IMT School for Advanced Studies, Lucca

Lucca, Italy

A Quantitative Analysis of News Consumption on Facebook

PhD Program in Complex Networks

XXX Cycle

By

Ana Lucía Schmidt

2017

To my family

Contents

List of Figures

List of Tables

Acknowledgements

The thesis is based on co-authored published or about to be published articles:

1. Chapter 4 is the reproduction of the paper *Anatomy of news consumption on Facebook*(SZDV$^+$17), published in the Proceeding of the National Academy of Science and co-authored with Fabiana Zollo, Michela Del Vicario, Antonio Scala, Alessandro Bessi, Guido Caldarelli, H. Eugene and Stanle and Walter Quattrociocchi.
2. Chapter 5 is the reproduction of the paper *A Study on European News Consumption on Facebook*(SZSQ17), currently in process, co-authored with Fabiana Zollo, Antonio Scala and Walter Quattrociocchi.
3. Chapter 6 is the reproduction of the paper *Polarization of the Vaccination Debate on Facebook*(SZS$^+$18), currently in process, co-authored with Fabiana Zollo, Antonio Scala, Cornelia Betsch and Walter Quattrociocchi.

Vita

July 22, 1989	Born, Bahía Blanca, Argentina
2014	Engineering Degree in Computer Systems Final mark: 8.03/10 Universidad Nacional del Sur, Bahía Blanca, Argentina
2017	Visiting Scholar Institute for Biocomputation and Physics of Complex Systems (BIFI) Universidad de Zaragoza, Spain
2017	Currently PhD Program IMT - School for Advanced Studies Lucca

Publications

1. Schmidt, A.L., Zollo, F., Del Vicario, M., Bessi, A., Scala, A., Caldarelli, G., Stanley, H.E. and Quattrociocchi, W., 2017. *Anatomy of news consumption on Facebook*. Proceedings of the National Academy of Sciences, p.201617052.
2. Schmidt, A.L., Zollo, F., Scala, A. and Quattrociocchi, W., 2017. *A Study on European News Consumption on Facebook*. In process.
3. Schmidt, A.L., Zollo, F., Scala, A., Betsch, C. and Quattrociocchi, W., 2018. *Polarization of the Vaccination Debate on Facebook*. arXiv preprint arXiv:1801.02903.

Abstract

The advent of social media and microblogging platforms has radically changed the way we consume information and form opinions. Despite the importance given by the scientific community to these changes in news production and consumption, not much is certain about the spreading and consumption patterns through social media platforms. This thesis does an in-depth quantitative analysis of news consumption and polarization on Facebook, taking into account factors such as trust in news, confirmation bias and cognitive dissonance.

First, we explore the anatomy of the English news by characterizing on a global scale the news consumption patterns of 376 million users over a time span of six years. Second, we compare the selective exposure and polarization of France, Germany, Italy and Spain (86M users), and present a model of selective exposure that considers *trust* in the emergence of communities. Finally, we analyze the polarization on the vaccination debate over time (2.6M users).

Our findings show that confirmation bias and homophily are present in the users' consumption of news, which leads to polarization and the creation of homogeneous communities. We also find that the preferences of users and news providers differ. By tracking how Facebook pages like each other and examining their geolocation, we find that users have a more cosmopolitan perspective of the information space than news providers. In addition, we devised two simple models of selective exposure that reproduce the observed connectivity patterns.

Chapter 1

Introduction

The emergence of social media reshaped the relationship between news organizations and their audiences, causing a radical shift from the traditional content production paradigm where media organizations dictated the content, the frequency, the timing and the medium of communications. Online readers are no longer confined to word-of-mouth communication after consumption because social media platforms significantly amplify the readers ability to communicate with each other and provide feed-back, thus changing the way users get informed as well as the quality of information (QCS14). This paradigm shift also affected the agenda setting power of traditional media, where there was a strong correlation between the news reported and the issue perceived as critical to the public opinion (MS72).

The advent of social media and microblogging platforms has significantly changed the way we consume information and form opinions. Sharing news online has now become central to the way people experience news in their everyday lives. According to the Pew Research Center (MHP13)(BSGM15), there has been a substantial increase in the number of users who consume news through social media. From 2013 to 2015 users who got their news from Facebook went from 47% to 63%, and a similar growth was seen among Twitter users (from 52% to 63%). A significant number of social media users rely on their social network to tell them what they need to know rather than turning solely to news organizations (HFKL12). The experience of news as a shared social experience seems to be a central motivation: around two thirds of social media users said they value the ability to share content with their peers and that one

of the main reasons they use social media sites, was to obtain firsthand information about important events (HFKL12).

Social media increases the number of news sources a user is exposed to through the interactions with their network (ACGC11). This exposure, however, is influenced by homophily: the tendency of people to connect with others who validate their core beliefs rather than be exposed to opposing viewpoints. Notwithstanding, social media users tend to believe that the variety of content they consume through social media is broad (HFKL12) and neutral (BSGM15). This opinion radically clashes with the tendency of users to process information through a shared system of meaning (MRZ$^+$15)(BZDV$^+$17).

Social media platforms like Facebook or Twitter have created a direct path for users to easily access information tailored to their tastes, with no regards to veracity. However, the more customized the information consumption process is, the more we are facing against our own cognitive limitations. People do not select information rationally, they choose information that adheres to their system of beliefs and ignore any dissenting information. This scenario is not just limited to conspiracy theories, it is related to all issues that are perceived as critical by the population, from geopolitics to brand loyalty.

Our own cognitive biases, added to the wide availability of information through social media, foster polarization and the emergence of information bubbles, i.e. clusters of like-minded people that rarely interact outside their own groups. Despite the great importance given by the scientific community to the recent changes in news production and consumption, not much is certain about the spreading patterns and consumption customs of official news through social media platforms.

This dissertation provides an in-depth quantitative analysis of social traces from online social media. The focus is on the consumption of online news among Facebook users and the polarized communities of news outlets that emerge from the users' consumption patterns. We take into account factors such as trust in news, confirmation bias and cognitive dissonance. The results provide invaluable information on how official sources of news, in parallel with conspiracy theories, help shape mass opinion.

Thesis Advancement

The present dissertation focuses on the analysis of news consumption and polarization on social media, more specifically, on Facebook. It considers the effect confirmation bias, cognitive dissonance and selective exposure have in the users' consumption process. Quantitative methods are used to measure how people's cognitive limitations affect the selection of the information they consume, which leads to the emergence of echo chambers. Different scopes where considered when analyzing the news consumption process on Facebook: global, examining English news sources and their influence around the globe; national, comparing the consumption patterns of four European countries; and topical, addressing the vaccination debate.

The dissertations starts by studying a dataset of 376 million Facebook users spanning 920 English news pages on a global scale. Through quantitative analysis we reveal that the consumption of news on Facebook is influenced by confirmation bias, leading to the polarization of the users and the emergence of echo chambers. The results also show that selective exposure plays an important role in the way news are consumed online. A simple model based on the Bounded Confidence Model is also introduced, that accurately reproduces the observed user dynamics.

Then we compare the consumption habits of $87M$ Facebook users considering the top news sources of four European countries: France, Germany, Italy and Spain. We once again show evidence that selective exposure is central in the online news consumption process and that confirmation bias in the users' consumption habits leads to the emergence of polarized communities, regardless of the country analyzed, and that different users from different countries exhibit different levels of selective exposure. We also measure the polarization of the users within those communities and establish a ranking of countries, showing that Italian users are very polarized, followed by the French, the German and finally the Spanish users. Another model based on the Bounded Confidence Model is presented, this time incorporating the *trust* the users have on the news outlets.

We finish by analyzing the evolution of the public debate on vaccines on Facebook, taking into account two opposing narratives: *anti-vaccines* and *pro-vaccines*. Considering the liking and commenting behavior of $2.6M$ users, we validate the existence of these two polarized echo-chambers. Additionally, we find evidence that while selective exposure plays an essential role in the way users consume content online, differ-

ent echo chamber display different rates at which the variety of news sources diminishes for increasing levels of user activity. Finally, we study the evolution of the two communities over time, taking into account the number of users, the number of pages and the cohesiveness of the communities.

The dissertation is structured as follows. Chapter 2 includes a review of the literature on user behavior on social media, with an emphasis on information consumption and polarization. Chapter 3 introduces the Facebook datasets, as well as the concepts and materials, used throughout this dissertation. Chapter 4 describes in detail the results published in (SZDV$^+$17) on the anatomy of news consumption on Facebook on a global scale. Chapter 5 compares the way users from France, Germany, Italy and Spain interact with Facebook news. Chapter 6 studies the polarization of users on Facebook regarding the topic of vaccines. The results in Chapters 5 and 6, come from the papers (SZSQ17) and (SZS$^+$18). The Appendix A enumerates the list of Facebook pages that compose the different datasets used in this work.

Chapter 2

Background and Literature Review

Social media and microblogging platforms have changed the way we access information and form opinions (BBL07; KK04; QCS14; KMM10). Communication has become increasingly personalized, both in the way messages are framed and how they are shared across social networks.

Nowadays, news organizations have also adopted social media platforms, like Facebook and Twitter, as a mean to distribute news and connect with their audiences. This increased interaction with the audience contrasts significantly with the traditional journalism, particularly print, where interactive communication with the audience was limited.

The adoption of social platforms, however, was mostly done to take advantage of the user generated content and obtain more profit (LLH12; HT08; MF09). While news organizations are increasing the interactions with their audiences, they still moderate them in an effort to retain the traditional gate-keeping role of the past.

The concepts of agenda setting, priming and framing are essential to the study of news consumption. Agenda setting refers to the idea that there is a strong correlation between the importance that news organizations place on certain events and the importance attributed to these events by the audience (MS72). Priming occurs, as an extension of agenda setting, when the content of the news suggests to the audience that they should use specific aspects as benchmarks for evaluating the performance of leaders and governments. Thus, by making some issues more promi-

nent in people's minds (agenda setting), news organizations can also shape the way people make judgments about political figures or events (priming).

Unlike agenda setting and priming, framing refers to how different presentations of the same news event can influence the audience to perceive the event in a different manner. For journalists in particular, framing is a tool to adapt the presentation of the news in a way that resonates with their audience. Although agenda setting and framing may appear similar processes, an important distinction must be made: while mere exposure might be sufficient for an agenda setting effect to take place, attention to the message is required for the framing to be successful (ST07).

This change brought by social media in the production and consumption on news has also affected the agenda setting power of traditional media, as well as their ability to prime and frame the information to suit their needs. In (Mer09) the agenda setting power of traditional media and blogs is compared, showing that while traditional media still has considerable power, the aggregated effect of user generated content can affect it to some extent. Traditional media seems to no longer have a monopoly when it comes to agenda setting, it is now one among many forces fighting to gain attention in the unmediated context of social media.

In (SBS+10), the authors studied the agenda setting power of YouTube when it came to driving, and in some cases leading, the public discourse on California's Proposition 8. They found that the connections between social media and mainstream media in the case of Proposition 8 shifted dramatically over the course of the time studied. Before the voting, mainstream media was leading the discussion, while after, YouTube videos actually increased in number as the newspapers' attention to the issue faded.

In a study on the agenda setting influence of online wired services and newspapers in Korea (Lim06), the author showed that online newspapers strongly influence the agenda of the online wire services, while the reverse was not true. The study also showed that the main online newspaper also influenced other secondary newspapers, leading to content homogenization. A similar study (BDS07) analyzed agenda setting in Argentina's main online news, noticing that as the day unfolds, content homogenization increases, most likely due to the monitoring and mimicking among media during the day.

Research has also been done in order to analyze how social media platform influenced journalist (Her10; HL11; LLH12; Hon12) and cul-

tural habits (HT08).

(HL11) analyzed the tweets made by famous journalists on Twitter to see how humor affects the topic and engagement of their conversations. (LLH12) studied how journalists are normalizing and fitting Twitter in their existing professional practices. They found that on Twitter, journalists expressed their opinion more freely, in spite of the impartial nature of journalistic writing. (Hon12) showed that the adoption of Twitter by news organization is positively associated with an increase in their online readership and that, in 2009, online traffic generated by social media sites was less concentrated than the one generated by search engines.

There are also studies on the diverse effects social media had on news readers, from the variety of their news diet to the increasingly personalization of the consumption process.

According to a recent study (NLN15), approximately 63% of users acquire their news from social media and these news are subject to the same popularity dynamics as other forms of content. (KLPM10) did a massive quantitative analysis of the Twitter space and found that over 85% of the trending topics were headlines or persistent news.

In (SdWL05), the authors found that the quick consumption of online news might increase awareness of the most relevant events, but does not really broaden the variety of topics. (ABS+14) found that incivility on comments among blog users can increase polarization.

(HSP+06) found evidence that media dissociation can drive online news consumption. They found that opponents to the War in Iraq who felt their views differed from the mainstream media, sought out online sources for perspectives about the war.

Selective Exposure and Polarization

Selective exposure has had a variety of definitions throughout the years, but today it is mostly understood as the people's tendency to seek out narratives that conform to their current beliefs, and avoid those that challenge them (SF67; Mut06; Mut01). While this self-selection of information is an active and deliberate process, it can also occur unconsciously (BS64).

A piece of news can present diverse information, mixing viewpoints and submitting evidence that supports multiple narratives. News outlets aggregate pieces of news allowing the users to select from a wide variety of information. This selection process has been further simpli-

fied with the advent of social media, permitting the users to efficiently choose those in line with their system of beliefs. For example, in (Mut01) and (Mut06), the authors show that people with more control over their news exposure to tend select news outlets that reinforce their existing opinions, and become more familiar with evidence that supports them.

There is substantial evidence that people are interested in opinion-reinforcing political information (Fre86). (Str08) found evidence that people's political beliefs motivate their media use patterns and that cable news audiences became increasingly politically divided over the course the election of 2004.

One common example of selective exposure, is that of cognitive dissonance theory (Fes62). This theory is based on the assumptions that people are sensitive to inconsistencies between actions and beliefs and that the recognition of such inconsistencies will cause dissonance, motivating them to find a resolution. The attachment to the belief and the degree of inconsistency with their behavior will, of course, vary the dissonance and with it the motivation to resolve it.

A person can resolve their cognitive dissonance either by changing their beliefs, their actions or the way the actions are *perceived*. While the first option is straightforward, people's beliefs are generally deeply rooted and sudden changes to them is neither easy nor common. Also, while people can always choose to change or stop the behavior that causes the dissonance, self-interest and lack of consequences might prove motivation enough against it. Thus, the re-framing or rationalizing of the behavior in such a way that is consistent with their beliefs. This way people can keep their beliefs and actions unchanged and the cognitive dissonance is resolved, taking the discomfort with it.

Cognitive bias also factors in this process (Nic98), as people have a natural tendency to not look for evidence that their beliefs are wrong. Selective exposure and cognitive bias, added to the ability granted by social media to easily and efficiently self-tailor our news consumption, could produce highly homogeneous audiences that do not interact among each other. These homogeneous audiences have appeared in the literature under many different names, such as sphericules (Git02), red-media blue-media (IH09) and echo-chambers (Sun02b), as they are more generally known today.

A large body of research has addressed news consumption on online social media and its polarizing effect on public opinion. Recent works (QSS16) provide empirical evidence of the pivotal role of confirmation bias and selective exposure in online social dynamics. Users tend to fo-

cus on specific narratives and join polarized groups – i.e., echo chambers –(BPDV+15; DVBZ+16) where they end up reinforcing their worldview, even if pieces of the content consumed are deliberately false (MRZ+15; BCD+15), and dismissing contradictory information (ZBDV+15). Discussion and elaboration of narratives in such a segregated environment elicits group polarization and negatively influences user emotion (Sun02a; ZNDV+15; YB10; BMA15).

Other studies focused on how to stop misinformation diffusion with algorithms (CSR+15; WAL+14) and fact checking (SGATM+13; HMKW14). However, empirical results pointed out the inefficacy, and backfire effect, of such approaches on online social networks (BCDV+14; ZBDV+15).

On (BCD+15) they performed a study that examines both scientific and conspiracy Facebook pages. They found that users are polarized in separate communities, each matching to one of the two types of pages considered. The authors showed that the consumption patterns for both type of information are similar, but the users consuming content from the conspiracy pages are more prone to interact with posts and pages of their category, and especially more prone to share the information.

(MRZ+15) did a detailed study on how Facebook users consume information considering different kind of pages: mainstream media, political activism and alternative sources. The authors provide evidence that qualitatively different information is consumed in a similar way, with unsubstantiated claims spreading similarly to the content produced by mainstream media. They also found that users who are prominently interacting with conspiracy-like sources are more prone to interact with intentional false claims.

In (DVGQ+17), the authors focus on the debate surrounding the Italian Constitutional Referendum by conducting a quantitative analysis on Facebook and Twitter. The work provides evidence of the spontaneous emergence of well-defined communities on both platforms, and that the users restrict their consumption to a specific set of pages/accounts.

(DVZC+17) studies the news consumption process around Brexit on Facebook considering official news outlets, and shows that the consumption patterns elicit the emergence of two distinct communities. The authors also introduce a new technique which combines automatic topic extraction and sentiment analysis.

Chapter 3

Methods, Concepts and Materials

This chapter provides a brief description of the data, main methods and concepts, as well as the tools used in the acquisition and analysis of the data.

3.1 Data Collection and Ethics Statement

The entire data collection process was performed exclusively by means of the publicly available Facebook Graph API (Fac13). The work abides the terms, conditions, and privacy policies of Facebook and hence does not require ethical approval as the data was preexisting and public. Users with any kind of privacy restrictions are not included in the datasets.

In the following chapters three different datasets will be presented:

- **European Media Monitor Dataset.** This dataset was generated from a list of news sources provided by the European Media Monitor. It is restricted to Facebook pages of websites that report news in English. It includes all posts made from the 1st January 2010 to 31st December 2015, as well as all likes and comments that have been made on those posts. This dataset also includes the Facebook pages liked by our set of pages in order to create a network of favorite pages.

The European Media Monitor list also includes the country and region of each news source. Nevertheless, for an accurate mapping on the globe, the geographical location of each page was collected manually. A table with the pages is provided in Appendix A.1.

- **EU Countries Dataset.** This dataset contains the top news sources, in their official languages, for France, Germany, Italy and Spain. The list for each country was generated manually and verified by considering the Reuters Digital News Report of 2015 (NLN15) and 2016 (NLN16). It includes all posts made from 1st January 2015 to 31st December 2016, as well as all likes and comments that have been made on those posts. A table with the set of pages found for each country can be found in Appendix A.2.

- **Vaccine Dataset.** This dataset was generated from requests to Facebook for pages containing the keywords *vaccine*, *vaccines* or *vaccination* in their name or description. These pages were then filtered to keep only those that are in English and have at least 10 posts. Pages whose topic did not pertain to the vaccination debate were also removed. This last step is essential, as the existence of one of the keywords in the description does not necessarily mean the page's topic is about vaccines.

  The dataset includes all posts made from 1st January 2010 to 31st May 2017, as well as all likes and comments made on those posts. Considering their content, the Facebook pages were also manually classified into two groups: *pro-vaccines* and *anti-vaccines*. A table with the Facebook pages, as well as their group label, can be found in Appendix A.3.

3.2 Graphs, Communities and Polarization

In order to analyze the polarization and consumption patterns of Facebook users, the datasets have to be manipulated into useful structures that represent the existing relationship between the users and the pages.

3.2.1 Projection of Bipartite Graphs

A bipartite graph is a triple $\mathcal{G} = (A, B, E)$ where $A = \{a_i \,|\, i = 1 \dots n_A\}$ and $B = \{b_j \,|\, j = 1 \dots n_B\}$ are two disjoint sets of vertices, and $E \subseteq A \times B$

is the set of edges, i.e. edges that exist only between vertices of sets A and B. The bipartite graph $\mathcal{G}$ is described by the rectangular matrix M where $M_{ij} = 1$, if an edge exists between a_i and b_j, and $M_{ij} = 0$ otherwise.

For this work, the bipartite graphs consist of two disjointed set of nodes: Facebook users and Facebook pages. That is, $\mathcal{G} = (P, U, E)$ where P is a set of Facebook pages and U is the set of users active on those pages. Edges represent interactions among users and pages, that is, either likes or comments.

As an example, a like given to a post on page p constitutes a link between the user u and the page p so $M_{p,u} = 1$. We can then build the co-occurrence matrices $C^P = MM^T$ and $C^U = M^T M$ that quantify, respectively, the number of common neighbors between two vertices of P or U.

Only two graphs per dataset will be relevant for the analyses: $\mathcal{G}_L$ and $\mathcal{G}_C$. These are the result from the projection C^P of two bipartite graphs: one considering the users' liking activity ($\mathcal{G}_L$) and another considering the comments ($\mathcal{G}_C$).

3.2.2 Community Detection and Comparison

Once we have the projection of our bipartite graphs, $\mathcal{G}_L$ and $\mathcal{G}_C$, we want to discover the communities considering the users' activity. That is, we want to group our pages in such a way that the internal connections within a group are dense —a high number of common users— and the external connections are sparse. To do this we take an unsupervised approach and utilize existing community detection algorithms implemented in R (CN06).

Throughout this dissertation, five different community detection algorithms are used:

- **FastGreedy (FG).** It optimizes the modularity score in a greedy manner[1]. This algorithm takes an agglomerative bottom-up approach: initially each vertex belongs to a separate community and, at each iteration, the communities are merged in a way that yields the largest increase in the current value of modularity (CNM04).

[1]One of the main strategies for community detection relies on the modularity of the network. The modularity is a benefit function that measures the quality of a particular division of a network into communities. The general approach is to achieve a high modularity score which corresponds to a dense connectivity between nodes inside a cluster and sparse connections between clusters.

The algorithm stops when it is no longer possible to further increase the modularity. Due to its speed and its lack of parameters in need of tuning, this algorithm will be the main reference to compare against the partitions resulting from the application of other community detection algorithms.

- **MultiLevel (ML).** It uses a multi-level optimization procedure for the modularity score (BGLL08). It takes a bottom-up approach where each vertex initially belongs to a separate community and in each step, unlike FastGreedy, vertices are *reassigned* in order to achieve the highest modularity.

- **WalkTrap (WT).** It uses random walks within the graph to detect communities (PL06). It exploits the fact that a random walker tends to become trapped in the denser parts of a graph, i.e., in communities.

- **LabelPropagation (LP).** It uses a simple approach where each vertex is assigned a unique label, which is updated according to majority voting in the neighboring vertices (RAK07). Dense node groups quickly reach a consensus on a common label.

- **SpinGlass (SG).** This algorithm uses as spin-glass model and simulated annealing to find the communities inside a network. SpinGlass is an approach from statistical physics (RB06)(NG04)(TB09), where each particle (i.e. vertex) can be in one of c spin states, and the interactions between the particles (i.e. the edges of the graph) specify which pairs of vertices would prefer to stay in the same spin state (i.e. belong to the same community) and which ones prefer to have different spin states. The model is then simulated for a given number of steps, and the spin states of the particles in the end define the communities.

Finally, to compare the various community partitions and the similarity between different clustering methods, we use the Rand index (Ran71), where a comparison between two partitions yields a value between 0 and 1, such that 0 indicates that there is no agreement on any vertex between the two partitions, whereas 1 indicates that the partitions are exactly the same.

3.2.3 User Polarization

Having the network of pages, both $\mathcal{G}_L$ and $\mathcal{G}_C$, and their community partitions, we want to calculate the polarization of the users with respect to those communities. That is, the propensity of the users to interact with more than one community. Depending on the number of existing communities we can consider two different cases.

Bi-community Polarization

If there are two communities, C_1 and C_2, we just have to measure the activity of the users in each of them. Assuming that a user u has performed y and x likes (or comments) on two communities C_1 and C_2, respectively, we can calculate their polarization $\rho(u)$ as:

$$\rho(u) = \frac{y - x}{y + x} \tag{3.1}$$

Thus, a user u for whom $\rho(u) = -1$ is polarized towards C_2, meaning all their likes (or comments) belong to community C_2, whereas a user whose $\rho(u) = 1$ is polarized towards C_1, meaning their activity is confined to community C_1.

After calculating the polarization of each user we can visualize the distribution of the $\rho(u)$ values by plotting the Probability Density Function. Figure 1 shows an example with a sharply bi-modal distribution where the users are highly polarized.

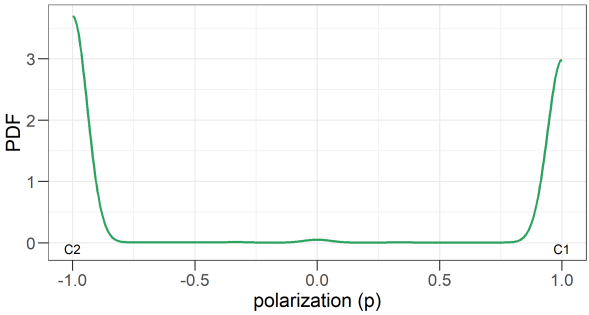

Figure 1: Example of Bi-community polarization.

Mutli-community Polarization

For the cases where we want to determine the polarization of users between more than two communities we can utilize two different methods.

Radial Scatter Plot. To visualize the activity of the users between more than two communities we can use a radial scatter plot (also known as polar scatter plot). To do this, we calculate for each community c_i, where $i \in (1, \dots N)$ and N equals the number of communities to be considered, their respective polarization centers (x_i, y_i) as:

$$x_i = \cos(\frac{2i\pi}{N}) \qquad y_i = \sin(\frac{2i\pi}{N}) \tag{3.2}$$

The polarization centers (x_i, y_i) will be the vertices of a regular polygon of N sides inscribed in a circle of radius 1.

Once we have the polarization centers, we calculate for each user u, a position (x, y) such that the closeness of the point to a polarization center c_i represents their tendency to like posts from that community. So, for example, if user u only likes posts belonging to community c_i their position (x, y) will be equal to the polarization center of c_i. On the other hand, if user u likes posts from two communities c_i and c_j their position will be somewhere in the line that joins the polarization centers of c_i and c_j.

If for user u we have:

- $L_u = \{k_1, \dots k_N\}$, where L_u is the liking activity of user u and each k_i is the number of likes given to community c_i.
- $\sum_{i=1}^{N} k_i = K$, where K is the total number of likes of user u.
- $K \geq 10$, user u has more than 10 likes.

Then, the position (x, y) of user u will be:

$$x = \frac{1}{K} \sum_{i=1}^{N} k_i \cos(\frac{2i\pi}{N}) \qquad y = \frac{1}{K} \sum_{i=1}^{N} k_i \sin(\frac{2i\pi}{N}) \tag{3.3}$$

$user$	c_1	c_2	c_3
u_1	15	0	0
u_2	0	23	0
u_3	0	1	27
u_4	10	20	0
u_5	7	15	2
u_6	6	6	6

Table 1: Example of the liking activity of 6 users within 3 communities.

For example, let us consider three communities ($N = 3$) and a set of 6 users whose liking activity can be seen in Table 1.

As we can see in Figure 2, for each of our three communities c_i we plot three polarization centers that are the vertices of an equilateral triangle inscribed in a circle of radius 1. The users will also be positioned according to the liking activity as provided in Table 1. Thus, users u_1 and u_2 are placed exactly in each of the polarized centers as they are both strictly polarized to one community. The users u_3 and u_4 are active in two communities and are plotted in the line that unites the corresponding polarization centers, c_2 and c_3 in the case of u_3, and c_1 and c_2 for u_4. The user u_5 has likes in all three communities so it's plotted inside the triangle, closer to the polarization center of the community in which the liking activity is higher. Finally, user u_6 shows no preference between the communities and is thus placed in the center of the triangle, equidistant from all 3 polarization centers.

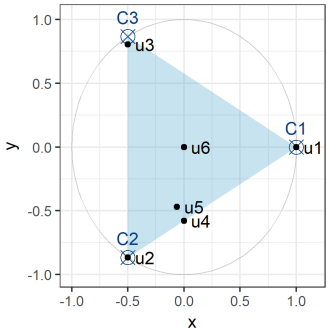

Figure 2: Example of a polarization radial plot with 3 communities.

Users' Localization. Another approach to understand polarization is to measure the number of communities a user is active on. For a user with K likes with $\sum_i k_i = K$ such that each k_i belongs to the i^{th} community ($i = 1 \dots N$, where N equals the number of communities). The probability ϕ_i that the user belongs to the i-th community will then be $\phi_i = k_i/K$. We can define the localization order parameter L as:

$$L\left[\phi\right] = \frac{\left(\sum_i \phi_i^2\right)^2}{\sum_i \phi_i^4} \tag{3.4}$$

Thus, in the case in which the user only has likes in one community, $L = 1$. If a user, on the other hand, interacts equally with all the communities ($\phi_i = 1/N$) then $L = N$; hence, L *counts* the communities.

Since we are considering many users, each with their likes k_i and their frequency ϕ_i, we can plot the probability distribution of L along the entire user set. Figure 3 shows an example, with $N = 3$, where the distribution is well behaved (single peak). The orange line marks the mean value of L.

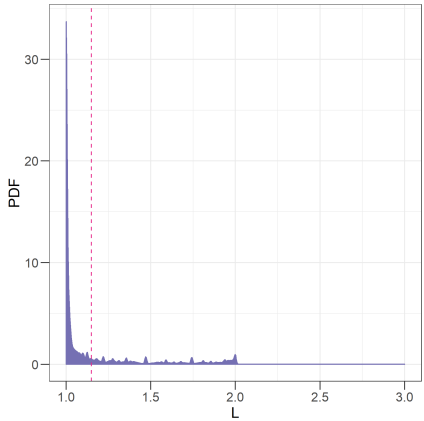

Figure 3: Example of the localization of the users activity with 3 communities.

3.3 The Bounded Confidence Model

In the following chapters, two models based on the Bounded Confidence Model (BCM) (DNAW00; Lor07) will be presented. The BCM considers a set of N agents arranged in a complex network G. Each agent initially holds and opinion x_i, $i \in \{1, \ldots, N\}$, and the values x_i are uniformly distributed in $[0, 1]$. At each time step any two randomly chosen agents i and j, connected in G, interact only when their difference in opinion is smaller in magnitude than a threshold ϵ, that is, $|x_i - x_j| < \epsilon$ for $\epsilon \in [0, 1]$. If they interact, their opinions change according to equation 3.5:

$$\begin{cases} x_i = x_i + \mu(x_j - x_i) \\ x_j = x_j + \mu(x_i - x_j) \end{cases} \tag{3.5}$$

where the convergence parameter μ is taken in the interval $(0, 0.5)$.

The threshold ϵ is in place to simulate real social dynamics where people with different opinions are not usually open to debate. Thus, the threshold conditions the interaction between agents so that they only do it when their opinion are similar enough.

The final opinions can be calculated for small enough values of ϵ. Density variations $\delta\rho\,(x)$ of opinions x obeys the following dynamics:

$$\delta\rho\,(x) = \frac{\epsilon^3}{2}\mu\,(\mu - 1)\,\frac{\partial^2\left(\rho^2\right)}{\partial x^2} \tag{3.6}$$

This indicates that starting from an initial distribution of opinions in the population, any local higher opinion density is amplified. Peaks of opinions increase and valleys are depleted until very narrow peaks remains among a desert of intermediate opinions.

Chapter 4

Anatomy of News Consumption on Facebook

All the results shown in this chapter refer to the article (SZDV$^+$17)[1], published in the Proceedings of the National Academy of Sciences. Here we explore the anatomy of the information space on Facebook by characterizing on a global scale the news consumption patterns of 376 million users over a time span of 6 years (January 2010 to December 2015).

Social media heavily changed the way we consume information and shape our opinions. Communication has become increasingly personalized, both in the way the information is framed and the way it is shared across social networks. According to a recent study 63% of users acquire their news from social media, and these news are subject to the same popularity dynamics as other forms of content (NLN15).

Recent works (QSS16) provide empirical evidence of the central role confirmation bias and selective exposure play in online social dynamics. Users focus on specific narratives and join polarized groups (BZDV$^+$15; BPDV$^+$15; DVBZ$^+$16). There, contradicting information is dismissed (ZBDV$^+$15) and their existing beliefs are reinforced, even if the information seen is intentionally false (MRZ$^+$15; BCD$^+$15).

[1]The results shown in this chapter are all part of the paper (SZDV$^+$17). It is a joint work with Fabiana Zollo, Michela Del Vicario, Alessandro Bessi, Antonio Scala, Guido Caldarelli, H. Eugene Stanley and Walter Quattrociocchi. ALS, AS and WQ designed the research. ALS, FZ, MDV, HES and WQ performed the research. FZ, MDV and AB contributed new reagents and analytic tools. ALS, FZ, MDV, AS, GC, HES and WQ analyzed the data. ALS, AS, HES and WQ wrote the paper.

In this chapter, we use quantitative analysis to better understand how echo chambers emerge. We first focus on how Facebook posts from 920 news outlets are consumed and how users' activity causes connectivity patterns to appear. We find evidence that users' polarization seem to dominate news consumption on Facebook and that selective exposure plays a pivotal role in the shaping of news consumption online. Furthermore, our results suggest that the perspectives of the news outlets and the users differ. By taking into account the geo-location of the pages, we find that users have a more cosmopolitan perspective of the information space than news providers. Finally, we present a simple model of selective exposure that well reproduces the observed dynamics.

4.1 Data Description

In this chapter we use the European Media Monitor Dataset described in Chapter 3.1. As explained before, it was generated from a list of news sources on a global scale provided by the European Media Monitor (SPVDG13). The list was then reduced to keep only those news sources that report in English and have an official Facebook page. We then used the Facebook Graph API (Fac13) to download all the posts made on those pages from 1st January 2010 to 31st December 2015, with their respective likes and comments. For each Facebook page we also downloaded the list of their favorite pages, that is, the list of pages they liked on Facebook.

A breakdown of the dataset can be seen in Table 13, and the list of news sources with their respective Facebook pages and locations can be found in the Appendix A.1.

	Total
Pages	920
Posts	12,825,291
Likes	3,621,383,495
Comments	366,406,014
Users	376,320,713
Likers	360,303,021
Commenters	60,115,975

Table 2: European Media Monitor Dataset Numbers. Users is the number of people that gave likes and/or comments. Likers is the number of people that gave likes. Commenters is the number of people that gave comments.

The European Media Monitor list includes the country and region of each news source, however, for a more accurate mapping on the globe we manually collected the geographical location, latitude and longitude, of each page.

4.2 Attention Patterns

Posts on Facebook can be liked[2], commented or shared by users. A *like* represents positive feedback from a user. A *share* indicates the user's desire to spread the information in the post to their social circle. A *comment*, on the other hand, can have multiple features and meanings and can generate collective debate.

To understand how users behave on Facebook we explore the way they interact with the Facebook posts. The left panel of Figure 4 shows the distribution of the number of likes, comments and shares given to Facebook posts. As seen on the plot, all the distributions are heavy-tailed.

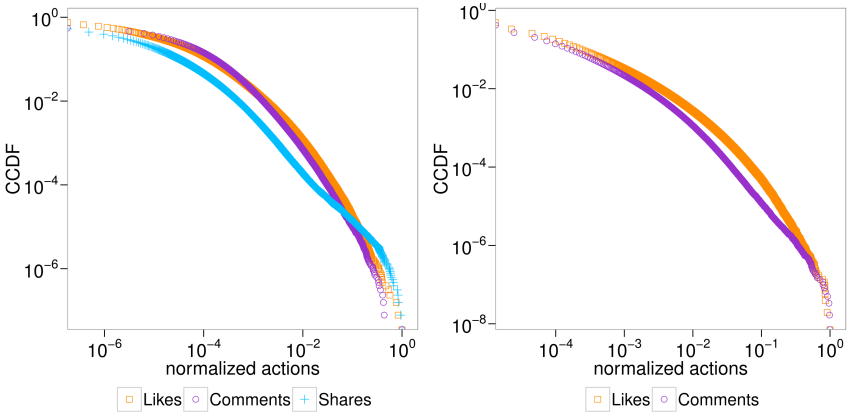

Figure 4: Complementary Cumulative Distribution Function of the posts and the users. CCDF of the likes, comments and shares given to the posts (left), and of the likes and comments given by the users (right).

[2]As Facebook didn't introduce *reactions* —*love*, *wow*, *haha*, *sad* and *angry*— until February 2016, this study limits the features used to the likes.

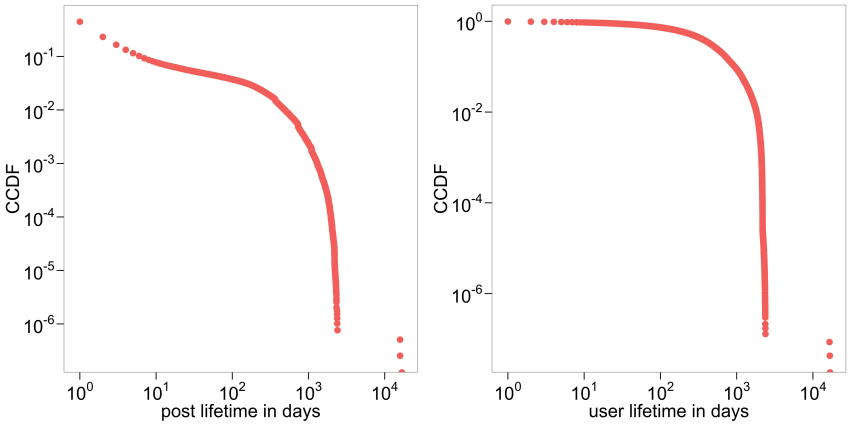

Figure 5: Complementary Cumulative Distribution Function of the posts' persistence (left) and the users' lifetime (right).

The right panel of Figure 4 shows the distribution of likes and comments given by users. The Complementary Cumulative Distribution Function for the users' activity also presents heavy tails. Hence, users activity is heterogeneous, the number of likes and comments they give range from very few, the majority, to hyperactivity.

The persistence of a post can be measured by the time difference between the time-stamp of its first and last comment. Also, the lifetime of a user would be the time difference between the time-stamp of their first and last given comment. Figure 14 shows the CCDF of the posts' persistence and users' lifetime considering all $366M$ comments. Thus, the persistence of posts online indicates that the majority are usually relevant for a few days, while only a select few gain notoriety for long periods of time. Similarly, the main bulk of the users display short periods of activity (from 1 to 100 days).

4.3 Selective Exposure

The overall number of likes given by each user is a good proxy of their level of *engagement* with the Facebook news pages. The *lifetime* of a user, meaning the period of time where the user started and stopped interacting with our set of pages, can be approximated by the time difference

between the time-stamp of their latest and earliest liked post. These measures could provide important insights about news consumption patterns, specifically, the heterogeneity of the user activity.

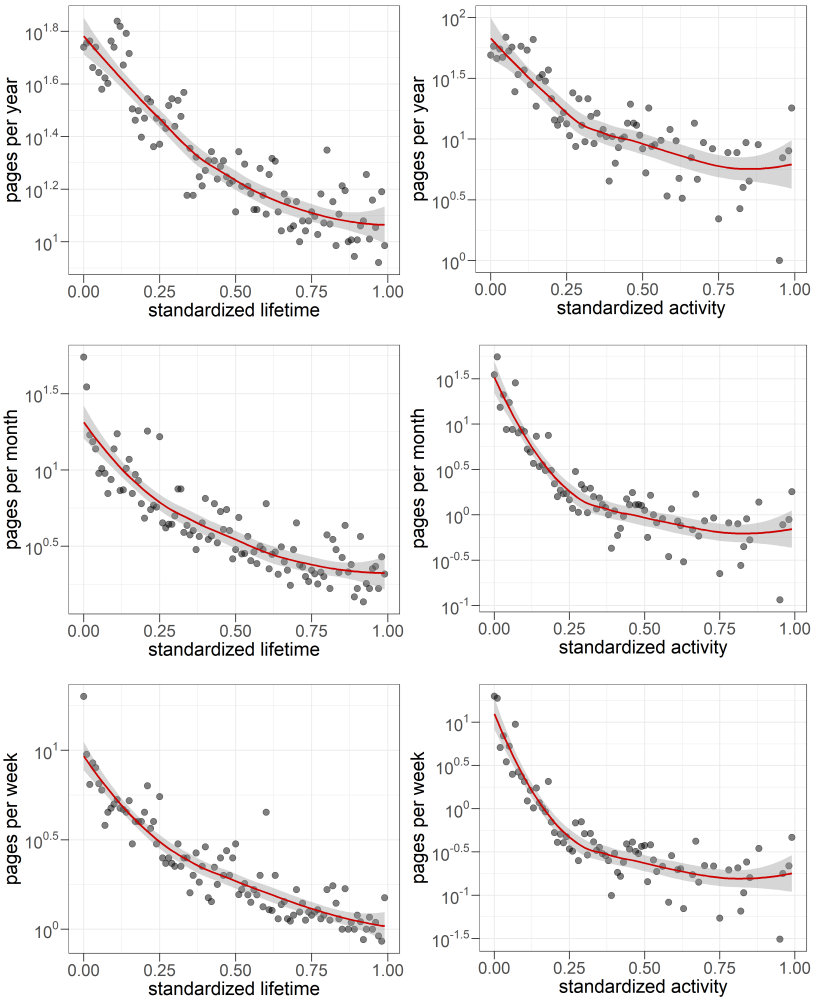

Figure 6: Selective Exposure. Maximum number of unique news sources that users with increasing levels of standardized lifetime (left) and standardized activity (right) interacted with yearly, monthly and weekly.

We say that a user has consumed a page in a given time window, if the user has at least one positive interaction with that page in that period, that is, the user *liked* a post made by that page. We do not consider *comments* as a valid interaction for regular consumption because they have very diverse meanings and, dissimilar from the *likes*, they do not unambiguously represent positive feedback. Thus, we measure the number of different news pages each user consumed in a weekly, monthly and yearly basis.

Figure 6 shows the maximum number of news sources users consumed considering longer periods of user activity and increasing levels of user engagement. For a comparative analysis, both lifetime and engagement are standardized between 0 and 1 over the entire user set. Only users who have given at least 10 likes in their entire lifetime are considered. As mentioned above, the results were calculated considering the yearly (top), monthly (middle) and weekly (bottom) rates.

In spite of the wide availability of news outlets, users only interact with a small number of pages. Higher levels of users' activity (number of likes) and longer lifetimes (active time on Facebook) correspond to fewer news sources being consumed, regardless of the time window considered. We can thus say that the users have a natural tendency to confine themselves to a limited set of pages. News consumption on Facebook is dominated by selective exposure.

4.4 Community Detection

Users' tendency to interact with few news sources might elicit communities of pages. To test this hypothesis, we characterize the emergent community structure of pages according to the users' activity.

As described in Chapter 3.2, the users' page likes can be projected to derive the weighted graph $\mathcal{G}_L$ (and $\mathcal{G}_C$) in which nodes are pages and two pages are connected if a user likes (or comments on) both of them. Thus, the weight of a link on a projected graph is determined by the number of users the two pages have in common.

Figure 7 shows the global geo-location of the news pages of the backbone of $\mathcal{G}_L$ ($\alpha = 0.001$). The opacity of the links is proportional to the number of users two pages have in common. The size of the nodes is proportional to the activity of that page in the entire period between 1st January 2010 to 31st December 2015, that is, the total number of posts made by the page.

Figure 7: Geo-location of the pages in the backbone of $\mathcal{G}_L$.

Once the projections are complete, we apply the Fast-Greedy (FG) community detection algorithm to detect clusters of pages within the networks. To validate the community partition, we compare with the Rand index (Ran71) the FG membership against the membership of other community detection algorithms. As seen in Table 3, there's a high level of similarity between the different partitions. This indicates that the users' activity leads to the emergence of stable clusters of news sources.

	WT	ML	LP
$\mathcal{G}_L$-FG	0.7552564	0.9669315	0.8448607
$\mathcal{G}_C$-FG	0.7463154	0.9749981	0.6731904

Table 3: Comparison of the Fast-Greedy communities of $\mathcal{G}_L$ and $\mathcal{G}_C$ against the community partitions of the other algorithms. FG, FastGreedy; WT, WalkTrap; ML, MultiLevel; LP, LabelPropagation.

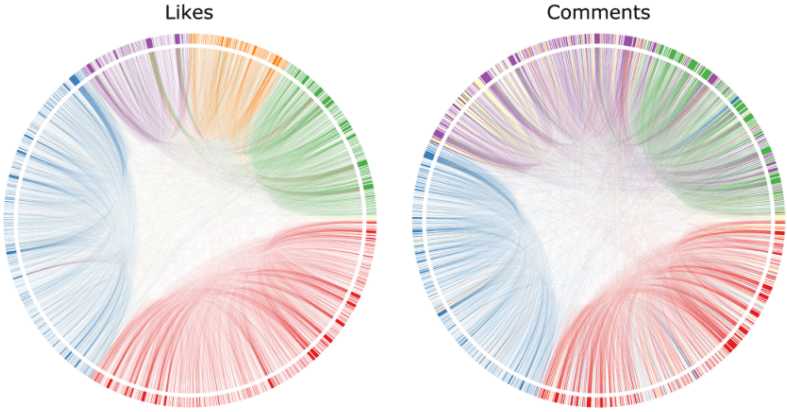

Figure 8: Community Partition. Backbone of the FastGreedy community partition of $\mathcal{G}_L$ (left) and $\mathcal{G}_C$ (right).

Figure 8 shows the community partitions of $\mathcal{G}_L$ and $\mathcal{G}_C$ according to the FastGreedy community detection algorithm. Each node represents a page, and the color of the node indicates the community that page belongs to. For a visual comparison, nodes in $\mathcal{G}_L$ are ordered according to the detected communities, whereas in $\mathcal{G}_C$, the nodes follow the same order established by $\mathcal{G}_L$. Note that, for visualization purposes, the plot shows the backbone structure of the networks (SBV09).

	FG	WT	ML	LP
$\mathcal{G}_L$ - $\mathcal{G}_C$	0.6230107	0.850695	0.6344089	0.614519

Table 4: Comparison of the partitions of $\mathcal{G}_L$ against those of $\mathcal{G}_C$.

Finally, considering the different community detection algorithms, we compare against each other the various partitions of $\mathcal{G}_L$ and $\mathcal{G}_C$. As shown in Table 4, there is some level of similarity between the communities generated by the users' likes and the comments. The passive consumption of news and the debate around them generate similar communities, however, there are slight differences that require further study.

4.5 User Polarization

The above results show that the users' activity leads to the emergence of communities of news pages. As outlined in Chapter 3.2.3, we can measure and visualize in a radial scatter plot the spread of the users' activity across different clusters.

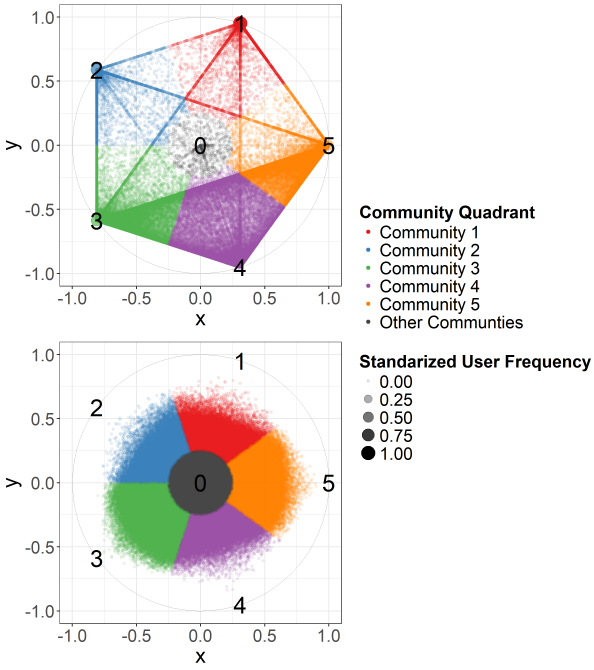

Figure 9: Users Polarization. Liking activity of the users across the five largest communities (top) and null model where users' activity is randomly distributed (bottom).

Figure 9 shows the activity of users across the five largest communities (top) and compares this with a null model (bottom) in which the users' activity is randomly distributed. Vertices of the pentagon represent the five largest communities, and the central point all of the remaining ones. The position of each dot is determined by the number of communities the users interacts with, whereas, the size and transparency indicate the number of users in that position. Only users with at least 10 likes in their liking activity were considered.

As shown in Figure 9, users are strongly polarized and their attention is generally confined to a single community of pages. Thus, we can conclude that users' interactions with Facebook news outlets form a dominant community structure with sharply-identified groups. The news sphere on Facebook is clustered and the users are polarized in distinct echo-chambers.

4.6 Users and News Outlets Perspectives

Facebook pages can like other pages. This pattern of favorite pages can be represented as a graph, $\mathcal{N}_P$, that reproduces the news outlets preferences. Thus, $\mathcal{N}_P$ is the network in which nodes are pages, and links are pages liking each other As in the previous section, we apply the FG community detection algorithm and compare the resulting partition against those obtained from the other algorithms. The Rand index value when comparing the FG communities of $\mathcal{N}_P$ is 0.860 for WT, 0.891 for ML and 0.646 for LP. Then, in the pages' favorites network ($\mathcal{N}_P$) the pages cluster together establishing solid community structures.

We can then compare the preferences of the users and the news providers by measuring the similarity of the community partitions of $\mathcal{N}_P$ against those of the bipartite projections $\mathcal{G}_L$ and $\mathcal{G}_C$. For a fair comparison $\mathcal{G}_L$ and $\mathcal{G}_C$ has been reduced to the set of pages P that appear in $\mathcal{N}_P$. As seen in Table 5, there is some similarity between the communities of $\mathcal{N}_P$ and those of $\mathcal{G}_L$ and $\mathcal{G}_C$. This indicates that there's some overlap between the preferences of the users and the news outlets.

	FG	**WT**	**ML**	**LP**
$\mathcal{G}_L$ - $\mathcal{N}_P$	0.6765225	0.7331109	0.7120639	0.5040556
$\mathcal{G}_C$ - $\mathcal{N}_P$	0.6932035	0.7804737	0.70506	0.5072296

Table 5: Comparison of the communities of $\mathcal{N}_P$ against $\mathcal{G}_L$ and $\mathcal{G}_C$.

Figure 10: FG communities and geographical locations of $\mathcal{G}_L$ and $\mathcal{N}_P$. AF, Africa; AS, Asia; CA, Central America; EU, European Union; EU-C, EU Candidate; EU-O, EU Other; GL, Global; ME, Middle East; NA, North America; OC, Oceania; SA, South America.

Another comparison between the users and the news providers can be made by taking into account the geographical location of each page. Thus, news pages can be clustered by their respective countries, information provided by the European Media Monitor, and then compared to the communities that emerge from the users' activity or the pages' preferences.

The Rand similarity index of the geographical partition of the pages against the FG community partitions is 0.71 for $\mathcal{G}_L$, 0.72 for $\mathcal{G}_C$ and 0.84 for $\mathcal{N}_P$. This suggests that the communities emerging from the news outlets preferences are more locally confined that those emerging from the users' consumption habits, which can span across nations and continents. It seems that despite selective exposure and the agenda setting power of news outlets, users consume from a more international range of sources.

Figure 10 shows the communities of $\mathcal{G}_L$ (top) and $\mathcal{N}_P$ (bottom) represented by taking into account the geographical location of the pages. Order of the nodes in both cases is done by region, nation and community, in that order. In the graph, the external bundle groups the pages by region and the middle bundle by nation, as established by the European Media Monitor. The inner nodes represent the news pages, and their color indicates their community membership as identified by the FG algorithm. Note that in the plot of $\mathcal{G}_L$ we use the backbone structure of the networks to simplify the visualization (SBV09) and thus the graph lacks most of its smaller isolated communities.

By looking at the FG communities and their presence in the different countries, for both $\mathcal{G}_L$ and $\mathcal{N}_P$, we can see that the user communities are more likely to span multiple countries. Some cases jump up, like Ireland (IE), Belgium (BE) and Switzerland (CH) where the communities that emerge from the news sources preferences are only one (IE and BE) or two (CH), while if we look at the user communities present in those countries outlets we find more of them.

4.7 The Model

Users on Facebook tend to focus on a limited set of news sources, on a macro scale. This mechanism of selective exposure generates a clustered and polarized structure. In this section we provide a simple model of users' preferential attachment to specific sources that reproduces the observed community structure.

The entities of our model are pages $p \in P$ and users $u \in I$. We then define for each page p a set of opinions c_p, modeled as a real number that ranges $[0..1]$. This will represent the an editorial line of the news outlet. For each user u we define an initial opinion θ_u, also modeled as a real number and ranging $[0..1]$. We assume that both, the c_p and the θ_u values, are uniformly distributed.

We suppose c_p and θ_u to be homogeneous such that the quantity $|c_p - \theta_u|$ is the distance between the opinion of user u and the editorial line of page p. We mimic confirmation bias by assuming that if user u interacts with a page p and the opinion distance $|c_p - \theta_u|$ is less than a given tolerance parameter Δ, the preference of user u will converge toward the editorial line of page p according to the BCM (DNAW00) equation:

$$\theta'_u = (1 - \mu) \cdot \theta_u + \mu \cdot c_p \tag{4.1}$$

where μ is a simple convergence parameter.

To simulate the user activity, each user u will have an activity coefficient a_u that represents the number of pages the user can visit. To mimic the long tail distribution of our data we set the activity distribution to be power law distributed $p(a) \sim a^{-\gamma}$ with exponent $\gamma = 3$. Thus, the final opinion of a user will average the editorial lines of the pages the user liked.

If Ω is the set of $|\Omega|$ pages that matches the preferences of user u, then the average opinion will be:

$$\overline{\theta}_u = (1 - \mu)\, \overline{\theta}_u + \mu\, |\Omega|^{-1} \sum_{p \in \Omega} c_p \tag{4.2}$$

We then use numerical simulation to study our model. A user randomly selects a subset of P with which to interact. The user likes a page only when the opinion distance is below the tolerance parameter, that is, $|c_p - \theta_u| < \Delta$. When this occurs, the feedback mechanism of the equation 4.1 reinforces the users' page preferences.

Once a user's opinion converges, we build the bipartite graph $\mathcal{B}_{\text{sim}} = (I, P, E_{\text{sim}})$ where the set of edges E_{sim} are the pairs (u, p) where user u liked page p. Hence $\mathcal{B}_{\text{sim}}$ will represent users interacting with their favorite pages, and from $\mathcal{B}_{\text{sim}}$ we can build the projected graph G^p_{sim} that links the pages according their common users.

Figure 11 shows an analysis of G^p_{sim} as a function of the tolerance parameter Δ. Each point of the simulation is averaged over 50 iterations. The left panel shows both the size of the largest connected component

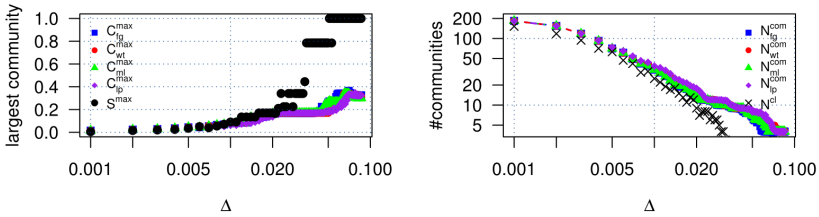

Figure 11: Analysis of the synthetic pages-to-pages graph $G^p_{\rm sim}$ generated according to the extension of the BCM model. The left panel shows the size of the largest component ($S_{\rm max}$) and the largest community ($|C^{\rm alg}_{\rm max}|$) according to several algorithms versus the tolerance parameter Δ. The right panel shows the number of connected components ($N_{\rm clu}$) and the number of detected communities ($N^{\rm alg}_{\rm com}$). Notice that by definition the number of communities must be $N^{\rm alg}_{\rm com} \geq N_{cl}$.

$S_{\rm max}$ and the size of the largest community $|C^{\rm alg}_{\rm max}|$ as detected by several algorithms on $G^p_{\rm sim}$. The right panel shows the number of connected components $N_{\rm clu}$ and the number of communities $N^{\rm alg}_{\rm com}$ detected by the different algorithms of $G^p_{\rm sim}$.

At $\Delta \sim 0$, $G^p_{\rm sim}$ breaks down into disconnected pieces. The number of connected components, $N_{\rm clu}$, is of the order of the number of nodes $|P|$ of $G^p_{\rm sim}$, and the size of the largest component, $S_{\rm max}$, is extremely small. Although his regime is unlike real online social networks that are usually dense and strongly connected, $S_{\rm max}$ increases rapidly to $|P|$ as Δ departs from zero, indicating that $G^p_{\rm sim}$ becomes a single connected graph at $\Delta \sim 0.03$. On the other hand, the size of the largest communities as detected by the various algorithms are consistently smaller than $|P|$, as seen on the left of Figure 11, and the number of communities is consistently greater than one and decreases slowly with increasing Δ, as seen on the right of Figure 11.

Thus, the synthetic page-to-page graph $G^p_{\rm sim}$ shows a stable, non-trivial community structure induced by user preferences even when it is a dense, connected graph like real online communities. We can then say that the model replicates the observed behavior seen in the previous sections.

4.8 Concluding Remarks

Using quantitative analysis we show that the more active a user is, the more the user tends to focus on a small number of news sources. Looking at the page clusters generated by user activity, we find a distinct community structure and strong user polarization. We provide evidence that preferences of users and news outlets differ in that communities established by page creators are more locally confined than communities identified by the users' activity, which can span across international borders. This segregation in distinct communities can be reproduced by a simple model that mimics the selective exposure of users.

Content consumption on Facebook is strongly affected by the tendency of users to limit their exposure to a few sites. Despite the wide availability of content and heterogeneous narratives, there is major segregation and growing polarization in online news consumption where news undergo the same popularity dynamics as videos of kittens and selfies. The spreading of fake news and unsubstantiated rumors motivated major corporations like Google and Facebook to provide solutions to the problem. Google news decided to flag fact-checked information and to penalize providers of fake news, others are proposing to use black lists of sources in order to automatically limit their spread.

Debates, however, especially on socially relevant issues, are often based upon conflicting narratives. Our findings suggest that probably the main driver of misinformation diffusion is the polarization of users on specific narratives rather than the lack of fact-checked certifications. Probably, the main problem behind misinformation is polarization of users online.

Chapter 5

A Study on European News Consumption on Facebook

All the results shown in this chapter refer to the paper (SZSQ17)[1]. Here we use quantitative analysis to characterize and contrast the news consumption patterns on Facebook of four European countries: France, Germany, Italy and Spain.

The advent of social media and microblogging platforms has significantly changed the way we consume information and form opinions. A substantial number of users consuming news through social media (MHP13; BSGM15). From 2013 to 2015 users who got their news from Facebook went from 47% to 63%. The growth, however, was not equal all over the globe, with different countries displaying varying rates of acclimatization to these technologies (NLN15; NLN16; NFK+17).

Spurred by the sharing of news through social media platforms and news customization, tailored to the user's specific tastes and beliefs, false and misleading narratives are currently rampant, with no clear end in sight (AL17). Social media does increases the number of news sources a user is exposed to compared to the consumption through traditional news channels (ACGC11). However, while the quick consumption of

[1]The results shown in this Chapter are all part of the paper (SZSQ17). It is a joint work with Dr. Fabiana Zollo, Dr Antonio Scala and Dr. Walter Quattrociocchi. ALS, WQ and AS outlined the research question. ALS and AS provided the analysis tools. ALS, FZ, AS and WQ performed the analysis and interpreted the results. ALS, AS and WQ contributed equally to the writing and reviewing of the manuscript.

online news might increase awareness of the most relevant events, it does not really broaden the variety of topics seen (SdWL05).

Recent works (QSS16) provide empirical evidence of the pivotal role of confirmation bias and selective exposure in online social dynamics. Users tend to focus on specific narratives and join polarized groups where they end up reinforcing their beliefs (BPDV$^+$15; DVBZ$^+$16), even if pieces of the content consumed are deliberately false (MRZ$^+$15; BCD$^+$15), and dismissing contradictory information (ZBDV$^+$15). Discussion and elaboration of narratives in such a segregated environment elicits group polarization and negatively influences user emotion(Sun02a; ZNDV$^+$15; YB10; BMA15).

In this paper we compare the consumption patterns of news on Facebook of four European countries: France, Germany, Italy and Spain. First, we compare the pages' posting behavior and the users' interacting patterns across countries. Second, we explore the selective exposure of the users and the polarized communities generated by their consumption patterns. Then, we measure the polarization of the users of each country and rank them accordingly. Finally, we present a variation on the Bounded Confidence Model (DNAW00) to simulate the emergence of these communities by considering the users' trust on the news.

5.1 Data Description

In this chapter we use the EU Countries described in Chapter 3.1. As explained before, it was generated from a list of top news sources, in their official language, of France, Germany, Italy and Spain. The list for each country was compiled considering the Reuters Digital News Report of 2015 (NLN15) and 2016 (NLN16).

We then obtained the official Facebook page of each news outlet and proceeded to download all the posts made from 1st January 2015 to 31st December 2016, as well as all likes and comments that have been made on those posts. A breakdown of the dataset can be seen in Table 6, and the list of news sources with their respective Facebook pages and country of origin can be found in the Appendix A.2.

In February 2016 Facebook expanded the possible user responses from just a *like* to a range of emoticons (*love*, *wow*, *haha*, *sad* and *angry*). Considering this new feature, we downloaded for all the posts made from 1st July 2016 to 31st December 2016 all the public user *reactions*. While not used in this study, this data will prove useful in future works where we

	France	Germany	Italy	Spain
Pages	65	49	54	57
Posts	1,008,018	749,805	1,554,817	1,372,805
Likes	419,371,366	183,599,003	409,243,176	333,698,985
Likers	21,647,888	14,367,445	14,012,658	32,812,007
Comments	47,225,675	31,881,407	51,515,121	34,336,356
Commenters	5,755,268	5,338,195	4,086,351	6,494,725
Users	22,560,889	15,564,360	14,587,622	34,383,820
Population	$66M$	$81M$	$62M$	$46M$

Table 6: EU Countries Dataset Numbers. Population according to (NFK$^+$17). Users is the number of people that gave likes and/or comments. Likers is the number of people that gave likes. Commenters is the number of people that gave comments.

plan to analyze the emotional response of users considering their country.

5.2 Attention Patterns

As a first step we characterize how different countries consume news on Facebook. We focus particularly on the allowed users' actions through the entire period of the data collection: *likes*, *shares* and *comments*. Naturally, each action has prescribed meaning. A *like* represents a positive feedback to a post; a *share* expresses the user's desire to increase the visibility of a given piece of information; and a *comment* is the way in which online collective debates take form. Therefore, comments may contain negative or positive feedback with respect to the post.

In Figure 12 we show the distribution of the number of likes, comments and shares received by the posts belonging to each country. As seen from the plots, all the distributions are heavy-tailed, that is, they are best fitted by power laws (as seen in Table 7) and posses similar scaling parameters (Table 8) with some notable differences when looking at the number of comments and likes.

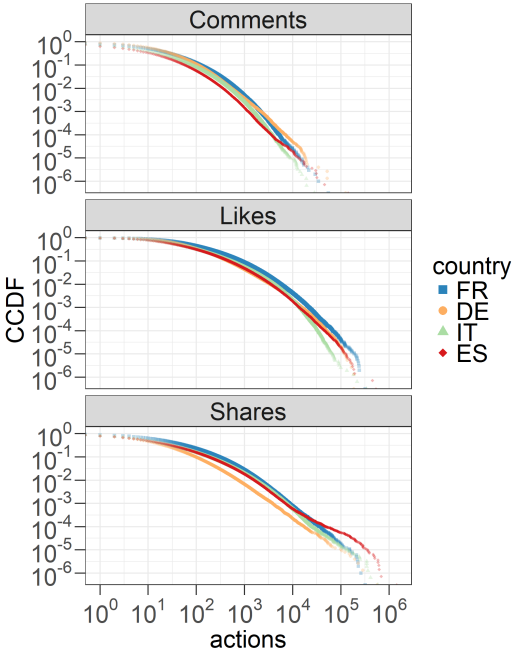

Figure 12: Complementary Cumulative Distribution Function of the comments, likes and shares received by the posts of each country.

	Action	Poisson	Log-Normal	Exponential	PowerLaw
FR	comment	$-81,474,887$	$-4,094,569$	$-5,042,908$	$-10,467$
DE	comment	$-53,857,610$	$-3,208,655$	$-3,692,816$	$-124,780$
IT	comment	$-92,959,791$	$-5,353,204$	$-7,158,219$	$-9,815$
ES	comment	$-64,633,469$	$-4,192,227$	$-5,896,527$	$-28,449$
FR	like	$-716,163,037$	$-6,463,931$	$-7,135,298$	$-9,793$
DE	like	$-336,233,651$	$-4,429,366$	$-4,906,736$	$-243,938$
IT	like	$-732,132,678$	$-9,034,577$	$-10,276,500$	$-12,514$
ES	like	$-625,371,478$	$-7,905,112$	$-8,978,996$	$-34,532$
FR	share	$-302,119,999$	$-5,029,592$	$-6,102,954$	$-68,981$
DE	share	$-100,787,846$	$-2,972,740$	$-3,809,317$	$-37,466$
IT	share	$-399,573,409$	$-6,760,982$	$-8,902,324$	$-24,265$
ES	share	$-456,628,686$	$-5,852,126$	$-7,960,407$	$-128,667$

Table 7: Maximum-Likelihood fit of the posts different actions by country. FR: France, DE: Germany, IT: Italy, ES: Spain.

	Comments		Likes		Shares	
	$\hat{X}_{min}$	$\hat{\alpha}$	$\hat{X}_{min}$	$\hat{\alpha}$	$\hat{X}_{min}$	$\hat{\alpha}$
FR	$1,929$	3.44	$23,338$	3.09	$2,498$	2.63
DE	315	2.63	$1,132$	2.25	$1,084$	2.45
IT	$1,736$	3.63	$15,519$	3.71	$5,753$	2.79
ES	733	3.10	$8,491$	2.89	$1,508$	2.47

Table 8: Powerlaw fit of posts' attention patterns.

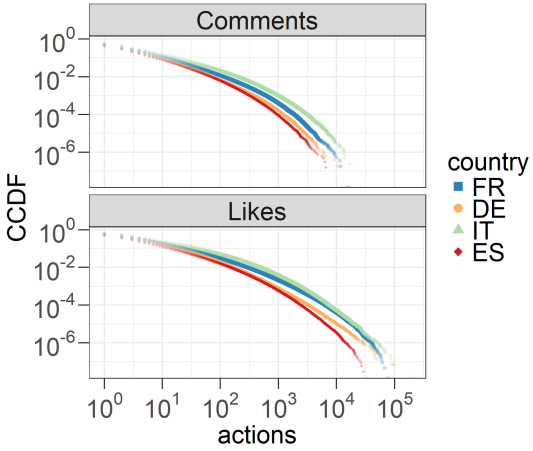

Figure 13: Complementary Cumulative Distribution Function of the likes and comments given by the users of each country.

We continue the analysis by examining how users from each country interact with the pages. In Figure 13, we show the distribution of the number of likes and comments given by the users according to each country. Once again, all the distributions are heavy-tailed, as seen in Table 10, with some notable differences in their scaling parameters (Table 9) when considering the commenting activity of the users of the different countries.

	Comments		Likes	
	$\hat{X}_{min}$	$\hat{\alpha}$	$\hat{X}_{min}$	$\hat{\alpha}$
FR	$2,378$	4.07	648	2.45
DE	18	2.17	$3,156$	3.02
IT	529	2.70	$5,473$	3.26
ES	1	1.90	$1,876$	3.24

Table 9: Power law fit of users' attention patterns.

	Action	Poisson	Log-Normal	Exponential	PowerLaw
FR	comment	$-96,157,561$	$-13,916,586$	$-17,869,083$	$-2,433$
DE	comment	$-57,870,795$	$-11,951,470$	$-14,878,272$	$-1,268,430$
IT	comment	$-114,865,937$	$-10,972,733$	$-14,442,076$	$-81,786$
ES	comment	$-62,141,913$	$-13,638,119$	$-17,309,835$	$-11,920,701$
FR	like	$-1,042,576,644$	$-63,945,214$	$-85,808,958$	$-643,618$
DE	like	$-377,979,910$	$-40,766,648$	$-50,972,666$	$-13,788$
IT	like	$-985,441,955$	$-45,609,241$	$-61,296,249$	$-26,385$
ES	like	$-720,112,905$	$-83,156,334$	$-108,917,647$	$-48,326$

Table 10: Maximum-Likelihood fit of the users' different actions by country. FR: France, DE: Germany, IT: Italy, ES: Spain.

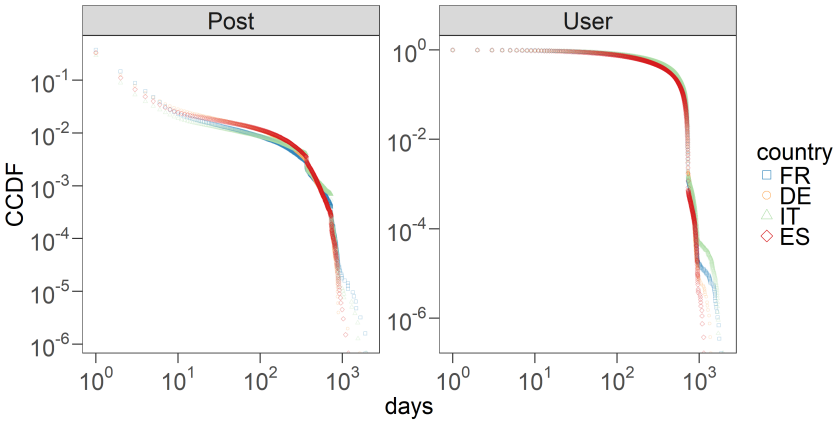

Figure 14: Complementary Cumulative Distribution Function of the posts' persistence (left) and the users' lifetime (right) for each country.

Finally, we analyze for each country the persistence of the posts (the time difference between the time-stamp of their first and last comment) and the lifetime of the users (the time difference between the time-stamp of their first and last given comment). Figure 14 shows the CCDF of the posts' persistence and users' lifetime of each country, considering all their respective comments. For all countries, the persistence of posts online is quite similar, with the majority of them relevant for a few days, and only a select few gaining notoriety for long periods of time. There's no notable differences between the countries when considering their users' lifetime.

5.3 Selective Exposure

As in the previous chapter, we will say that a user has consumed a page in a given time window, if the user has at least one positive interaction with that page in that period, that is, the user *liked* a post made by that page. For the users of each country, we then measure their *activity* (total number of likes) and their *lifetime* (time difference of their first and last liked post). Thus, we can measure the collection of pages consumed in a weekly, monthly and quarterly basis while taking into account the activity and engagement of the users of each country.

Figure 15 shows the number of news sources a user interacts with considering their lifetime and for increasing levels of engagement for each country. For a comparative analysis, we standardized between 0 and 1 the number of pages present in each country, as well as the lifetime and engagement over the entire user. The results were calculated considering the quarterly (top), monthly (middle) and weekly (bottom) rates.

Note that for all countries, users usually interacts with a small number of news outlets and that higher levels of activity and longer lifetime correspond to a smaller variety of news sources being consumed. We can also observe clear differences between the countries. When considering the users' lifetime, France has clearly a more varied news consumption diet than the rest; and when considering the users' activity Germany users consume consistently the less diverse set of news sources.

We can conclude that there is a natural tendency of the users to confine their activity to a limited set of pages, news consumption on Facebook is indeed dominated by selective exposure (SZDV$^+$17) and users from different countries display different rates for the decreasing variety

of news outlets they consume.

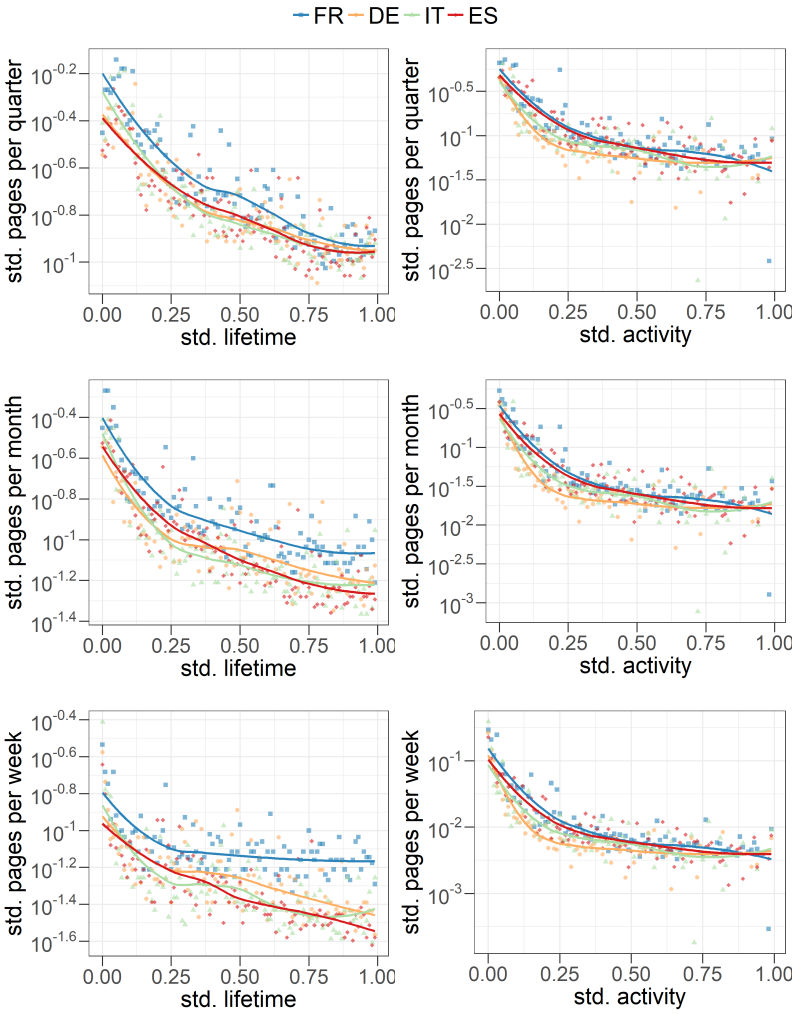

Figure 15: Selective Exposure. Maximum number of unique news sources that users with increasing levels of standardized lifetime (left) or standardized activity (right) interact with quarterly, monthly and weekly for each country.

5.4 Emerging Communities

User tendency to interact with few news sources might elicit page clusters. To test this hypothesis, we first characterize the emergent community structure of pages according to the users' activity for each country κ with $\kappa = \{FR, DE, IT, ES\}$. We project the user page likes to derive the weighted graph $\mathcal{G}_L^\kappa$ (and $\mathcal{G}_C^\kappa$) in which nodes are pages and two pages are connected if a user likes (or comments on) both of them. The weight of a link on a projected graph is determined by the number of users the two pages have in common.

We then apply the FastGreedy community detection algorithm to see if there are well-defined communities for each case. To validate the community partitioning, we then compare the membership found by the other community detection algorithms using the Rand method (Ran71) and find high level of similarity for all four countries, as seen in Table 11).

$\mathcal{G}$	κ - **Country**	**Type**	**ML**	**SG**
$\mathcal{G}_L^{FR}$	France	Likes	0.795	0.796
$\mathcal{G}_L^{DE}$	Germany	Likes	0.771	0.838
$\mathcal{G}_L^{IT}$	Italy	Likes	0.982	0.851
$\mathcal{G}_L^{ES}$	Spain	Likes	0.923	0.981
$\mathcal{G}_C^{FR}$	France	Comments	0.918	0.969
$\mathcal{G}_C^{DE}$	Germany	Comments	0.836	0.925
$\mathcal{G}_C^{IT}$	Italy	Comments	0.871	0.903
$\mathcal{G}_C^{ES}$	Spain	Comments	0.828	0.817

Table 11: Algorithm comparison. Comparison between the FastGreedy (FG) communities against the MultiLevel (ML) and SpinGlass (SG) communities for both the likes and comments projections of every country.

We also compared the communities of $\mathcal{G}_L^\kappa$ and $\mathcal{G}_C^\kappa$ against each other using different community detection algorithms and find, overall, low levels of similarity (see Table 12). This indicates that, for all four countries, the set of pages users generally approve of (like), differ from the set of pages where they often debate (comment).

Comparing	κ - Country	FG	ML	SG
$\mathcal{G}_L^{FR}$-$\mathcal{G}_C^{FR}$	France	0.514	0.522	0.545
$\mathcal{G}_L^{DE}$-$\mathcal{G}_C^{DE}$	Germany	0.528	0.537	0.518
$\mathcal{G}_L^{IT}$-$\mathcal{G}_C^{IT}$	Italy	0.562	0.560	0.619
$\mathcal{G}_L^{ES}$-$\mathcal{G}_C^{ES}$	Spain	0.555	0.554	0.625

Table 12: Comparison of the projection of likes and comments for every country. Comparison of the communities detected in $\mathcal{G}_L^{\kappa}$ and $\mathcal{G}_C^{\kappa}$, of each country, with FastGreedy (FG), MultiLevel (ML) and SpinGlass (SG).

5.5 User Polarization

By examining the activity of users across the various clusters and measuring how they span across news outlets, we find that most users remain confined within specific groups of pages. To understand the relationship between page groupings and user behavior, we measure the polarization of users with respect to the communities found for each country κ where $\kappa = \{FR, DE, IT, ES\}$.

As outlined in Chapter 3.2.3, we can measure and visualize in a radial scatter plot the spread of the users' activity across different clusters for all countries. Considering the users with at least 10 likes, we calculated their position by considering their liking activity. Figure 16 shows the activity of the users of each country across all the communities found for each country with the FastGreedy algorithm. The vertices of the regular polygons represent the communities found for each country, France and Italy have 4, Germany has 3 and Spain has 5.

As seen in Figure 16, users are strongly polarized and their attention is generally confined to a single community of pages. Irregardless of the country, users' interactions with Facebook news outlets form a dominant community structure with sharply-identified groups. The polarization of the users, however, might be different when considering the country.

For each user u with a total of K likes, where $\sum_i k_i = K$ such that each k_i belongs to the i^{th} community ($i = 1 \dots N$), we can calculate their localization order parameter L which effectively *counts* the number of communities they are active in. Since we are considering many users, we plot the probability distribution of L along the entire user set for each country κ. This would allow for a fair comparison of the polarization of the users between countries.

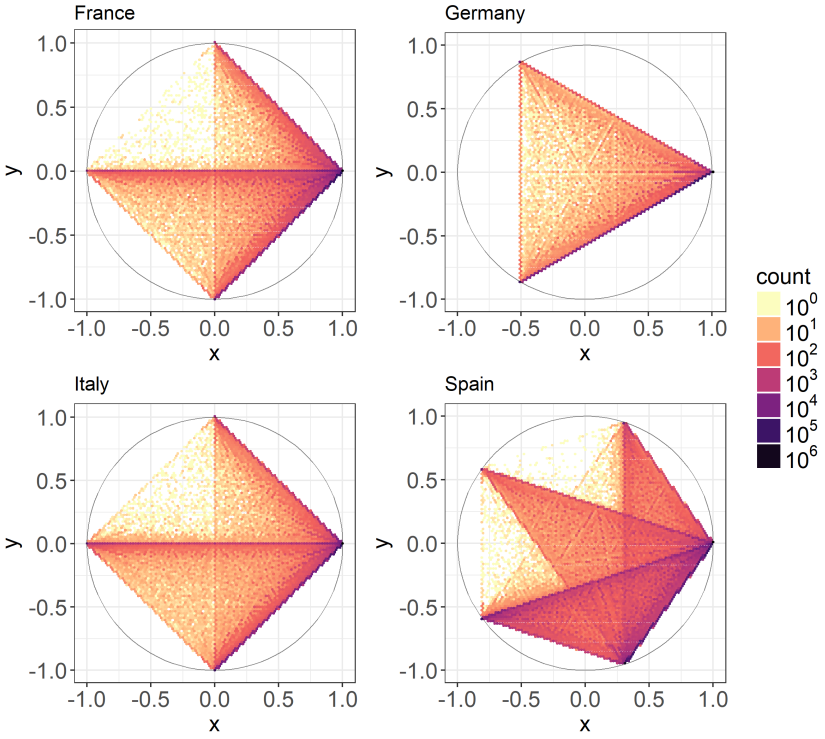

Figure 16: Polarization of the users of each country κ by considering all the communities as found with FastGreedy. The users are grouped in 100 bins.

Figure 17 shows, for each country, the Probability Density Function of the localization L of all users with at least 10 likes. As we can see, the density is well behaved, that is, presents a single peak. The mean value for each country allows us to rank them from the one with least polarized users to the one with the most: Spain (1.199), Germany (1.125), France (1.106) and Italy (1.099).

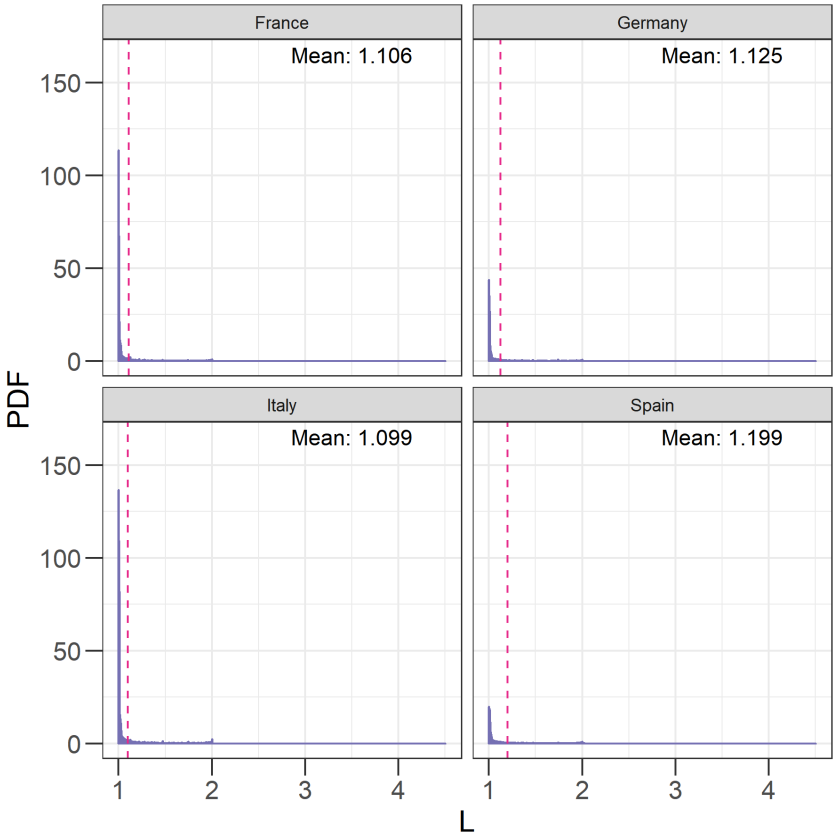

Figure 17: Probability Density Function of L for each country κ. The orange vertical line indicates the mean value.

5.6 The Model

In this section we provide a simple model of users' preferential attachment to specific sources that considers the users' *trust* on the media as a parameter and reproduces the observed community structure.

The entities of our model are pages $p \in P$ and users $u \in U$. Each page p is characterized by a set of opinions (an editorial line) modeled as a real number c_p that ranges $[0..1]$. We assume that the c_p values are uniformly distributed. Each user u has an initial opinion that is modeled as a real number θ_u, which ranges between $[0..1]$ and it is uniformly distributed. Each users u also has a measure of trust in the media modeled by the real number τ_u, which ranges between $[0..1]$. User's trust will follow a truncated normal distribution.

We suppose c_p and θ_u to be homogeneous such that the quantity $|c_p - \theta_u|$ is the distance between the opinion of user u and the editorial line of page p. We mimic confirmation bias by assuming that if user u interacts with a page p and the opinion distance $|c_p - \theta_u|$ is less than a given tolerance parameter Δ, the preference of user u will converge toward the editorial line of page p according to the modified BCM (DNAW00) equation:

$$\theta'_u = (1 - \tau_u) \cdot \theta_u + \tau_u \cdot c_p \tag{5.1}$$

To mimic user activity we give each user u an activity coefficient a_u that represents the number of pages a user can visit. Thus, the final opinion of a user will average the editorial lines of the pages the user likes. If Ω is the set of $|\Omega|$ pages that matches the preferences of user u, then the average opinion will be $\overline{\theta}_u = (1 - \tau_u)\, \overline{\theta}_u + \tau_u\, |\Omega|^{-1} \sum_{p \in \Omega} c_p$, i.e., $\overline{\theta}_u = |\Omega|^{-1} \sum_{p \in \Omega} c_p$. To mimic the long tail distribution of our data we set the activity distribution to be power law distributed $p(a) \sim a^{-\gamma}$ with exponent $\gamma = 3$.

We use numerical simulation to study our model. A user randomly selects a subset of P with which to interact. The user likes a page only when $|c_p - \theta_u| < \Delta$. When this occurs, the feedback mechanism reinforces the user's page preference using the trust parameter τ_u to control the extent of the feedback. Thus the final opinion of a user will be the average of the editorial lines of the pages the user likes.

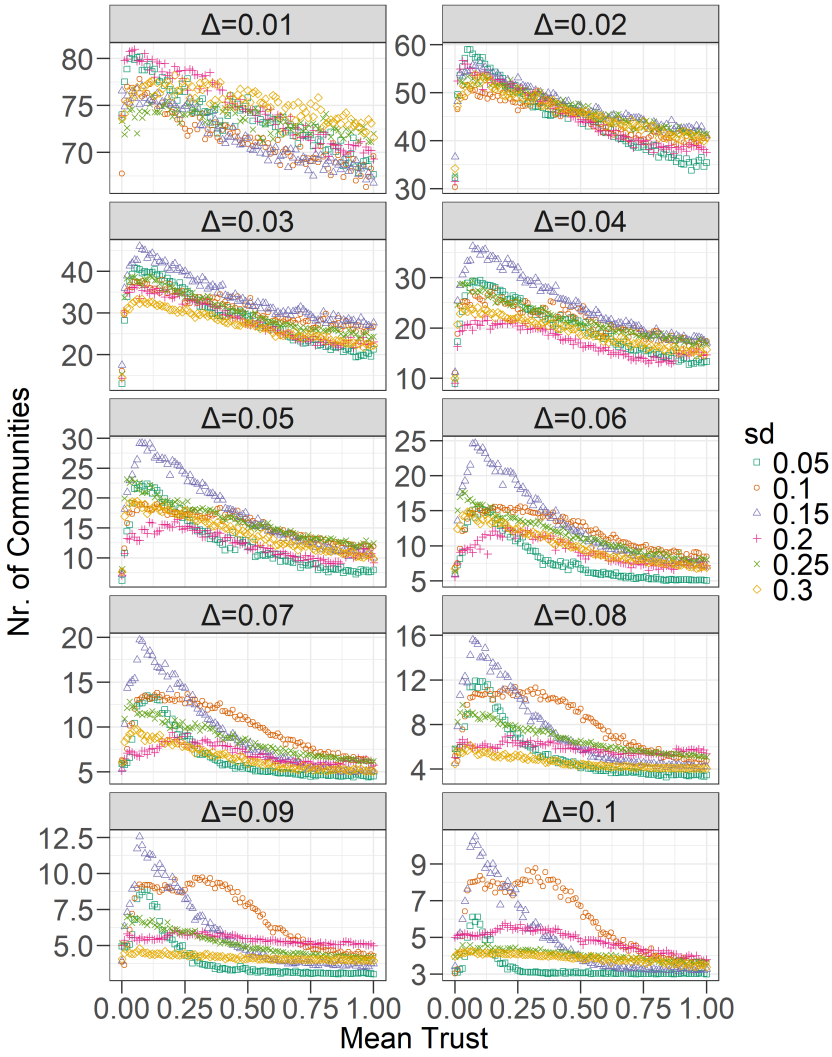

Figure 18: Analysis of the synthetic pages-to-pages graph G^{p}_{sim}. It shows the number of communities as a function of the mean user trust.

When a user's opinion converges, we build in the bipartite graph $\mathcal{G}_{\text{sim}} = (I, P, E_{\text{sim}})$ where the set of edges E_{sim} are the couplings (u, p) with which user u likes page p. Hence, $\mathcal{G}_{\text{sim}}$ represents users interacting with their favorite pages, and from $\mathcal{G}_{\text{sim}}$ we can build the projected graph G^p_{sim} that links the pages according their common users.

Figure 18 shows an analysis of G^p_{sim} as a function of the mean values used for the truncated normal distribution that models the trust τ, with different standard deviations and tolerance. Each point of the simulation is averaged over 100 iterations.

We can see that increasing the tolerance Δ leads to the reduction of the number of communities, that is, agreement is reached faster and polarization takes place. Very low, and higher values of the user trust also display similar behavior. No trust or absolute trust in the media leads to fast polarization, either the user will trust what they read fully and change their opinion accordingly, or they won't.

The simulation displays an interesting behavior at $\tau = 0.1$ where the number of communities formed by the users' consumption habits seem to peak. This indicates that some skepticism might actually factor against polarization. Users' who distrust the news they interact with, even when their opinions were similar, are more reluctant to further change their own beliefs. Perhaps a solution for the issue of false and misleading narratives could be found by fostering critical readers.

5.7 Concluding Remarks

In this chapter we use quantitative analysis to understand and compare the news consumption patterns of four European countries: France, Germany, Italy and Spain. We show that while there are similarities in the consumption behaviours between the four countries, the posting and consumption behavior is not universal.

The results also show that all users, regardless of country, display selective exposure, that is, the more active a user is on Facebook the less variety of news sources they tend to consume. This behavior is seen in all four countries, with different rates of selective exposure for each case. News consumption on Facebook is dominated by selective exposure.

Additionally, we studied the cluster of news pages that emerge from the user's activity, and found that users, regardless of their nationality, are polarized. We then measure the polarization of the users of each country, and ranked them accordingly, finding that Italy contains the

most polarized users, followed by France, Germany and finally Spain. Further studies might gain insights into the reasons behind the slight variations in consumption habits.

Finally, we present a variation on the Bonded Confidence Model (DNAW00) that mimics the users' behavior of selective exposure taking into account user trust. The simulation seems to indicate that users' who have some distrust of the news they interact with, even when the narrative presented conforms to their beliefs, are more reluctant to further change their own beliefs. Thus, a tentative solution for the issue of user polarization might be found by fostering critical readers.

Chapter 6

Polarization of the Vaccination Debate on Facebook

All the results shown in this chapter refer to the paper (SZS$^+$18)[1]. Here we use quantitative analysis on Facebook data to further study the evolution of the vaccination debate.

Undeterred by the scientific consensus that vaccines are safe and effective, unsubstantiated claims against their safety still occur to this day. Perhaps the most famous case is the multiple times disproved (CLSF04; DeS07; oP13) myth that the MMR vaccine causes autism. However, outbreaks and deaths resulting from objections to vaccines continue to happen (ZWH$^+$15; Boc17), with the anti-vaccination movement gaining media attention as a result. Mandatory vaccination policies only seem to add to the controversy (BB15; Lor17).

Recent studies outline that spreading misinformation is a consequence of the shift of paradigm in the consumption of content induced by the advent of social media. Social media platforms like Facebook or Twitter have created a direct path for users to produce and consume content,

[1]The results shown in this Chapter are all part of the paper (SZS$^+$18). It is a joint work with Dr. Fabiana Zollo, Dr. Antonio Scala, Dr. Cornelia Betsch and Dr. Walter Quattrociocchi. ALS, FZ, AS, WQ and CB outlined the research question. ALS and CB provided the analysis tools. ALS, CB and WQ performed the analysis and interpreted the results. ALS, CB and WQ contributed equally to the writing and reviewing of the manuscript.

reshaping the way people get informed (BBL07; KK04; QCS14; KMM10; SZDV+17; BBB+12).

Like for other misinformation campaigns, Facebook provides an ideal medium for the diffusion of anti-vaccination ideas. Users can access a wide amount of information and narratives and selection criteria are biased toward personal viewpoints (BCD+15; BPDV+16; MRZ+15). Online users select information adhering to their system of beliefs and tend to ignore dissenting information and to join polarized groups that cooperated to reinforce and frame a shared narrative (DVBZ+16; QSS16; DVVB+16). The interaction with content dissenting the shared narrative is mainly ignored or might even foment segregation of users, heated debating and thus bursting polarization of opinions (ZBDV+15). Such a scenario is not limited just to conspiracy theories, but it is related to all issues that are perceived as critical by the users such as geopolitics and health (BZDV+15). This effect allows for the emergence of polarized groups (SZDV+17), i.e. clusters of users with opposing views that rarely interact with one another.

In this chapter we use quantitative analysis to understand the evolution of the debate about vaccines on the entire public sphere of Facebook, taking into account two opposing views: *anti-vaccines* and *pro-vaccines*. By measuring the liking and commenting behavior of $2.6M$ users, we study the evolution of the two communities over time, considering the number of users, the number of pages and the cohesiveness of the communities. The analyses confirm the existence of two polarized communities that barely interact with each other. Additionally, we find evidence that while selective exposure plays an essential role in the way users consume content online (SZDV+17), different echo chambers display different rates at which the variety of news sources diminishes for increasing levels of user activity. This is consistent with the results obtained in Chapters 4 and 5.

6.1 Data Description

In this chapter we use the Vaccine Dataset described in Chapter 3.1. As explained before, the Facebook pages in the dataset were the result of requests to Facebook for pages that contain in their name or description the keywords *vaccine*, *vaccines* or *vaccination*. We then cleaned the raw Facebook results. Inclusion criteria were language (English), a minimum level of activity (at least 10 posts), and relation of the page to the topic of

vaccination. This last step was essential, as having one of the keywords in the title or description of the Facebook page does not necessarily mean the page's topic is about vaccines. Some examples of those false positive search results are the pages *The Vaccines* (an UK music band) and *Arthur D'vaccine* (a comedian).

From the resulting set of Facebook pages we downloaded all the posts made from 1st January 2010 to 31st May 2017, as well as all likes and comments made on those posts. Considering the content of the posts made on the pages, all the Facebook pages were also manually classified by two raters into two groups: *pro-vaccines* and *anti-vaccines*. The Cohens kappa inter-agreement between both raters is 0.966, showing nearly perfect agreement.

	Total	**Anti-vaccines**	**Pro-vaccines**
Pages	243	98	145
Posts	298,018	189,759	108,259
Likes	24,155,735	12,696,440	11,459,295
Comments	2,101,048	1,351,839	749,209
Users	2,640,513	1,277,170	1,388,677
Likers	2,456,755	1,145,650	1,325,511
Commenters	410,062	271,598	146,196

Table 13: Vaccination Dataset Numbers. Users is the number of people that gave likes and/or comments. Likers is the number of people that gave likes. Commenters is the number of people that gave comments.

A breakdown of the dataset can be seen in Table 13 and the list of Facebook pages and with their respective community label can be found in the Appendix A.3.

6.2 Communities and Polarization

In order to validate the manual partition of the pages into two communities, as explained in Chapter 3.2, we generate the projections of the bipartite networks considering the user likes, $\mathcal{G}_L$, and the user comments, $\mathcal{G}_C$. We then applied the community detection algorithms to extract the communities of pages according to the users' behavior and compare those to the manual partition. A high similarity between the manual and the different unsupervised approaches would indicate that the manual par-

tition accurately classified the pages into communities that emerge from the users' consumption habits.

Table 14 shows the comparison between a random partition of the pages, the manual partition and the FastGreedy partition against those resulting from the different algorithms. We can see that the manual classification matches well against the unsupervised approaches, even more so for $\mathcal{G}_L$, and that the FastGreedy results have a high agreement with the other algorithms. This indicates that the users' behavior generates well defined communities of pages and that these communities are similar to the *anti-vaccines* and *pro-vaccines partition* as manually tagged.

$\mathcal{G}$	**Communities**	**FG**	**WT**	**ML**	**LP**
$\mathcal{G}_L$	Random	0.496	0.497	0.495	0.497
$\mathcal{G}_L$	Manual	0.774	0.721	0.738	0.714
$\mathcal{G}_L$	FastGreedy	1	0.935	0.950	0.901
$\mathcal{G}_C$	Random	0.497	0.499	0.495	0.496
$\mathcal{G}_C$	Manual	0.590	0.610	0.567	0.570
$\mathcal{G}_C$	FastGreedy	1	0.909	0.876	0.824

Table 14: Validation of the community partition. FG, FastGreedy; WT, WalkTrap; ML, MultiLevel; LP, LabelPropagation.

Thus, the pages cluster together according to the users' activity. In a next step, we analyzed the polarization of the users. As explained in Chapter 3.2.3, we can calculate for each user u their bi-community polarization $\rho(u)$ by considering the number of likes or comments given to the two main communities. Thus, we measure the polarization of all users considering the manual classification of pages, *pro-vaccines* and *anti-vaccines*, and the two biggest communities as detected with FastGreedy, C_1 and C_2.

Figure 19 shows the Probability Density Function of $\rho(u)$ for all users who have given at least 10 likes in their lifetime. The PDF for the polarization of all users is sharply bi-modal, meaning the majority of the users are either at -1 or at 1. This indicates there's a strong polarization among the communities, that is, the majority of the users are active either in the *pro-vaccines* or *anti-vaccines* community, not both.

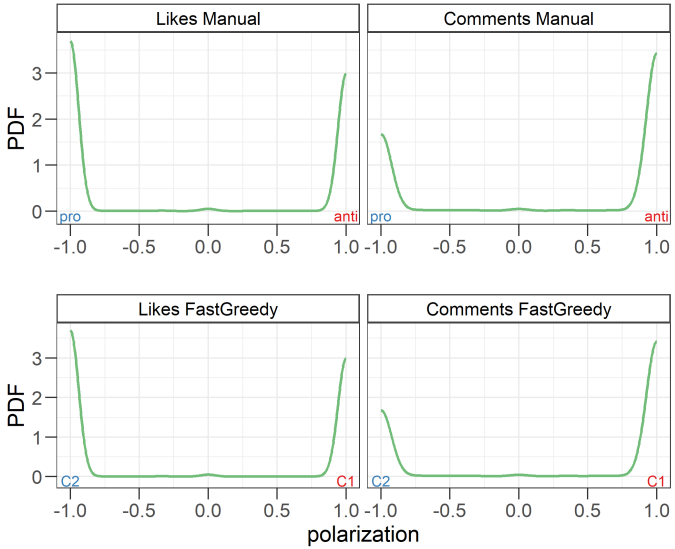

Figure 19: Probability Density Function of $\rho(u)$. Manual partition of the pages into *pro-vaccines* and *anti-vaccines* (top) and the two largest communities detected with FastGreedy (bottom).

6.3 Attention Patterns

To understand the way users behave with respect to the *pro-vaccines* and *anti-vaccines* communities we explore how they interact with their Facebook posts. Figure 20 shows the distribution of the number of likes (left), comments (middle) and shares (right) given to Facebook posts of each community, as manually classified. As seen on the plot, all the distributions are heavy-tailed.

We can see, however, differences in behavior when taking into account the communities. Posts from the *pro-vaccines* community tend to have more likes, while posts from the *anti-vaccines* community are partial to longer debates, i.e., more comments. It's worth noting that *anti-vaccines* are more likely to be shared, up to a point, and the *pro-vaccines* are more capable of reaching wider audiences.

The persistence of a post can be measured by the time difference between the time-stamp of its first and last comment. Also, the lifetime of

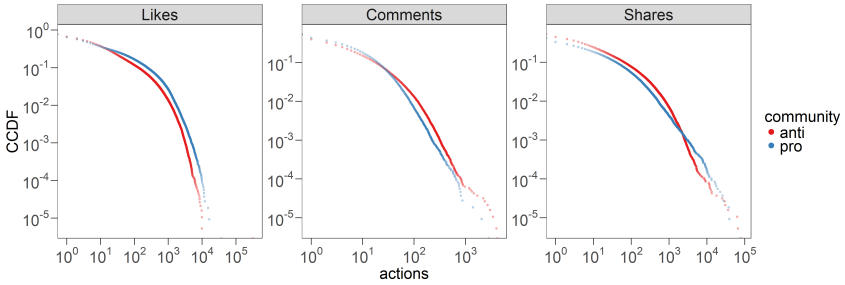

Figure 20: Complementary Cumulative Distribution Function of the likes, comments and shares received by the posts of the anti-vaccines and pro-vaccines community.

a user would be the time difference between the time-stamp of their first and last given comment. Figure 21 shows the CCDF of the posts' persistence and users' lifetime, considering the community of the pages as manually classified.

We can see that the posts made on the *anti-vaccines* pages are more likely to persist in time than those made on the *pro-vaccines* pages. Similarly, the users that comment on the *anti-vaccines* pages display longer periods of activity than those that comment in the *pro-vaccines* pages. The *anti-vaccines* community displays higher user engagement than the *pro-vaccines* community.

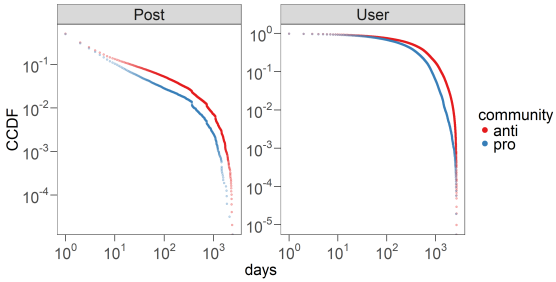

Figure 21: Complementary Cumulative Distribution Function of the posts' persistence (left) and users' lifetime (right) considering the anti-vaccines and pro-vaccines community.

6.4 Selective Exposure

As in the previous chapters, we use the number of likes given by each user to measure their engagement with different Facebook pages considering their community, *pro-vaccines* or *anti-vaccines*, within different time windows. Once again we calculate the *lifetime* of the users, the time difference between their latest and earliest liked post, and their *activity*, the number of likes given in their entire lifetime. Thus, we can measure the variety of pages consumed in a weekly, monthly and yearly basis while taking into account their position in the vaccination debate.

Figure 22 shows the maximum number of unique pages users from the *anti-vaccines* (red) and *pro-vaccines* (blue) communities interact with, considering increasing levels of lifetime and activity for different time windows. The communities correspond to those tagged manually and we consider only users who have given at least 10 likes in their entire lifetime. For a comparative analysis, we standardized lifetime and activity to range between 0 and 1, both over the entire user set of each community, and the number of pages. The results were calculated considering the yearly (top), monthly (middle) and weekly (bottom) rates.

Note that for both communities, users usually interacts with a small number of Facebook pages. Longer lifetime and higher levels of activity correspond with less number of pages being consumed. This is consistent with the results from the previous chapters that show that news consumption on Facebook is dominated by selective exposure and, over time, users personalize their information sources accordingly with their tastes which results in a smaller number of sources being consumed.

We can also observe that users in the *anti-vaccines* community consume information from a more diverse set of pages than those in the *pro-vaccines* community, regardless of the time window considered. So while there is a natural tendency of users to confine their activity to a limited set of pages (SZDV$^+$17), the two communities display different rates of selective exposure. The fringe *anti-vaccines* community shows more commitment to the consumption of their posts, which is consistent with (BCD$^+$15).

Pro-vaccine users interact with M = 1.42 pages (SD = 0.79), anti-vaccine users with 2.45 (SD = 2.13). This difference is displayed in Figure 22: users in the *anti-vaccines* community (red line) consume information from a more diverse set of pages than those in the *pro-vaccines* community, regardless of the time window considered. Grey shades are 95% CI of the local regression of the data, indicating significant differences between

the groups at any time. So while there is a natural tendency of users to confine their activity to a limited set of pages (SZDV+17), the two communities display different rates of selective exposure. The anti-vaccine community shows more commitment to the consumption of their posts.

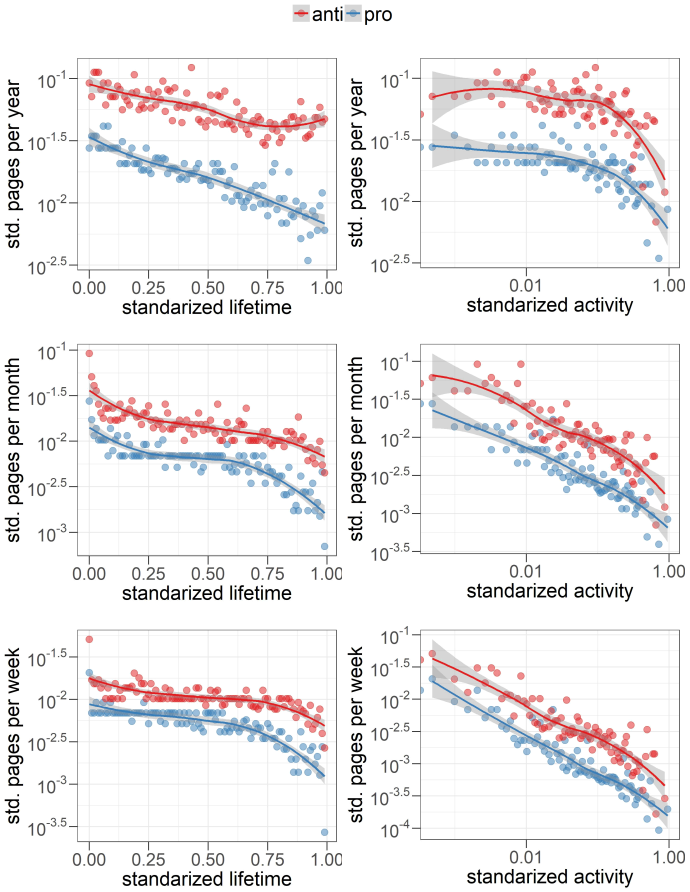

Figure 22: Selective Exposure. Maximum number of unique pages that users with increasing levels of standardized lifetime (left) or standardized activity (right) interact with yearly, monthly and weekly for each community.

6.5 Growth Over Time

We also analyzed the growth of the two communities over time, considering the variety of pages and the number of users that interact with them. Figures 23 and 24 show the evolution of the communities over the years in quarterly increments. The communities in these figures correspond to those manually tagged as *anti-vaccines* and *pro-vaccines*.

Figure 23 plots the number of active users in each community. We define a user as active in a community on a given quarter, if they gave a like (or comment) to any page of that community in the given quarter. The plot shows that while both communities gain users throughout the entire period, the *anti-vaccines* community has, until the end of 2015 and beginning of 2016, more users than the *pro-vaccines* community. After that, this relation reverses (interaction effect in a MANOVA with sentiment (pro, anti) and time (until 2015Q4 vs. following) as factors and comments and likes as dependent variables $F(2,55) = 12.218$, $p < 0.001$; eta2 = 0.31; both main effects are highly significant).

Figure 24 plots the number of active pages in each community. We define a page as active in a specific quarter if it made a post (bottom panel), received a like (middle panel) or comment (upper panel) in that period. Overall, the number of active pages in both communities increases at similar rates, with slight variations when we consider the action that marks a page as *active*. If we use the pages' posting activity or the likes they received to determine whether they were active in a given quarter, we can see that, from 2013, the *pro-vaccines* community consistently shows a higher number of active pages than the *anti-vaccines* community (interaction effect in a MANOVA with sentiment (pro, anti) and time (until 2012Q4 vs. following) as factors and posts and likes as dependent variables $F(2,55) = 2.708$, $p = 0.076$; eta2 = 0.09; both main effects are highly significant). On the other hand, if we focus on the *comments*, the *anti-vaccines* community shows a boost in activity starting in 2015 (interaction effect in an ANOVA with sentiment (pro, anti) and time (until 2014Q4 vs. following) as factors and comments as dependent variable $F(1,56) = 5.053$, $p = 0.029$; eta2 = 0.08; both main effects are significant). This indicates more debates started taking place on the *anti-vaccines* pages since 2015.

Another important factor to consider is the cohesiveness of the *pro-vaccines* and *anti-vaccines* communities. We want to see if the growth of the communities is deponent on the emergence of isolated pages or if it's done in a tightly linked manner. To do this, we split each of the

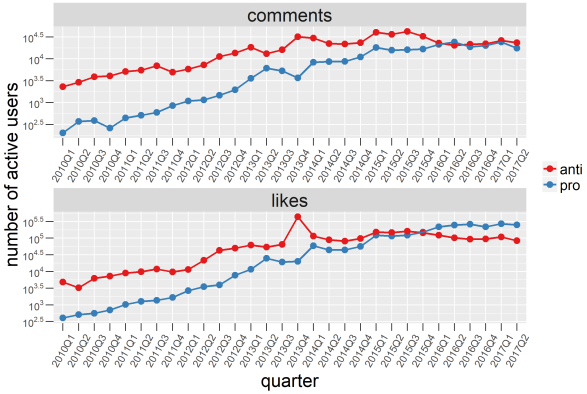

Figure 23: Number of active users in each community considering the likes and the comments given by the users.

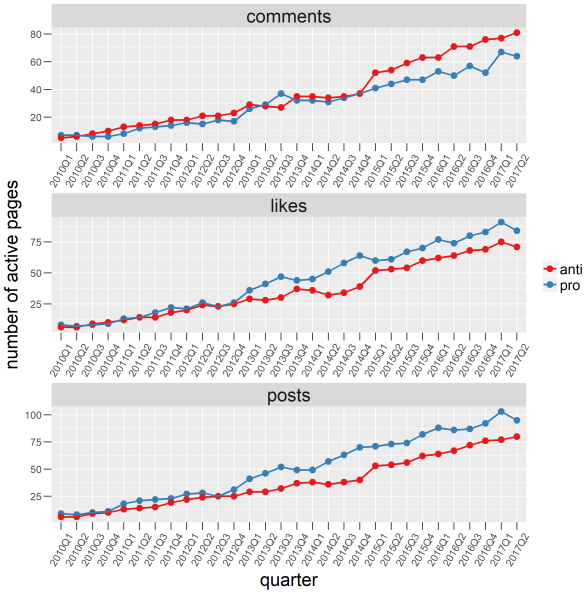

Figure 24: Number of active pages in each community considering the posting activity and the likes and comments received.

projections, $\mathcal{G}_L$ and $\mathcal{G}_C$, into two sub-graphs by considering the community of the pages. This way we obtain $\mathcal{G}_L^{anti}$, $\mathcal{G}_L^{pro}$, $\mathcal{G}_C^{anti}$ and $\mathcal{G}_C^{pro}$ where the pages in each sub-graph belong to either the *anti-vaccines* ($\mathcal{G}_L^{anti}$ and $\mathcal{G}_C^{anti}$) or *pro-vaccines* community ($\mathcal{G}_L^{pro}$ and $\mathcal{G}_C^{pro}$). We can then calculate the fragmentation of each community by applying the community detection algorithms and obtaining their partition.

Figure 25 shows the number of pages in the largest connected component in each sub-graph $\mathcal{G}_L^{anti}$, $\mathcal{G}_L^{pro}$, $\mathcal{G}_C^{anti}$ and $\mathcal{G}_C^{pro}$ in a given quarter. The black line represents the total number of pages in the sub-graphs, that is, the maximum possible size for the largest connected component to take in that quarter. The closer the size of the largest connected component is to the total number of pages in the system, the more tightly linked that community will be in that moment in time.

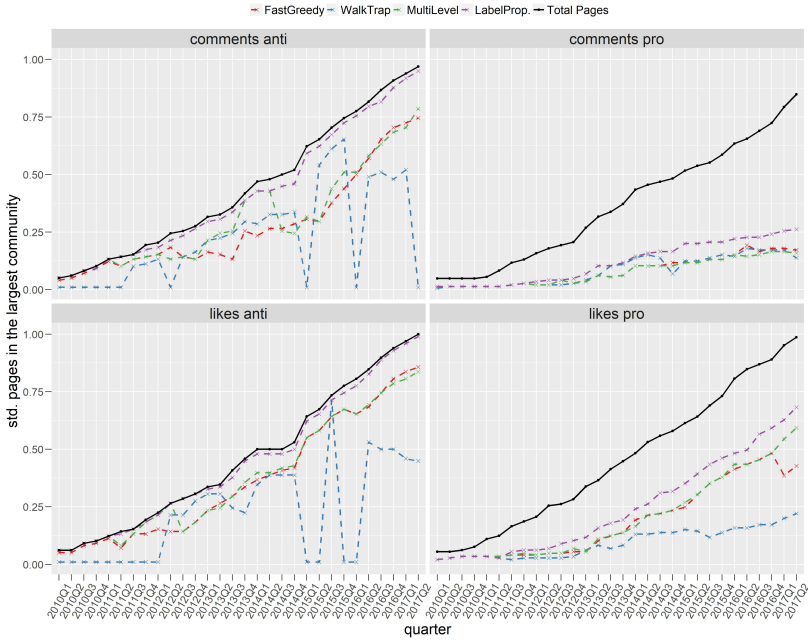

Figure 25: Size of the largest connected component of the *anti-vaccines* and *pro-vaccines* sub-graphs of $\mathcal{G}_C$ (top) and $\mathcal{G}_L$ (bottom). The black line indicates the total number of pages in each sub-graph in that moment in time.

The plots show that in the *anti-vaccines* community the number of pages in the largest component remains close to the total number of pages in the system. In the case of the *pro-vaccines* sub-graphs, however, the size of the largest community does not increase closely with the number of pages in the system. This signifies that the *anti-vaccines* community grows in a more cohesive manner, with pages tightly linked by their users' activity, while the *pro-vaccines* community is more fragmented.

6.6 Concluding Remarks

In this chapter we use quantitative analysis to understand the evolution of the debate about vaccines on the entire public sphere of Facebook, taking into account the emergence of echo chambers around two opposing system of belief: *anti-vaccines* and *pro-vaccines*. By analyzing the consumption patterns of $2.6M$ users, i.e., their likes and comments, we validate the existence of two opposing narratives regarding the vaccination debate on Facebook. We show that the communities emerge from the users consumption habits and that users are highly polarized, that is, the majority of users only consumes and produces information in favor or against vaccines, not both.

We also showed that both narratives are subjected to selective exposure, and that the more active a user is on Facebook the less variety of sources they tend to consume. We note, however, that the users from the *anti-vaccines* community consume information from a wider variety of sources than the *pro-vaccines* users. This is consistent with previous studies (BCD$^+$15) that show that people in conspiracy-like groups show display higher engagement with their community.

We also analyzed the communities' evolution over time, from 2010 to mid 2017. While the *pro-vaccines* community is generally more active, the *anti-vaccines* community concentrates the majority of the debate, receiving more comments from users. Additionally, we show that the *anti-vaccines* community had a more active user base until the end of 2015, where the activity seems to stall. Further studies are needed to determine if this is due to algorithmic changes in the Facebook newsfeed, a saturation of the system or an increase in the vaccination campaigns as a result of the outbreak of measles at Disneyland (ZWH$^+$15), which put the *anti-vaccines* movement in the spotlight and gained the attention of mainstream media (Sal15; Pea15; Hab15; HP15; Kri15; Bar15; Gum15).

Finally, we find evidence that while both narratives have grown on

Facebook over time, the *anti-vaccines* community displays a more cohesive growth than the *pro-vaccines* community. The *anti-vaccines* users are quick to engage with the Facebook pages as soon as they emerge.

Chapter 7

Conclusions

The works collected in this dissertation address news consumption and polarization on Facebook, studying the problem on different scales. The first work considers the issue on a global scale, analyzing the consumption patterns of $376M$ English speaking users on a six year time span. The second one, focuses on four European countries and the differences that emerge from the consumption habits of their users. Finally, the last work reduces the scope of the problem to a single topic, that of *vaccination*.

With quantitative analysis on three massive Facebook datasets the research provides empirical evidence that, in spite of the wide availability of news outlets, users only interact with a small number of pages. Users display a natural tendency to confine themselves to a limited set of pages. Higher levels of users' activity (number of likes) and longer lifetimes (active time on Facebook) correspond to fewer news sources being consumed, regardless of the scope of the study considered.

This dissertation also shows that users from different countries (like France) and communities (*anti-vaccines*) are more committed to the consumption of a wider set of outlets, displaying different rates of selective exposure. However, this commitment to the consumption process does not exempt the users from the tendency to limit the variety of their news diet with the increase in their activity. Thus, we can say that news consumption on Facebook is dominated by selective exposure.

The results also revealed that the users' consumption habits, subject to confirmation bias and selective exposure, lead to their polarization and the emergence of distinct communities of news sources, i.e., echo

chambers. Users rarely interact with news sources outside their own information bubbles. By measuring the polarization of the users from different European countries a ranking established, showing that Italian users are the most polarized, followed by the French, the German and finally the Spanish.

Furthermore, our results suggest that the perspectives of the news outlets and the Facebook users differ. By taking into account the geolocation of the pages, we find that users have a more cosmopolitan perspective of the information space than news providers, who are more locally confined. It seems that despite selective exposure and the agenda setting power of news outlets, users consume from a more international range of sources.

We present two simple models based on the Bounded Confidence Model, one taking into account the users' *trust*, that accurately reproduce the observed user dynamics.

Finally, we study the evolution of the public debate around vaccines on Facebook, taking into account the two opposing narratives: *anti-vaccines* and *pro-vaccines*. Considering the liking and commenting behavior of $2.6M$ users, we validate the existence of these two polarized communities and show their evolution over time. We found that the *pro-vaccines* community is generally more active and the *anti-vaccines* community concentrates the majority of the comments. While both narratives gain users throughout the entire period, the *anti-vaccines* community has more users than the *pro-vaccines* community. This is true until the end of 2015 and beginning of 2016, where the growth of both communities becomes stagnate. Further studies are needed to determine if this is due to algorithmic changes in the Facebook newsfeed, a saturation of the system or an increase in the vaccination campaigns as a result of the outbreak of measles at Disneyland (ZWH$^+$15), which put the *anti-vaccines* movement in the spotlight and gained the attention of mainstream media (Sal15; Pea15; Hab15; HP15; Kri15; Bar15; Gum15). The results also show that the *anti-vaccines* community displays a more cohesive growth than the *pro-vaccines* community.

Future Works

The data gathered for the studies in this dissertation can still provide further insights into how news are consumed on Facebook. Further studies should be done by considering the *content* of the posts and comments,

and not just the consumption patterns that emerge from the users' activity. The posts could be analyzed in order to obtain notable events and entities, topics and recurrent issues, and the comments could be studied to gain insights into the surrounding debate.

The emotional response of the users could also be studied. This could be done considering a topic, the countries or the echo chambers, and would allow for a comparison of different emotional dynamics taking place at various levels of user engagement. The downloaded users' *reactions* from the posts done in 2017 would be particularly relevant in this case.

Appendix A

List of Pages in the Datasets

A.1 European Media Monitor Dataset

The Table 15 contains the 920 news sources in the European Media Monitor Dataset. It contains the name, website, country and region as given by the European Media Monitor and the Facebook ID of the corresponding Facebook page. The countries are indicated with their ISO Alpha-2 international code. The regions are indicated as follow:

- AF - Africa.
- AS - Asia.
- EU - European Union.
- EU-C - EU Candidate.
- EU-O - EU Other.
- ME - Middle East.
- OC - Oceania.
- NA - North America.
- CA - Central America.
- SA - South America.
- GL - Global.

Table 15: Pages in the European Media Monitor Dataset.

	Name and Website	Facebook ID	Country	Region
1	24 Tanzania 24tanzania.com	142633892541250	TZ	AF
2	680 News 680news.com	204410527704	US	NA
3	7 Days UAE 7days.ae	134378466585520	AE	ME
4	7 News Belize 7newsbelize.com	523592704341166	BZ	NA
5	+972 Magazine 972mag.com	148081438555256	IL	ME
6	9 News Australia 9news.com.au	107637365950776	AU	OC
7	Australian Broadcasting Corporation abc.net.au	194764094549	AU	OC
8	abc7NY 7online.com	31160214090	US	NA
9	ABC News abcnews.go.com	86680728811	US	NA
10	ABS-CBN News abs-cbnnews.com	27254475167	PH	AS
11	ABS TV/Radio abstvradio.com	558698930896002	AG	CA
12	AFP News Agency afp.com/en	155857464452265	FR	EU
13	Africa Intelligence africaintelligence.com	1600441596851782	FR	EU
14	African Christian Democratic Party acdp.org.za	379138726819	ZA	AF
15	African Brains africanbrains.net	137238649646766	ZA	AF
16	African Business Magazine africanbusinessmagazine.com	114117578656259	GB	EU
17	African Mining africanmining.com	413345415394061	ZA	AF
18	African Review africanreview.com	507239115959583	GB	EU
19	Ahram Online english.ahram.org.eg	138561829527411	EG	AF

20	Airforce Tecnology airforce-technology.com	376588539031515	GB	EU
21	Air Transport World atwonline.com	98452961409	US	NA
22	The Atlanta Journal-Constitution ajc.com	13310147298	US	NA
23	AJW - Asia & Japan Watch ajw.asahi.com	157466287640348	JP	AS
24	Al-Ahram Weekly weekly.ahram.org.eg	444452792278401	EG	AF
25	Al Bawaba News albawaba.com/en	145270758819050	JO	ME
26	Al Jazeera english.aljazeera.net	7382473689	QA	ME
27	AllAfrica allafrica.com	98946450029	US	NA
28	All Ghana News allghananews.com	110299945719263	GH	AF
29	Almal News en.almalnews.com	275916392585253	EG	AF
30	Al Manar almanar.com.lb/english	567297210065647	LB	ME
31	AlterNet alternet.org	17108852506	US	NA
32	Amandala Newspaper amandala.com.bz/news	100548070049565	BZ	NA
33	Ambergris Today ambergristoday.com	150068098371942	BZ	NA
34	Center for American Progress americanprogress.org	6072343558	US	NA
35	Ammon News - English en.ammonnews.net	124114467624068	JO	ME
36	Amnesty International amnesty.org/en	111658128847068	GB	EU
37	Athens News Agency - Macedonian Press Agency amna.gr/english	324281057673335	GR	EU
38	Anadolu Agency - English aa.com.tr/en	1469323633339182	TR	EU-C
39	An Garda Síochána garda.ie	167613868000	IE	EU

40	The Anglo-Celt anglocelt.ie	125313887068	IE	EU
41	Associated Press ap.org	249655421622	US	NA
42	Arab Times arabtimesonline.com	185411171491289	KW	ME
43	Oratert oratert.com/news	219526301441279	AM	AS
44	Army Technology army-technology.com	273161589364941	US	NA
45	ArtMatters.Info artmatters.info	156915596222	KE	AF
46	Asbarez asbarez.com	30695878200	AM	AS
47	Americas Society - Council of the Americas as-coa.org	22309058308	US	NA
48	Asharq Al-Awsat English Edition aawsat.net	135349813159197	LB	ME
49	Asia News Network asianewsnet.net	291155398585	TH	AS
50	Asian Tribune asiantribune.com	85012674846	TH	AS
51	Asia Times atimes.com	186423081422922	CN	AS
52	Australia Network News australianetworknews.com	1569497949997753	AU	OC
53	Aviation Week aviationweek.com	16067432199	US	NA
54	Awate awate.com	211930818821450	ER	AF
55	The Arizona Republic azcentral.com	50978409031	US	NA
56	AzerNews azernews.az	138411436293144	AZ	AS
57	Azo Mining azomining.com	195005930530874	GB	EU
58	B92 English b92.net/eng	294076883992251	RS	EU-O
59	Bakhtar News Agency bakhtarnews.com.af/eng	128677440577060	AF	ME

60	Balkan Insight balkaninsight.com/en	97956845505	GB	EU
61	The Baltimora sun baltimoresun.com	9299179711	US	NA
62	Bangkok Post bangkokpost.com	133643127712	TH	AS
63	Banknet India banknetindia.com	156945944397481	IN	AS
64	Barents Observer barentsobserver.com	196717396635	NO	EU-O
65	Barron's barrons.com	64579042740	US	NA
66	BBC iPlayer Radio bbc.co.uk/radio	1470145583204829	GB	EU
67	BBC News bbc.com/news	228735667216	GB	EU
68	Beat 102-103 beat102103.com	117022648342887	IE	EU
69	BelTA - Belarusian Telegraph Agency eng.belta.by	159485357443987	BY	EU-O
70	News Letter newsletter.co.uk	117370764948881	GB	EU
71	Belfast Telegraph belfasttelegraph.co.uk	237692023818	GB	EU
72	The Bermuda Sun bermudasun.org	12844875188	BM	NA
73	BetaNews betanews.com	167777169963870	US	NA
74	Better Diamond Initiative betterdiamondinitiative.org	474528435937430	US	NA
75	BioFuels Journal biofuelsjournal.com	785884781480346	US	NA
76	Beijing Review bjreview.com.cn	164344500263690	CN	AS
77	Bloomberg News bloomberg.com	266790296879	US	NA
78	Bahrain News Agency bna.bh/portal/en	155290664516690	BH	ME
79	BNO News bnonews.com	60109657413	US	NA

80	Boston boston.com	6879409364	US	NA
81	Boston Haitian Reporter bostonhaitian.com	132342263485363	HT	CA
82	Boston Herald bostonherald.com	197211981599	US	NA
83	Botswana Guardian botswanaguardian.co.bw	261592817198239	BW	AF
84	Breaking News breakingnews.ie	120689931275023	IE	EU
85	The Argus theargus.co.uk	57197526698	GB	EU
86	Brisbane Times brisbanetimes.com.au	95683517460	AU	OC
87	The Brookings Institution brookings.edu	137459917707	CO	SA
88	The Broad Street Journal broadstreetjournalbarbados.com	306701519754	BB	CA
89	Budapest Business Journal bbj.hu	162210567199344	HU	EU
90	The Budapest Times budapesttimes.hu	473754729371730	HU	EU
91	Buenos Aires Herald buenosairesherald.com	333788759997981	AR	SA
92	Bulatlat bulatlat.com	63467907915	PH	AS
93	Bulawayo 24 News bulawayo24.com	128990327163597	ZW	AF
94	Bulgarian Telegraph Agency bta.bg/en	649620041801768	BG	EU
95	Business Daily Africa businessdailyafrica.com	111397575568693	KE	AF
96	Business Day Live businessday.co.za	286638794693284	ZA	AF
97	Business & Human Rights Resource Centre business-humanrights.org/en	10510773883	IE	EU
98	Business Wire businesswire.com	82442677571	US	NA
99	BusinessWorld Online bworldonline.com	115158345163134	PH	AS

100	Republic of Botswana gov.bw	148228411926492	BW	AF
101	The Belize Times belizetimes.bz	325514113131	BZ	NA
102	Cafébabel cafebabel.co.uk	357343795001	GB	EU
103	The Cameroon Daily Journal cameroonjournal.com	1497726193822330	CM	AF
104	Cameroon Online cameroononline.org	287988637383	CM	AF
105	AKIpress Central Asian News Service en.ca-news.org	138634076266590	KG	ME
106	Caperi capitaleritrea.com	257893353570	ER	AF
107	Cape Verde capeverde.com	328585463869159	CV	AF
108	Capital FM capitalfm.co.ke	178342827608	KE	AF
109	CARE care.org	30139072158	US	NA
110	Caribbean360 caribbean360.com	147441385288431	BB	CA
111	Carlow Nationalist carlow-nationalist.ie	233149060080354	IE	EU
112	Carnegie Endowment for International Peace carnegieendowment.org	62935143720	US	NA
113	Carnegie Europe carnegieeurope.eu	74341131538	BE	EU
114	Carnegie Moscow Center carnegie.ru/?lang=en	111281202235293	RU	AS
115	Cato Institute cato.org	26668999076	US	NA
116	Caucasian Knot eng.kavkaz-uzel.ru	129869803756134	RU	AS
117	Cayman Compass caycompass.com	237869980023	KY	NA
118	CBC News cbc.ca	5823419603	CA	NA
119	Caribbean Broadcasting Corporation cbc.bb/index.php/en	1447053148908724	BB	CA

120	CBS Baltimore baltimore.cbslocal.com	261323008476	US	NA
121	CBS News cbsnews.com	131459315949	US	NA
122	CBS Radio cbsradio.com	165760583475583	US	NA
123	CCN TV6 tv6tnt.com	192696260772574	TT	SA
124	Center on Budget and Policy Priorities cbpp.org	42002969352	US	NA
125	Centre for Asia Pacific Aviation centreforaviation.com	212338015492650	GB	EU
126	CERN home.web.cern.ch	169005736520113	CH	EU-O
127	CFO cfo.com	146601585370995	US	NA
128	Channel 4 News channel4.com/news	6622931938	GB	EU
129	Channel Africa channelafrica.org	287474612524	ZA	AF
130	Channel NewsAsia channelnewsasia.com	93889432933	SG	AS
131	Chatham House chathamhouse.org	202798122688	GB	EU
132	Chicago Tribune chicagotribune.com	5953023255	US	NA
133	China.com english.china.com	623480657770927	CN	AS
134	China Daily chinadaily.com.cn	191347651290	CN	AS
135	China.org.cn china.org.cn	371171589575669	CN	GL
136	China Plus News - CRI english.cri.cn	223495844457800	CN	AS
137	The China Post chinapost.com.tw	143632722314869	CN	AS
138	The Christian Science Monitor csmonitor.com	14660729657	US	NA
139	The Houston Chronicle chron.com	12852567813	US	NA

140	CIP Americas Program cipamericas.org	113343085368711	US	NA
141	City AM cityam.com	213682385348579	GB	EU
142	Civil.ge civil.ge/eng	154006094635924	GE	AS
143	Clare FM clare.fm	76448532790	IE	EU
144	CNBC cnbc.com	97212224368	US	NA
145	CNC3 cnc3.co.tt	126206357995	TT	SA
146	CNET News news.cnet.com	7155422274	US	NA
147	CTV ctntworld.com/cnews2	377693645609	TT	SA
148	CNN International edition.cnn.com	18793419640	US	NA
149	CNNMoney news money.cnn.com	6651543066	US	NA
150	Colombia Reports colombiareports.co	64680372407	CO	SA
151	European Union - Committee of the Regions cor.europa.eu/en	527351247377965	BE	EU
152	Common Ground News Service commongroundnews.org	60149902595	US	NA
153	Congress Rental Network congressrentalnetwork.com	554092937948891	MT	EU
154	The Connaught Telegraph con-telegraph.ie	187794613261	IE	EU
155	Construction Week Online constructionweekonline.com	150848527337	US	NA
156	The Copenhagen Post cphpost.dk	38409311585	DK	EU
157	Cork's 96FM 96fm.ie	194278177271109	IE	EU
158	Cornwall Seaway News cornwallseawaynews.com	120383281376364	CA	NA
159	Courier Journal courier-journal.com	69357466992	CR	CA

160	The Courier Mail couriermail.com.au	90004797701	AU	OC
161	CPAC - Cable Public Affairs Channel cpac.ca/en	8260742627	CA	NA
162	CPI Financial cpifinancial.net/news	509446489112881	US	NA
163	International Crisis Group crisisgroup.org	341675908125	BE	EU
164	The Critical Threats Project criticalthreats.org	131736233536920	US	NA
165	CSIS - Center for Strategic and International Studies csis.org	118399079197	US	NA
166	CTV News ctvnews.ca	194553860586548	CA	NA
167	Cyprus Expat cyprusexpat.co.uk	357342727764507	GB	EU
168	Cyprus Mail cyprus-mail.com	308790590111	CY	EU
169	Cyprus Traveller cyprustraveller.com	363334110438460	CY	EU
170	The Daily Caller dailycaller.com	182919686769	US	NA
171	The Daily and Sunday Express express.co.uk	129617873765147	GB	EU
172	Daily Finance dailyfinance.com	65750045740	US	NA
173	Independent Newspapers Nigeria dailyindependentnig.com	344063115645237	NG	AF
174	Daily Mail dailymail.co.uk	164305410295882	GB	EU
175	Daily Maverick dailymaverick.co.za	171090380838	ZA	AF
176	The Mirror Online mirror.co.uk	6149699161	GB	EU
177	Daily Mirror dailymirror.lk	129843276524	LK	AS
178	Daily Monitor monitor.co.ug	105583497196	UG	AF
179	Daily Nation nation.co.ke	105983259496	KE	AF

180	The Daily Nation zambiadailynation.com	520665064730150	ZM	AF
181	Daily News dailynews.co.zw	129826597051376	ZW	AF
182	The Daily Observer observer.gm	155754624520664	GM	AF
183	Daily Post Nigeria dailypost.ng	247156892008627	NG	AF
184	The Daily Star - Lebanon dailystar.com.lb	43930085355	LB	ME
185	The Daily Star thedailystar.net	100117146754830	BD	AS
186	Daily Star Nigeria dailystar.com.ng	475906012461633	NG	AF
187	Daily Times dailytimes.com.pk	710854355609434	PK	ME
188	Daily Times of Nigeria dailytimes.com.ng	482749575105792	NG	AF
189	Daily Trust dailytrust.com.ng	96160950863	NG	AF
190	Dalje dalje.com/en	161625400547781	HR	EU
191	The Dallas Morning News dallasnews.com	20946638799	US	NA
192	Dawn dawn.com	86398345441	PK	ME
193	Dayton Daily News daytondailynews.com	168815400507	US	NA
194	DCist dcist.com	258755510899692	US	NA
195	Deccan Herald deccanherald.com	318318083127	IN	AS
196	Defense News defensenews.com	70531852030	US	NA
197	Dehai Eritrean News dehai.org	224118577679888	ER	AF
198	The Denver Post denverpost.com	6181619439	US	NA
199	Derby Telegraph thisisderbyshire.co.uk	142370589115824	GB	EU

200	Derry Journal Newspaper derryjournal.com	104291312965731	IE	EU
201	The Des Moines Register desmoinesregister.com	8031989578	US	NA
202	Detroit Free Press freep.com	13642915529	US	NA
203	Deutsche Welle dw.de	24369314439	DE	EU
204	Diamond Development Initiative ddiglobal.org/media	118224424879645	GB	EU
205	The Israeli Diamond Industry israelidiamond.co.il/english	215033265188341	IL	ME
206	Digital Journal digitaljournal.com	67175109350	US	NA
207	Digjitale digjitale.com	247722835253293	FR	EU
208	DNA India dnaindia.com	154284380440	IN	AS
209	Dominican Today dominicantoday.com	154894428290	DO	CA
210	Donegal Democrat donegaldemocrat.ie	112739298752474	IE	EU
211	Defence and Security Alert dsalert.org	197617058395	IN	AS
212	Dublin People dublinpeople.com	144107362283722	IE	EU
213	Dundalk Democrat dundalkdemocrat.ie	62703437305	IE	EU
214	Dunmow Broadcast dunmowbroadcast.co.uk	181182540669	GB	EU
215	DutchNews dutchnews.nl	314156250291	NL	EU
216	East Anglian Daily Times eadt.co.uk	6478299951	GB	EU
217	In2EastAfrica in2eastafrica.net	179227395457307	TZ	AF
218	Earthquake Report earthquake-report.com	145927268753491	CA	NA
219	European Asylum Support Office easo.europa.eu	1449052665318446	MT	EU

220	ECHO - Humanitarian Aid and Civil Protection ec.europa.eu/echo	146955527905	BE	EU
221	eCanada Now ecanadanow.com	270230853108003	CA	NA
222	Eco-Business eco-business.com	136176823413276	SG	AS
223	The Economic Times economictimes.com	21540067693	IN	AS
224	Edmonton Journal edmontonjournal.com	100429659640	CA	NA
225	Egypt Independent egyptindependent.com	244154415645700	EG	GL
226	89.7 Bay Network bay.com.mt	186470373746	MT	EU
227	EITI - Extractive Industries Transparency Initiative eiti.org	182289264273	NO	EU-O
228	Kathimerini English Edition ekathimerini.com	142088502518488	GR	EU
229	Elections Canada elections.ca	633812813408377	CA	NA
230	Electoral Commission of Jamaica eoj.com.jm	229947173716650	JM	NA
231	El País elpais.com/elpais/inenglish.html	279465748829565	ES	EU
232	Executive Mansion - Government the Republic of Liberia emansion.gov.lr	310591779028211	LR	AF
233	Emirates 247 emirates247.com	136509933034734	US	NA
234	Trend News Agency en.trend.az	187192942493	AZ	AS
235	eNCA - e-News Channel Africa enca.com	160836574053016	ZA	AF
236	EnergyWorld Magazine energyworldmag.com	201095326580070	RO	GL
237	Engadget engadget.com	5738237369	US	NA
238	The Chosun Ilbo english.chosun.com	129892740363306	KR	AS

239	KBS World Radio world.kbs.co.kr/english	170517169668683	KR	AS
240	Enough Project enoughproject.org	252083412738	US	NA
241	Cincinnati Enquirer cincinnati.com	36808884697	US	NA
242	European Policy Centre epc.eu	115769411816768	BE	EU
243	EPP Group eppgroup.eu/home/en	291699857688	BE	EU
244	El Paso Times elpasotimes.com	115612465136194	MX	NA
245	ERR News news.err.ee	147712425255536	EE	EU
246	The Express Tribune tribune.com.pk	111457038871331	PK	AS
247	EUbusiness eubusiness.com	215108901846669	GB	EU
248	Council of the European Union consilium.europa.eu/en	147547541961576	BE	EU
249	EUobserver euobserver.com	227715390929	BE	EU
250	EurActiv euractiv.com/en	15299247059	GB	EU
251	EurasiaNet eurasianet.org	61048477840	US	NA
252	Eurasia Review eurasiareview.com	339585248573	US	NA
253	EU Reporter eureporter.co	172730922764856	BE	EU
254	Euromoney euromoney.com	192279900885723	GB	EU
255	Euronews euronews.com	101402598109	FR	EU
256	European Commission europa.eu	107898832590939	BE	EU
257	European Railway Review europeanrailwayreview.com	404359882930504	GB	EU
258	ESA - European Space Agency esa.int	54912575666	NL	EU

259	ESRF - European Synchrotron Radiation Facility esrf.eu	116961611670251	FR	EU
260	Europe's World europesworld.org/feed	143299008908	BE	EU
261	EURweb eurweb.com	45752458150	US	NA
262	Evening Echo eecho.ie	137101942966920	IE	EU
263	London Evening Standard standard.co.uk	165348596842143	GB	EU
264	Expatica expatica.com/uk	206982432584	GB	EU
265	Trinidad Express trinidadexpress.com	134545155813	TT	SA
266	Eyewitness News ewn.co.za	168892509821961	ZA	AF
267	Pambazuka News pambazuka.net/en	210266369010114	ZA	AF
268	FairWarning fairwarning.org	307820806867	US	NA
269	Famagusta Gazette famagusta-gazette.com	156070247783260	CY	EU
270	Farmers Weekly fwi.co.uk	17050705902	GB	EU
271	Farming Life farminglife.com	243070359106664	GB	EU
272	Ferghana News enews.fergananews.com	140679709325246	RU	AS
273	The Fiji Times Online fijitimes.com	7130088677	FJ	OC
274	Financial Mirror financialmirror.com	112107185515325	CY	EU
275	Financial Times ft.com	8860325749	GB	EU
276	Finland Times finlandtimes.fi	189776234502564	FI	GL
277	Fish Information and Services fis.com	228637974386	JP	AS
278	Fars News Agency in English english.farsnews.com	199676426877938	IR	ME

279	Focus Taiwan focustaiwan.tw	196543097045548	TW	AS
280	Forbes forbes.com	30911162508	US	NA
281	Foreign Affairs foreignaffairs.org	35640742015	US	NA
282	4-traders 4-traders.com	570518562988663	FR	EU
283	Fox Business foxbusiness.com	12795435237	US	NA
284	Fox News foxnews.com	15704546335	US	NA
285	France 24 france24.com/en	176585044433	FR	EU
286	FriedlNews friedlnews.com	139471392786597	AT	EU
287	FrontPage Magazine frontpagemag.com	296396490870	US	NA
288	This is Africa thisisafricaonline.com	779213412106756	GB	EU
289	Gabz-FM gabzfm.com	510261369119743	BW	AF
290	Gambia News Online gambianewsonline.com	108203649260707	GM	AF
291	The Gazette thegazette.com	7711594722	CA	NA
292	Geo News geo.tv	101059326616167	PK	AS
293	Georgia Today georgiatoday.ge	358827420882252	GE	AS
294	GhanaWeb ghanaweb.com	349823335095071	GH	AF
295	GINA - Government Information Agency gina.gov.gy	275385732567266	GY	SA
296	Barbados Government Information Service gisbarbados.gov.bb	271334199698066	BB	CA
297	Gizmodo gizmodo.com	5718758966	US	NA
298	Jamaica Gleaner jamaica-gleaner.com	116087685213012	JM	NA

299	Global Issues globalissues.org	116446965860	HT	CA
300	GlobalPost globalpost.com	35930083446	US	NA
301	Global Research - Centre for Research on Globalization globalresearch.ca	200870816591393	CA	NA
302	Global Security globalsecurity.org	130707690322552	US	NA
303	Tax-News tax-news.com	375456009146619	GB	EU
304	Global Times globaltimes.cn	115591005188475	CN	AS
305	Global Voices globalvoicesonline.org	6110663875	NL	EU
306	The Boston Globe bostonglobe.com	5637143257	US	NA
307	INTA - International Trademark Association inta.org	132681243408881	US	NA
308	Government of Ghana ghana.gov.gh	395672803872128	GH	AF
309	South African Government gov.za	194109891221	ZA	AF
310	Government of Prince Edward Island gov.pe.ca	122466107790811	CA	NA
311	Great Lakes Voice greatlakesvoice.com	123665291025561	RW	AF
312	Greenpeace greenpeace.org	7297163299	NL	EU
313	The Greens - European Free Alliance in the European Parliament greens-efa.eu	150527751674751	DE	EU
314	Grenada Informer thegrenadainformer.com	444053475702112	GD	CA
315	Oxford International oxfam.org	197021747008439	US	NA
316	Pacific Daily News guampdn.com	90329609573	GU	OC
317	The Guardian guardian.bz	625050694259788	BZ	NA

318	GUE/NGL - Confederal Group of the European United Left/Nordic Green Left guengl.eu	164355333623231	BE	EU
319	Gulf Daily News gulf-daily-news.com	382277502795	BH	ME
320	Gulf News gulfnews.com	143017562431177	AE	ME
321	Gulf-Times gulf-times.com	274209762705885	QA	ME
322	Gurtong Trust gurtong.net	105221526186803	SD	AF
323	Guyana Chronicle guyanachronicle.com	607416296042745	GY	SA
324	Haaretz haaretz.com	64588666340	IL	ME
325	Haveeru haveeru.com.mv	1619525378288076	MV	AS
326	Hellenic Shipping News Worldwide hellenicshippingnews.com	336846839674977	GR	EU
327	The Herald herald.ie	167882646558086	IE	EU
328	Herald Scotland heraldscotland.com	271154343382	GB	EU
329	The Daily Herald thedailyherald.sx	842786335748518	AN	NA
330	The Heritage Foundation heritage.org	21375324480	US	NA
331	Hiawatha World Online hiawathaworldonline.com	140225006013485	US	NA
332	Highland Radio highlandradio.com	104685286233423	IE	EU
333	Hindustan times hindustantimes.com	111445058579	IN	AS
334	History News Network historynewsnetwork.org	187220577957886	US	NA
335	Homeland Security Newswire homelandsecuritynewswire.com	355645114549300	US	NA
336	HTS St. Lucia htsstlucia.org	110158905668844	LC	CA
337	The Huffington Post US huffingtonpost.com/?country=US	18468761129	US	NA

338	L'Humanité humaniteinenglish.com	211118048927823	FR	EU
339	Human Rights Watch hrw.org	42940254353	US	NA
340	Hürriyet Daily News hurriyetdailynews.com	353247165122	TR	EU-C
341	IAEA - International Atomic Energy Agency iaea.org	96699782061	AT	EU
342	iafrica iafrica.com	95285567334	ZA	GL
343	BIO - Biotechnology Innovation Organization bio.org	37791069588	US	NA
344	IBTimes - International Business Times US ibtimes.com	128026713884051	US	NA
345	IBTimes - International Business Times India ibtimes.co.in	134755750024728	IN	AS
346	IBTimes - International Business Times Australia au.ibtimes.com	125955890795363	AU	OC
347	IBTimes - International Business Times UK ibtimes.co.uk	224377357631653	GB	EU
348	IceNews icenews.is	147438545270363	IS	EU-O
349	International Consortium of Investigative Journalists icij.org	99609391511	US	NA
350	International Catholic Migration Mission icmc.net	124780317689542	CH	EU-O
351	International Committee of the Red Cross icrc.org/eng	336620687262	CH	EU-O
352	IFEX ifex.org	164693013619204	TN	AF
353	Intergovernmental Authority on Development igad.org	247268655431305	DZ	AF
354	IHS Janes's 360 janes.com	118688984881608	GB	EU

355	Africa Review africareview.com	265970097056	KE	AF
356	IMEC imec.be/be_en/home.html	48799938690	BE	EU
357	International Monetary Fund imf.org	629825573770523	US	NA
358	IMPEL impel.eu	1554126881544220	BE	EU
359	In-Cyprus incyprus.com.cy	160324974064965	CY	EU
360	Independent.ie independent.ie	96796398469	IE	EU
361	The Independent independent.co.ug	555291447866478	UG	AF
362	Reuters India in.reuters.com	27917365629	IN	GL
363	IndyStar indystar.com	9812439851	US	NA
364	The Indian Express expressindia.com	163648403825	IN	AS
365	InfoMine infomine.com	298200071890	CA	NA
366	Informanté informante.web.na	183476948385940	NA	AF
367	Inkatha Freedom Party ifp.org.za	1679769598931438	ZA	AF
368	Inmarsat inmarsat.com	317156988374684	GB	EU
369	InSight Crime insightcrime.org	147437478639179	CO	SA
370	Institute for War and Peace Reporting iwpr.net	27283822672	US	NA
371	Institut Laue-Langevin ill.eu	148452651846298	FR	EU
372	International Alert international-alert.org	23799263071	GB	EU
373	International Civic Aviation Organization icao.int	1509413712605916	CA	NA
374	International IDEA idea.int	196055247078923	SE	EU

375	Internews internews.org	54282384157	US	NA
376	Interpol interpol.int	282736625084940	FR	EU
377	IntraFish intrafish.com	117849188285395	NO	EU-O
378	Investor's Business Daily investors.com	18107421115	US	NA
379	IOL iol.co.za	21993963624	ZA	AF
380	International Organization for Migration iom.int	161303029020	CH	EU-O
381	International Peace Information Service ipisresearch.be	380249695352799	BE	EU
382	Inter Press Service ipsnews.net	41883361077	DZ	AF
383	Iran Daily iran-daily.com	233108793428420	IR	ME
384	UN Integrated Regional Information Networks irinnews.org	49134323939	US	NA
385	Irish Abroad irishabroad.com	104716382899827	IE	EU
386	Irish Examiner irishexaminer.com	175322202485450	IE	EU
387	Irish Farmers Journal farmersjournal.ie	208480693834	IE	EU
388	Irish Sun irishsun.com	138617672855411	IE	EU
389	The Irish Times irishtimes.com	45441411157	IE	EU
390	Iran-Va-Jahan iranvajahan.net/en	1498587017035003	IR	ME
391	International Relations and Security Network isn.ethz.ch	141505796320	CH	EU-O
392	Israel National News israelnationalnews.com	102510354100	IL	ME
393	Institute for Security Studies issafrica.org	204054676287947	DZ	AF

394	ITV News itv.com/news	148007467671	GB	EU
395	Indian Television indiantelevision.com	517483194951446	IN	AS
396	The Jakarta Post thejakartapost.com	102060486500863	ID	AS
397	Jamaica Information Service jis.gov.jm	129366111456	JM	NA
398	The Jamaica Observer jamaicaobserver.com	95746458800	JM	NA
399	Jamaica Star jamaica-star.com	94678696229	JM	NA
400	Japan Today japantoday.com	206382352722433	JP	AS
401	Japan Update japanupdate.com	193676320645525	JP	AS
402	JCK - Jewelry Industry News - Trends and Analysis jckonline.com	72079325608	US	NA
403	Journal of Commerce joc.com	73172116398	US	NA
404	Jollof News jollofnews.com	264329360246554	GM	AF
405	Journal Sentinel jsonline.com	16511263815	US	NA
406	Justice in Mexico justiceinmexico.org	204554056223931	MX	NA
407	Kaieteur News kaieteurnewsonline.com	126977184000816	GY	SA
408	Kansas City Star kansascity.com	81580834093	US	NA
409	KCLR 96 FM kclr96fm.com	146296768437	IE	EU
410	Khaleej Times khaleejtimes.com	211983032863	AE	ME
411	Kilkenny People kilkennypeople.ie	148324858612545	IE	EU
412	KIMT News kimt.com	89871599854	US	NA
413	Kippreport kippreport.com	115801885129398	AE	ME

414	Korea.net korea.net	181274814520	KR	AS
415	Korea JoongAng Daily koreajoongangdaily.joinsmsn.com	190108377775039	KR	AS
416	Krakow Post krakowpost.com	82857764783	PL	EU
417	KUAM News kuam.com	39567613741	GU	OC
418	Kuensel kuenselonline.com	140918395943068	BT	AS
419	Ekurd Daily ekurd.net	299681230060006	GB	EU
420	Kuwait Times news.kuwaittimes.net	80073822064	KW	ME
421	Kyiv Post kyivpost.com	218630044858401	UA	EU-O
422	The Labour Spokesman labourspokesman.com	376337209232404	KN	CA
423	La Crosse Tribune lacrossetribune.com	66774556378	US	NA
424	Los Angeles Daily News dailynews.com	55107646487	US	NA
425	Latin American Herald Tribune laht.com	198630266978389	VE	SA
426	Colombo Page colombopage.com	164230806937385	LK	AS
427	La Strada International lastradainternational.org	107095159376651	NL	EU
428	Los Angeles Times latimes.com	5863113009	SA	AS
429	Las Cruces Sun-News lcsun-news.com	148251148521543	US	NA
430	Leinster Leader leinsterleader.ie	445456162136874	IE	EU
431	Leitrim Observer leitrimobserver.ie	223377501087566	IE	EU
432	Lesotho Times lestimes.com	263874940420105	LS	AF
433	Limerick Leader limerickleader.ie	242208309132671	IE	EU

434	Limerick's Live 95FM live95fm.ie	102779203091115	IE	EU
435	Lincoln Journal Star journalstar.com	12493817252	US	NA
436	LMFM 95.8 lmfm.ie	185885781448435	IE	EU
437	Longford Leader longfordleader.ie	342214075801273	IE	EU
438	Lydian International lydianinternational.co.uk	186900121339682	GB	EU
439	Ma'an News maannews.net/en	66674009455	IL	ME
440	Mail&Guardian Online mg.co.za	161428670566653	ZA	AF
441	Malaysiakini malaysiakini.com	47298465905	MY	AS
442	Maldives Independent maldivesindependent.com	295327362366	MV	AS
443	Malta Star maltastar.com	456049707755135	MT	EU
444	Malta Today maltatoday.com.mt	21535456940	MT	EU
445	Managing Intellectual Property managingip.com	133642376683214	US	NA
446	Manica Post manicapost.com	152535288239280	ZW	AF
447	The Maravi Post maravipost.com	145297758853971	MW	AF
448	Marine Link marinelink.com	283894011630279	US	NA
449	Market Watch marketwatch.com	131043201847	US	NA
450	McClatchy DC mcclatchydc.com	27177163800	US	NA
451	Meath Chronicle meathchronicle.ie	107690035817	IE	EU
452	Media Monitoring Africa mediamonitoringafrica.org	247186148894	ZA	AF
453	MercoPress en.mercopress.com	139220750082	BR	SA

454	The Mercury News mercurynews.com	63095136336	US	NA
455	The Messenger messenger.com.ge	155667667914073	GE	AS
456	Messenger-Inquirer messenger-inquirer.com	127426175386	KY	NA
457	Metro Éireann metroeireann.com	183768441657310	IE	EU
458	Metro metro.co.uk	117118184990145	GB	EU
459	Miami Herald miamiherald.com	38925837299	US	NA
460	Mid-Day mid-day.com	101982830708	IN	AS
461	Midlands 103 midlandsradio.fm	101869576532842	IE	EU
462	The Midland Tribune midlandtribune.ie	106650586336938	IE	EU
463	Midwest Radio midwestradio.ie	115609018461867	IE	EU
464	The Milli Gazette milligazette.com	108910729126586	IN	AS
465	Mindanao Examiner mindanaoexaminer.com	197032943653361	PH	AS
466	Mineweb mineweb.com	133899146653638	ZA	AF
467	MINING.com mining.com	170783659631727	CA	NA
468	Mining News miningnews.net	145316085622369	AU	OC
469	Mining Technology mining-technology.com	326019370778750	GB	EU
470	Mining Weekly miningweekly.com	112406545484603	ZA	AF
471	The Government of the Bahamas bahamas.gov.bs	361099777333587	BS	CA
472	Mizzima mizzima.com	150773224985493	IN	AS
473	Mmegi Online mmegi.bw	121039846823	BW	AF

474	MNI News mninews.marketnews.com	124576787245	US	NA
475	Modern Ghana modernghana.com	366939671198	US	NA
476	InfoPak - Ministry of Information infopak.gov.pk	712150158816432	PK	ME
477	Le Monde Diplomatique mondediplo.com	198560474895	FR	EU
478	Mondo Visione mondovisione.com	169767016460715	GB	EU
479	MoneyWeek moneyweek.com	110326662354766	GB	EU
480	Monsters and Critics monstersandcritics.com	193326863118	GB	EU
481	Montreal Gazette montrealgazette.com	273805206181	CA	NA
482	Morningstar morningstar.com	428809707142935	US	NA
483	The Moscow Times themoscowtimes.com	203688324765	RU	AS
484	Dayton Most Metro mostmetro.com	216346418455890	US	NA
485	MPR News minnesota.publicradio.org/features	99142348590	US	NA
486	Doctors Without Borders msf.org	33110852384	CH	EU-O
487	MSNBC msnbc.com	273864989376427	US	NA
488	Muscatine Journal muscatinejournal.com	55294244167	US	NA
489	Marianas Variety mvariety.com	266082993493810	NP	OC
490	Joy Online news.myjoyonline.com	115080018529590	GH	AF
491	Naharnet naharnet.com	7227980682	LB	ME
492	Namibia Economist economist.com.na	538225062922656	NA	AF
493	NASA nasa.gov	54971236771	US	NA

494	National Mirror nationalmirroronline.net	296803303684069	NG	AF
495	The Nation Barbados nationnews.com	270969375988	BB	CA
496	Natural Gas Europe naturalgaseurope.com	1425903867642047	CA	NA
497	NBC News nbcnews.com	155869377766434	US	NA
498	NBC Radio St Vincent and the Grenadines nbcsvg.com	119263408099925	VC	CA
499	New Civil Engineer nce.co.uk	166793706822441	GB	EU
500	Nepal News nepalnews.com	195499967173234	NP	AS
501	New Economics Foundation neweconomics.org	110275553302	GB	EU
502	New Era newera.com.na	131485450303317	NA	AF
503	New Europe neurope.eu	111051292072	NL	EU
504	New Poland Express newpolandexpress.pl	164066290543	PL	EU
505	news.com.au news.com.au	111416688885713	AU	OC
506	Kenya Today kenya-today.com	293248774116017	KE	AF
507	news24.com news24.com	10227041841	ZA	AF
508	News 5 Belize edition.channel5belize.com	472013519586191	BZ	NA
509	NEWS.am news.am/eng	160111214030709	AM	AS
510	NewsBlogged newsblogged.com	253819121296131	US	NA
511	NewsChannel5 newschannel5.com	91345192547	GB	EU
512	New Scientist newscientist.com	235877164588	GB	EU
513	NewsDay newsday.co.zw	215170571826981	ZW	AF

514	Durham Region durhamregion.com	115912925116255	US	NA
515	News from Africa newsfromafrica.org	157591600945944	KE	AF
516	The News-Gazette news-gazette.com	120713712889	CA	NA
517	Newstalk newstalk.ie	70425092906	IE	EU
518	Newstalk ZB newstalkzb.co.nz	171762839533606	NZ	OC
519	New Statesman newstatesman.com	100959719644	GB	EU
520	Thanh Nien News thanhniennews.com	108754382559490	VN	AS
521	Daily Newswatch mynewswatchtimesng.com	469757943065522	NG	AF
522	New Zimbabwe newzimbabwe.com	125973854117415	ZW	AF
523	The Guardian ngrguardiannews.com	176502461635	NG	AF
524	Business News businessnews.com.ng	211032168937402	NG	AF
525	Masterweb Reports nigeriamasterweb.com	291985020925887	NG	AF
526	The Nigerian Observer nigerianobservernews.com	148046368615205	NG	AF
527	Nippon News nipponnews.net	277898815655	JP	AS
528	The Times-Picayune nola.com	99597577059	US	NA
529	The North Africa Journal north-africa.com	110960715505	US	NA
530	The Northern Miner northernminer.com	276490482462006	CA	NA
531	The Northern Standard northernstandard.ie	168433759858012	IE	EU
532	Noseweek noseweek.co.za	116905531677008	ZA	AF
533	NOW. now.mmedia.me/lb/en	171247176237894	LB	ME

534	NPR npr.org	10643211755	US	NA
535	Natural Resource Governance Institute resourcegovernance.org	219848484745333	US	NA
536	Nova Scotia novascotia.ca	161817467263	CA	NA
537	New Straits Times nst.com.my	135289458465	MY	AS
538	Nyasa Times nyasatimes.com	250647473519	MW	AF
539	NY Daily News nydailynews.com	268914272540	US	NA
540	New York Post nypost.com	134486075205	US	NA
541	The New York Times nytimes.com	5281959998	US	NA
542	NZ Herald nzherald.co.nz	34497296301	NZ	OC
543	The Daily Observer antiguaobserver.com	126646997376452	AG	CA
544	Ocean FM oceanfm.ie	115071508505532	IE	EU
545	Nuclear Energy Agency oecd-nea.org	161503327202946	FR	EU
546	Offaly Express offalyexpress.ie	207799065511	IE	EU
547	Offaly Independent offalyindependent.ie	121886488053	IE	EU
548	The Spectator spectator.co.uk	111263798903232	GB	EU
549	Olu Famous olufamous.com	292824824087658	NG	AF
550	Jamaicans.com jamaicans.com	128078147433	JM	NA
551	OneWorld oneworld.org	106968052697581	GB	EU
552	Government of Ontario news.ontario.ca	367410119963157	CA	NA
553	Online Nigeria onlinenigeria.com/uk	143914392334873	NG	AF

554	OpEdNews.com opednews.com	200583909956972	US	NA
555	Open Europe Today openeurope.org.uk	321253057971308	GB	EU
556	Organisation of Islamic Cooperation oic-oci.org	176887235707491	PS	ME
557	Organization for Security and Co-operation in Europe osce.org	59678478979	AT	EU
558	Ovation International ovationinternational.com	58902011402	GH	AF
559	Oxford Analytica oxan.com	160525917321265	GB	EU
560	Pacific Islands Report pidp.eastwestcenter.org	184980038221527	US	NA
561	Pakistan Tribune pakistantribune.com.pk	556680544487686	PK	ME
562	Pakistan Observer pakobserver.net	143465663100	PK	ME
563	The Palestinian Information Centre english.palinfo.com	233417303344412	PS	ME
564	Pan European Networks paneuropeannetworks.com	230201663697109	GB	EU
565	The Patriot Post patriotpost.us	51560645913	US	NA
566	People's Daily english.peopledaily.com.cn	188625661189259	CN	AS
567	People's Daily Online peoplesdaily-online.com	228444807376	NG	AF
568	People's Review peoplesreview.com.np	191277000903079	NP	AS
569	The Phnom Penh Post phnompenhpost.com	154245617928723	KH	AS
570	Phys phys.org	47849178041	US	NA
571	Twin Cities twincities.com	276896490404	US	NA
572	Pittsburgh Post-Gazette post-gazette.com	184142654825	US	NA
573	Planet Tonga planet-tonga.com	335939203143941	TO	OC

574	POLITICO politico.com	62317591679	US	NA
575	POLITICO Europe politico.eu	587266261407195	BE	EU
576	Portfolio portfolio.hu/en	546626485410734	HU	EU
577	The Portugal News theportugalnews.com	322352061097	PT	EU
578	Marianas Variety Guam mvguam.com	269007167849	GU	OC
579	The Post postzambia.com	1512926715599838	ZM	AF
580	Prague Daily Monitor praguemonitor.com	299876763374822	CZ	EU
581	The Prague Post praguepost.com	128602892226	CZ	EU
582	The Presidency of The Republic of South Africa thepresidency.gov.za	8383028996	ZA	AF
583	PR-Inside pr-inside.com	161831927209456	AT	EU
584	PR Newswire prnewswire.com	26247320522	US	NA
585	PTV World ptvworldnews.com.pk	406940679374569	PK	ME
586	Publish What You Pay publishwhatyoupay.org	176624229034172	GB	EU
587	The Punch punchng.com	206270189411151	NG	AF
588	Rabble rabble.ca	115517676270	CA	NA
589	American Renaissance amren.com	19051993499	US	NA
590	Radio Australia radioaustralia.net.au/international	128522530498035	AU	OC
591	Dabanga dabangasudan.org/en	82615736748	SD	AF
592	RNZ International rnzi.com	7759768730	NZ	OC
593	Radio Poland thenews.pl	121677301182864	PL	EU

594	Radio Prague radio.cz/en	183496134175	CZ	EU
595	Radio Caribbean International rcistlucia.com	386558457078	LC	CA
596	Red Pepper redpepper.co.ug	131452373539377	UG	AF
597	ReliefWeb reliefweb.int	57397818992	CH	EU-O
598	Repeating Islands repeatingislands.com	928983033797008	US	NA
599	Reporter 365 reporter365.com	278327162277225	US	NA
600	Sahara Reporters saharareporters.com	96184337702	NG	AF
601	Investing News Network investingnews.com	66284580629	CA	NA
602	Reuters reuters.com	114050161948682	GB	EU
603	Reuters UK uk.reuters.com	208314602512037	GB	EU
604	The Reykjavík Grapevine grapevine.is	8969907778	IS	EU-O
605	Radio Free Europe/Radio Liberty rferl.org	122264309574	CZ	EU
606	Radio France International english.rfi.fr	28764872018	FR	EU
607	RNW Media rnw.org	42251765664	NL	EU
608	Routes Online routesonline.com	126251777434574	GB	EU
609	Radio St. Lucia 97 rslonline.com	204248159604926	LC	CA
610	RTÉ News rte.ie	257558294273180	IE	EU
611	Russia Today rt.com	326683984410	RU	AS
612	RTT News rttnews.com	166456801229	US	NA
613	Rwandinfo rwandinfo.com/eng	107926999225610	RW	AF

614	The Salt Lake Tribune sltrib.com	35281584398	US	NA
615	Samoa News samoanews.com	446910992109668	WS	OC
616	Samoa Observer samoaobserver.ws	278391252189695	MP	OC
617	Security Assistance Monitor securityassistance.org	134409343259972	US	NA
618	Syrian Arab News Agency sana.sy/en	166379316760326	SY	ME
619	The San Diego Union-Tribune utsandiego.com	133508396112	US	NA
620	South Africa News sanews.gov.za	612864685431270	ZA	AF
621	The San Pedro Sun sanpedrosun.com	246187688749182	BZ	NA
622	Saudi Gazette saudigazette.com.sa	137192489702943	US	NA
623	Save the Children savethechildren.org	8047221596	US	NA
624	Santa Barbara Independent independent.com	30883294835	US	NA
625	SBS News sbs.com.au/news	125982670754724	AU	OC
626	SciDev.Net scidev.net/	109375082451665	GB	EU
627	Science Alert sciencealert.com	7557552517	US	NA
628	Science sciencemag.org	96191425588	US	NA
629	Scientific American scientificamerican.com	22297920245	US	NA
630	Searchlight searchlight.vc	152771864791320	VC	CA
631	Seatrade Maritime News seatrade-global.com	470795739645931	GB	EU
632	The Seattle Times seattletimes.com	38472826214	US	NA
633	Security Council Report securitycouncilreport.org	131442826906850	US	NA

634	The Sentinel thisisstaffordshire.co.uk	11878899813	GB	EU
635	SFGate sfgate.com	105702905593	US	NA
636	Shabait shabait.com	455303841165054	ER	AF
637	Shannon Side shannonside.ie	109597305738451	IE	EU
638	ShippingWatch - English shippingwatch.com	606618229391196	DK	EU
639	Silicon Republic siliconrepublic.com	166431264240	IE	EU
640	Sioux City Journal siouxcityjournal.com	51119691219	US	NA
641	SIPRI sipri.org	309816241933	SE	EU
642	Sky News news.sky.com	164665060214766	GB	EU
643	Sligo Weekender sligoweekender.ie	139420522800336	IE	EU
644	The Smithville Herald smithvilleherald.com	79136059123	US	NA
645	Socialists and Democrats in the European Parliament socialistsanddemocrats.eu	127925800618165	BE	EU
646	Mining Engineering Online me.smenet.org	43749334577	US	NA
647	The Sofia Echo sofiaecho.com	538279649571477	BG	EU
648	Novinite novinite.com	59362904867	BG	EU
649	South China Morning Post scmp.com	355665009819	CN	AS
650	SouthernMinn southernminn.com	376746812416329	US	NA
651	South Sudan News Agency southsudannewsagency.com	124568540913185	SD	AF
652	Asia One asiaone.com	121790674546188	SG	AS
653	SpiceIslander spiceislander.com	147014315348275	GD	CA

654	Spiegel spiegel.de/international	39205942284	DE	EU
655	Sport Fishing sportfishingmag.com	14417384293	US	NA
656	Sputnik en.rian.ru	357990416180	RU	AS
657	Stabroek News stabroeknews.com	130548565052	GY	SA
658	Standard Digital standardmedia.co.ke	88201339429	KE	AF
659	Honolulu Star-Advertiser staradvertiser.com	112299605447935	US	NA
660	Star Tribune startribune.com	42739463017	US	NA
661	State House Uganda statehouse.go.ug	158559710878311	UG	AF
662	Stuff stuff.co.nz	21253884267	NZ	OC
663	Sudanese Media Center smc.sd/eng	328263710712432	SD	AF
664	Sudan Tribune sudantribune.com	147157358639460	SD	AF
665	Sudan Vision Daily news.sudanvisiondaily.com	100247370113145	SD	AF
666	Sunday BusinessPost businesspost.ie	811773748881458	IE	EU
667	Sunday Standard sundaystandard.info	404354156357611	BW	AF
668	Sunday World sundayworld.com	175466346599	IE	EU
669	Survival survivalinternational.org	19668531552	GB	EU
670	Swazi Observer observer.org.sz	253248861365174	SZ	AF
671	swissinfo swissinfo.org	81049933496	CH	EU-O
672	SW Radio Africa swradioafrica.com	105987640597	ZW	AF
673	The Sidney Morning Herald smh.com.au	104598631263	AU	OC

674	Taipei Times taipeitimes.com	210998785327	TW	AS
675	Talk City 91.1 talkcity91fm.wordpress.com	133446376708654	TT	SA
676	Radio Tamazuj radiotamazuj.org	298524290184928	SD	AF
677	TamilNet tamilnet.com	129476557086822	LK	AS
678	Tampa Bay Times tampabay.com	9924394837	US	NA
679	The Afro News theafronews.com	115007838514142	CA	NA
680	TASS en.itar-tass.com	221338351211505	RU	AS
681	TCPalm tcpalm.com	62811590881	US	NA
682	Teagasc teagasc.ie	124025717676271	IE	EU
683	TechCrunch techcrunch.com	8062627951	US	NA
684	MIT Technology Review technologyreview.com	17043549797	US	NA
685	Tehran Times tehrantimes.com	265237380156809	IR	ME
686	The Telegraph telegraph.co.uk	143666524748	GB	EU
687	Television Jamaica televisionjamaica.com	135996373112459	JM	NA
688	Intelligence and Terrorism Information Center terrorism-info.org.il	228183363874247	IL	ME
689	The Africa Report theafricareport.com	10750083165	FR	EU
690	The Anguillian theanguillian.com	108369529266234	AI	NA
691	Arab News arabnews.com	10250877124	CA	NA
692	The Athens News athensnews.com	94790271234	US	NA
693	The Atlantic theatlantic.com	29259828486	CA	NA

694	The Australian theaustralian.com.au	45388134977	AU	OC
695	The Bahama Journal jonesbahamas.com	234368746664613	BS	CA
696	The Baltic Times baltictimes.com	224870291024597	LV	EU
697	The Brunei Times bt.com.bn	119608528105659	BN	AS
698	The Buffalo News buffalonews.com	181362508150	US	NA
699	The Business Times businesstimes.com.sg	288031731262011	SG	AS
700	The Charlotte Observer charlotteobserver.com	42580340317	US	NA
701	Chicago Sun-Times chicago.suntimes.com	47864940833	US	NA
702	The Citizen thecitizen.co.tz	234582629937266	TZ	AF
703	The Clare Champion clarechampion.ie	137725366369881	IE	EU
704	The Costa Rica News thecostaricanews.com	214498872385	CR	CA
705	The Courier thecourier.co.uk	325681791214	GB	EU
706	The Daily Tribune tribune.net.ph	253124567195	PH	AS
707	The Dominican thedominican.net	102945123139643	DO	CA
708	The EastAfrican theeastafrican.co.ke	155073814515056	KE	AF
709	The Economist economist.com	6013004059	GB	EU
710	The Financial Express financialexpress.com	157671354275436	IN	AS
711	The Financial Gazette financialgazette.co.zw	288500807952079	ZW	AF
712	The Fishing Website fishing.net.nz	147044425346125	NZ	OC
713	The Freeport News freeport.nassauguardian.net	46726844467	BS	CA

714	Fremont Tribune fremonttribune.com	108066434972	US	NA
715	The Frontier Post thefrontierpost.com	125681824173980	PK	ME
716	The Globe and Mail theglobeandmail.com	140961138903	CA	NA
717	The Grio thegrio.com	75928194876	US	NA
718	The Guardian theguardian.com/uk	10513336322	GB	EU
719	Harborough Mail harboroughmail.co.uk	219817851378553	GB	EU
720	The Herald herald.co.zw	380838785328009	ZW	AF
721	The Himalayan Times thehimalayantimes.com	166920243347320	NP	AS
722	The Hindu thehindu.com	163974433696568	IN	AS
723	The International Institute for Strategic Studies iiss.org	29840385993	GB	EU
724	The Independent independent.co.uk	13312631635	GB	EU
725	The Insider insiderzim.com	122033327855067	ZW	AF
726	The Irish World theirishworld.com	49341828952	IE	EU
727	The Irrawaddy irrawaddy.org	112882212089978	TH	AS
728	Jarkata Globe thejakartaglobe.beritasatu.com	26578392579	ID	AS
729	The Jamestown Foundation jamestown.org	106920576037705	IQ	ME
730	The Japan Times japantimes.co.jp	6321018343	JP	AS
731	The Jerusalem Post jpost.com	159050394216641	IL	ME
732	The Jordan Times jordantimes.com	68601553125	JO	ME
733	TheJournal.ie thejournal.ie	137576076262825	IE	EU

734	Daily Post kenyan-post.blogspot.it	271776752878863	KE	AF
735	Kenya Star kenyastar.com	143728442343553	KE	AF
736	The Korea Times koreatimes.co.kr	227456724028836	KR	AS
737	Herald Argus heraldargus.com	565599370262091	US	NA
738	The Leader theleader.info	238743816156394	ES	EU
739	The Local Austria thelocal.at	1426552570931475	AT	EU
740	The Local Denmark thelocal.dk	1433835800214625	DK	EU
741	The Local France thelocal.fr	258002227555924	FR	EU
742	The Local Germany thelocal.de	214435206012	DE	EU
743	The Local Italy thelocal.it	384349881653734	IT	EU
744	The Local Norway thelocal.no	514495608666410	SE	EU
745	The Local Spain thelocal.es	181292335348466	ES	EU
746	The Local Sweden thelocal.se	220353389618	SE	EU
747	The Local Switzerland thelocal.ch	381700821924600	SE	EU
748	The Malta Independent independent.com.mt	480288348662981	MT	EU
749	The Manila Times manilatimes.net	111026632011	PH	AS
750	The Mayo News mayonews.ie	64555193331	IE	EU
751	The Middle East Media Research Institute memrijttm.org	14310874716	US	NA
752	The Montserrat Reporter themontserratreporter.com	203080105851	MS	NA
753	Mumbai Mirror mumbaimirror.com	114157215335068	IN	AS

754	The Namibian namibian.com.na	284922901537221	NA	AF
755	The National thenational.ae	148788988477345	AE	ME
756	The Nationalist nationalist.ie	189189374491255	IE	EU
757	The Nation nation.com.pk	57383301711	PK	AS
758	The Nation thenation.com	7629206115	US	NA
759	The New Age thenewage.co.za	138238556209769	ZA	AF
760	New Republic newrepublic.com	161419311535	US	NA
761	The News thenews.com.pk	131257086910180	PK	ME
762	The New Sudan Vision newsudanvision.com	105477586161630	SD	AF
763	The New Times newtimes.co.rw	301148803327544	RW	AF
764	New Vision newvision.co.ug	329423169077	UG	AF
765	The Norway Post norwaypost.no	185753021499903	NO	EU-O
766	Nottingham Post thisisnottingham.co.uk	309833935716287	GB	EU
767	The Nassau Guardian thenassauguardian.com	131498596889079	BS	CA
768	The Oregonian oregonlive.com	6321831972	US	NA
769	Ottawa Citizen ottawacitizen.com	6533373917	CA	NA
770	The Palestine Chronicle palestinechronicle.com	302592203117857	PS	ME
771	The Parliament theparliamentmagazine.eu	471876160233	BE	EU
772	The Peninsula thepeninsulaqatar.com	111063551581	QA	ME
773	The Post thepost.co.ls	1151120021571319	LS	AF

774	The Railway Magazine railwaymagazine.co.uk	135345903226042	GB	EU
775	The Reporter reporter.bz	236081746489917	BZ	NA
776	The Republic therepublic.com	54499120758	US	NA
777	The Russian Navy rusnavy.com	110378652324684	RU	AS
778	The Scotsman scotsman.com	293226174987	GB	EU
779	Daily Record and Sunday Mail dailyrecord.co.uk	187523381277554	GB	EU
780	The Scottish Government scotland.gov.uk	200786289976224	GB	EU
781	The Slovak Spectator spectator.sme.sk	59260989565	SK	EU
782	The Sofia Globe sofiaglobe.com	445561718804553	BG	EU
783	The Source thesourceng.com	416043501758998	NG	AF
784	The Southern Star southernstar.ie	310855455632145	IE	EU
785	The Standard thestandard.com.hk	140973382599494	HK	AS
786	The Standard thestandard.co.zw	103305193039107	ZW	AF
787	The Standard thestandard.com.ph	835114793210549	PH	AS
788	The Star the-star.co.ke	224865787558662	KE	AF
789	The Star thestar.com.my	11450527254	MY	AS
790	The St. Kitts and Nevis Observer thestkittsnevisobserver.com	476828099098371	KN	CA
791	The Straits Times straitstimes.asiaone.com	129011692114	SG	AS
792	The Sun thesun.co.uk	161385360554578	GB	EU
793	The Sun sunnewsonline.com	139709372821659	NG	AF

794	The Vincentian thevincentian.com	370649629649050	VC	CA
795	The Vindicator vindy.com	77780874764	US	NA
796	The Visitor thevisitor.co.uk	68554461041	GB	EU
797	The Voice thevoicebw.com	202343559085	BW	AF
798	The Voice thevoiceslu.com	160460070659226	LC	CA
799	The Washington Times washingtontimes.com	35994014410	US	NA
800	The Weather Channel climate.weather.com	118071565920	US	NA
801	The Yemen Times yementimes.com	130936023624588	YE	ME
802	The Zambezian thezambezian.com	117965765017340	BW	AF
803	The Zimbabweans thezimbabwean.co.uk	217203691648702	ZW	AF
804	The Zimbabwe Independent theindependent.co.zw	331288056957552	ZW	AF
805	The Zimbabwe Mail thezimbabwemail.com	467260923305980	ZW	AF
806	This Day thisdaylive.com	142936439094106	NG	AF
807	Thomson Reuters Foundation trust.org	31301735406	GB	EU
808	Tico Times ticotimes.net	124823954224180	CR	CA
809	The Sunday Times thesundaytimes.co.uk	147384458624178	GB	EU
810	Times Colonist timescolonist.com	50465429712	CA	NA
811	Times Daily timesdaily.com	121598674610038	US	NA
812	Times LIVE timeslive.co.za	136956534616	ZA	AF
813	The Times of India timesofindia.indiatimes.com	26781952138	IN	AS

814	Times of Malta timesofmalta.com	160227208174	MT	EU
815	Times Of Oman timesofoman.com	136413806438283	OM	ME
816	Times Of Swaziland times.co.sz	142322529190260	SZ	AF
817	Times of Zambia times.co.zm	312705315487393	ZM	AF
818	Tipp FM tippfm.com	221599401189017	IE	EU
819	Tipperary Star tipperarystar.ie	138048782943147	IE	EU
820	Tobago News thetobagonews.com	335414633195243	TT	SA
821	Today FM todayfm.com	36108807567	IE	EU
822	TODAY todayonline.com	147858757571	SG	AS
823	Today's Zaman todayszaman.com	173873526003430	TR	EU-C
824	Topix topix.net	19062931201	CA	NA
825	The Toronto Star thestar.com	184906186150	CA	NA
826	Toronto Sun torontosun.com	189526659635	CA	NA
827	TorrentFreak torrentfreak.com	9087497371	NL	EU
828	Trade Arabia tradearabia.com	439634269411506	BH	ME
829	Trade Winds tradewindsnews.com	132240771277	US	NA
830	Trading Markets tradingmarkets.com	151413724902563	US	NA
831	TribLIVE triblive.com	55863814979	US	NA
832	Tribune242 tribune242.com	33564008678	BS	CA
833	Nigerian Tribune tribune.com.ng	77912203638	NG	AF

834	Trinidad and Tobago Newsday newsday.co.tt	311107925748485	TT	SA
835	T&T Guardian guardian.co.tt	78081948066	TT	SA
836	TV3 tv3.ie/news.php	100183534537	IE	EU
837	United Arab Emirates Interact uaeinteract.com	193093937374424	GB	EU
838	United Democratic Movement udm.org.za	494801983919865	ZA	AF
839	The Observer observer.ug	267688186591432	UG	AF
840	UK Government gov.uk	408582579294175	GB	EU
841	MSN News msn.com/en-gb/news	358837740527	GB	EU
842	United Nations Mission in Darfur unamid.unmissions.org	164650630228650	SD	AF
843	UN Dispatch undispatch.com	6069582499	US	NA
844	United Nations in Azerbaijan un-az.org	281319152013943	AZ	AS
845	Food and Agriculture Organization of the United Nations fao.org	46370758585	IT	EU
846	UNHCR The UN Refugee Agency unhcr.org	13204463437	CH	EU-O
847	UNIAN unian.info	1476822619259011	UA	EU-O
848	United Nations un.org/en	54779960819	US	NA
849	United Nations Mission in South Sudan unmiss.unmissions.org	160839527325060	SD	AF
850	United Press International upi.com	101911273177707	US	NA
851	UN Office for the Coordination of Humanitarian Affairs unocha.org	135156639833927	US	GL
852	UN Office on Drugs and Crime unodc.org	43559937330	AT	EU
853	United Nations Radio unmultimedia.org/radio/english	235134190239	US	NA

854	The Finnish Institute of International Affairs fiia.fi/en	117351301649740	FI	EU
855	The Post postnewsline.com	146662198720301	CM	AF
856	US Agency for International Development usaidlandtenure.net	62690599685	US	NA
857	USA Today usatoday.com	13652355666	US	NA
858	US Department of State state.gov	15877306073	US	NA
859	US News and World Report usnews.com	5834919267	US	NA
860	UTV u.tv	115860925163321	IE	EU
861	Estonian Ministry of Foreign Affairs vm.ee/en	57904691979	EE	EU
862	The Vancouver Sun vancouversun.com	7116517082	CA	NA
863	Vanguard vanguardngr.com	135140476511057	NG	AF
864	Vietnam News vietnamnews.vn	254068371381722	VN	AS
865	Vibe Ghana vibeghana.com	102329089882316	GH	AF
866	viEUws vieuws.eu	117994898215654	FR	EU
867	The Virgin Islands Daily News virginislandsdailynews.com	132196636944018	VI	NA
868	Voice of Barbados vob929.com	1511527652404540	BB	CA
869	Voice of America voanews.com	36235438073	US	NA
870	Voice of Russia sputniknews.com/voiceofrussia	842603639086940	RU	EU-O
871	Flanders News deredactie.be/cm/vrtnieuws.english	212796021897	BE	EU
872	The Register theregister.co.uk	206419956048907	GB	EU
873	Wales Online walesonline.co.uk	21226447182	GB	EU

874	Emirate News Agency wam.ae/en	124222007613925	AE	ME
875	Wandsworth Guardian wandsworthguardian.co.uk	113349742029506	GB	EU
876	Warsaw Business Journal wbj.pl	91558833313	PL	EU
877	Washington Examiner washingtonexaminer.com	40656699159	US	NA
878	The Washington Post washingtonpost.com	6250307292	US	NA
879	Watching America watchingamerica.com	118258794866306	US	NA
880	Waterford News and Star waterford-news.ie	120792148007585	IE	EU
881	WBUR wbur.org	9427513649	US	NA
882	Waterloo Cedar Falls Courier wcfcourier.com	212364145478265	US	NA
883	Blitz weeklyblitz.net	242506829624	BD	AS
884	Western People westernpeople.ie	192666647731	IE	EU
885	Western Telegraph westerntelegraph.co.uk	180521675319022	GB	EU
886	Independent Westmeath westmeathindependent.ie	251005360424	IE	EU
887	Wexford Echo wexfordecho.ie	111308662280849	IE	EU
888	The White House whitehouse.gov	63811549237	US	NA
889	Wired wired.com	19440638720	US	NA
890	World News wn.com	229101503845879	GB	EU
891	The World Bank documents.worldbank.org	153371894688575	US	NA
892	Worldcrunch worldcrunch.com	271986174770	IE	EU
893	World Finance worldfinance.com	725598290845413	US	NA

894	World Fishing and Aquaculture worldfishing.net	552321618120006	GB	EU
895	UN World Food Programme wfp.org	28312410177	IT	EU
896	Omaha World-Herald omaha.com	6445219629	US	NA
897	World Maritime News worldmaritimenews.com	309171835876244	NL	EU
898	Luxemburger Wort wort.lu/en	174116812644146	LU	EU
899	The Wall Street Journal online.wsj.com	8304333127	US	NA
900	WWF panda.org	20373776304	CH	EU-O
901	WXOW wxow.com	231408475203	US	NA
902	China Xinhua News xinhuanet.com/english	338109312883186	CN	AS
903	Yahoo! Indian News in.news.yahoo.com	131747896861126	IN	AS
904	Yahoo! News news.yahoo.com	338028696036	US	NA
905	Ya Libnan yalibnan.com	793949857311259	LB	ME
906	Yemen Post yemenpost.net	500685219968202	YE	ME
907	YLE yle.fi	192534820828660	FI	EU
908	Ynetnews ynetnews.com	129653250402500	IL	ME
909	The Yorkshire Post yorkshirepost.co.uk	316795048375439	GB	EU
910	Zambia Daily Mail daily-mail.co.zm	172219889538699	ZM	AF
911	Zambian Watchdog zambianwatchdog.com	129987587052000	ZM	AF
912	Zambia Reports zambiareports.com	208755685909131	ZM	AF
913	Thomson Reuters Zawya zawya.com	112684488765578	PS	ME

914	Zodiak Online zodiakmalawi.com	120942427951729	MW	AF
915	ZDNet zdnet.com	5953112932	US	NA
916	Zero Hedge zerohedge.com	116467201763793	BG	EU
917	ZF English zfenglish.com	130883803643584	RO	EU
918	Zimbabwe Broadcasting Corporation zbc.co.zw	126155317511807	ZW	AF
919	ZimEye zimeye.org	150154425045764	ZW	AF
920	ZNS Network znsbahamas.com	248934135146821	BS	CA

A.2 EU Countries Dataset

The Table 17 contains the 225 news pages that form the EU Countries Dataset. It includes the name, website and country of the news sites, as well as the Facebook ID of their corresponding Facebook page at the time. The countries are indicated with their ISO Alpha-2 international code.

Table 17: Pages in the EU Countries Dataset.

	Name and Website	Facebook ID	Community
1	ARD ard.de	48219766388	DE
2	Augsburger Allgemeine Zeitung augsburger-allgemeine.de	121104385783	DE
3	Badische Zeitung badische-zeitung.de	177670301122	DE
4	Berliner Morgenpost morgenpost.de	46239931235	DE
5	Berliner Zeitung berliner-zeitung.de	137267732953826	DE
6	Bild bild.de	25604775729	DE
7	B.Z. bz-berlin.de	57187632436	DE
8	Das Erste daserste.de	176772398231	DE
9	Der Spiegel spiegel.de	38246844868	DE
10	Der Tagesspiegel tagesspiegel.de	59381221492	DE
11	Der Westen derwesten.de	243001859426137	DE
12	Die Tageszeitung taz.de	171844246207985	DE
13	Die Welt welt.de	97515118114	DE
14	Die Zeit zeit.de	37816894428	DE
15	Express express.de	172718036608	DE

16	Focus focus.de	37124189409	DE
17	Frankfurter Allgemeine Zeitung faz.net	346392590975	DE
18	Frankfurter Rundschau fr.de	134100583282150	DE
19	Freie Presse freiepresse.de	375109771472	DE
20	Freitag freitag.de	313744767921	DE
21	GMX gmx.net	187741777922914	DE
22	Hamburger Abendblatt abendblatt.de	121580125458	DE
23	Hamburger Morgenpost mopo.de	196072707519	DE
24	Handelsblatt handelsblatt.com	104709558232	DE
25	Hannoversche Allgemeine Zeitung haz.de	198530121257	DE
26	Huffington Post DE huffingtonpost.de	366193510165011	DE
27	Junge Freiheit jungefreiheit.de	13479664941	DE
28	Kölner Stadt-Anzeiger ksta.de	141063022950	DE
29	Leipziger Volkszeitung lvz.de	114360055263804	DE
30	Mitteldeutsche Zeitung mz-web.de	141558262607	DE
31	n-tv online n-tv.de	126049165307	DE
32	Ostsee-Zeitung ostsee-zeitung.de	374927701107	DE
33	ProSieben Newstime prosieben.de/tv/newstime	64694257920	DE
34	Rheinische Post rp-online.de	50327854366	DE
35	RTL aktuell rtluell.de	119845424729050	DE

36	SAT1 Nachrichten sat1.de/news	171663852895480	DE
37	Schleswig-Holsteinischer Zeitungsverlag shz.de	248528847673	DE
38	Stern stern.de	78766664651	DE
39	Stuttgarter Nachrichten stuttgarter-nachrichten.de	144537361776	DE
40	Stuttgarter Zeitung stuttgarter-zeitung.de	129349103260	DE
41	Süddeutsche Zeitung sueddeutsche.de	215982125159841	DE
42	tagesschau tagesschau.de	193081554406	DE
43	t-online t-online.de	24897707939	DE
44	WAZ waz.de	117194401183	DE
45	WEB.DE web.de	56488242934	DE
46	Wirtschafts Woche wiwo.de	93810620818	DE
47	Yahoo News DE de.nachrichten.yahoo.com	166721106679241	DE
48	ZDF zdf.de	154149027994068	DE
49	ZDF heute heute.de	112784955679	DE
50	20 MINUTOS 20minutos.es	38352573027	ES
51	ABC abc.es	7377874895	ES
52	Antena 3 antena3.com	55353596297	ES
53	Cadena Ser cadenaser.com	15658775846	ES
54	Canarias 7 canarias7.es	85160277321	ES
55	Cinco Días cincodias.elpais.com	36280712574	ES

56	COPE cope.es	15829535820	ES
57	Cuatro news cuatro.com/noticias	96876562265	ES
58	Diario de Cádiz diariodecadiz.es	128335533904779	ES
59	Diario de Ibiza diariodeibiza.es	255177630236	ES
60	Diario de Mallorca diariodemallorca.es	155352736257	ES
61	Diario de Navarra diariodenavarra.es	103384039711468	ES
62	El Comercio elcomercio.es	64673887657	ES
63	El Confidencial elconfidencial.com	63830851925	ES
64	El Confidencial Digital elconfidencialdigital.com	202726949863885	ES
65	El Correo elcorreo.com	280982578099	ES
66	El Correo Gallego elcorreogallego.es	152802838075123	ES
67	El Día eldia.es	165210860204301	ES
68	ElDiario.es eldiario.es	417471918268686	ES
69	El Diario Montañés eldiariomontanes.es	109434489075314	ES
70	El Diario Vasco diariovasco.com	91085818678	ES
71	El Economista eleconomista.es	56760767000	ES
72	El Español elespanol.com	693292367452833	ES
73	El Mundo elmundo.es	10407631866	ES
74	El Norte de Castilla elnortedecastilla.es	98474974005	ES
75	El País elpais.com	8585811569	ES

76	El Periódico elperiodico.com	93177351543	ES
77	Expansión expansion.com	93983931918	ES
78	Faro de Vigo farodevigo.es	123746764304270	ES
79	Heraldo de Aragón heraldo.es	130012437016272	ES
80	Hoy hoy.es	85593393832	ES
81	Ideal ideal.es	64258697112	ES
82	Información diarioinformacion.com	410523955526	ES
83	La Gaceta de Salamanca lagacetadesalamanca.es	319669591452311	ES
84	La Nueva España lne.es	51837272861	ES
85	La Opinión de Málaga laopiniondemalaga.es	80999977105	ES
86	La Opinión de Murcia laopiniondemurcia.es	106647502704110	ES
87	La Opinión de Tenerife laopinion.es	112238345503995	ES
88	La Provincia laprovincia.es	124641092828	ES
89	La Razón larazon.es	113080018770027	ES
90	La Sexta lasexta.com	39172614918	ES
91	Las Provincias lasprovincias.es	20810574989	ES
92	La Vanguardia lavanguardia.com	156552584408339	ES
93	La Verdad laverdad.es	120857625399	ES
94	La Voz de Asturias lavozdeasturias.es	101351926940208	ES
95	La Voz De Galicia lavozdegalicia.es	350393845757	ES

96	Levante-EMV levante-emv.com	106329485190	ES
97	Libertad Digital libertaddigital.com	141423087721	ES
98	MSN España msn.com/es-es	35966491049	ES
99	Onda Cero ondacero.es	99040469027	ES
100	Público publico.es	75084861845	ES
101	QUE! que.es	97090259641	ES
102	RTVE rtve.es	133623265400	ES
103	Sur diariosur.es	52107727250	ES
104	Telecinco telecinco.es	50353113909	ES
105	Última Hora ultimahora.es	114680095225282	ES
106	Yahoo News ES es.noticias.yahoo.com	284428852938	ES
107	20 Minutes 20minutes.fr	51555073310	FR
108	Agence France-Presse afp.com/fr	114100038626559	FR
109	BFMTV bfmtv.com	43896752783	FR
110	Canal+ canalplus.fr	144056732332683	FR
111	Challenges challenges.fr	79566127213	FR
112	Charente Libre charentelibre.fr	144375072241306	FR
113	Charlie Hebdo charliehebdo.fr	106626879360459	FR
114	CNES Matin cnewsmatin.fr	181111805243991	FR
115	CNEWS cnews.fr	76952916976	FR

116	Corse Matin corsematin.com	107249929306302	FR
117	Courrier international courrierinternational.com	142114104887	FR
118	Dernieres Nouvelles d'Alsace dna.fr	19004867327	FR
119	FranceInfo francetvinfo.fr	135112586936434	FR
120	France Soir francesoir.fr	53638966652	FR
121	France Télévisions francetelevisions.fr	179086202130933	FR
122	Huffington Post FR huffingtonpost.fr	284129444969978	FR
123	La Croix la-croix.com	108828257010	FR
124	La Dépêche du Midi ladepeche.fr	271219815470	FR
125	L'Alsace - Le Pays lalsace.fr	181480351879611	FR
126	La Montagne lamontagne.fr	146949065315655	FR
127	La Nouvelle République du Centre Ouest lanouvellerepublique.fr	87693933163	FR
128	La Provence laprovence.com	119213845538	FR
129	La République des Pyrennées larepubliquedespyrenees.fr	148446219817	FR
130	La République du Centre larep.fr	211082695569481	FR
131	La Tribune latribune.fr	18950434380	FR
132	La Voix du Nord lavoixdunord.fr	76635774021	FR
133	Le Bien Public bienpublic.com	106094599409	FR
134	Le Courrier Picard courrier-picard.fr	58080584133	FR
135	Le Dauphiné Libéré ledauphine.com	122601757780987	FR

136	Le Figaro lefigaro.fr	61261101338	FR
137	Le Journal du Dimanche lejdd.fr	246577183385	FR
138	Le Monde lemonde.fr	14892757589	FR
139	Le Monde Diplomatique monde-diplomatique.fr	34398236687	FR
140	Le Nouvel Observateur tempsreel.nouvelobs.com	198508090036	FR
141	Le Parisien leparisien.fr	36550584062	FR
142	Le Point lepoint.fr	49173930702	FR
143	Le Populaire du Centre lepopulaire.fr	240500052515	FR
144	Le Progrès leprogres.fr	104985642868265	FR
145	Le Républicain Lorrain republicain-lorrain.fr	142638581774	FR
146	Les Échos lesechos.fr	123440511000645	FR
147	L'Est Républicain estrepublicain.fr	190366851765	FR
148	Le Télégramme letelegramme.fr	97539957978	FR
149	L'Express lexpress.fr	9359316996	FR
150	L'Humanité humanite.fr	254585183694	FR
151	Libération liberation.fr	147126052393	FR
152	L'Indépendant lindependant.fr	52697519148	FR
153	L'internaute linternaute.com	156569814356922	FR
154	L'Opinion lopinion.fr	445890365491209	FR
155	L'Union lunion.fr	100163350071823	FR

156	Marianne marianne.net	369717525444	FR
157	Mediapart mediapart.fr	116070051527	FR
158	Metro France lci.fr	411124728976705	FR
159	Midi Libre midilibre.fr	183518182558	FR
160	MSN France msn.com/g00/fr-fr	136932803018290	FR
161	Nice-Matin nicematin.com	388223307574	FR
162	Nord-Littoral nordlittoral.fr	344969675415	FR
163	Ouest France ouest-france.fr	270122530294	FR
164	Paris Match parismatch.com	117714667328	FR
165	Paris Normandie paris-normandie.fr	195238257180091	FR
166	Révolution Permanente revolutionpermanente.fr	732277203520737	FR
167	Sud Oest sudouest.fr	58305334711	FR
168	Télérama telerama.fr	109520835773096	FR
169	TF1 news tf1.fr/news	34610502574	FR
170	Var Matin varmatin.com	365009223614	FR
171	Yahoo News FR fr.news.yahoo.com	138207559575213	FR
172	Alto Adige altoadige.gelocal.it	447795960541	IT
173	Ansa ansa.it	158259371219	IT
174	Avvenire avvenire.it	128533807252295	IT
175	Corriere Adriatico corriereadriatico.it	431943793507773	IT

176	Corriere della Sera corriere.it	284515247529	IT
177	Corriere del Mezzogiorno corrieredelmezzogiorno.corriere.it	84805991975	IT
178	Gazzetta di Modena gazzettadimodena.gelocal.it	131613613524326	IT
179	Gazzetta di Reggio gazzettadireggio.gelocal.it	102328739818445	IT
180	Giornale di Brescia giornaledibrescia.it	352193836938	IT
181	Giornale di Sicilia gds.it	211307618890745	IT
182	Huffington Post IT huffingtonpost.it	276376685795308	IT
183	Il Blog di Beppe Grillo beppegrillo.it	56369076544	IT
184	Il Centro ilcentro.gelocal.it	261504285205	IT
185	Il Fatto Quotidiano ilfattoquotidiano.it	132707500076838	IT
186	Il Foglio ilfoglio.it	61703722992	IT
187	Il Gazzettino ilgazzettino.it	154142713068	IT
188	Il Giornale ilgiornale.it	323950777458	IT
189	Il Giornale di Vicenza ilgiornaledivicenza.it	154836331469	IT
190	Il Manifesto ilmanifesto.info	61480282984	IT
191	Il Mattino ilmattino.it	210639995470	IT
192	Il Mattino di Padova mattinopadova.gelocal.it	189556995002	IT
193	Il Messaggero ilmessaggero.it	124918220854917	IT
194	Il Messaggero Veneto messaggeroveneto.gelocal.it	195905383236	IT
195	Il Piccolo ilpiccolo.gelocal.it	341809745380	IT

196	Il Resto del Carlino ilrestodelcarlino.it	200174860861	IT
197	Il Secolo XIX ilsecoloxix.it	36493277214	IT
198	Il Sole 24 Ore ilsole24ore.com	38812693516	IT
199	Il Tirreno iltirreno.gelocal.it	75980429042	IT
200	LA7 la7.it	252449503661	IT
201	L'Adige ladige.it	134572506600855	IT
202	La Gazzetta del Mezzogiorno lagazzettadelmezzogiorno.it	184749620911	IT
203	La Gazzetta di Mantova gazzettadimantova.gelocal.it	62769612287	IT
204	La Gazzetta di Parma gazzettadiparma.it	309928567597	IT
205	La Nazione lanazione.it	87812020989	IT
206	La Nuova di Venezia e Mestre nuovavenezia.gelocal.it	338049475695	IT
207	La Nuova Sardegna lanuovasardegna.gelocal.it	226626114877	IT
208	La Provincia Pavese laprovinciapavese.gelocal.it	57687391957	IT
209	L'Arena larena.it	108431819182401	IT
210	La Repubblica repubblica.it	179618821150	IT
211	La Stampa lastampa.it	63873785957	IT
212	La Tribuna di Treviso tribunatreviso.gelocal.it	243933437208	IT
213	L'Eco di Bergamo ecodibergamo.it	197197145813	IT
214	L'Espresso espresso.repubblica.it	259865949240	IT
215	Libero Quotidiano liberoquotidiano.it	188776981163133	IT

216	L'Unione Sarda unionesarda.it	231465552656	IT
217	L'Unità unita.tv	292449724097	IT
218	MSN Italia msn.com/it-it	232690009759	IT
219	Nuovo Quotidiano di Puglia quotidianodipuglia.it	119992291359480	IT
220	RAI News rainews.it	124992707516031	IT
221	Rai.TV raiplay.it	88988179171	IT
222	Sky TG24 tg24.sky.it	215275341879427	IT
223	TgCom24 tgcom24.mediaset.it	40337124609	IT
224	Trentino trentinocorrierealpi.gelocal.it	82383189226	IT
225	Yahoo News IT it.notizie.yahoo.com	81262596234	IT

A.3 Vaccine Dataset

The Table 19 contains the 243 Facebook pages that form the Vaccine Dataset. The pages in the table correspond to the English filtered set. The table includes the name of the Facebook page as well as the ID, and the community they belong to after the manual classification. The pages that belong to the *anti-vaccines* community are tagged as C_1, and those belonging to the *pro-vaccines* community as C_2

Table 19: Pages in the Vaccine Dataset.

	Facebook Name	Facebook ID	Community
1	Animals Talk: Vaccine Damage	579549288803619	C_1
2	Anti Vaccination Saskatoon	620976781291717	C_1
3	Australian Vaccination-skeptics Network Inc. - AVN	1594559574167102	C_1
4	Autism & Vaccine Injury Recovery With Homeopathy	160119314016596	C_1
5	Californians for Vaccine Choice	1374879262831019	C_1
6	Carroll Troberman Vaccine Injury Law	1415251895436029	C_1
7	Childhood Vaccine Safety	1560012117349488	C_1
8	Council for Vaccine Safety	661205827357302	C_1
9	Dangers of Vaccines	174614699699238	C_1
10	Dear Jill Hennessy - From the Anti-Vaccine Community	1626492777643266	C_1
11	Dissolving Illusions - Disease, Vaccines, and the Forgotten History	264612703591313	C_1
12	Doctors & Parents Are Speaking Out Against HPV Vaccines	539737672763787	C_1
13	Dr. Tenpenny on Vaccines and Current Events	171964245890	C_1
14	EFVV - European Forum for Vaccine Vigilance	919754961423177	C_1
15	Great Mothers (and others) Questioning Vaccines	199382473435132	C_1
16	Healthy Alternatives to Vaccinations	904259436307108	C_1
17	Hendra, the vets, the virus and the vaccine	126110434593574	C_1

18	HIV Vaccine	148921448644869	C_1
19	Homeoprophylaxis:The Vaccine Alternative	445735848802935	C_1
20	Indiana Coalition for Vaccination Choice	318619221553484	C_1
21	Inside Vaccines	312424130103	C_1
22	International Medical Council on Vaccination	121591387888250	C_1
23	International Protest Against Mandatory Vaccinations	185642288624228	C_1
24	Irish vaccination awareness	395440297214176	C_1
25	Jacob's Vaccine Story #jacobthebrave	1280103762035540	C_1
26	Kentucky Vaccine Rights Coalition	435134176638311	C_1
27	Light a Candle for the Vaccine-damaged	287010891636074	C_1
28	Louisiana Parents for Vaccine Rights	1650023171986221	C_1
29	Malaysian Vaccines Exposed	350017531795403	C_1
30	Michigan Vaccine Freedom PAC	448357925360257	C_1
31	Mississippi Parents for Vaccine Rights	141170989357307	C_1
32	My Vaccine Lawyer	662025280493214	C_1
33	National Vaccine Information Center	143745137930	C_1
34	National Vaccine Injury Attorney	1479038925701608	C_1
35	National Vaccine Injury Compensation Program - VICP Awareness	291826954275107	C_1
36	New York Alliance for Vaccine Rights	134385976923719	C_1
37	New York Coalition for Vaccination Choice	605350919532403	C_1
38	No Vaccines Australia	1419580101615748	C_1
39	North Carolina for Vaccine Rights	627567714037556	C_1
40	Ohio Vaccine Freedom	413575935476246	C_1
41	Oklahomans for Vaccine and Health Choice-PAC	1672900919606699	C_1
42	Oregonians for Vaccine Truth and Healthcare Choice	337316776389390	C_1
43	Over Vaccination Nation	1357818757569011	C_1
44	Parents Against Vaccines	143842215711000	C_1
45	Personal Stories of Vaccine Injuries	646458042147284	C_1

46	RAGE Against The Vaccines	341081422696435	C_1
47	Real Vaccine information	921875794578736	C_1
48	Reid Thomas Englehart : Death by routine Vaccinations	1522103818093606	C_1
49	Rethink Vaccines	136451986492761	C_1
50	RI against mandated HPV vaccine #NoHpv-mandateri	670134063087477	C_1
51	Sallie O. Elkordy for Mayor, Vaccine Free NYC	587347554614900	C_1
52	Tennessee Coalition for Vaccine Choice	1592571291060774	C_1
53	Texans For Vaccine Choice	1557722217817482	C_1
54	The Center for Vaccine Shoulder Pain Recovery	1724838497773162	C_1
55	The Dangers of Gardasil (HPV/Cervical Cancer Vaccine)	184833174871567	C_1
56	The Truth About Vaccines	133579170019140	C_1
57	The Truth About Vaccines Docu-Series	411516962540551	C_1
58	The Vaccine-Friendly Plan	1572605193045445	C_1
59	The Vaccine Machine	59188273283	C_1
60	Thinktwice Global Vaccine Institute	179833531264	C_1
61	Truth About Vaccines	168105386951066	C_1
62	UK Association of HPV Vaccine Injured Daughters	1637034089889113	C_1
63	Uniting Healthcare Professionals Against Mandatory Vaccination	405930626223420	C_1
64	V for Vaccines	719623851517180	C_1
65	Vaccination Choice Judy Wilyman PhD	530702127033360	C_1
66	Vaccination Information Network - UK (VINE UK)	242241312482649	C_1
67	Vaccination Information Network (VINE)	69667273997	C_1
68	Vaccination Lawyer	518992891532451	C_1
69	Vaccination Risk Awareness	294475453904145	C_1
70	Vaccine-Injury Survivors & Families Venting	626985363998850	C_1
71	Vaccine Awareness Movement	198813146959	C_1

72	Vaccine Awareness South Africa - VASA	398819000228205	C_1
73	Vaccine Choice Canada	330700720307290	C_1
74	Vaccine Corruption Today	463771403668839	C_1
75	Vaccine Dangers	246293915457536	C_1
76	Vaccine Epidemic	190754844273581	C_1
77	Vaccine Freedom for Washington State	916858375002999	C_1
78	Vaccine Freedom Utah	1748610515416671	C_1
79	Vaccine Injury Help Center	290768197788	C_1
80	Vaccine Injury Law Project	295977950440133	C_1
81	Vaccine Injury Lawyers - Maglio Christopher & Toale, P.A.	370902196382833	C_1
82	Vaccine Injury Network	496842373860237	C_1
83	Vaccine Injury Reports	464659256897483	C_1
84	Vaccine Injury Team at Rawls McNelis	1671142199839686	C_1
85	Vaccine Research Library	1430234513855060	C_1
86	Vaccine Resistance Movement	314338291923261	C_1
87	Vaccine Truth Education	938899099538600	C_1
88	VaccineImpact	783513531728629	C_1
89	Vaccines Cripple and Kill Children	1735864636640546	C_1
90	Vaccines News	477752302405379	C_1
91	Vaccines: Assault with a Deadly Weapon	697276820289153	C_1
92	VaccineSafety	366270656889324	C_1
93	VaccinesRevealed	534632916742245	C_1
94	Vaccinetruthmovement	1505157789775446	C_1
95	Vermont Coalition for Vaccine Choice	380959335251497	C_1
96	VISA - Vaccination Information Serving Australia	241297122629194	C_1
97	Voices of the Vaccine Injured	176383812524158	C_1
98	Watch The Truth About Vaccines Free	203576920131483	C_1
99	Advantage Vaccination Services	230405923638822	C_2
100	Affinity Vaccines	134177516651933	C_2
101	AIDS Vaccine 200	59663328616	C_2

102	Alabama Vaccine Research Clinic at UAB	243970235077	C_2
103	All Valley Vaccines	159956407691961	C_2
104	Amson Vaccines And Pharma	365860713456667	C_2
105	Anti-Cocaine Vaccine	1458856820804937	C_2
106	ARMM Diagnostic Laboratory And Medical Services	853297521370804	C_2
107	Baby Vaccination Center	1426130021049278	C_2
108	Bachpan Child Care Clinic & Vaccination Centre	1740923012845877	C_2
109	Bachpan Child Care Clinic & Vaccination Centre	507200072814690	C_2
110	Bharat Serums and Vaccines Limited	651335845030475	C_2
111	BHIMANI Children's Hospital & Vaccination Centre	204319143320631	C_2
112	Binag-Lee Medical and Vaccination Clinic	353329021717705	C_2
113	Branded and Generic Vaccine	171109169730168	C_2
114	CDG Gujranwala EPI Vaccination Program	201772836626875	C_2
115	Center for Vaccine Development, University of Maryland	439783256101612	C_2
116	Centre d'étude de vaccins du CUSM - MUHC Vaccine Study Centre	226688234021202	C_2
117	Child Care Clinic & Vaccination Centre	1702662259964466	C_2
118	Community Protest of Drs. Wolfson's Stance on Vaccines	295454707572711	C_2
119	Crohns MAP Vaccine	427832267319217	C_2
120	D.R Vaccination Clinic-Namugongo	375222335972654	C_2
121	Dendritic Cell & Vaccine Science	150811194977851	C_2
122	Doctors for Less Vaccines	559957857426165	C_2
123	Dubai International Conference on Infectious Diseases and Vaccination	1689648447961126	C_2
124	Duke Human Vaccine Institute	190286527664169	C_2
125	Family Vaccine and Specialty Clinics	389618337808965	C_2
126	FCN Medical & Vaccination Center	1423880664534477	C_2
127	Flu Vaccination Jammu	876673845773901	C_2

128	Flublok Influenza Vaccine	347171345372021	C_2
129	Flujabs.org - Flu Vaccinations From Fleet Street Clinic	157794600956499	C_2
130	G.M. Pharma & Vaccines House, Muzffarabad	349106601871044	C_2
131	GAIA Vaccine Foundation	438436612893386	C_2
132	Gavi, the Vaccine Alliance	92572781406	C_2
133	Get Vaccinated Ohio	159704247440990	C_2
134	Global Travel Vaccinations	297994225192	C_2
135	Global Travel Vaccines & Wellness	157842514740404	C_2
136	Global Vaccination for Cancer Prevention, Inc.	270433786405009	C_2
137	Global Vaccine House	1044323705627816	C_2
138	Global Vaccines Inc.	258657790890838	C_2
139	Go4Vaccine	1572360509755907	C_2
140	Hampstead Heath Pharmacy, Travel Health & Vaccination Clinic	391053871075535	C_2
141	HBL - IVC	1007278239336571	C_2
142	Health clinic	352476351547691	C_2
143	HealthSmartVaccines	316931354280	C_2
144	HIV Vaccine Trials Network	55636956839	C_2
145	Hope children's clinic and vaccination centre	854165088057428	C_2
146	Immunization Action Coalition	456742707709399	C_2
147	Imran Pediatric Clinic & Vaccination Centre	532530633473981	C_2
148	Infectious Diseases and Vaccines	447758915290250	C_2
149	International AIDS Vaccine Initiative	10288724013	C_2
150	International Vaccine Access Center	142642355753567	C_2
151	International Vaccine Institute	210349259126884	C_2
152	Jalali Vaccination Centre	458206047622381	C_2
153	Jenny McCarthy Body Count	180783538575	C_2
154	JustVaccines	379143562264657	C_2
155	Killol Children Hospital,NICU and Vaccination center,Surat	144805359009583	C_2

156	Krishnas Vaccine MART	218837158220715	C_2
157	Life Surgicals & Vaccines	958756494250383	C_2
158	Limerick Travel Vaccinations	1818386771760994	C_2
159	London Vaccination Clinic	724393447586258	C_2
160	Luton Travel Vaccination Clinic	1520772364884966	C_2
161	LYF Vaccination & Medical Services, Inc.	108847186008	C_2
162	M/s Mahaveer Vaccine and Drug Agency	954118494626992	C_2
163	Malaria Vaccine Project	1229016103861192	C_2
164	Marinus Vaccination Center	237829803007996	C_2
165	MASTA Travel Health Advice & Vaccinations	168005550009353	C_2
166	Measles and Rubella Vaccination Campaign	687320011474677	C_2
167	Meningitis b Vaccination	566137536858938	C_2
168	MNM Vaccines	1146934755429140	C_2
169	My Travel Vaccine	1481575865467341	C_2
170	MyKidsVaccines	112286665556521	C_2
171	NEOMAT Medical Clinic & Rabies Vaccination Center	453006521570475	C_2
172	New HIV Vaccine and Microbicide Advocacy Society (NHVMAS)	283240308415581	C_2
173	Northern BC Travel Health and Vaccination Clinic	191446951236279	C_2
174	Northern Rivers Vaccination Supporters	150057145203234	C_2
175	Nurses Who Vaccinate	201552619898653	C_2
176	Nursing Vaccination Specialists, Inc.	1437104503205464	C_2
177	On the Fence About Vaccines	171299949904818	C_2
178	Onsite Flu Vaccination Services	137272513280777	C_2
179	Orlando Vaccines	108701379190142	C_2
180	Oxford Vaccine Group, Department of Paediatrics, University of Oxford	923619751020270	C_2
181	Panipat Vaccines	845103652278786	C_2
182	Pravin Child Clinic & Vaccination Centre	428911333952700	C_2
183	Pro-Vaccine Shills for Big Pharma, the Illumanati, Reptilians, and the NWO	709431502441281	C_2

184	Pro Vaccinations	366571850109848	C_2
185	Protecting Children and Communities through Vaccination - Global Network	125879120944528	C_2
186	PROUDLY VACCINE	174408251772	C_2
187	Rainbow Children's Clinic and Vaccination Center,pune.	814036381979335	C_2
188	Refutations to Anti-Vaccine Memes	414643305272351	C_2
189	Regional Diagnostic & Vaccination Centre	349998218443004	C_2
190	Sabin Vaccine Institute	205183899497146	C_2
191	Saint Louis University Center for Vaccine Development	149667925136336	C_2
192	SF Screening and Vaccinations	1457694604246151	C_2
193	Shemitah Vaccines	459426124238786	C_2
194	South African Tuberculosis Vaccine Initiative	1435423513388186	C_2
195	Stop the Australian (Anti)Vaccination Network	143367983587	C_2
196	The Institute for Cancer Vaccines and Immunotherapy	415137961844924	C_2
197	The SEED Polyclinic	722882724450019	C_2
198	The Vaccine Center	233660419991678	C_2
199	The Vaccine Center-Chicago	910833702328754	C_2
200	The Vaccine Center and Travel Medicine Clinic	488044771271220	C_2
201	The Vaccine Clinic	475401869229486	C_2
202	The Vaccine Mom	509076135873314	C_2
203	The Vaccine Page	198633933629643	C_2
204	Travel Bug Vaccination Clinic	155296344547634	C_2
205	Travel Clinic at the Boardwalk	1694038320813654	C_2
206	Travel Medicine and Vaccination Center	133384863347657	C_2
207	Travel Vaccination Center	159396664125337	C_2
208	Travel Vaccination Center	188931388295859	C_2
209	Travel Vaccination Clinic, Sydney CBD	418419044979284	C_2
210	Travel Vaccines & Wellness Solutions	611012639104994	C_2

211	TravelDoc Vaccination Clinic	1641459086101561	C_2
212	Travellers Health & Vaccination Clinic	682764408455291	C_2
213	TravelMed Vaccine Centre	1779030195698389	C_2
214	Travelwise Vaccination Services	268614193177792	C_2
215	Tumor Vaccine Group (TVG) - University of Washington	178581138867172	C_2
216	UDontGetIt	774074382685932	C_2
217	University of Iowa Vaccine Research and Education Unit	413045925384178	C_2
218	University of Pittsburgh Center for Vaccine Research	269582359774038	C_2
219	UTMB Sealy Center for Vaccine Development - Clinical Trials	172917279424524	C_2
220	VacAware	1296039703814935	C_2
221	Vaccinations: The Truth	868034953238504	C_2
222	Vaccine Ambassadors	159021630913926	C_2
223	Vaccine care	414702652070437	C_2
224	Vaccine Centre at LSHTM	693339580689291	C_2
225	Vaccine Evaluation Center	165411290533657	C_2
226	Vaccine House	498992026843143	C_2
227	VACCINE INDIA	135487179800211	C_2
228	Vaccine Point	286646574723128	C_2
229	Vaccine Science and Information Archive	497150723772893	C_2
230	Vaccine Technology VI: Albufeira, Portugal, June 12-17, 2016	1682981045297046	C_2
231	Vaccines for Life	266912780340142	C_2
232	Vaccines for the Philippines	136534869735064	C_2
233	Vaccines for travels	1585819705036363	C_2
234	Vaccines From Anti-Vaxxers	942679192449055	C_2
235	Vaccines M.D.	410430106003373	C_2
236	Vaccines On Line	136962443308433	C_2
237	Vaccines Today	607391652607909	C_2
238	Vaccines411	109834275754225	C_2

239	Vanderbilt HIV Vaccine Trials Unit	311611039730	C_2
240	VBI Vaccines Inc.	807959829234756	C_2
241	Verwood Vaccination	1943400282611030	C_2
242	Voices for Vaccines	279714615481820	C_2
243	We Love GMOs and Vaccines	1380693538867364	C_2

References

[ABS+14] Ashley A Anderson, Dominique Brossard, Dietram A Scheufele, Michael A Xenos, and Peter Ladwig. The nasty effect: online incivility and risk perceptions of emerging technologies. *Journal of Computer-Mediated Communication*, 19(3):373–387, 2014. 7

[ACGC11] Jisun An, Meeyoung Cha, P Krishna Gummadi, and Jon Crowcroft. Media landscape in twitter: A world of new conventions and political diversity. In *ICWSM*, 2011. 2, 34

[AL17] Janna Anderson and Rainie Lee. The future of truth and misinformation online. *Washingtown, Pew Research Center*, 2017. 34

[Bar15] Fred Barbash. Disneyland measles outbreak strikes in antivaccination hotbed of california. *The Washington Post*, Jan 2015. 61, 64

[BB15] Cornelia Betsch and Robert Böhm. Detrimental effects of introducing partial compulsory vaccination: experimental evidence. *The European Journal of Public Health*, 26(3):378–381, 2015. 50

[BBB+12] Cornelia Betsch, Noel T Brewer, Pauline Brocard, Patrick Davies, Wolfgang Gaissmaier, Niels Haase, Julie Leask, Frank Renkewitz, Britta Renner, Valerie F Reyna, et al. Opportunities and challenges of web 2.0 for vaccination decisions. *Vaccine*, 30(25):3727–3733, 2012. 51

[BBL07] Jo Brown, Amanda J Broderick, and Nick Lee. Word of mouth communication within online communities: Conceptualizing the online social network. *Journal of interactive marketing*, 21(3):2–20, 2007. 5, 51

[BCD+15] Alessandro Bessi, Mauro Coletto, George Alexandru Davidescu, Antonio Scala, Guido Caldarelli, and Walter Quattrociocchi. Sci-

ence vs conspiracy: collective narratives in the age of misinformation. *PloS one*, 10(2):02, 2015. 9, 19, 35, 51, 56, 61

[BCDV+14] Alessandro Bessi, Guido Caldarelli, Michela Del Vicario, Antonio Scala, and Walter Quattrociocchi. Social determinants of content selection in the age of (mis) information. In *International Conference on Social Informatics*, pages 259–268. Springer, 2014. 9

[BDS07] Pablo J Boczkowski and Martin De Santos. When more media equals less news: Patterns of content homogenization in argentina's leading print and online newspapers. *Political Communication*, 24(2):167–180, 2007. 6

[BGLL08] Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment*, 2008(10):P10008, 2008. 13

[BMA15] Eytan Bakshy, Solomon Messing, and Lada A Adamic. Exposure to ideologically diverse news and opinion on facebook. *Science*, 348(6239):1130–1132, 2015. 9, 35

[Boc17] Elvira Bocci, Michele e Naselli. Morbillo, muore una bambina di 9 anni a roma. *La Repubblica*, Jun 2017. 50

[BPDV+15] Alessandro Bessi, Fabio Petroni, Michela Del Vicario, Fabiana Zollo, Aris Anagnostopoulos, Antonio Scala, Guido Caldarelli, and Walter Quattrociocchi. Viral misinformation: The role of homophily and polarization. In *Proceedings of the 24th International Conference on World Wide Web*, pages 355–356. ACM, 2015. 9, 19, 35

[BPDV+16] Alessandro Bessi, Fabio Petroni, Michela Del Vicario, Fabiana Zollo, Aris Anagnostopoulos, Antonio Scala, Guido Caldarelli, and Walter Quattrociocchi. Homophily and polarization in the age of misinformation. *The European Physical Journal Special Topics*, 225(10):2047–2059, 2016. 51

[BS64] Bernard Berelson and Gary A Steiner. Human behavior: An inventory of scientific findings. 1964. 7

[BSGM15] Michael Barthel, Elisa Shearer, Jeffrey Gottfried, and Amy Mitchell. The evolving role of news on twitter and facebook. *Pew Research Center*, 14, 2015. 1, 2, 34

[BZDV+15] Alessandro Bessi, Fabiana Zollo, Michela Del Vicario, Antonio Scala, Guido Caldarelli, and Walter Quattrociocchi. Trend of narratives in the age of misinformation. *PloS ONE*, 10(8):e0134641, 2015. 19, 51

[BZDV+17] Alessandro Bessi, Fabiana Zollo, Michela Del Vicario, Antonio Scala, Fabio Petroni, Bruno Gonçcalves, and Walter Quattrociocchi. Everyday the same picture: popularity and content diversity. In *Workshop on Complex Networks CompleNet*, pages 225–236. Springer, 2017. 2

[CLSF04] W Chen, S Landau, P Sham, and Eric Fombonne. No evidence for links between autism, mmr and measles virus. *Psychological medicine*, 34(3):543–553, 2004. 50

[CN06] Gabor Csardi and Tamas Nepusz. The igraph software package for complex network research. *InterJournal*, Complex Systems:1695, 2006. 12

[CNM04] Aaron Clauset, Mark EJ Newman, and Cristopher Moore. Finding community structure in very large networks. *Physical review E*, 70(6):066111, 2004. 12

[CSR+15] Giovanni Luca Ciampaglia, Prashant Shiralkar, Luis M Rocha, Johan Bollen, Filippo Menczer, and Alessandro Flammini. Computational fact checking from knowledge networks. *PloS one*, 10(6):e0128193, 2015. 9

[DeS07] Frank DeStefano. Vaccines and autism: evidence does not support a causal association. *Clinical Pharmacology & Therapeutics*, 82(6):756–759, 2007. 50

[DNAW00] Guillaume Deffuant, David Neau, Frederic Amblard, and Gérard Weisbuch. Mixing beliefs among interacting agents. *Advances in Complex Systems*, 3(01n04):87–98, 2000. 18, 31, 35, 46, 49

[DVBZ+16] Michela Del Vicario, Alessandro Bessi, Fabiana Zollo, Fabio Petroni, Antonio Scala, Guido Caldarelli, H Eugene Stanley, and Walter Quattrociocchi. The spreading of misinformation online. *Proceedings of the National Academy of Sciences*, 113(3):554–559, 2016. 9, 19, 35, 51

[DVGQ+17] Michela Del Vicario, Sabrina Gaito, Walter Quattrociocchi, Matteo Zignani, and Fabiana Zollo. Public discourse and news consumption on online social media: A quantitative, cross-platform analysis of the italian referendum. *arXiv preprint arXiv:1702.06016*, 2017. 9

[DVVB+16] Michela Del Vicario, Gianna Vivaldo, Alessandro Bessi, Fabiana Zollo, Antonio Scala, Guido Caldarelli, and Walter Quattrociocchi. Echo chambers: Emotional contagion and group polarization on facebook. *Scientific reports*, 6, 2016. 51

[DVZC+17] Michela Del Vicario, Fabiana Zollo, Guido Caldarelli, Antonio Scala, and Walter Quattrociocchi. Mapping social dynamics on facebook: The brexit debate. *Social Networks*, 50:6–16, 2017. 9

[Fac13] Facebook. Using the graph api. Website, 8 2013. last checked: 19.01.2014. 10, 20

[Fes62] Leon Festinger. *A theory of cognitive dissonance*, volume 2. Stanford university press, 1962. 8

[Fre86] Dieter Frey. Recent research on selective exposure to information. *Advances in experimental social psychology*, 19:41–80, 1986. 8

[Git02] Todd Gitlin. 10 public sphere or public sphericules? *Media, ritual and identity*, page 168, 2002. 8

[Gum15] Andrew Gumbel. Us states face fierce protests from anti-vaccine activists. *The Guardian*, Apr 2015. 61, 64

[Hab15] Clyde Haberman. A discredited vaccine studys continuing impact on public health. *The New York Times*, Feb 2015. 61, 64

[Her10] Alfred Hermida. Twittering the news: The emergence of ambient journalism. *Journalism practice*, 4(3):297–308, 2010. 6

[HFKL12] Alfred Hermida, Fred Fletcher, Darryl Korell, and Donna Logan. Share, like, recommend: Decoding the social media news consumer. *Journalism Studies*, 13(5-6):815–824, 2012. 1, 2

[HL11] Avery E Holton and Seth C Lewis. Journalists, social media, and the use of humor on twitter. *Electronic Journal of Communication*, 21(1/2), 2011. 6, 7

[HMKW14] Aniko Hannak, Drew Margolin, Brian Keegan, and Ingmar Weber. Get back! you don't know me like that: The social mediation of fact checking interventions in twitter conversations. In *ICWSM*, 2014. 9

[Hon12] Sounman Hong. Online news on twitter: Newspapers social media adoption and their online readership. *Information Economics and Policy*, 24(1):69–74, 2012. 6, 7

[HP15] Jack Healy and Michael Paulson. Vaccine critics turn defensive over measles. *The New York Times*, Jan 2015. 61, 64

[HSP+06] Hyunseo Hwang, Michael Schmierbach, Hye-Jin Paek, Homero Gil de Zuniga, and Dhavan Shah. Media dissociation, internet use, and antiwar political participation: A case study of political dissent and action against the war in iraq. *Mass Communication & Society*, 9(4):461–483, 2006. 7

[HT08] Alfred Hermida and Neil Thurman. A clash of cultures: The integration of user-generated content within professional journalistic frameworks at british newspaper websites. *Journalism practice*, 2(3):343–356, 2008. 5, 7

[IH09] Shanto Iyengar and Kyu S Hahn. Red media, blue media: Evidence of ideological selectivity in media use. *Journal of Communication*, 59(1):19–39, 2009. 8

[KK04] Richard Kahn and Douglas Kellner. New media and internet activism: from the battle of seattleto blogging. *New media & society*, 6(1):87–95, 2004. 5, 51

[KLPM10] Haewoon Kwak, Changhyun Lee, Hosung Park, and Sue Moon. What is twitter, a social network or a news media? In *Proceedings of the 19th international conference on World wide web*, pages 591–600. ACM, 2010. 7

[KMM10] Ravi Kumar, Mohammad Mahdian, and Mary McGlohon. Dynamics of conversations. In *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 553–562. ACM, 2010. 5, 51

[Kri15] Nicholas Kristof. The dangers of vaccine denial. *The New York Times*, Feb 2015. 61, 64

[Lim06] Jeongsub Lim. A cross-lagged analysis of agenda setting among online news media. *Journalism & Mass Communication Quarterly*, 83(2):298–312, 2006. 6

[LLH12] Dominic L Lasorsa, Seth C Lewis, and Avery E Holton. Normalizing twitter: Journalism practice in an emerging communication space. *Journalism studies*, 13(1):19–36, 2012. 5, 6, 7

[Lor07] Jan Lorenz. Continuous opinion dynamics under bounded confidence: A survey. *International Journal of Modern Physics C*, 18(12):1819–1838, 2007. 18

[Lor17] Beatrice Lorenzin. L'italia del "free vaxx" in piazza: da cagliari a torino si manifesta per chiedere libert di scelta. *La Repubblica*, Jun 2017. 50

[Mer09] Sharon Meraz. Is there an elite hold? traditional media to social media agenda setting influence in blog networks. *Journal of Computer-Mediated Communication*, 14(3):682–707, 2009. 6

[MF09] W Glynn Mangold and David J Faulds. Social media: The new hybrid element of the promotion mix. *Business horizons*, 52(4):357–365, 2009. 5

[MHP13] Amy Mitchell, Jesse Holcomb, and Dana Page. News use across social media platforms. *Washingtown, Pew Research Center*, 2013. 1, 34

[MRZ+15] Delia Mocanu, Luca Rossi, Qian Zhang, Marton Karsai, and Walter Quattrociocchi. Collective attention in the age of (mis) information. *Computers in Human Behavior*, 51:1198–1204, 2015. 2, 9, 19, 35, 51

[MS72] Maxwell E McCombs and Donald L Shaw. The agenda-setting function of mass media. *Public opinion quarterly*, 36(2):176–187, 1972. 1, 5

[Mut01] Diana C Mutz. Facilitating communication across lines of political difference: The role of mass media. *American Political Science Review*, 95(1):97–114, 2001. 7, 8

[Mut06] Diana C Mutz. *Hearing the other side: Deliberative versus participatory democracy*. Cambridge University Press, 2006. 7, 8

[NFK+17] Nic Newman, Richard Fletcher, Antonis Kalogeropoulos, David AL Levy, and Rasmus Kleis Nielsen. Reuters institute digital news report 2017. *Reuters Institute for the Study of Journalism*, 2017. x, 34, 36

[NG04] Mark EJ Newman and Michelle Girvan. Finding and evaluating community structure in networks. *Physical review E*, 69(2):026113, 2004. 13

[Nic98] Raymond S Nickerson. Confirmation bias: A ubiquitous phenomenon in many guises. *Review of general psychology*, 2(2):175, 1998. 8

[NLN15] Nic Newman, David AL Levy, and Rasmus Kleis Nielsen. Reuters institute digital news report 2015. *Reuters Institute for the Study of Journalism*, 2015. 7, 11, 19, 34, 35

[NLN16] Nic Newman, David Levy, and Rasmus Kleis Nielsen. Reuters institute digital news report 2016. *Reuters Institute for the Study of Journalism*, 2016. 11, 34, 35

[oP13] American Academy of Pediatrics. Vaccine safety: Examine the evidence. *American Academy of Pediatrics*, 2013. 50

[Pea15] Brendan Pease. Fear vs. fact: The modern anti-vaccination movement. *Harvard Science Review*, Mar 2015. 61, 64

[PL06] Pascal Pons and Matthieu Latapy. Computing communities in large networks using random walks. *J. Graph Algorithms Appl.*, 10(2):191–218, 2006. 13

[QCS14] Walter Quattrociocchi, Guido Caldarelli, and Antonio Scala. Opinion dynamics on interacting networks: media competition and social influence. *Scientific reports*, 4:4938, 2014. 1, 5, 51

[QSS16] Walter Quattrociocchi, Antonio Scala, and Cass R Sunstein. Echo chambers on facebook. *Available at SSRN*, 2016. 8, 19, 35, 51

[RAK07] Usha Nandini Raghavan, Réka Albert, and Soundar Kumara. Near linear time algorithm to detect community structures in large-scale networks. *Physical review E*, 76(3):036106, 2007. 13

[Ran71] William M Rand. Objective criteria for the evaluation of clustering methods. *Journal of the American Statistical association*, 66(336):846–850, 1971. 13, 25, 42

[RB06] Jrg Reichardt and Stefan Bornholdt. Statistical mechanics of community detection. *Physical Review E*, 74(1):016110, 2006. 13

[Sal15] Steven Salzberg. Anti-vaccine movement causes worst measles epidemic in 20 years. *Forbes*, Feb 2015. 61, 64

[SBS+10] Ben Sayre, Leticia Bode, Dhavan Shah, Dave Wilcox, and Chirag Shah. Agenda setting in a digital age: Tracking attention to california proposition 8 in social media, online news and conventional news. *Policy & Internet*, 2(2):7–32, 2010. 6

[SBV09] M. Angeles Serrano, Marian Boguna, and Alessandro Vespignani. Extracting the multiscale backbone of complex weighted networks. *Proceedings of the National Academy of Sciences*, 106(16):6483–6488, 2009. 26, 30

[SdWL05] Klaus Schoenbach, Ester de Waal, and Edmund Lauf. Research note: Online and print newspapers: Their impact on the extent of the perceived public agenda. *European Journal of Communication*, 20(2):245–258, 2005. 7, 35

[SF67] David O Sears and Jonathan L Freedman. Selective exposure to information: A critical review. *Public Opinion Quarterly*, 31(2):194–213, 1967. 7

[SGATM+13] Harald Schoen, Daniel Gayo-Avello, Panagiotis Takis Metaxas, Eni Mustafaraj, Markus Strohmaier, and Peter Gloor. The power of prediction with social media. *Internet Research*, 23(5):528–543, 2013. 9

[SPVDG13] Ralf Steinberger, Bruno Pouliquen, and Erik Van Der Goot. An introduction to the europe media monitor family of applications. *arXiv preprint arXiv:1309.5290*, 2013. 20

[ST07] Dietram A Scheufele and David Tewksbury. Framing, agenda setting and priming: The evolution of three media effect models. *Journal of Communication*, 57(1):9–20, 2007. 6

[Str08] Natalie Jomini Stroud. Media use and political predispositions: Revisiting the concept of selective exposure. *Political Behavior*, 30(3):341–366, 2008. 8

[Sun02a] Cass R Sunstein. The law of group polarization. *Journal of political philosophy*, 10(2):175–195, 2002. 9, 35

[Sun02b] Cass R Sunstein. *Republic. com*. Princeton University Press, 2002. 8

[SZDV+17] Ana Lucía Schmidt, Fabiana Zollo, Michela Del Vicario, Alessandro Bessi, Antonio Scala, Guido Caldarelli, H Eugene Stanley, and Walter Quattrociocchi. Anatomy of news consumption on facebook. *Proceedings of the National Academy of Sciences*, page 201617052, 2017. xii, 4, 19, 40, 51, 56, 57

[SZS+18] Ana Lucia Schmidt, Fabiana Zollo, Antonio Scala, Cornelia Betsch, and Walter Quattrociocchi. Polarization of the vaccination debate on facebook. *arXiv preprint arXiv:1801.02903*, 2018. xii, 4, 50

[SZSQ17] Ana Lucía Schmidt, Fabiana Zollo, Antonio Scala, and Walter Quattrociocchi. A study on european news consumption on facebook. 2017. xii, 4, 34

[TB09] Vincent A Traag and Jeroen Bruggeman. Community detection in networks with positive and negative links. *Physical Review E*, 80(3):036115, 2009. 13

[WAL+14] You Wu, Pankaj K Agarwal, Chengkai Li, Jun Yang, and Cong Yu. Toward computational fact-checking. *Proceedings of the VLDB Endowment*, 7(7):589–600, 2014. 9

[YB10] Sarita Yardi and Danah Boyd. Dynamic debates: An analysis of group polarization over time on twitter. *Bulletin of Science, Technology & Society*, 30(5):316–327, 2010. 9, 35

[ZBDV+15] Fabiana Zollo, Alessandro Bessi, Michela Del Vicario, Antonio Scala, Guido Caldarelli, Louis Shekhtman, Shlomo Havlin, and Walter Quattrociocchi. Debunking in a world of tribes. *arXiv preprint arXiv:1510.04267*, 2015. 9, 19, 35, 51

[ZNDV+15] Fabiana Zollo, Petra Kralj Novak, Michela Del Vicario, Alessandro Bessi, Igor Mozetič, Antonio Scala, Guido Caldarelli, and Walter Quattrociocchi. Emotional dynamics in the age of misinformation. *PloS one*, 10(9):e0138740, 2015. 9, 35

[ZWH+15] Jennifer Zipprich, Kathleen Winter, Jill Hacker, Dongxiang Xia, James Watt, Kathleen Harriman, Centers for Disease Control, Prevention (CDC), et al. Measles outbreakcalifornia, december 2014-february 2015. *MMWR Morb Mortal Wkly Rep*, 64(6):153–154, 2015. 50, 61, 64