher quality, the difficulty of shifting models of analysis from one platform to another, ethical issues on how sensitive data should be treated, and finally the computer resources and knowledge that is harnessed when applying these methods, which can create a new digital divide in the sphere of research, since the possession of computing knowledge and power becomes a limitation that is difficult to overcome (Boyd and Crawford, 2011). At the same time, we cannot ignore the fact that despite the large amount of information that is accessible via these tools and mining techniques, only the corporations that are behind each of the social media platforms – and probably some state espionage services – have access to all of the information generated by their users (Manovich, 2011).

DIGITAL METHODS

We previously discussed the specific features of new media based on their status as digital objects and computer processes that intervene in the process of producing and disseminating contents. Beyond identifying their unique features in order to categorize and define them more precisely, their existence opens the door to methodological approaches that more intensely exploit the specificities of new media and their status as digital objects. This is the case of Digital Methods, an initiative based at the University of Amsterdam which, instead of drawing on common research methodologies in the social sciences and adapting (digitalizing) them to study new media, proposes to draw from digital research methods that repurpose the specific computer data and social media processes to study the medium based on its objects, formats, devices and platforms (Rogers, 2013). Digital methods are aligned with Science, Technology and Society Studies (STS): they both assume the existence of a technological mediation of social and cultural practices.

Rogers contrasts ‘natively digital’ methods from ‘digitized’ methods. An understanding of the former is based on the analysis of objects, contents, devices and environments born in the digital environment. Digitized methods are those that are based on the migration of standard practices from the social and human sciences to the new medium, the observation of browsing habits, the management of surveys and interviews over the Internet.

Meanwhile, the analysis of the results of a search engine, links and content published on a social platform are ‘natively digital’ methods.

In other words, the specificity of new media also becomes a specificity of the method through the use and study of digital-native objects like websites, links, search engine algorithms and social media on the Internet and the processes that articulate them. Thus, Digital Methods suggest appropriation of the methods and processes of new media and making them useful for performing research not *about* social media but *with* social media.

Follow the methods of the medium as they evolve, learn from how the dominant devices treat natively digital objects, and think along with those objects treatments and devices so as to recombine or build on top of them. Strive to repurpose the methods of the medium for research that is not primarily or solely about online culture (Rogers, 2013)

In parallel to the opportunities offered by these methods, there are also risks, such as the introduction of methodological assumptions that are foreign to social research and the limitation represented by the fact that these data are preformatted according to the operative needs of the platform from which they are obtained, which often entails predetermined analytical bias (Marres and Weltevrede, 2013). Certainly, the status of digital-native methods is a factor that limits the research questions we can tackle and the conclusions we can draw from them, given that it is difficult to transfer them outside of the digital environment. To this, we must add the fact that not all data are accessible and that the most relevant information for the researcher is often beyond our reach. In addition, the treatment and analysis of these data often requires a level of technological skill – and even access to computer resources – that is not available to research teams in the field of social sciences and humanities. Finally, the phenomena that we want to investigate might not resonate with the necessary force on social networks.

The use of digital methods, therefore, enables digital platforms and devices to be used to conduct research in the social sciences, taking advantage of the analytical and empirical possibilities of the new media which thus also participate in the research process. This idea, research as a task redistributed among different actors, gains momentum in the context of digital methods, since in addition to digital devices and platforms, users can also participate by labelling, providing visibility to and, in short, generating the corpus of analysis and the interpretative framework. Viewing research as an activity that involves a broad range of actors, such as researchers, research subjects, financiers, infrastructures, amateurs, etc., refers us to the discipline (STS), which has repeatedly upheld the notion of scientific research as a shared pursuit by different actors (Marres, 2012).

This paper is based on Netvizz, a tool developed for the extraction and analysis of Facebook data to investigate social and cultural phenomena with digital methods

(Rieder, 2013). Netvizz is a Facebook application and, as such, it uses the Application Programming Interface (API), a development environment designed for other computer programs to be able to operate with Facebook by following a protocol that regulates the conditions and methods of access to the metadata that move through the layer that is invisible to users, whereby Facebook and other platforms allow third-party applications to post contents, offer new functionalities, articulate commercial services and, of course, access the data for research purposes (Giglietto *et al.*, 2012). Generally, an API specifies from a technical, legal and logistical point of view: what data can be collected (some fields are inaccessible or incomplete), how much data (most APIs have limits); what temporary window (there are usually access limits) and the ‘subjectivity’ of the sample (privacy parameters and personalization can bias access to the data).

As we have explained, Netvizz is an extraction software that enables data to be collected in an automated and non-intrusive manner. This process is carried out from some of the areas into which Facebook organizes its contents: groups, pages, URLs and events. For each of these items, it is possible to obtain different data and reports that, to date, serve to structure the extraction modules:

- Group Data. This grants access to the posts published in open groups together with engagement statistics, user comments and network files that connect users, interactions and posts.
- Page data: This offers the possibility of downloading the published posts accompanied by engagement statistics, user comments and network files that relate users, interactions and posts.
- Like Network: Provides a network that shows the relationships between pages.
- Page Timeline Images: This provides a file in which the images published on a page are displayed along with engagement statistics.
- Search: This offers the possibility to search pages, groups and events, based on keywords, and download the results.
- Links: This can be used to obtain the number of shares of a list of URLs.

RESEARCH QUESTIONS

The aim of this article is to explore the possibilities offered by digital methods to retrieve and analyze content created by Facebook users in order to reconstruct the narrative and explanatory frameworks that are created and disseminated through this social platform. To do this, we will take the representation of the Spanish Civil War as a case study and we are going to use an analysis and extraction software (Netvizz) that will serve to exemplify and propose different techniques to investigate the processes of creation, dissemination and representation of historical knowledge on Facebook. Our intention, therefore, is not to deepen the case study or exhaustively reconstruct the narratives of the Spanish Civil War, but to offer tools for

scholars in the field of human sciences to use digital methods to carry out their research.

To accomplish this goal, this study will be based on three research questions:

RQ1: How can we explore and detect the spaces in which different narratives unfold around a topic on Facebook?

RQ2: How can we extract and collect the different elements that make up the narratives on Facebook?

RQ3: To what extent do the data extracted from Facebook help us to reconstruct and analyze these narratives

EXPLORING THEMATIC SPACES

Search Interface

The Search module is a very useful tool to start exploring a specific thematic ecosystem of pages and groups. It “provides an interface to Facebook’s search functions for pages, groups, places and events” (Netvizz, 2017) in which one can enter a query composed of one or more words and select the type of Facebook object we are looking (for -pages, groups or events). After this request is made, Netvizz provides information on the match criteria, and will show these results on an HTML table and a downloadable tabular file (.tab) that can be opened using spreadsheet software. Both types of output allow at-a-glance exploration of a list of match results with additional fields of information such as category (Facebook provides different pre-established categories that can be selected by the administrator of the page to describe the topic), description (a short text used to describe the content or purpose of the page or group), fan counter (the number of users who follow each fan page) or privacy (a field only available for groups that define whether it is open or closed administrator validation is required to join the group). Interestingly, the Search function can retrieve results not only for public groups, but also for private ones.

Examples

If we take as a case study the presence and representation of the Spanish Civil War on Facebook, it is easy to see how this module can be used to discover and explore pages and communities that contribute to the presence, interpretation and knowledge of this historical event. So, if we use the search string “Guerra Civil Española” (Spanish Civil War) to locate pages that deal with this topic, we quickly obtain a list of 70 results that match the search criteria. The table 1 reproduces only the name, category and number of fans (leaving out the rest of the information fields) of the 10 pages with the most followers:

In comparison to Facebook’s user search interface, the Search Module is a more effective method to explore how presence around an issue is articulated, discard all irrelevant results and identify and make first contact with the most relevant social spaces on a certain topic. When using the

TABLE 1. Top 10 of the most followed pages.

Page name	Category	Fans
Arqueología de la Guerra Civil Española	College & University	11817
Plataforma Memoria Histórica - Guerra Civil Española	Nonprofit Organization	9498
Guerra Civil Española	Interest	2685
Guerra Civil Española - Spanish Civil War - Guerra Civile Spagnola	Community	1157
La Guerra Civil Española/ The Spanish Civil War	Community	987
GRH Guerra Civil Española	Education	745
Archivo General de la Guerra Civil Española	Landmark & Historical Place	515
Red de Museos y restos de la Guerra Civil Española	Museum/Art Gallery	476
Hispanoamericanos en la Guerra Civil Española	College & University	410
Guerra Civil Española	Community	240

search module, it is important to take into account the way in which Facebook's search interface makes its choices, since it only returns as results those pages whose information fields contain the exact words that make up our search string. So, if we want to complete our search and make it more exhaustive, it is advisable to repeat the operation with small changes, as would be the case of "Guerra Civil Española" (Civil War Spain), which returns 6 additional pages, or make radically different searches. For example, if we use words referring to a more specific episode, such as "batalla ebro" (Battle of the Ebro), we get 74 new pages and "exilio republicano" (republican exile) provides 8 additional results. Likewise, resorting to more generic queries such as "memoria histórica" (historical memory) would allow us, after filtering out the pages that are not related to our case study, to expand the sample. It is therefore crucial to have an in-depth knowledge of the case study in order to develop a sufficiently complete list of keywords to provide a more complete view of the ecosystem of pages and groups that are articulated around the chosen study object.

Page Like Networks

Facebook allows pages to like other pages, thus creating networks of contents related to each other based on themes, interests, causes, etc. If we consider that the act of liking a certain page has a certain cultural significance, these networks allow us to observe both shared interests and tastes and ideological and organizational connections (Ben-David and Matamoros, 2016). In addition to analyzing the nature and composition of these networks, this module can be used to explore and discover new pages that integrate the thematic ecosystem studied and increase the variety and exhaustiveness of our data collection.

To start using this functionality, one has to select a page that will serve as a starting point to retrieve all the liked pages predefining the crawling depth. By default, depth is set to 1, which means that pages liked from the seed page will be collected. By setting crawl depth to 2, the retrieval process has additional reach, also gathering pages liked by pages liked by the seed page. The greater the depth, the

longer the extraction time, the amount of data and the complexity of the interconnections. The output is a network file in .gdf format, where nodes represent unique pages, and edges are the relation (in this case, a like) between two nodes. These network files can be processed with a graph display software, such as Gephi, a software for analyzing and displaying networks (Bastian *et al.*, 2009). This type of display allows for contextualization and analysis of a page in an ecosystem of social spaces that can be connected through different structures and patterns.

Examples

In Figure 1, we can observe the 'like' network graph taking as a seed the page "Guerra Civil Española – Spanish Civil War – Gerra Civil Spagnola"¹ with crawl depth 1. Nodes are unique Facebook pages, and edges are created between two nodes when the seed page 'likes' another page.

If the same operation is repeated with crawl depth 2, a network with 228 nodes (pages) with 711 connections (likes) emerges. Figure 2 shows how the network of interconnected pages now constitutes a richer and more complex ecosystem of pages. In the display, the colors serve to show communities of pages that are more densely connected to each other (Blondel *et al.*, 2008). The size of the nodes and the titles of the pages reflect the number of times a page has been liked by another page in the same network, thus signaling its referential character. To distribute the nodes in space, we used the spatial distribution algorithm Force Atlas 2 that organizes the elements, so that the edges act as a force of attraction and the nodes as a force of repulsion (Jacomy *et al.*, 2011), so that the most interconnected pages are grouped and structural features and patterns can be observed.

In the network that concerns us (fig. 2), there are two large groupings of pages with 4 easily distinguishable communities. On the left we find pages of international scope such as 'Humans Rights-Watch', 'Democracy Now', 'Abraham Lincoln Brigade Archives', etc. On the right, at the state level, we find, for example, 'Humor Indignado 99%', 'Revista Mongolia', 'Spanish Revolution', 'Diario Público', etc. Going into details of the communities, note

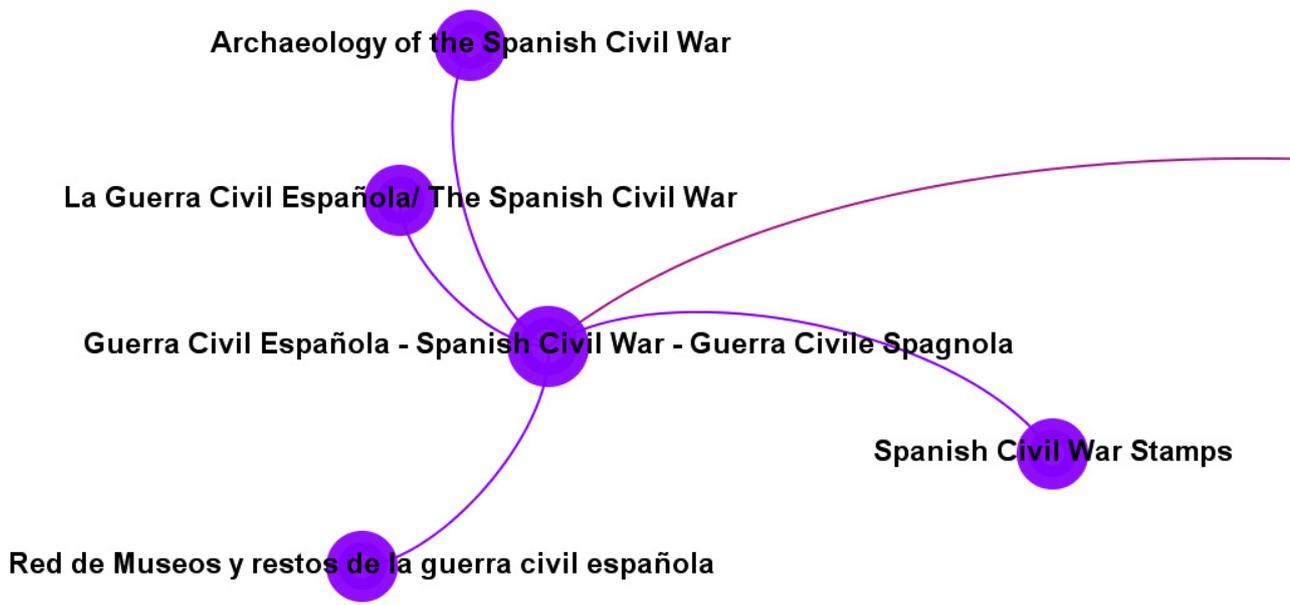


FIGURE 1. Depth 1 like network. Source: Own Elaboration.

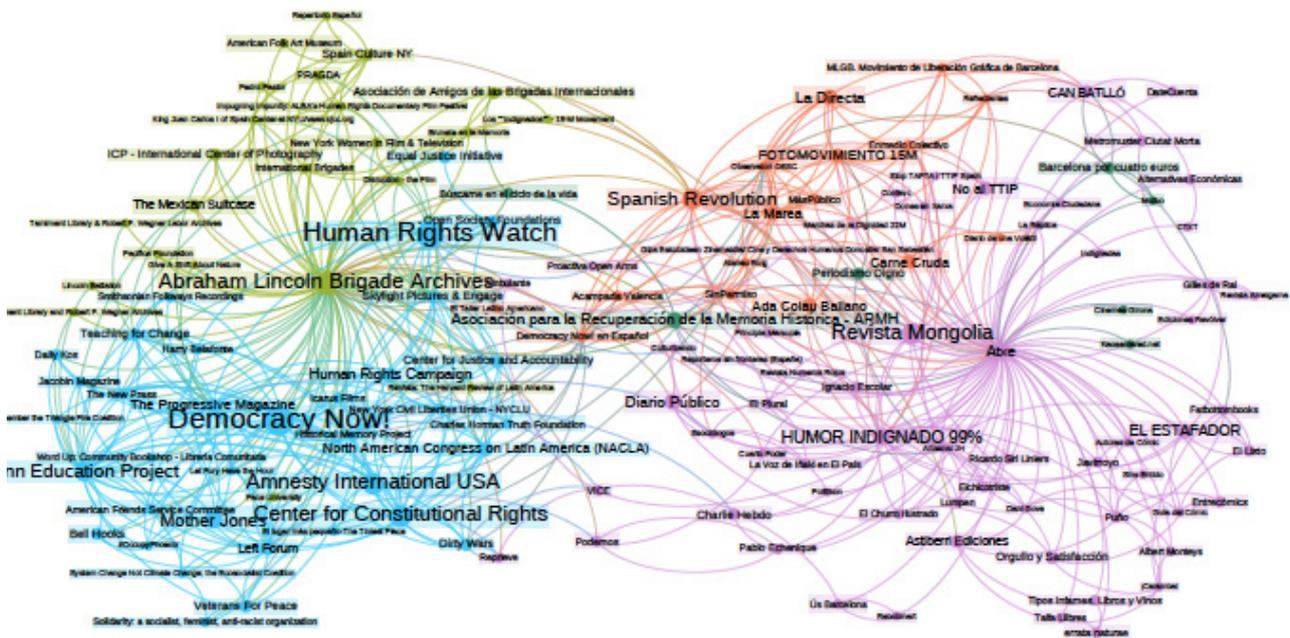


FIGURE 2. Depth 2 like network. Source: Own Elaboration.

how the blue agglutinates non-governmental entities as opposed to civil rights. The green also brings together organizations from the non-governmental sphere but this can be with a more direct link to the memory of the Lincoln Brigade in particular. At the other extreme, the pink identifies a media network close to the left wing of the Spanish political spectrum and the orange identifies movements close to activism and to alternative media such as La Directa. It is also true that the deeper the crawl, the more probable the appearance of contents that are thematically different

from the narratives of the Spanish Civil War, and also that this type of network has the virtue of showing us actors who share common interests, agendas and ideologies and that on some occasion, or more consistently, participate in the narrative of certain issues of the Spanish Civil War.

CONTENT ANALYSIS

As we said before, Netvizz is structured in different extraction modules for different Facebook entities. In-

stead of describing one by one the type of information that we can access in each of the modules, we have structured this section from 3 objects that we consider especially relevant to perform content analysis: posts, images and comments.

Posts

The Page and Group Data modules enable automated extraction of content and metadata published on Facebook pages and open groups. The module request will result in some Tabular separated values (.tsv) files for analysis, which can be opened with spreadsheet software. Some API restrictions come into play:

- Groups must be open, otherwise it is impossible to drain data.
- Technical limitations: The extraction process may collapse if extraction is too big and the machine runs out of memory.
- A maximum of 999 posts can be tracked.
- Users will be made anonymous when seeking for comments in posts.
- For Page and Group extractions, Facebook permits grabbing of the 600 most recent posts in a given year.

In order to filter the amount of data that will be requested by the API, and prevent application collapse, it is possible to limit the request to a specific number of the latest posts, or to only extract posts between certain dates.

As regards content, we recover the text, images, links, etc. that make up each post. In the metadata section, we obtain additional information about the content, such as the date of publication, the number of reactions, shares, comments and also a classification based on the publication formats available on Facebook. Thus, the ‘status’ category corresponds to 100% textual content, ‘photo’ to text plus image, ‘link’ to text accompanied by a link, ‘music’ to an audio file plus text, video to audiovisual content accompanied by text, etc. This information can be especially useful for content analysis based on the formal qualities used in each page for the deployment of one’s narrative.

The number of likes, reactions, comments and shares can also be very useful for analyzing the contents. These indicators represent the engagement, understanding as such the interactions of the users with the contents that constitute the mechanism of publication and visibility of the platform. We should not lose sight of the fact that each of these ‘social’ interactions are something more than indicators of popularity, they all have consequences within the platform: likes and comments activate notifications, and the shares are published on the users’ walls. And also for the user, the engagement entails a cognitive and/or affective link with the content and its topic (Mollen and Wilson, 2010). With the introduction of reactions in February 2016, Facebook gave us the opportunity of knowing about the state of mind of our audiences, which

is qualitative information that allows researchers to introduce new variables, something that goes beyond the traditionally neutral ‘like’. In fact, since this feature was introduced to Facebook, the surface content (feed) algorithm has been giving more visibility to posts with reactions, assuming that a reaction implies greater attention to the content and a decision made by the user and, therefore, should have more value than a mere like (Stewart, 2017).

Another option that Netvizz offers to analyze how users interact with content is the network files that relate posts and users based on reactions and comments. This functionality makes it possible to obtain a network file that can be analyzed with Gephi. With this graph, it is possible to see how users interact with a certain post, and detect behavior patterns in comments. For example, whether certain users only comment on certain topics, revealing greater interest, generating opinion or defining the explanatory framework of shared content.

Examples

Table 2 reflects the format used in the last 500 posts published on the “Arqueología de la Guerra Civil Española”²² page, a page that stands out for publishing content related to the topic on a regular basis. It is clearly observed that the preferred publication format is the ‘link’ (text + link), followed by a ‘photo’ and that plain text and video, on the other hand, are used much more residually. These results clearly show that the main activity of this page is articulated around the exchange of information, but they also suggest a special role of images in the construction of the narratives of the Civil War that could be investigated by analyzing the content of the posts more exhaustively.

Table 3 shows the contents that have generated most interaction on the same page “Arqueología de la Guerra Civil Española”. Here, the beginning of the textual content of the post is reproduced together with the figures of the engagement metrics, so it is possible to focus the analysis on the contents that have been most prominent in the narrative that this page dedicated to the recovery of the memory of this historical event.

As well as contributing to the dissemination and visibility of the contents, reactions provide special information on the emotions and feelings expressed by the users of the platform and the characteristics of the content. Allowing us to analyze different types of stories based on the kind of emotions they arouse. Table 4 reproduces the 5

TABLE 2. Number of posts by format.

Format	Publications
Link	306
Photo	185
Status	3
Video	6

TABLE 3. Top 10 of the most reacted content.

Text	Comments	Shares	Likes	Love	Wow	Haha	Sad	Angry
Hoy hemos recibido la visita de la policía...	47	94	307	6	22	1	4	79
Encontrado un arsenal de la Guerra Civil en un convento...	24	107	251	23	32	2	0	1
Un crucifijo en las trincheras republicanas...	26	34	252	7	10	0	5	0
Hoy hemos estado en Madrid con uno de los dos últimos brigadistas vivos...	16	51	221	40	0	0	0	0
Arqueología de las torturas de ayer arqueología de las torturas de hoy...	6	125	241	6	2	0	6	1

TABLE 4. Top 5 of the most sad posts.

Image	Text	Reactions
	CÁRCELES Y ASESINATOS COLECTIVOS DE PRESOS A partir del 18 de julio de 1935 en la parte de España que quedó dominada por el Gobierno del Frente (...)	15
	“EDIFICIO INCAUTADO POR LA GENERALITAT”. LA REPRESIÓN EN CATALUÑA CONTRA LOS CATÓLICOS EN TIEMPOS DEL GOLPISTA COMPANYYS. “Iglesia incautada” y “Edificio propiedad del estado”. En español y en catalán. Dos carteles en la fachada de una iglesia catalana que no dejan lugar a dudas sobre la actuación represiva de las autoridades republicanas (...)	10
	Barruecopardo (Salamanca) a sus Caídos. La placa ha sido eliminada el recuerdo a los asesinados quiere pero no puede ser borrado.	6
	29 de octubre de 2017 81er aniversario del asesinato de Ramiro Ledesma Ramos. ¡PRESENTE!	2
	Monóvar (Alicante) a sus Caídos. Hoy ya por desgracia desaparecido	2

posts that have caused most ‘sadness’ out of the last 100 to be published on the page “Familiares y amigos de los represaliados por la 2ª República (1931-1939)”³, which has almost 12000 followers. Of these 100 messages, we have filtered the images that are related to the memory of the Civil War. Finally, we wanted to find out which 5 images caused the most ‘sad’ reactions. Apart from clearly

viewing the ideological alignment of the page and its followers, it helps us to identify the discourse, themes and expressive resources that make up the narrative of a particular interpretation of the Civil War. In the case at hand, the imprisonment of opponents, the expropriation of a church and the murder of one of the ideologues of the Falangists in the Republican side. As well as photographs of

war memorials built during the Franco dictatorship that have been removed from public space.

Figure 3 is a network that relates users to the last 50 posts published on the “Arqueología de la Guerra Civil Española” page on the basis of their interaction. The size of the nodes reflects the number of interactions that each of the posts receives, thus signalling the most popular content. Edges are created when there is an interaction between two nodes, for example a comment, a like or a reaction to a post. The more interactions are made between two nodes, the greater the edge strength, and the closer the two nodes are. The colouring, as in the network in figure 3, identifies the communities that are created through the interaction. We observe how 3 large communities emerge, each represented by a different color that indicates how certain user groups especially interact with and express interest in a type of content. This type of analysis can be relevant when we want to know how communities are structured, the degree of user participation and the detection of those actors that favour interaction within Facebook groups and pages.

Images

In table 4 we have seen how visual content plays an important role in the construction of discourse in pages and groups. By extracting data from pages and groups, it is possible to recover images from posts. But if it is clear that the analysis to be carried out will be based on only the images on a Facebook Page, Netvizz has a specific module (Timeline Images) that enables the collection of visual content together with some engagement indicators. As with the posts, the type of reactions provoked by the images allows us to approach them on the basis of their emotional load.

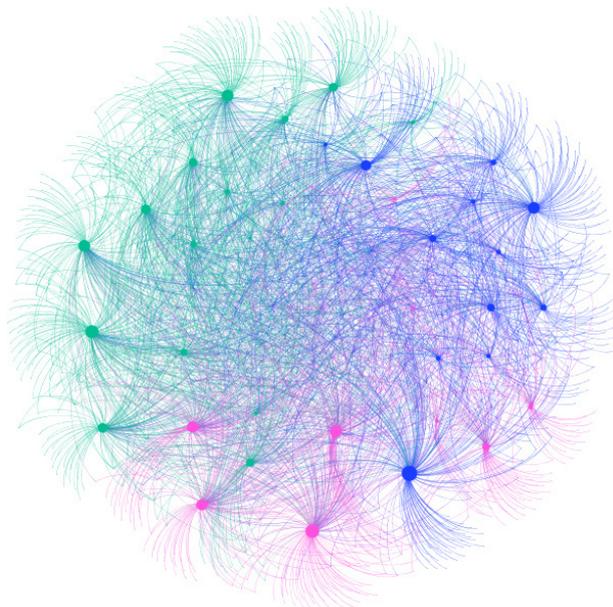


FIGURE 3. Post and users network. Source: Own Elaboration.

Examples

Table 5 was prepared by extracting Post data from the Facebook group “Guerra Civil Española”⁴, which has more than 4000 members and an average of 10 publications a day. It shows the three images with most love, sad and angry reactions from the last 500 publications. As is observed, the nature of the images is totally different in each of the reaction categories and it is easy to associate them with different narratives, whereby the images that accumulate the most hearts show a militiaman survivor of the Spanish Civil War, a Cuban militia volunteer in the war, and a propaganda poster to encourage women to perform surveillance work in the war’s rearguard. Meanwhile, sadness and anger are concentrated on photographs that show scenes of violence and repression during and after the armed conflict.

Comments

The posts published on pages and groups allow users to leave comments that can be used to complete, criticize or applaud their content. In this regard, users’ participation must be taken into account as an element that extends and gives depth to the narratives that are developed in these spaces. Netvizz facilitates the extraction of content (text, stickers, emojis, etc.) from these comments in a format compatible with spreadsheet software so that it can be analyzed using different techniques and methodologies.

TABLE 5. Top 3 images with more love, sad and angry reactions.

Love 1	Love 2	Love 3
Sad 1	Sad 2	Sad 3
Angry 1	Angry 2	Angry 3

Examples

In creating the following example, we used the group “Guerra Civil Española”, from which we have downloaded the latest publications and their respective comments. There are different tools that can be used to count the repetitions of words in a text and eliminate words without a relevant meaning, such as adverbs, articles and prepositions. By applying this technique to 900 comments left by the members of said group, it is possible to display a tag cloud in which the size of each word reflects its prominence in the comments and allows observation of the dominant tone of the conversation generated around the words. Obviously, with this operation we lose much of the context and meaning that only careful reading of the comments can provide, but it is very useful to make first contact with the themes and language present. In the case that concerns us, figure 4 refers to episodes such as the Battle of the Ebro and the locations where the events occurred.

CONCLUSION

As we explained in the introduction, our methodological approach is aligned with digital methods. Therefore, we have addressed Facebook, not only as a source of data, but we have also adopted the methods built into its API to extract and analyze the contents. For this, we have used Netvizz, a software specifically designed for academic research in the field of social sciences and humanities. This methodology has proved especially useful when it comes to exploring and detecting groups and Facebook pages in which contents related to the Spanish Civil War are published and play an important role in the representation of this historic event on Facebook. As we have observed interrogating the search function and visualizing the Page Like Network, Netvizz allows a quick identification of the discursive spaces on facts related to the Spanish Civil War. In addition, we can quantify the scope of the contents published in pages and groups based on the number of users that are part of these spaces.



FIGURE 4. Word cloud. Source: Own Elaboration.

The quantity and diversity of the spaces found allows, in an automated and non-intrusive way, to collect a large amount of information organized in a database format. The information fields in which the data are structured facilitate the creation of a data collection on which we can analyze content with traditional methods. The easy and fast access to information, the automated creation of a data collection structured in information fields that facilitate the analysis of content alone could justify the adoption of Digital Methods. However, the fact that the information collected goes beyond the media objects that make up the narrative (text, photos, videos, links, etc.) and includes the interactions of the users (reactions, comments, shares) offers new dimensions of analysis. This is the case of the characterization of the narratives based on the emotions expressed by the users, which we have been able to observe in the set of images and reactions shown in table 4 and table 5. Or the content of the comments, which we have exemplified in the word cloud in figure 4, where we observe the predominance of certain terms, and can identify the central frames of the debate that takes place in a Facebook group.

At the same time, resorting to digital methods involves assuming some limitations that should not be overlooked. In this sense, it is important not to lose sight of the fact that access to information is limited to the contents published in open groups and pages. Therefore, everything that is published on the timelines of personal profiles or in closed groups is invisible to our scrutiny. This fact is especially relevant since it means that, independently of how exhaustive our exploration of the thematic spaces and the process of data extraction is, the analysis and reconstruction of the narrative will always be partial.

Another consequence from the use of digital methods derives from the fact that the extraction and analysis software adds new layers of technological mediation when it comes to reproducing the narratives that are carried out on Facebook. In fact, extracting the contents and organizing them from information fields is, in essence, a process of deconstruction since it involves analytically disaggregating each of the elements that constitute the stories on Facebook. It is after this process that we proceed to represent and visualize the contents organized in tables, graphs and networks, that have the virtue of facilitating the analysis and interpretation of the narratives. But at the same time, they take us away from the visual appearance and the context in which they are represented and perceived by Facebook users.

NOTES

1. Facebook Page “Guerra Civil Española – Spanish Civil War – Gerra Civil Spagnola”: <https://www.facebook.com/guerracivil.espanola/>
2. Facebook Page “Arqueología de la Guerra Civil Española”: <https://www.facebook.com/guerraenlauniversidad/>
3. Facebook Page “Familiares y amigos de los represaliados por la 2ª república(1931-1939)”: <https://www.facebook.com/FAMILIARES-Y-AMIGOS-DE-LOS-REPRESALIADOS-POR-LA-2%C2%AA-REPUBLICA-1931-1939-373480367011/>

4. Facebook Group “Guerra Civil Española”: <https://www.facebook.com/groups/378972102146174/>

REFERENCES

- Agre, P. E. (1994) “Surveillance and capture”. *Information Society*, 10 (2), 101-127.
- Bastian, M.; Heymann, S.; Jacomy, M. (2009) “Gephi: an open source software for exploring and manipulating networks”. *ICWSM* (8), 361-362.
- Ben-David, A.; Matamoros-Fernandez, A. (2016) “Hate speech and covert discrimination on social media: Monitoring Facebook pages of extreme-right political parties in Spain”. *International Journal of Communication* (10), 1167-1193.
- Blondel, V. D.; Guillaume, J.; Lambiotte, R.; Lefebvre, E. (2008) “Fast unfolding of communities in networks”. *Journal of statistical mechanics: theory and experiment*, 2008, 10.
- Boyd, D.; Crawford, K. (2011) “Six provocations for Big Data”. *A decade in Internet time: symposium on the dynamics of the Internet and society*. Rochester: Social Science Research Network.
- Coromina, Óscar (2017) “Pugna por el relato en los contenciosos políticos. El caso del proceso participativo del 9N de 2014”. *El profesional de la información*, 26 (5): 884-893.
- Ellison, N. B.; Vitak, J.; Gray, R. and Lampe, C. (2014). “Cultivating social resources on social networking sites: Facebook relationship maintenance behaviors and their role in social capital processes”. *Journal of Computer-Mediated Communication*, 19 (4), 855-870.
- Ellison, N.; Boyd, D. (2013) “Sociality through social network sites. In W. H. Dunton”. *The Oxford Handbook of Internet Studies* (151-172). Oxford: Oxford University Press.
- Facebook Inc. (2017). *Facebook Newsroom*. Retrieved 10/November/2017 from Facebook: <https://es.newsroom.fb.com/company-info/>
- Facebook Inc. (2017). *Anuncios de Facebook*. Retrieved 29/november/2017 from Facebook para empresas: <https://www.facebook.com/business/products/ads>
- Fernández Peña, E.; Coromina, Ó.; Gila, J. M. (2017) “Nature of engagement on Facebook during London 2012 olympic games: insight inot public participation in terms of language and gender”. *South African Journal for Research in Sport, Physical Education and Recreation*, 39 (1:2): 135-151.
- Gerlitz, C.; Helmond, A. (2013) “The like economy: social buttons and the data-intensive web”. *New media & society*, 15, 1349-1365.
- Giglietto, F.; Rossi, L.; Bennato, D. (2012). “The open laboratory: limits and possibilities of using Facebook, Twitter, and Youtube as a Research Data Source”. *Journal of technology in human services*, 30 (3-4), 145-149.
- Hille, S.; Bakker, P. (2014) “Engaging the social news use”. *Journalism Practice*, 8 (5): 563-572. doi: 10.1080/17512786.2014.899758
- Jacomy, M.; Heymann, S.; Venturini, T.; Bastian, M. (2011). “Forceatlas2, a continous graph layout algorithm for handy network visualization”. *Medialab center of Research*, 560.
- Jenkins, Henry. (2006). *Convergence Culture*. New York: New York University Press.
- Kaplan, Andreas M.; Haenlein, M. (2010) “Users of the world unite! The challenges and oportunities of Social Media”. *Business horizons*, 53.1, 59-68.
- Kim, W.; Jeong, O. R.; Lee, S. W. (2010). “On social web sites” *Information System*, 35 (2).
- Langlois, G.; Elmer, G. (2013). “The research politics of social media platforms”. *Culture Machine*, 14.
- Latour, Bruno (2007) “Beware, your imagination leaves digital traces”. *Times Higher Literary Supplement*, 6.4.
- Latour, Bruno. (2010). “Tarde’s idea of quantification”. In M. Candea, *The social after Gabriel Tarde: debates and assessments* (pp. 145-162). New York: Routledge.
- Latour, B.; Jensen, P.; Venturini, T.; Grauwijn, S.; Boulier, D. (2012). “The whole is always smaller than its parts”. *British Journal of sociology*, 590-615.
- Lazer, D.; Pentland, A.; Adamic, L.; Sinan, A.; Laszlo, A.; Brewer, D.; Christakis, N.; Contractor, N.; Fowler, J.; Gutmann, M.; Jebara, T.; King, G.; Macy, M.; Roy, D.; Van Alstyne, M. (2009) “life in the network: the coming age of computational social science”. In *Science* 323(5915): 721–723.
- Lenhart, Amanda (2009) *Adults and social network sites*. Pew Research Center.
- Manovich, L. (2011) “Trending: the promises and the challenges of big social data”. In M. K. Gold, *Debates in digital humanities* (pp. 460-475). Minneapolis: The university of Minnesota Press.
- Marres, N. (2012) “The redistribution of methods: on interactions in digital social research, broadly conceived”. *The sociological review*, 60:SI, 139-165.
- Marres, N.; Weltevrede, E. (2013) ”Scraping the social? issues in live social research”. *Journal of cultural economy*, 6 (3), 313-335.
- McQuail, Denis (1994). *Mass communication theory: An introduction*. Sage Publications.
- Meraz, S.; Papacharissi, Z. (2013) “Networked Gatekeeping and Networked Framing on #Egypt”. *The International Journal of Press/Politics*, XX(X), 1-29.
- Mollen, A.; Wilson, H. (2010) ”Engagement, telepresence and interactivity in online consumer experience: Reconciling scholastic and managerial perspectives”. *Journal of business research*, 63(9), 919-925.
- Musial, K.; Kazienko, P. (2012) “Social networks on the Internet”. *World Wide Web* 16, 16 (1), 31-72.
- Netvizz (2017). Amsterdam: Digital Methods Initiative, Bernhard Rieder.
- Papacharissi, Zizi. (2015). *Affective Publics*. New York: Oxford University Press.
- Poell, T.; Abdulla, R.; Rieder, B.; Woltering, R.; Zack, L. (2015). “Protest leadership in the age of social media”. In *Information, Communication & Society*, 19 (7): 994-1014.
- Puschmann, C.; Burgess, J. (2013) “The Politics of Twitter Data”. In K. Weller; A. Bruns; J. Burgess; M. Mahrt; C. Puschmann, *Twitter and Society* (pp. 43-54). New York: Peter Lang Publishers.
- Rader, E.; Gray, R. (2015) “Understanding user beliefs about algorithmic curation in the facebook news feed”. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (173-182). Seoul: ACM.
- Rains, S. A.; Brunner, S. R. (2015) “What can we learn about social networking sites by Studying Facebook? A call and recommendations for research on social network sites”. *New Media & Society*, 17 (1), 114-131.
- Rieder, Bernhard. (2013) “Studying Facebook via data extraction: the Netvizz Application”. *Proceeding of the 5th Annual ACM Web Science Conference* (pp. 346-355). ACM.
- Rogers, Richard (2013). *Digital Methods*. Cambridge, Massachusetts: MIT Press.
- Stewart, Rebecca. (28 de february de 2017). *Facebook tweaks its algorithm to give more prominence to posts with “reactions”*. Retrieved 18 de november de 2017 from Business Insider: <http://www.businessinsider.com/facebook-tweaks-algorithm-to-give-more-value-to-posts-with-reactions-2017-2>
- Thelwall, Mike (2009). “Social network sites: Users and uses”. In M. (. Zerkowitz, *Advances in Computers* (pp. 19-73). Amsterdam: Elsevier.
- Van Dijck, Jose. (2013). “‘You have on identity’: performing the self on Facebook and LinkedIn”. *Media, Culture & Society*, 35 (2), 199-215.