

Article

# Networks and Stories. Analyzing the Transmission of the Feminist Intangible Cultural Heritage on Twitter

Jordi Morales-i-Gras \*, Julen Orbeago-Terradillos, Ainara Larrondo-Ureta and Simón Peña-Fernández 

Journalism Department, University of the Basque Country, 48940 Leioa, Spain; julen.orbeago@ehu.eus (J.O.-T.); ainara.larrondo@ehu.eus (A.L.-U.); simon.pena@ehu.eus (S.P.-F.)

\* Correspondence: info@jordimorales.com

**Abstract:** Internet social media is a key space in which the memorial resources of social movements, including the stories and knowledge of previous generations, are organised, disseminated, and reinterpreted. This is especially important for movements such as feminism, which places great emphasis on the transmission of an intangible cultural legacy between its different generations or waves, which are conformed through these cultural transmissions. In this sense, several authors have highlighted the importance of social media and hashtivism in shaping the fourth wave of feminism that has been taking place in recent years (e.g., #metoo). The aim of this article is to present to the scientific community a hybrid methodological proposal for the network and content analysis of audiences and their interactions on Twitter: we will do so by describing and evaluating the results of different research we have carried out in the field of feminist hashtivism. Structural analysis methods such as social network analysis have demonstrated their capacity to be applied to the analysis of social media interactions as a mixed methodology, that is, both quantitative and qualitative. This article shows the potential of a specific methodological process that combines inductive and inferential reasoning with hypothetico-deductive approaches. By applying the methodology developed in the case studies included in the article, it is shown that these two modes of reasoning work best when they are used together.

**Keywords:** feminism; hashtivism; Twitter; social network analysis; Machine Learning



**Citation:** Morales-i-Gras, J.; Orbeago-Terradillos, J.; Larrondo-Ureta, A.; Peña-Fernández, S. Networks and Stories. Analyzing the Transmission of the Feminist Intangible Cultural Heritage on Twitter. *Big Data Cogn. Comput.* **2021**, *5*, 69. <https://doi.org/10.3390/bdcc5040069>

Academic Editors: Manolis Wallace, Vassilis Pouloupoulos, Angeliki Antoniou and Martín López-Nores

Received: 18 October 2021  
Accepted: 17 November 2021  
Published: 24 November 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

This article is part of a broader research project dedicated to the analysis of social movements through digital conversations in social networks, taking as a reference certain public controversies of high impact in the online and offline public debate. Our methodological proposal is framed within the new research currents within the Sociology of Communication [1], which employ the digital footprint that millions of Internet users leave at the disposal of the scientific community through their interactions and actions. It is therefore a matter of using massive data and processing them through certain methodological processes to obtain information that helps the scientific community to describe and put into their interpretative context the social phenomena that take place around us, with the aim of better understanding the dynamics and changes in the logics of collective action. It is also a matter of understanding the consequences of these dynamics in the shaping of social movements, necessarily anchored in their own immaterial cultural heritage, and projecting themselves towards a future that each generation defines based on its own aspirations.

In this context, this methodological proposal offers a research perspective to the scientific community interested in social movements, the logics of collective action, contemporary public debate, and deliberative processes, among others. It does so, moreover, using the big data provided by a microblogging network such as Twitter, and with a method that not only describes, but also explains, interprets, and helps to understand how and why social networks are used and what effects and what social and democratic transformations they promote.

Social media are, after all, conversation tools of our society in the contemporary digital context and, as Castillo [2] argues, examining the conversation tools of a culture is an excellent way to understand it, and to understand its links with the past and with the future. The social media that have emerged alongside the web 2.0 have created spaces for communication and citizen participation that foster cooperation and mutual aid [3]. Those media are one of the main open mechanisms of public conversation, fundamental for the creation of the public agenda and deserving of in crescendo attention from the scientific community. Information technologies have given rise to what authors such as Dery [4], Joyanes [5] or Lévy [6] have baptized as “cyberculture” or “culture of connectivity” [7]. There is no doubt that the expansion of the main online platforms such as Facebook, Twitter, Flickr, Youtube, or Wikipedia reinforce the idea that contemporary society is facing a constantly evolving technocultural ecosystem and a phase of sociability that has online interaction as one of its main exponents. In such an ecosystem, meanings are permanently negotiated and reconsidered in a multilateral situation in which different generations participate and in which they reinterpret and construct themselves.

The methodological proposal contained in this article to observe and analyze public debate through digital network conversations reinforces the scientific production on the phenomenon that the sociologist Javier Toret calls the “connected multitude”. This is defined as “the ability to connect, group and synchronize, through technological and communicative devices, and around objectives, the brains and bodies of a large number of subjects in sequences of time, space, emotions, behaviors and languages” [8] (p. 23). According to Toret, this would be one of the many structural conditions in the Network Society [9]. The connected crowd, then, emerges in the new paradigm of Mass Self-Communication [10] in the Network Society and as one of the main characteristics of what researchers such as Melucci [11], Candón-Mena [12] and Romero [13] call “New Social Movements”. In this context, the demonstrations against the World Trade Organization summit in Seattle (1999), the Black Lives Matter movement, the Arab Spring, the Spanish 15-M, the movements fighting for degrowth or the feminist movements that are re-emerging in the new political, social, and communicative context.

In this organizational context, the concept of a “social network” ceases to be a metaphor and becomes pure metonymy, and therefore, all those that understand that the relationship between agents is the minimum unit of social analysis emerge as privileged perspectives of analysis: we will see that Social Network Analysis—or simply, SNA—is particularly fertile in these contexts. Activist networks or networks of social movements, sometimes defined with uncomfortably cybernetic references, are bundles of interactions, communicative and action spaces where experiences of struggle and self-organization are shared, where a certain reflexivity lives and a shared sense of protests is built through current and virtual dialogues with past generations that embody the different stages or waves of the movements themselves, thus managing their immaterial cultural heritage. Beyond a social morphology, networks have become a model for emergent forms of politics [14] (p. 92). In our opinion, this also applies to the politics of collective memory and of the intangible cultural heritages of political and social movements.

In this article we take this metonymic conception of social movements as social networks as our starting point. We intend, firstly, to present the main characteristics of the research with which our epistemology is connected, and secondly, to detail the methodological proposal that we have articulated in other research and make it available to the scientific community for discussion and improvement.

We are going to present a methodological proposal designed for the study of massive conversations in social media, through which to generate knowledge about a particular object of study, which is hashtag feminism and its importance for the configuration of the so-called fourth wave of feminism, understanding such a process of self-definition as an exercise of transmission and management of an immaterial cultural heritage [15]. At the heart of the proposal lies the will to contribute to the necessary hybridization between perspectives linked to Computer Science and Social Science, between Data Engineering and

content analysis, between quantitative and qualitative analysis techniques, and between inductive and hypothetico-deductive reasoning. We strongly believe, and we will try to argue, that such hybrid approaches are today more necessary than ever.

In this article we will focus on the following issues:

- We will argue the generational importance of hashtag feminism and the fourth feminist wave, and we will especially focus on explaining why working with Twitter data gives us access to a privileged vantage point from which to observe the dynamics of self-definition of the movement itself that are taking place during these same years.
- We will outline the main characteristics of big data as a socio-technical paradigm and highlight the opportunities it offers to social scientists who approach it with a hybrid analytical perspective: mathematically and technically solvent as well as phenomenologically grounded;
- We will also discuss the universe of technical opportunities and the legal limits we face when we want to work with massive data from social media to understand the dynamics that occur in them;
- We will present a particular methodological proposal based on Network Analysis and Machine Learning techniques. We will argue in favor of the implementation of a series of unsupervised algorithms to provide analytical context to big data, and then we will defend the articulation of Data Engineering techniques to facilitate further analysis;
- We will present four different investigations that we have already developed in the framework of the broad project that also gives rise to this article with a more methodological orientation. In these four investigations we have deployed the methodology we present, giving epistemological priority to the analysis of the context by means of inductive logics, without renouncing causal reasoning and hypothetico-deductive logic;
- Finally, we will elaborate a series of concluding reflections.

## 2. Objectives

### 2.1. *Analyzing the Shaping of the Current Feminist Wave through Twitter*

The aim of this article is double. On the one hand, we want to bring to the table a specific methodology for the analysis of Twitter conversations that can be applied to the study of the shaping of the contemporary feminist movement. This involves the assumption, in line with Deborah Withers' work on the politics of transmission of the feminist intangible cultural heritage in the digital age [15] (p. 5), that each feminist generation defines and generates itself through practices that involve the transmission of an intangible cultural heritage that connects and enables dialogue between generations. This is precisely what a metaphor as beautiful as that of the "waves" tells us when characterizing such generations. On the other hand, we would also like to present a series of empirical works that we have developed and reflect on them in these same keys of transmission of a feminist intangible cultural heritage.

This methodological proposal focuses especially on the most well-known microblogging platform at a global level, which is Twitter. This is so, among other reasons, because Twitter is the source of information that best allows segmenting users, discovering how citizens participate in the political debate and how they are grouped by ideological affinity [16]. Likewise, Twitter has become a consolidated medium for communicating issues related to politics, having since its birth in 2006 a growing importance in political contexts and having been used by virtually all actors interacting in the public-political space [17]. Five years after its creation in 2006, Rodríguez and Ureña [18] already pointed out Twitter as the social network that had acquired the greatest relevance among the political and journalistic class. For Piscitelli [19] (p. 15) at that time it also constituted "one of the most powerful communication mechanisms in history".

Subsequently, from various scientific perspectives authors such as Pariser [20], Page [21], Carr [22], Marwick [23], or Fuchs [24] lowered the most encouraging expectations around the use of social networks and social or political mobilization. As summarized by Giraldo-

Luque, Fernández-García, and Pérez-Arce [25] in their research on the mobilization that emerged around the hashtag #Niunamenos, Twitter is a means of dissemination and a space for expression around certain public controversies, but its scope for building consensus scenarios or transforming preconceived imaginaries is limited.

That said, Twitter has also been defined during its decade of existence as a space for social interaction, dialectical exchange, and as a sphere of deliberation in which much of the activism of social movements and contemporary social mobilization is [26–28]. In fact, this social network has aroused great interest in the academic community in recent years due to the specific type of conversation that takes place on it. Twitter is undoubtedly the most popular network for discussing political issues and current news, and has had a great impact on all the political and social mobilizations that have taken place in the world in recent years: from the Arab Spring to the Black Lives Matter movement that emerged during the pandemic following the spread of the COVID-19 virus in the U.S. For this reason, social and political movements have been a privileged object of analysis through Twitter data and SNA techniques.

In this sense, the scientific field attaches particular importance to the observation of the changing communication paradigm to further elucidate the dichotomy between the dichotomy of social media and their social function. To this end, big data from social media interaction is an immense source of information with great potential to explain social processes from multidisciplinary perspectives such as sociology, communication, or politics.

In the specific case of feminism, we believe that the analysis of conversations established on Twitter allows us to understand several dynamics that are established for the transmission of the feminist intangible cultural heritage, and even for the conformation of the “waves” that characterize the extension and temporal evolution of the movement. Several authors have already pointed out the importance of feminist hashtivism in shaping the fourth wave [28–30]. In this regard, over the past few years many studies have proliferated around the #metoo movement and its aftershocks beyond the initial scandals linked to the Hollywood film industry [31–34]. In our view, all these analyses and meta-analyses pivot around a series of generational phenomena that are intimately linked to the transmission of the feminist intangible cultural heritage, and even, to the controversies that can develop between generations of activists.

## 2.2. Related Works

The perspective we will develop in the following section is certainly innovative. However, it should also be acknowledged that we are also underpinned by a growing scholarly interest in social movements on Twitter, and more specifically, in feminism on Twitter. Several authors have already contributed to framing fourth-wave feminism as a connected or networked feminism, which was internationally raised by the strength of protests such as #MeToo [34,35]. There has also been a strong recent interest in the particularity of Spanish feminism on Twitter [36–38], which is the subject of several of the papers that follow.

A trend that has advanced in parallel to the academic interest in feminism is the interest in the social consequences of artificial intelligence. Here, a small yet increasingly important number of feminist articles around the concept of algorithmic injustice, data justice or data feminism are noteworthy [39,40]. In addition to this, there is a small group of research with which, in addition to sharing an object of study such as fourth-wave feminism on Twitter, we have important methodological links. This is research that uses Social Network Analysis to investigate relationships and discourse [41,42]. Undoubtedly, our methodological proposal should be considered within this general paradigm.

### 3. Methodology

#### 3.1. Big Data and Interpretative Perspectives

The kind of challenges that have shaped the big data paradigm have largely been technical and technological challenges. In his famous 2001 article—in which big data is not yet referred to as such—technologist Doug Laney [43] mentioned the three “Vs” (i.e., volume, velocity, and variety) that would become crucial in the field of data management over the next few years. All of Laney’s Vs referred to different technical aspects of data storage and processing infrastructures. Later, other authors [44–46] would go on to add more Vs to characterize the paradigm, such as “variety”, “veracity”, “validity”, “volatility”, “virtuality”, or “visualization”. It is at this point that the concept of “value” is presented as central, associated with the notion that data needs to be interpreted to generate return, whether economic or otherwise.

The predictions of some overconfident observers during the first decade of the 21st century invited us to think of a “post-analytical” world [47]. Instead, if anything has become clear over the last 20 years in reference to big data, it is that the analysis and interpretation of such data is a key aspect that can compromise the most sophisticated of automatic processing systems. Over the last few years, dozens of cases have come to light in which systems based on heavily automated massive data—many of them based on “black box” algorithms such as neural networks—have given rise to socially unacceptable situations. Among these situations or perverse effects, algorithms that reinforce human cognitive biases giving rise to echo chambers [48] or bubble filters [20], algorithms that discriminate socially vulnerable collectives [49,50] or, even, chatbots that acquire racist behaviors through community “training” [51] stand out.

Nowadays, large amounts of data flow through new channels becoming a valuable source of information [52]. At the same time, as evidenced by all the cases mentioned above, the most important and socially transcendent challenges faced by the big data paradigm are those related to the analysis and interpretation of data, and not so much to the technical capacity for its storage and processing. Such is the case that some of the most authoritative voices in the world of Artificial Intelligence [53] have already urged the community to abandon the use of black box algorithms (e.g., deep neural networks). These experts propose to redesign systems based on simpler and more transparent algorithms (e.g., regression or decision trees) that facilitate analytical and interpretative work.

This epistemological shift that is taking place among researchers and practitioners of big data, artificial intelligence, and data mining in general [54], represents a great opportunity for social scientists, and for communication scientists. The big data paradigm relies on enormously diverse data sources: hence the V for “variety”. Leaving aside exceptional sources such as genomic and biomedical data, meteorological, and environmental data, and some of the data from industry and mining, most big data is social, or has a large social component (e.g., financial, banking, GPS mobility, urban sensor, web browsing, e-commerce, or credit card consumption data). Among them, data from the so-called social media are particularly voluminous, as they come from a wide variety of user-platform and user-user interactions within the different platforms (e.g., posts, mentions, likes, swipes, or shares).

Social media data is a sociotechnological by-product generated jointly by platforms and users from the systematic recording of a series of interactions [55]. Therefore, given its interactive and relational nature, the most abundant data in social media is that which is easily computable as a matrix of relationships (e.g., mentions between users, friendship or follower relationships between users, or relationships established between users and content). This gives great centrality to structural analysis methods [56] such as Social Network Analysis (SNA). With a somewhat smaller but equally important presence, social media also includes data of an attributional nature (e.g., metadata associated with a post or a user). Unlike relational data, attributive data tend to be used in prediction and classification models using Machine Learning (ML) techniques [57].

It is around these two types of techniques (i.e., SNA and ML), that most social media data mining studies are framed, often combining aspects of both. These techniques are usually labeled as “quantitative” because of their mathematical and computational orientation. According to the view defended in this article, this is a more than questionable label, rooted in a dichotomy that is debatable to say the least (i.e., the difference between qualitative and quantitative perspectives). SNA has demonstrated on multiple and diverse occasions its ability to be applied as a mixed methodology [58,59]; on the other hand, ML is increasingly used as a supporting method in qualitative analyses, especially with data from social media [60,61].

In our view, both SNA and ML challenge the tension artificially established in Social Science between quantitative and qualitative techniques, inviting us to overcome this dichotomy. These techniques put on the table the need to articulate analytical strategies that combine the mathematical and computational rigor typical of quantitative approaches with the interpretative skills that characterize qualitative analysis. The type of perspective that we have tried to develop in the research reported in this article is intended to be a contribution to this way of understanding Social Science and big data.

### *3.2. Social Media as Relational and Textual Big Data Sources: Possibilities and Legal Limits*

Twitter is the social media with the most open data policy to date, compared to other platforms such as Facebook or Instagram. Twitter has a free API (Application Programming Interface) that allows data retrieval with a maximum of seven days of retroactivity, and allows, according to the information provided by the company on its website, real-time data capture, provided that no more than 1% of the platform’s global traffic is captured.

The data that the standard, free Twitter API can retrieve is quite extensive: tweets and retweets published, relationships between users, and even their metadata (e.g., their biography, number of followers or number of followers). As reported by the company itself, the standard API does not return 100% of the tweets issued, but it does return “the most important ones” since its API “is oriented towards relevance and not completeness” [62]. As such, the data we can retrieve from the free API represents an indeterminate portion of the total that, in principle, reflects the totality of the conversation very well. Twitter raises the possibility of acquiring 100% of the data and greater retroactivity in its payment plans.

Derived from these conditions of opportunity and the relational nature of the data that can be retrieved from the Twitter API, studies on Twitter using SNA techniques have proliferated during the second decade of the 21st century [63,64]. In this sense, as indicated above, social, and political movements have been a central object of analysis using Twitter data and SNA techniques.

It is possible to distinguish three different strategies of analysis through the conversations and digital interactions of this type of movements and other expressions of collective action developed on Twitter: (1) mention networks, (2) semantic networks, and (3) following relationship networks. All three types derive from a series of decisions that researchers make about the type of data to be represented in the graphs, and about the representation strategy itself. Likewise, the three types of analysis raise different possibilities to be transferred to the methodological processes applied to data captured in other social networks.

The first type consists of the analysis of dynamic relationships, formalized in networks of mentions, retweets, or replies between users [65–68]. This type of networks tends to be conceptualized as directed (i.e., the edges of the network have direction, they are emitted by a node and received by another node) and weighted (i.e., the edges of the network have weights, being able to represent a relationship of one or several mentions), due to the type of relationship they represent.

In general, these are networks with very low densities (i.e., most of the nodes in the network are not directly linked) and with very high “Modularity” figures obtained using the Louvain algorithm [69] (i.e., the communities reflect very strong intra-group association patterns, and very weak inter-group association patterns), which we will see in detail

later. Because of the type of data represented, this analysis can only be carried out on Twitter or other platforms where the mention-type relationship plays an analogous role: Mastodon, Gab, or Slack. This type of analysis is not directly transferable to networks such as Facebook or Instagram. On Facebook pages, it is normal to respond to the messages posted, and there is no analogous element to the retweet that is traceable between pages. On Instagram, people like rather than comment, and likes are not provided by the API at the level of each user.

Semantic networks, as a second line of research, have been explored in a complementary or alternative way to Topic Modeling algorithms [70–72]. Word networks tend to be conceptualized in an undirected and weighted way. That means that it is assumed, as a rule, that two words will co-occur symmetrically, or that they will be symmetrically linked, and also, that the number of times two words co-occur in a discourse is usually a relevant factor in the analysis.

The morphological characteristics of semantic networks are highly variable since they can represent different types of discourse with very different levels of lexical diversity. It is very common for semantic networks to be the result of a series of data processing operations using Natural Language Programming techniques, such as the segmentation or “tokenization” of a text (i.e., its division into words, sentences or paragraphs), the filtering of stopwords (i.e., the removal of particles that do not provide relevant information, such as articles, adverbs, or conjunctions) or “lemmatization” (i.e., the transformation of the words of a text into their canonical form, according to a pre-designed dictionary). In contrast to the previous case, this type of analysis is extremely versatile and transferable to any textual data source: social media, written and digital press, blogs, books, or scientific articles, among other cases of analysis.

Relationships between words, between sentences or between documents can be studied through SNA, which has yielded very good results in recent research [73]. To this end, several types of networks can be synthesized according to analytical needs: networks of words according to the number of times they appear together, networks of documents according to the number of words they share, networks of hashtags according to the frequency with which users have used them in their posts, and so on.

In any case, the most common approach to this type of analytical problems has been through heuristic rule processes from the fields of Natural Language Programming [74] or in combination with ML models [75], which usually imply a significant improvement in the predictive or classificatory capacity of such models. Due to the great complexity of the human language, black box analysis techniques such as embeds or embeddings have proliferated during the last few years [76]. These are deep neural networks which, as we have already seen, provide very good results in exchange for a great opacity in the internal processes of the algorithms. These technologies enjoy enormous popularity among computational scientists faced with problems such as word prediction in search engines or automatic text translation.

Finally, it is worth highlighting the third type of analysis, most likely the least employed, which is the one that consists of observing networks of established relationships and their effects or consequences [77,78]. The networks synthesized from the relationships established between social media users will be directed or undirected, depending on the platform (e.g., on Twitter they will be directed, since one user can follow another without being followed back by the other; whereas on Facebook or LinkedIn they should be undirected, since if there is no agreement between two users, they will not be “friends” or “contacts” on these networks), and, typically, they will be unweighted (i.e., it is not possible to follow anyone more than once on Twitter, nor to be friends with someone more than once on Facebook or LinkedIn).

In this type of analysis, a distinction can be made between egonets and socionets, which are fundamental categories of the SNA [79]. The first type of networks (i.e., egonets), in the context of social media, are those that reflect the links between a user’s followers or friends. The second type (i.e., socionets) represents the relationships established between

a group of nodes, without any of them constituting the center of the network. In the second case, the population of the network will have been designed according to some criterion external to the network itself (e.g., the network of relationships among the students of a course or among the journalists of a media outlet). The formal characteristics of the network will depend on the criteria according to which they have been constituted, although they will tend to be denser than the networks of mentions because of transitivity and homophily characteristic of personal networks: it is to be expected that someone's friends will end up knowing each other and establishing friendship as well [80]. Likewise, the Modularity figures derived from the Louvain algorithm will tend to be lower than for mention networks.

Twitter allows synthesizing egonets and socionets by retrieving data from its standard API. The other powerful networks, such as Facebook or LinkedIn, allowed egonets with their standard API before the Cambridge Analytica scandal [81]. Currently, these networks no longer provide these data, although they can be achieved through web scraping or web scraping techniques (i.e., techniques for the automatic extraction of data available on websites and social networks), increasingly popular, being used for a myriad of data mining operations (e.g., robots for indexing web content, flight, hotel, or insurance comparators, or for automatic alert systems). However, their legal basis is still somewhat unclear [82].

Web scraping can be implemented with completely legal tools, but its use may contravene the regulations of the social media platforms or websites from which the data is extracted. As a rule, therefore, these are operations that cannot be implemented by a logged-in user, but which the social media company will not be able to prevent if they occur from a non-logged-in user, since nothing that a social media platform makes available to a non-logged-in surfer can contravene the provisions of the data protection laws operating in the territories in which the platform operates. Although this is a swampy terrain with many issues, the type of judicial decisions that have been made over the last few years are favorable to web scraping of information available to non-logged-in users [83].

Following this doctrine, and if we stick to data that can be accessed by a non-logged-in user, web scraping is a very good alternative to API data access for research aimed at synthesizing semantic networks, or some hybrid models, such as networks between users and words, or between users and hashtags. One way or another, it will always be possible to apply web scraping techniques to obtain semantic data from social media, as well as from other sites on the Internet. In this way, the analyst will be able to rely on complete datasets rather than an indeterminate portion of the total and will have a greater temporal margin and retroactivity. However, web scraping is not feasible for research that focuses on the mention relationships between users of the major social media platforms, let alone their follower or friendship relationships. Obtaining this data is technically feasible, but it is necessary to violate the social media regulations, and in many cases, also the data protection laws in force in each territory.

## 4. Results

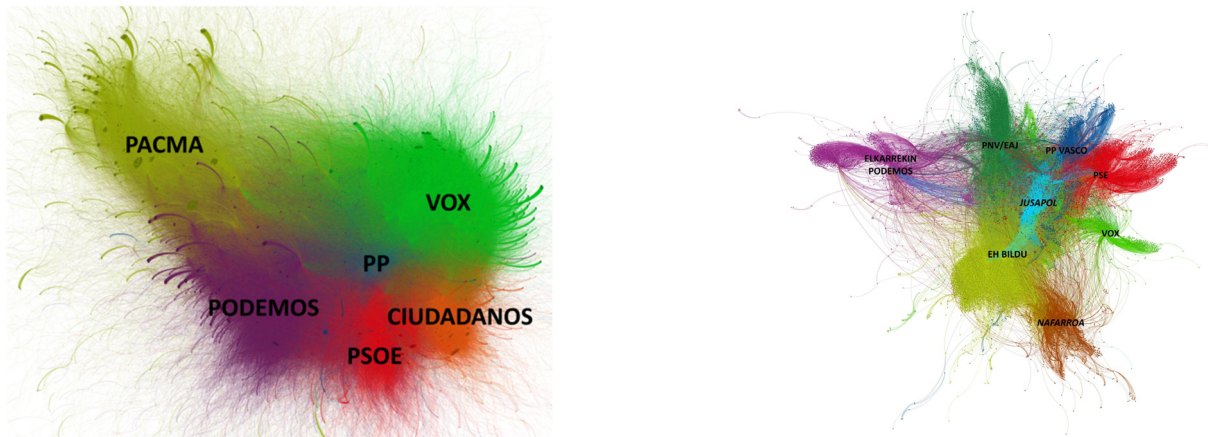
### 4.1. Network Analysis and Machine Learning as Assistants for the Interpretation of Dynamics in Virtual Networks

We have previously emphasized the need to articulate analytical and interpretative perspectives to overcome the artificial distinction between quantitative and qualitative analysis that has characterized social science in recent decades. The big data paradigm—and more specifically, techniques such as SNA and ML—exposes the obsolescence of this way of segmenting scientists based on their skill repertoires, while pointing to the need to generate new hybrid methodological frameworks that allow for simultaneous mathematical and phenomenological analyses.

In our opinion, one of the most effective ways to analyze and interpret the dynamics of virtual networks is to articulate SNA and content analysis techniques, through the development of workflows more typical of Data Engineering. This involves taking as a starting point the mention-type interactions (i.e., nominations of one user by another) on the social network Twitter, in the context of a series of digital conversations related to issues



of public and political debate, and then synthesizing networks or graphs from them (i.e., Figure 1). Thus, in the resulting massive graphs, each point or node represents a user of the Twitter social network (e.g., a personal account, a company, a media outlet, a political party, etc.) and each line represents an established mention from one user to another (e.g., a retweet, a reply, or a direct allusion). These are therefore directed and weighted networks, to which a series of algorithms are applied to generate value from the data.



**Figure 1.** Network of mentions to political parties on Twitter during an electoral campaign (April 2019, Spain) in two different spheres of influence. Source: Own elaboration with Gephi software.

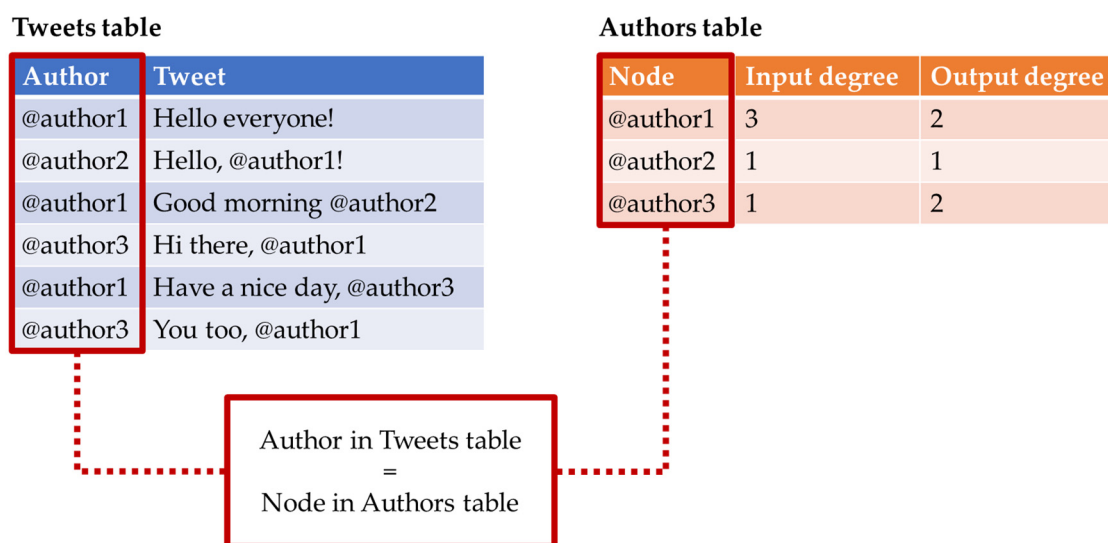
One of the most useful Machine Learning algorithms for a perspective such as the one detailed in this paper is the Louvain algorithm for community identification in massive graphs [69]. It is an unsupervised learning algorithm that performs a series of operations on the data in a recursive manner and allows the identification of clusters or sets of nodes that make up specific communities within a network. The process by which the Louvain algorithm identifies communities consists of randomly grouping nodes, and permanently evaluating the gain or loss of Modularity (i.e., a metric that evaluates the overall quality of the community partition of a network, comparing it with a randomly constituted network of equal size) [84] implementing only those groupings that result in gains in this metric (i.e., optimizing the quality of the community partition).

The output of the Louvain algorithm consists of a set of communities (i.e., a community partition) and a Modularity figure that allows us to evaluate its mathematical relevance. According to the creator of the Modularity metric, Mark Newman [84], values between 0.3 and 1.0 indicate a good quality of the community partition of a Network (i.e., it is assumed that the network is significantly different from the one that could have been constituted by chance). Despite their obvious similarities, modularity should not be confused with a hypothetical validation metric such as the “*p*-value” used in inferential statistics for the acceptance or rejection of the null hypothesis. Modularity is not used by the Louvain algorithm as a metric for hypothetical validation, but as an internal optimization mechanism. In other words, the algorithm is oriented to obtain the best possible Modularity figure. This feature, far from being a problem, is what allows the researcher to work with categories based on empirical data. This is a great example of how unsupervised algorithms facilitate qualitative readings of massive quantitative data.

In community identification, being an unsupervised process, the role of the analyst is not to train the algorithm to identify one or another type of groups, but to interpret the results of a node clustering process based on the patterns that the algorithm itself is able to identify in the data autonomously. Common SNA software (e.g., Gephi or Pajek) allows the analyst to establish community partitions at different resolutions [85], thus being able to choose between identifying more smaller groups or fewer larger groups. Thus, when it comes to analyzing social movements such as feminism itself, this type of approach allows conceptualizing the complexity of social identity (i.e., the diversity of

identifications available in the Self and its hierarchical structure) and intergroup relations in a privileged way and allows social analysts to move away from essentialist and reductionist conceptualizations [86].

In this analytical model, the cluster is the element that provides the context for the analysis of the rest of the data: the leaders in the network and its contents. Regarding the analysis of the leaders of a network (e.g., the most mentioned users, the most active in mentioning third parties, the best intermediaries, the ones that can most easily reach any other, etc.), it will be relevant to use metrics such as the input degree (i.e., the number of edges received by a node), the output degree (i.e., the number of outgoing edges) or the betweenness centrality degree (i.e., the number of shortest paths between pairs of nodes in a network that have to pass through each node). On the other hand, for content analysis, lists of tweets and lists of hashtags are elaborated by clusters, by means of Data Engineering strategies of crossing and combining data sources (i.e., Figure 2).



**Figure 2.** Example of combining data from two different tables. Source: Own elaboration with Power Point.

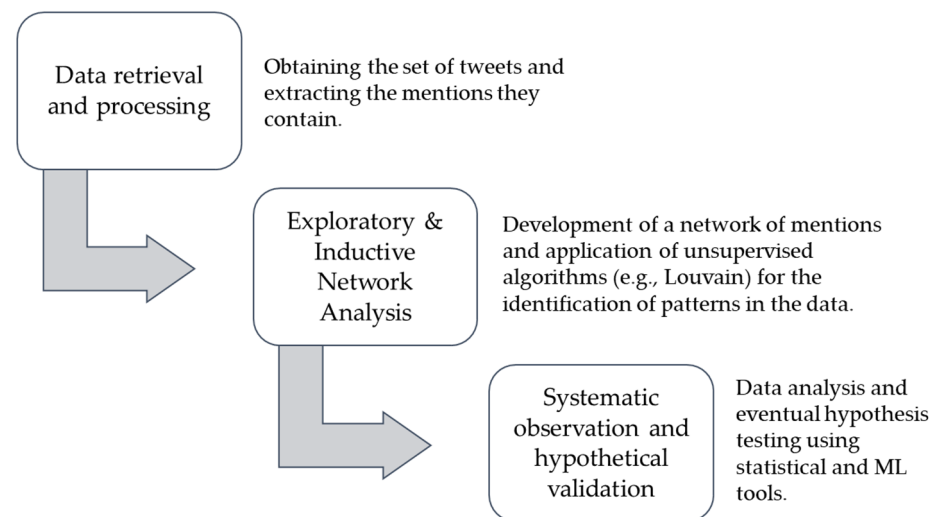
The next step in this proposal is to use certain Business Intelligence software such as Tableau, PowerBi, Google Data Studio, or Grafana to carry out the cross-referencing of data. All this without neglecting the desirability of also being able to count on SQL and NoSQL database technologies that allow to establish links between databases with different degrees of structuring, depending on specific categories, and to subsequently represent the combined fields in tables or graphs. Using this type of tools, it is possible to generate dashboards with linked visualizations that allow interactive navigation: selecting each cluster and visualizing its properties based on the indicators and key variables for each case of analysis. This type of approach therefore requires analytical profiles that are also familiar with some fundamental operations of Data Engineering, such as value transformation or table joining.

#### 4.2. Combining Induction and Deduction to Understand the Current Feminist Wave

Social media constitute a sort of public sphere in which different social movements deploy their strategies and define themselves through practices that involve the transmission of an immaterial legacy [15]. In the case of feminism, this has been recurrently expressed through the metaphor of “waves,” about which several authors suggest that we are currently facing the fourth [87]. The role of hashtivism and, very particularly that which has been developed on Twitter, is becoming very important in the self-definition of fourth wave feminism, a constructor of new political subjects that pivot around particular campaigns or hashtags [88]. The fourth wave of feminism thus dialogues with analogical activists who have endowed them with a whole tradition of struggle and a not inconsiderable set

of small victories that, taken together, have improved the living conditions of women in different parts of the world. It is precisely this dialogue, sometimes explicit and sometimes implicit, in which the practices of intangible cultural heritage transmission and political self-definition of fourth-wave feminism materialize. It is this dialogue that we wish to analyze here.

The guidelines and steps described above can be applied in various investigations to approach the object of study from different perspectives, depending on the research objectives. We will now look at four practical applications of the described methodology applied to the analysis of fourth-wave feminism and feminist hashtivism. In them, we start from a perspective that develops from a type of inductive and inferential reasoning that seeks theoretical synthesis from the observation of cases (i.e., Figure 3). We will also see how this type of approach is used in a complementary way with hypothetico-deductive approaches that seek the opposite: the validation of theories and hypotheses based on the observation of cases. In fact, by putting these practical and real examples on the table, it is argued that these two modes of reasoning work best when used together, and this is precisely one of the main strengths of the method of analysis described.



**Figure 3.** General workflow for the proposed methodology. Source: Own elaboration with Power Point.

In methodological terms, we argue that attributing epistemological priority to contextual analysis using unsupervised algorithms and an inductive logic (e.g., detecting communities with the Louvain algorithm) is an efficient strategy to overcome part of the most common problems in the analysis of massive social media data, such as, for example, uninformative and/or spurious sentiment analysis [89].

Below we cite four case studies of research published in scientific journals that employ this method and apply it to the observation of phenomena related to one of the most significant contemporary mobilizing currents such as feminisms and their presence in the online public debate. All these investigations are part of the general project whose methodological documentation we are carrying out in these pages. It is a project dedicated to the analysis of different conversations and controversies that occupy fourth wave feminism in the Spanish, Spanish-speaking and, eventually, also international sphere. All of them have in common, therefore, being investigations thought and executed from the same epistemological and methodological mentality, and furthermore, the fact of dealing with an object of study linked to the practices, transmitters of the intangible cultural heritage of feminism.

#### 4.2.1. Feminisms Outraged at Justice

The reaction of Spanish-speaking feminism on Twitter to the controversial sentence against the members of “la manada”—a group of men who raped a woman at a local

festival—is analyzed in this article using SNA techniques [90]. The sentence in question condemned the rapists for harassment, but not for rape, despite recognizing the materiality of the crime, also acquitting them of the crimes of recording the rape with a cell phone and robbery with intimidation.

Following network analysis and community identification, the five most disseminated tweets in each community are identified, along with their most prominent leaders by input degree (i.e., the most mentioned). From the content analysis, a total of three “macro-narratives” are identified around the Spanish judicial system, raised by the reaction to the sentence on Twitter. Two of these “macro narratives” projected a very negative evaluation of the Spanish judicial system, while the third narrative was one of defense of the system and, simultaneously, of criticism towards the feminist movement.

In this research it was possible to identify a series of practices strongly linked to the construction of the political subject of fourth-wave feminism, as well as a series of practices of identification and differentiation with respect to older generations and to the historical feminist movement.

#### 4.2.2. Digital Prospects of the Contemporary Feminist Movement for Dialogue and International Mobilization

An analysis is carried out based on the analysis of the conversation in Spanish and English around the 2018 international day against violence against women on Twitter [88]. The most shared contents of each cluster of the network, both messages and images, are inspected and the levels of intercommunity homophily (i.e., to what extent nodes tend to establish links with those who are part of the same group or with those who are part of the other group) [91] are examined.

Overall, the nuances between Latin feminist activism (i.e., generally more contentious) and Anglo-Saxon ones (i.e., more liberal and based on the support of individual cases) and the absence of shared transnational and translinguistic narratives are observed, which leads researchers to problematize the usefulness of feminist activism, at least, as far as its international coordination is concerned.

The results of this research open the door to consider the plurality of forms in which fourth-wave feminism materializes, as well as the different links established with the legacy of feminist activists of previous waves and the transmission of their immaterial cultural heritage.

#### 4.2.3. Feminist Hashtag Activism in Spain

An analysis is made of the contents of the clusters during the digital conversation that arose because of the sentence of “la manada” in April 2018 and continued for a few months, through a series of hashtags in solidarity with the victim [92]. On this occasion, attention is paid to the tone of their messages, and an interesting correlation is detected between those clusters with higher Input Degree Centralization (i.e., to what extent the reception of mentions is centralized or decentralized in the network) [93] and a greater banality in the contents (e.g., memes and other viral contents) and, in turn, between those clusters more decentralized in the reception of mentions and a greater seriousness, formality and anchorage with the historical feminist movement in the contents disseminated and the codes used.

In addition, a link is established between an external variable to the network (i.e., the degree of politicization of the users, according to the number of politicians’ accounts followed) and the belonging to a cluster of high or low input degree Centralization by means of a Machine Learning algorithm: a logistic regression that achieved a predictive capacity with an accuracy of 67.4%. In other words, it is inferred that the number of followed politician accounts is a very good predictor for the mode of political participation on Twitter.

This study invites us to consider the diversity of modes of participation and belonging to fourth-wave feminism, as well as the diversity of forms of linkage (i.e., stronger, or

weaker, more conscious, or more unconscious) with previous generations and with the feminist immaterial cultural heritage.

#### 4.2.4. Influence of Gender in Electoral Debates in Spain

The type and tone of the most shared messages in different clusters are identified, with the occasion of two televised electoral debates (April 2019 General Elections in Spain), which were starred entirely by men (#ElDebateDecisivo) and the other entirely by women (#L6Neldebate) [94].

It is observed, through the analysis of the most shared tweets in each cluster, and the strategy of identifying the nodes present in the female debate in the clusters of the male one, that the type of users who participate in the conversation of both debates do so with more serious codes, while those who participate only in the male one made use of banal and humorous resources more frequently.

In line with the findings of the previous article, in this one we can observe a certain alignment between the modes of articulation of the more self-conscious feminist hashtivism with the feminist agenda of political parties, while a greater alignment of spontaneous and depoliticized feminism with the irreverent and sometimes banal standards of virtual conversations.

### 5. Conclusions with Discussion

Social networks such as Twitter are a constant object of analysis by the various disciplines of social science, among which we highlight in this paper the discipline that focuses on the analysis of mass communication, transformed into mass interpersonal communication [95]. This field has evidenced an indisputable transformation that alters the traditional theoretical-methodological foundations of research on communicative production and reception. One of the main evidences of such an evolution would be the growth in the use of systematic and standardized data techniques in empirical works, as well as the tendency to put the research effort into the discursive and dialogic dimension of communication [96].

After presenting a specific workflow, which we believe is also replicable and applicable to many analytical objects, we have presented four different investigations around the same theoretical object. With variations, in the four investigations we have proceeded with the download of data from the Twitter API, with Network Analysis applied to the mention relationships between Twitter users, with community identification based on the network of mentions, and with the visualization of each community through a data modelling and visualization strategy. Not surprisingly, each investigation has led to a separate set of results. At the aggregate level, however, what we have seen is the usefulness of the methodology.

Our intention here has been to expose a replicable methodology for the generation of knowledge in massive data environments that require hybrid analytical perspectives and skills that can articulate rigorous mathematical analysis and relevant phenomenological interpretations. Our object of study has been fourth-wave feminism, and more specifically, we have focused on some reflections on the processes of transmission of the feminist intangible cultural heritage that take place on Twitter, in the thread of massive conversations, often chaotic, which nevertheless vertebrate and emplace the contemporary self-definition of the movement, always in dialogue with its recent and not so recent past, as well as projecting into the future. To this end, a series of specific analytical procedures have been defined to study the processes of communication on Twitter.

The distinctiveness of this methodological procedure consists not only in the use of analysis techniques and tools that are not very common in the social-scientific field, although increasingly so, but also in deploying a type of strategy that makes it possible to combine quantitative and qualitative analysis, as well as hypothetico-deductive and inductive perspectives. The paradigm of big data and the techniques associated with the computational sciences to which we have referred—SNA and ML—make it possible, to a large extent, to overcome old technical, methodological, and even epistemological

dichotomies or, at least, place us before new scenarios that provoke new metaphors and ways of looking at the classic objects of study of the social and communication sciences. These analyses not only serve to characterize the most recent research in communication 2.0, but also shape an analytical proposal based on the virtues of Twitter to weave individual messages with collective dialogic capacity around certain conversations or hashtags, this being a logic closely linked to feminist hashtivism, fourth-wave feminism, and the processes of transmission of immaterial cultural heritages in digital environments.

In any case, neither the proposed methodology nor the research we have used as an example obviate the existence of critical positions on the potential of the digital medium, on hashtivism in general, and on feminist hashtivism. We understand that, in addition, many of these critiques—traditionally established from the academy and from the feminist activism of previous generations—not only imply a form of real dialogue between feminists, but also imply in themselves practices of transmission of the immaterial cultural heritage: a transmission that never takes place in a linear and unidirectional way, but that necessarily passes through the self-constructive and creative filter of each generation or wave. Similarly, the proposals collected present not only the benefits, but also the methodological limitations of research on Twitter [97] based on the restrictions of the data collected, the bias of representation when making general assumptions and other problems, derived, for example, from the language of use of the users.

**Author Contributions:** Conceptualization, J.M.-i.-G. and J.O.-T.; methodology, J.M.-i.-G.; software, J.M.-i.-G.; validation, A.L.-U., J.O.-T. and S.P.-F.; formal analysis, A.L.-U.; investigation, J.O.-T.; resources, S.P.-F.; data curation, J.M.-i.-G.; writing—original draft preparation, J.M.-i.-G.; writing—review and editing, J.O.-T., A.L.-U. and S.P.-F.; visualization, J.M.-i.-G.; supervision, J.O.-T. and J.M.-i.-G.; funding acquisition, A.L.-U. and S.P.-F. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by Basque Government grant number IT-1112.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Written informed consent was not obtained due to Twitter API access policies.

**Data Availability Statement:** Data available upon request.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Fernández-Ferández, M.; Tardivo, V. La sociología de la Comunicación y la Mass Coomunication Research: Tradición y actualidad. *Espac. Abierto Cuad. Venez. Sociol.* **2016**, *25*, 133–142.
2. Castillo, A. *Neologismos y Sociedad del Conocimiento*; Ariel: Barcelona, Spain, 2007.
3. García-Estévez, N. La convergencia activista en Hong Kong: Del ciberactivismo de ‘Occupy Central’ al hacktivismo de ‘Operación Hong Kong’. In *Move.Net: Actas del I Congreso Internacional Move.Net Sobre Movimientos Sociales y TIC (139–156)*; Grupo Interdisciplinario de Estudios en Comunicación, Política y Cambio Social de la Universidad de Sevilla: Seville, Spain, 2015.
4. Dery, M. *Velocidad de Escape. La Cibercultura en el Final del Siglo*; Siruela: Madrid, Spain, 1998.
5. Joyanes, L. *Cibersociedad, los Retos Sociales Ante un Nuevo Mundo Digital*; McGraw-Hill: Madrid, Spain, 1997.
6. Lévy, P. *Cibercultura. La Cultura de la Sociedad Digital*; Anthropos: Mexico City, Mexico, 2007.
7. Van Dijck, J. *La Cultura de la Conectividad. Una Historia Crítica de las Redes Sociales*; Siglo Veintiuno Editores: Buenos Aires, Argentina, 2016.
8. Toret, J. *Tecnopolítica: La Potencia de las Multitudes Conectadas. El Sistema Red 15M, un Nuevo Paradigma de la Política Distribuida*; IN3 Working Paper Series; Universitat Oberta de Catalunya: Barcelona, Spain, 2013.
9. Castells, M. Comunicación, poder y contrapoder en la sociedad red (I). Los medios y la política. *Telos* **2008**, *74*, 1–19.
10. Castells, M. *Comunicación y Poder*; Alianza Editorial: Madrid, Spain, 2009.
11. Melucci, A. *Acción Colectiva, Vida Cotidiana y Democracia*; Centro de Estudios Sociológicos: Mexico City, Mexico, 1999.
12. Cándón-Mena, J.I. *Internet en Movimiento: Nuevos Movimientos Sociales y Nuevos Medios en la Sociedad de la Información*; Universidad Complutense de Madrid, Servicio de Publicaciones: Madrid, Spain, 2011.
13. Romero, U.P.M. *Movimientos Sociales y la Autocomunicación de Masas. Una Revisión del Movimiento 15-M*; Colegio San Luis: San Luis Potosí, Mexico, 2012.
14. Rovira-Sancho, G. Movimientos Sociales y Comunicación: La Red Como Paradigma. *Anàlisi: Quad. Comun. Cult.* **2012**, *45*, 91–104.

15. Withers, D. *Feminism, Digital Culture and the Politics of Transmission: Theory, Practice and Cultural Heritage*; Rowman & Littlefield: Lanham, MA, USA, 2015.
16. Congosto, M.L.; Fernández, M.; Moro Egido, E. Twitter y Política: Información, Opinión y ¿Predicción? *Cuad. Comun. Evoca* **2011**, *4*, 11–16.
17. Campos-Domínguez, E. Twitter y la comunicación política. *El Prof. Inf.* **2017**, *26*, 785–793. [[CrossRef](#)]
18. Rodríguez, R.; Ureña, D. Diez razones para el uso de Twitter como herramienta en la comunicación política y electoral. *Comun. Plur.* **2011**, *5*, 89–116.
19. Piscitelli, A. Prólogo: Twitter, la revolución y los enfoques ni-ni. In *Mundo Twitter: Una Guía Para Comprender y Dominar la Plataforma que Cambió la Red*; Orihuela, J.L., Ed.; Alianta: Barcelona, Spain, 2011; pp. 15–20.
20. Pariser, E. *The Filter Bubble: What the Internet Is Hiding from You*; Penguin: London, UK, 2011.
21. Page, R. The linguistics of self-branding and micro-celebrity in Twitter: The role of hashtags. *Discourse Commun.* **2012**, *6*, 181–201. [[CrossRef](#)]
22. Carr, D. Hashtag Activism and Its Limits. *The New York Times*, 25 March 2012.
23. Marwick, A. *Status Update: Celebrity, Publicity, and Branding in the Social Media Age*; Yale University Press: New Haven, CT, USA, 2013.
24. Fuchs, C. *Social Media. A Critical Introduction*; Sage: Thousand Oaks, CA, USA, 2014.
25. Giraldo-Luque, S.; Fernández-García, N.; Pérez-Arce, J.C. La centralidad temática de la movilización #NiUnaMenos en Twitter. *El Prof. Inf.* **2018**, *27*, 1–226.
26. Anduiza, E.; Cristancho, C.; Sabucedo, J.M. Mobilization through online social networks: The political protest of the indignados in Spain. *Inf. Commun. Soc.* **2014**, *17*, 750–764. [[CrossRef](#)]
27. Bennett, W.L.; Segerberg, A. The logic of connective action: Digital media and the personalization of contentious politics. *Inf. Commun. Soc.* **2012**, *15*, 739–768. [[CrossRef](#)]
28. Dixon, K. Feminist online identity: Analyzing the presence of hashtag feminism. *J. Arts Humanit.* **2014**, *3*, 34–40.
29. Zimmerman, T. #Intersectionality: The Fourth Wave Feminist Twitter Community. *Atlantis Crit. Stud. Gen. Cult. Soc. Justice* **2017**, *38*, 54–70.
30. Shiva, N.; Nosrat Kharazmi, Z. The fourth wave of feminism and the lack of social realism in cyberspace. *J. Cyberspace Stud.* **2019**, *3*, 129–146.
31. Mendes, K.; Ringrose, J.; Keller, J. # MeToo and the promise and pitfalls of challenging rape culture through digital feminist activism. *Eur. J. Women's Stud.* **2018**, *25*, 236–246.
32. Worthington, N. Celebrity-bashing or # MeToo contribution? New York Times Online readers debate the boundaries of hashtag feminism. *Commun. Rev.* **2020**, *23*, 46–65.
33. Zhou, Q.; Qiu, H. Predicting online feminist engagement after MeToo: A study combining resource mobilization and integrative social identity paradigms. *Chin. J. Commun.* **2020**, *13*, 351–369. [[CrossRef](#)]
34. Clark-Parsons, R. I see you, I believe you, I stand with you: # MeToo and the performance of networked feminist visibility. *Fem. Media Stud.* **2021**, *21*, 362–380.
35. Thelwall, M. International Women's Day 2009–2020 on Twitter: Postfeminist or Fourth Wave Feminism? *SSRN* **2021**, 3846542. [[CrossRef](#)]
36. Sanchez-Duarte, J.M.; Fernandez-Romero, D. Feminist sub-activism and digital collective repertoires: Cyberfeminist practices on Twitter. *El Prof. Inf.* **2017**, *26*, 894–902.
37. Navarro, C.; Coromina, Ó. Discussion and mediation of social outrage on Twitter: The reaction to the judicial sentence of “La Manada”. *Commun. Soc.* **2020**, *33*, 93–106. [[CrossRef](#)]
38. Idoiaga-Mondragon, N.; Berasategi-Sancho, N.; Beloki-Arztiti, N.; Belasko-Txertudi, M. # 8M women's strikes in Spain: Following the unprecedented social mobilization through twitter. *J. Gen. Stud.* **2021**, 1–16. [[CrossRef](#)]
39. Birhane, A. Algorithmic injustice: A relational ethics approach. *Patterns* **2021**, *2*, 100205. [[CrossRef](#)]
40. Graham, S.S.; Hopkins, H.R. AI for Social Justice: New Methodological Horizons in Technical Communication. *Tech. Commun. Q.* **2021**, 1–14. [[CrossRef](#)]
41. Kuo, R. Racial justice activist hashtags: Counterpublics and discourse circulation. *New Media Soc.* **2018**, *20*, 495–514. [[CrossRef](#)]
42. Almazor, M.G.; Canteli, M.J.P.; de la Cal Barredo, M.L. New approaches to the propagation of the antifeminist backlash on Twitter. *Investig. Fem.* **2020**, *11*, 221–237. [[CrossRef](#)]
43. Laney, D. 3D data management: Controlling data volume, velocity and variety. *META Group Res. Note* **2001**, *6*, 1.
44. Khan, M.A.; Uddin, M.F.; Gupta, N. Seven V's of Big Data understanding Big Data to extract value. In Proceedings of the 2014 Zone 1 Conference of the American Society for Engineering Education, Bridgeport, CT, USA, 3–5 April 2014; pp. 1–5.
45. Oguntimilehin, A.; Ademola, O. A Review of Big Data Management, Benefits and Challenges. *J. Emerg. Trends Comput. Inf. Sci.* **2014**, *5*, 433–438.
46. Patgiri, R.; Ahmed, A. Big data: The v's of the game changer paradigm. In Proceedings of the IEEE 18th International Conference on High Performance Computing and Communications, Sydney, Australia, 12–14 December 2016; pp. 17–24.
47. Anderson, C. The end of theory: The data deluge makes the scientific method obsolete. *Wired Mag.* **2008**, *16*, 16–17.
48. DiFonzo, N. The Echo-Chamber Effect. *The New York Times*, 22 April 2011.
49. Sweeney, L. Discrimination in online ad delivery. *Queue* **2013**, *11*, 10–29. [[CrossRef](#)]

50. Datta, A.; Tschantz, M.C.; Datta, A. Automated experiments on ad privacy settings: A tale of opacity, choice, and discrimination. *Proc. Priv. Enhancing Technol.* **2015**, *1*, 92–112. [CrossRef]
51. Metz, R. Why Microsoft Accidentally Unleashed a Neo-Nazi Sexbot. Available online: <https://www.technologyreview.com/2016/03/24/161424/why-microsoft-accidentally-unleashed-a-neo-nazi-sexbot/> (accessed on 23 November 2021).
52. Arcila-Calderón, C.; Barbosa-Caro, E.; Cabezuelo-Lorenzo, F. Técnicas big data: Análisis de textos a gran escala para la investigación científica y periodística. *El Prof. Inf.* **2016**, *25*, 623–631. [CrossRef]
53. Campolo, A.; Sanfilippo, M.; Whittaker, M.; Crawford, K. *AI Now 2017 Report*; AI Now Institute at New York University: New York, NY, USA, 2017.
54. Edizel, B.; Bonchi, F.; Hajian, S.; Panisson, A.; Tassa, T. FaiRecSys: Mitigating algorithmic bias in recommender systems. *Int. J. Data Sci. Anal.* **2020**, *9*, 197–213. [CrossRef]
55. Morales-i-Gras, J. *Datos Masivos y Minería de Datos Sociales, Conceptos y Herramientas Básicas*; Fundació Universitat Oberta de Catalunya: Barcelona, Spain, 2020.
56. Garcia-Alsina, M. *Big Data: Gestión y Explotación de Grandes Volúmenes de Datos*; Editorial UOC-El Profesional de la Información: Barcelona, Spain, 2017.
57. Morales-i-Gras, J. *Minería de Datos de los Social Media, Técnicas Para el Análisis de Datos Masivos*; Fundació Universitat Oberta de Catalunya: Barcelona, Spain, 2020.
58. Crossley, N. The social world of the network. Combining qualitative and quantitative elements in social network analysis. *Sociologica* **2010**, *4*. [CrossRef]
59. Edwards, G. *Mixed-Method Approaches to Social Network Analysis*; Routledge: Abingdon-on-Thames, UK, 2010.
60. Karamshuk, D.; Shaw, F.; Brownlie, J.; Sastry, N. Bridging big data and qualitative methods in the social sciences: A case study of Twitter responses to high profile deaths by suicide. *Online Soc. Netw. Media* **2017**, *1*, 33–43. [CrossRef]
61. Chen, N.C.; Drouhard, M.; Kocielnik, R.; Suh, J.; Aragon, C.R. Using machine learning to support qualitative coding in social science: Shifting the focus to ambiguity. *ACM Trans. Interact. Intell. Syst.* **2018**, *8*, 1–20. [CrossRef]
62. Twitter. *Standard Search—Twitter Developers*; Twitter Developers: San Francisco, CA, USA, 2018.
63. Ediger, D.; Jiang, K.; Riedy, J.; Bader, D.A.; Corley, C.; Farber, R.; Reynolds, W.N. Massive social network analysis: Mining Twitter for social good. In Proceedings of the 39th International Conference on Parallel Processing Workshops, San Diego, CA, USA, 13–16 September 2010; pp. 583–593.
64. Chatfield, A.; Brajawidagda, U. Twitter tsunami early warning network: A social network analysis of Twitter information flows. In Proceedings of the 23rd Australasian Conference on Information Systems, Melbourne, Australia, 3–5 December 2012; Volume 56.
65. Tremayne, M. Anatomy of protest in the digital era: A network analysis of Twitter and Occupy Wall Street. *Soc. Mov. Stud.* **2014**, *13*, 110–126. [CrossRef]
66. Congosto, M.L. Elecciones Europeas 2014: Viralidad de los mensajes en Twitter. *Redes Rev. Hisp. Para Análisis Redes Soc.* **2015**, *26*, 23–52.
67. Del Fresno García, M.; Daly, A.J.; Supovitz, J. Desvelando climas de opinión por medio del Social Media Mining y Análisis de Redes Sociales en Twitter. El caso de los Common Core State Standards. *Redes Rev. Hisp. Para Análisis Redes Soc.* **2015**, *26*, 53–75.
68. Morales-i-Gras, J. Desenredando las identidades soberanistas vasca y catalana: Un Análisis de Redes Sociales de las etiquetas de Twitter# BasquesDecide y #Up4Freedom. *Pap. CEIC Int. J. Collect. Identity Res.* **2015**, *2*, 1–37.
69. Blondel, V.D.; Guillaume, J.L.; Lambiotte, R.; Lefebvre, E. Fast Unfolding of Communities in Large Networks. *J. Stat. Mech. Theory Exp.* **2008**, *2008*, 10. [CrossRef]
70. Zhao, Y. Analysing Twitter data with text mining and social network analysis. In Proceedings of the 11th Australasian Data Mining and Analytics Conference, Camberra, Australia, 13–15 November 2013; Volume 146.
71. Gualda, E.; Borrero, J.D. La ‘Spanish Revolution’ en Twitter (2): Redes de hashtags (#) y actores individuales y colectivos respecto a los desahucios en España. *Redes Rev. Hisp. Para Análisis Redes Soc.* **2015**, *26*, 1–22.
72. Himelboim, I.; Smith, M.A.; Rainie, L.; Shneiderman, B.; Espina, C. Classifying Twitter topic-networks using social network analysis. *Soc. Media Soc.* **2017**, *3*, 2056305117691545. [CrossRef]
73. Gerlach, M.; Peixoto, T.P.; Altmann, E.G. A network approach to topic models. *Sci. Adv.* **2018**, *4*. [CrossRef] [PubMed]
74. Bird, S.; Klein, E.; Loper, E. *Natural Language Processing with Python: Analyzing Text with the Natural Language Toolkit*; O’Reilly Media Inc.: London, UK, 2009.
75. Ray, P.; Chakrabarti, A. A mixed approach of deep learning method and rule-based method to improve aspect level sentiment analysis. *Appl. Comput. Inform.* **2019**. [CrossRef]
76. Mikolov, T.; Chen, K.; Corrado, G.; Dean, J. Efficient estimation of word representations in vector space. *arXiv* **2013**, arXiv:1301.3781.
77. Myers, S.A.; Sharma, A.; Gupta, P.; Lin, J. Information network or social network? The structure of the Twitter follow graph. In Proceedings of the 23rd International Conference on World Wide Web, Seoul, Korea, 7–11 April 2014; pp. 493–498.
78. Grandjean, M. A social network analysis of Twitter: Mapping the digital humanities community. *Cogent Arts Humanit.* **2016**, *3*, 1171458. [CrossRef]
79. Lozares-Colina, C. La teoría de redes sociales. *Pap. Rev. Sociol.* **1996**, *48*, 103–126.
80. Louch, H. Personal network integration: Transitivity and homophily in strong-tie relations. *Soc. Netw.* **2000**, *22*, 45–64. [CrossRef]



81. Isaak, J.; Hanna, M.J. User data privacy: Facebook, Cambridge Analytica, and privacy protection. *Computer* **2018**, *51*, 56–59. [[CrossRef](#)]
82. Lykousas, N.; Patsakis, C. Large-scale analysis of grooming in modern social networks. *arXiv* **2020**, arXiv:2004.08205. [[CrossRef](#)]
83. Boulanger, M. Scraping the Bottom of the Barrel: Why It Is No Surprise That Data Scrapers Can Have Access to Public Profiles on LinkedIn. *Sci. Technol. Law Rev.* **2018**, *21*, 77.
84. Newman, M.E. Modularity and community structure in networks. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 8577–8582. [[CrossRef](#)]
85. Lambiotte, R.; Delvenne, J.C.; Barahona, M. Laplacian dynamics and multiscale modular structure in networks. *arXiv* **2008**, arXiv:0812.1770.
86. Morales-i-Gras, J. Soberanías Enredadas: Una Perspectiva Reticular, Constructural y Agéntica Hacia los Relatos Soberanistas Vasco y Catalán Contemporáneos en Twitter. Ph.D. Dissertation, Universidad del País Vasco-Euskal Herriko Unibertsitatea, Leioa, Spain, 2017.
87. Munro, E. Feminism: A fourth wave? *Political Insight* **2013**, *4*, 22–25. [[CrossRef](#)]
88. Larrondo Ureta, A.; Orbegozo Terradillos, J.; Morales i Gras, J. Digital Prospects of the Contemporary Feminist Movement for Dialogue and International Mobilization: A Case Study of the 25 November Twitter Conversation. *Soc. Sci.* **2021**, *10*, 84. [[CrossRef](#)]
89. Gaspar, R.; Pedro, C.; Panagiotopoulos, P.; Seibt, B. Beyond positive or negative: Qualitative sentiment analysis of social media reactions to unexpected stressful events. *Comput. Hum. Behav.* **2016**, *56*, 179–191. [[CrossRef](#)]
90. Orbegozo-Terradillos, J.; Morales-i-Gras, J.; Larrondo, A. Feminismos indignados ante la justicia: La conversación digital en el caso de La Manada. *IC Rev. Científica Inf. Comun.* **2019**, *16*, 249–283.
91. Krackhardt, D.; Stern, R.N. Informal networks and organizational crises: An experimental simulation. *Soc. Psychol. Q.* **1988**, *51*, 123–140. [[CrossRef](#)]
92. Larrondo, A.; Morales-i-Gras, J.; Orbegozo-Terradillos, J. Feminist hashtag activism in Spain: Measuring the degree of politicisation of online discourse on #YoSíTeCreo, #HermanaYoSíTeCreo, #Cuéntalo y #NoEstásSola. *Commun. Soc.* **2019**, *32*, 207–221.
93. Freeman, L.C. Centrality in social networks conceptual clarification. *Soc. Netw.* **1978**, *1*, 215–239. [[CrossRef](#)]
94. Orbegozo-Terradillos, J.; Larrondo-Ureta, A.; Morales-i-Gras, J. Influencia del género en los debates electorales en España: Análisis de la audiencia social en #ElDebateDecisivo y# L6Neldebate. *El Prof. Inf.* **2020**, *29*, 12.
95. Noguera, J.M. *Redes y Periodismo. Cuando las Noticias se Socializan*; Editorial UOC: Barcelona, Spain, 2012.
96. Martínez-Nicolás, M.; Sopera-Lapiedra, E.; Carrasco-Campos, A. La investigación sobre comunicación en España en los últimos 25 años (1990–2014). Objetos de estudio y métodos aplicados en los trabajos publicados en revistas españolas especializadas. *Empiria Rev. Metodol. Cienc. Soc.* **2019**, *42*, 37–69. [[CrossRef](#)]
97. Ruiz-Soler, J. Twitter research for social scientists: A brief introduction to the benefits, limitations and tools for analysing Twitter data. *Digitos* **2017**, *3*, 17–31.