EUMEPLAT
European Media Platforms:
assessing positive and negative
externalities for European culture

Deliverable 2.5

Anti-European Fake News and What to Do

Document information

Grant Agreement #:	**101004488**
Project Title:	**EUROPEAN MEDIA PLATFORMS: ASSESSING POSITIVE AND NEGATIVE EXTERNALITIES FOR EUROPEAN CULTURE**
Project Acronym:	**EUMEPLAT**
Project Start Date:	**01/03/2021**
Related work package:	WP2 – Fake news: Platformization of Journalism
Related task(s):	T2.4 – Anti-European fake news and what to do
Lead Organisation:	P10 – UNIVE
Author(s):	Alessandro Galeazzi, UNIVE Fabiana Zollo, UNIVE
Status	Final
Submission date:	18/11/2022
Dissemination Level:	Public

History

Date	Submitted by	Reviewed by	Version
29/10/2022	P10-UNIVE	Steering Committee	V1.0[Rev. 0]
09/11/2022	P10-UNIVE	Partnership	V1.1[Rev. 1]
18/11/2022	P10-UNIVE	Approval	Final, V1.2

Table of Contents

1. Introduction

The long-standing celebration of the Internet as a great tool for sharing information, disseminating knowledge, and promoting freedom, has more recently given way to a growing concern for misinformation spreading and polarisation phenomena. The Internet has totally transformed the information space, coupling traditional media with a heterogenous mass of news sources, often born as an alternative to mainstream outlets. The advent of social media took the information ecosystem to a whole new level, allowing users to play an active role not only in the consumption, but also in the production of content. The increasing popularity of social media platforms made them rapidly become the main information source for many of their users, who now prefer to access news through social media, search engines, or news aggregators. Moreover, despite the increasing quantity of content, quality may be poor, for issues ranging from content monetization to the persisting reduction of investments in the news production and distribution. This context has contributed to the loss of reputation and trust for traditional media, encouraging people to rely on alternative information sources, not always qualified. The spreading of unreliable information may influence public opinion and impact people's behaviour and choices, raising important concerns about the consequences of misinformation. Socio-cognitive factors play a fundamental role in how one selects and processes information, and individual differences are key in the evaluation of misinformation (Marie et al., 2020; Schaewitz et al., 2020). Confirmation bias is the human tendency to look for, select, and interpret information in a way that is coherent to one's system of beliefs; it can be crucial in information consumption patterns and the formation of the so-called echo chambers. Here, the term 'echo chambers' denotes groups of like-minded individuals who select information based on their pre-existing beliefs and tend to ignore opposing views. Echo chambers are nothing new and their impact on opinion formation is clear (Sunstein, 2017). The existence of echo chambers on social media is supported by scientific evidence; many researchers have investigated their effect in shaping public opinion and fostering polarisation and radicalisation phenomena (Cinelli et al., 2021; Joint Research Centre et al., 2020). The understanding of these dynamics is crucial to develop appropriate counterstrategies, especially in the case of misinformation and controversial issues.

Here, we quantitatively analyse news production and consumption of more than 25M tweets on European issues in the years 2019–2021 to identify the most debated topics and their engagement over time, as well as the presence of misinformation. To shed light on anti-EU misinformation, we use the case study of Brexit and present a technique to capture the polarisation of the public debate around controversial topics, studying the interaction patterns of more than 290k users. These results offer a

valuable contribution to the literature for the timely identification of potential misinformation targets and the development of appropriate and effective countermeasures.

2. Misinformation: definitions and classification

The colloquial term "fake news" is widely used in the public debate, however its definition is vague and blurred. The term was coined to describe the spreading of false information, although the concept refers to at least two different notions, i.e., misinformation and disinformation. Misinformation refers to unintentionally spreading unreliable or misleading information (Schaewitz et al., 2020), and may generally refer to any content that is 'initially processed as valid but that is subsequently retracted or corrected' (Lewandowsky et al., 2012). Disinformation is 'the deliberate creation and dissemination of false and/or manipulated information that is intended to deceive and mislead audiences, either for the purposes of causing harm, or for political, personal or financial gain' (GCS, 2020). This distinction is crucial for the definition of effective countermeasures. Nevertheless, in many cases it is not trivial to discriminate between the two cases: the two forms co-occur, and the intentionality is hard to be proved. For these reasons, here we use a definition of misinformation that is able to capture a variety of shades in the quality of information, that goes beyond the mere reliability of the content. To this aim, we rely on NewsGuard1, a tool that provides trust ratings for news and information websites. NewsGuard assesses the credibility of a news domain based on nine journalistic criteria:

1. Does not repeatedly publish false content

2. Gathers and presents information responsibly

3. Regularly corrects or clarifies errors

4. Handles the difference between news and opinion responsibly

5. Avoids deceptive headlines

6. Website discloses ownership and financing

7. Clearly labels advertising

8. Reveals who's in charge, including possible conflicts of interest

9. The site provides the names of content creators, along with either contact or biographical information

---

[1] https://www.NewsGuardtech.com

These criteria are individually evaluated and then combined to produce a single ``trust score'' from 0 to 100 for a given news media domain. Scores are assigned by a team of professional journalists. Scores are not given to platforms (e.g., Twitter, Facebook), individuals, or satire content. Using the news media trust scores, we can cross reference the domains found in individual tweets (see Section 3) with the corresponding trust score from the NewsGuard database. If the score is greater than or equal to 60, the source is considered by NewsGuard Reliable, otherwise, it is labelled as Questionable. Furthermore, a big advantage of NewsGuard ratings is that they are available for three EU countries (Italy, France, Germany) and the UK. This enables us to distinguish between Questionable and Reliable content through a third-party independent classification and to analyse misinformation in different EU countries using the same assessment criteria. Although the UK withdrew from the European Union in 2020, it had been a member state for over 40 years. Moreover, our analysis has its starting point in 2019, before the Covid-19 pandemic. Therefore, we deemed it appropriate to include the UK in the study.

3.   Data collection and processing

We collected Twitter data from news sources listed in the NewsGuard dataset, which includes 874 sources (see Table 1).

Country	Reliable sources	Questionable sources	Total
France	187	49	236
Germany	196	25	221
Italy	175	29	204
UK	191	22	213
Total	749	125	874

Table 1. Breakdown of the news sources per country and reliability

The process of data collection was performed exclusively using the Twitter API for academic research[2]. We used the list of unique usernames from the NewsGuard dataset to extract the Twitter timelines of media sources in Italy, Germany, France, and the UK. All gathered data is publicly available and data from

---

private accounts is not included in our dataset. The dataset includes all the tweets published on the selected accounts in the period from 01 January 2019 to 11 November 2021, resulting in 25+ Million tweets. Table 2 reports the breakdown of the data. The percentage of posts retained after filtering is shown in parenthesis.

Country	Number of tweets	Reliable tweets	Questionable tweets
France	7,083,659 (28.19%)	6,151,554 (26.57%)	932,105 (47.32%)
Germany	4,904,179 (19.52%)	4,689,186 (20.25%)	214,993 (10.91%)
Italy	4,936,407 (19.65%)	4,528,606 (19.56%)	407,801 (20.70%)
UK	8,201,352 (32.64%)	7,786,239 (33.62%)	415,113 (21.07%)
Total	25,125,597	23,155,585	1,970,012

Table 2. Breakdown of the data after filtering for EU issues.

4. The public debate on European issues

In this section, we provide a quantitative overview of the public debate on European issues over the last years, from 2019 to 2021. We analyse news content production and engagement over time, considering the number of tweets produced and the users interacting with content. Moreover, we study the evolution of the topics debated over time in the selected countries. The entire analysis also considers misinformation in terms of source reliability, as described in Section 2.

4.1 Content production and engagement

Figure 1 reports, for each country, the news production in terms of tweets over time. To facilitate a fair comparison across countries, we show the relative number of tweets with respect to the number of news sources present in the country.   We may notice that UK sources produce more content than their counterparts in other countries for Reliable sources, while France has the highest volume of questionable tweets, as reported in Table 1. However, when we take into account the number of sources active in each nation (figure 1), we find that the UK has the most active accounts for both categories until September 2020. After that date, Italy and then France produce a similar amount of Questionable content per news outlet with respect to the UK, with France dominating from March to September 2021. A clear increase in content production can be observed for France, Italy, and Germany towards the end of 2019, with peaks in correspondence of March 2020, when the Covid-19 pandemic started. The UK instead shows a later

increase for Questionable sources and a quite stable trend for reliable sources. Finally, towards the end of 2020 and during 2021, content production has been decreasing gradually for all countries and both categories. Figures 2 and 3 show, respectively, the Complementary Cumulative Distribution Functions (CCDFs) of the number of Likes, Quotes, Replies, and Retweets received by tweets of Reliable and Questionable sources. We see that Likes and Retweets are the top two types of interactions on tweets of both categories of sources, with Quotes being at the bottom. The overall engagement for Questionable sources of all metrics does not display significant differences across different countries. We see that sources in Italy and France are more likely to engage their users in discussions over their posts through Replies. Moreover, tweets posted by Questionable sources in France are more likely to be retweeted than the rest of the countries. Also, Questionable tweets in Germany are likely to receive lower interaction than in all other countries. Reliable sources in the UK are most likely to engage their users, followed by France.
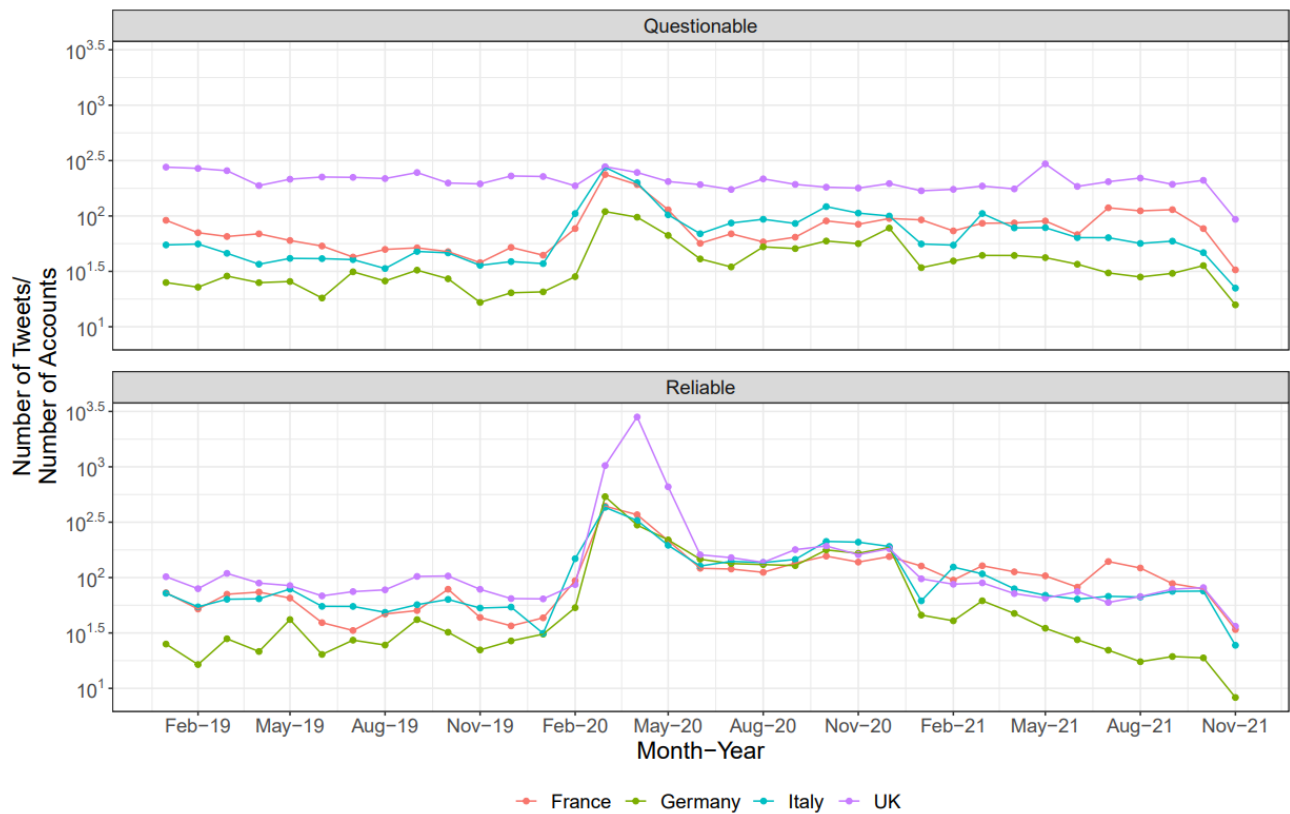

Figure 1. News production over time. For each month, we show the total number of tweets produced in each country divided by the total number of news sources in that country.
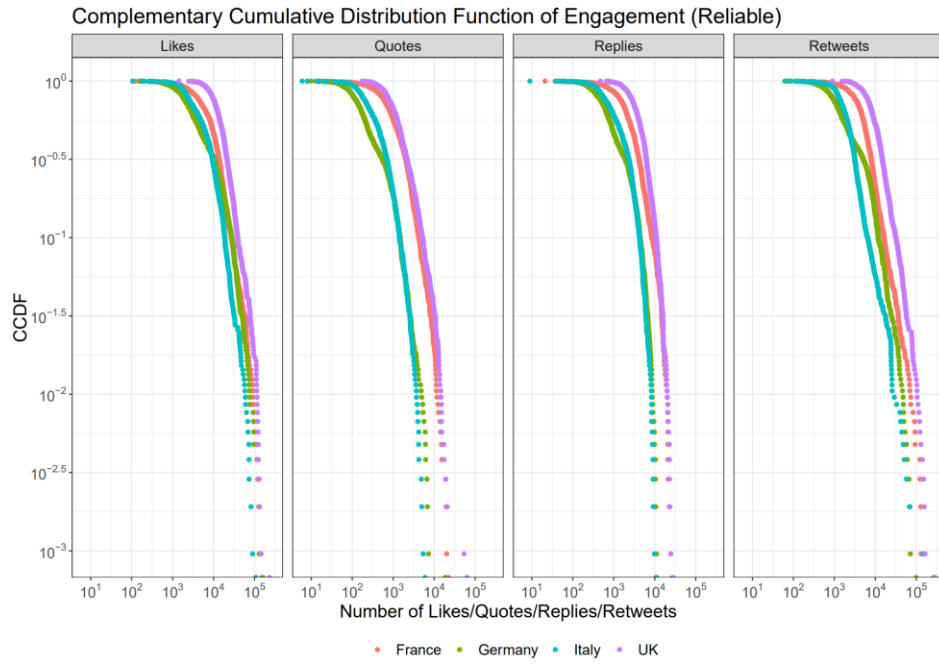
Figure 2. Complementary Cumulative Distribution Functions (CCDFs) of the number of Likes, Quotes, Replies, and Retweets received by tweets of Reliable sources.
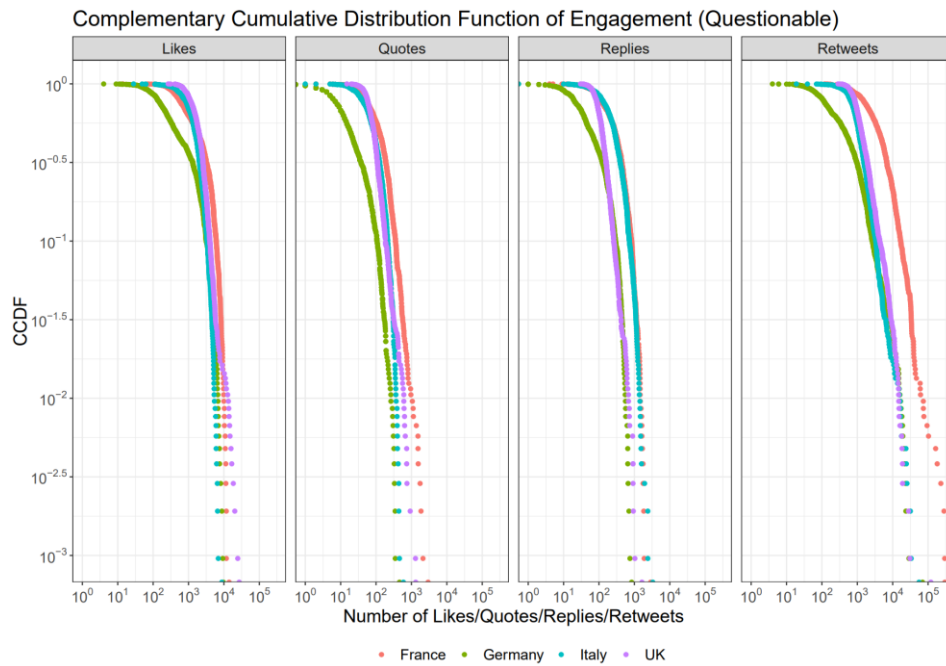

Figure 3. Complementary Cumulative Distribution Functions (CCDFs) of the number of Likes, Quotes, Replies, and Retweets received by tweets of Questionable sources.

4.2 Main topics of discussion and evolution over time

To identify the topics emerging in the public debate around European issues we use BERTopic (Grootendorst, 2022), a novel topic modelling tool that extracts latent topics from a collection of documents and is well-suited to analysing Twitter data. The base algorithm uses pre-trained transformer-based language models to build document embeddings and produces topic representations by clustering embeddings and applying a class-based TF-IDF procedure (Sammut and Webb, 2010).

In the following, we show the most debated topics found in each country, along with the relative engagement (i.e., sum of all the reactions divided by the number of tweets) received by the tweets belonging to each topic.

Figure 4 shows the most debated topics in France and the engagement received in terms of Likes, Replies, Quotes, and Retweets. The discussion in 2019 is dominated by the topics "EU - Emmanuel Macron" and "Brexit" in Reliable sources, whereas the former along with "Russie - Europe" in Questionable sources. Unsurprisingly, the discussion in both 2020 and 2021 is dominated by topics like "Coronavirus" and "Coronavirus vaccin" in both types of sources. "Coronavirus" remains the second most debated topic by Questionable sources in 2021 as well. In terms of engagement, "EU - Emmanuel Macron" in Reliable sources received significantly more interactions than in Questionable sources.
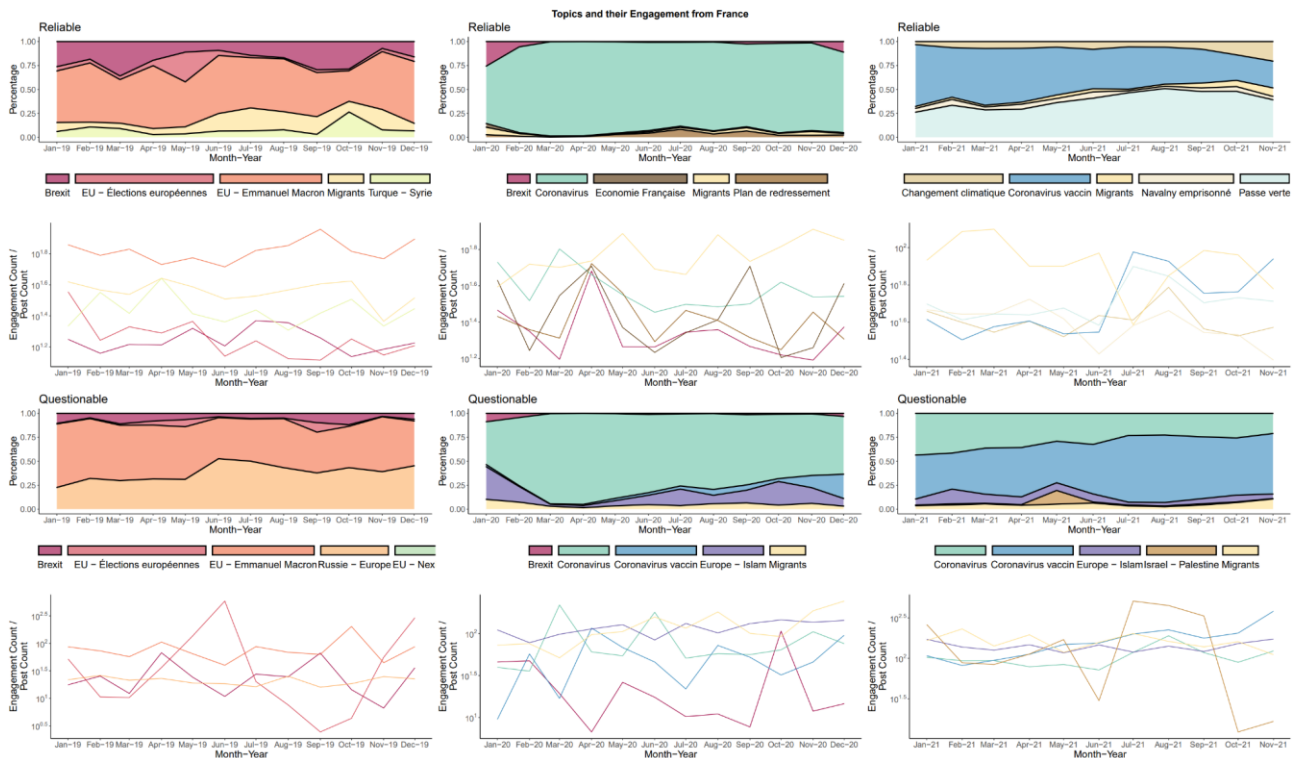

Figure 4. Most debated topics and their engagement in France.

Similarly, Figure 5 shows the most debated topics in Germany and the engagement received in terms of Likes, Replies, Quotes, and Retweets. "Klimawandel" (climate change) along with "Brexit" are the two most debated topics in 2019 in Reliable sources, whereas "Russland - Europa" is dominating in Questionable sources in 2019 and 2021. We can also observe a clear dominance of "Coronavirus" and "Covid Impfung" (covid vaccine), respectively, in 2020 and 2021 for both categories of sources. The presence of the topic "EU-Kommission" is consistent in Reliable sources in all years and is present in the first two years in Questionable sources as well. As in France, "Coronavirus" clearly dominated the rest of the topics in 2020. We may also notice that the discussion on "Brexit" reduced significantly towards the end of 2019 and was overshadowed by other topics in the following years. As for the engagement, there is no clear trend with fluctuating spikes over the years, except in 2019 for the topic "Russland - Europa" in Questionable sources.
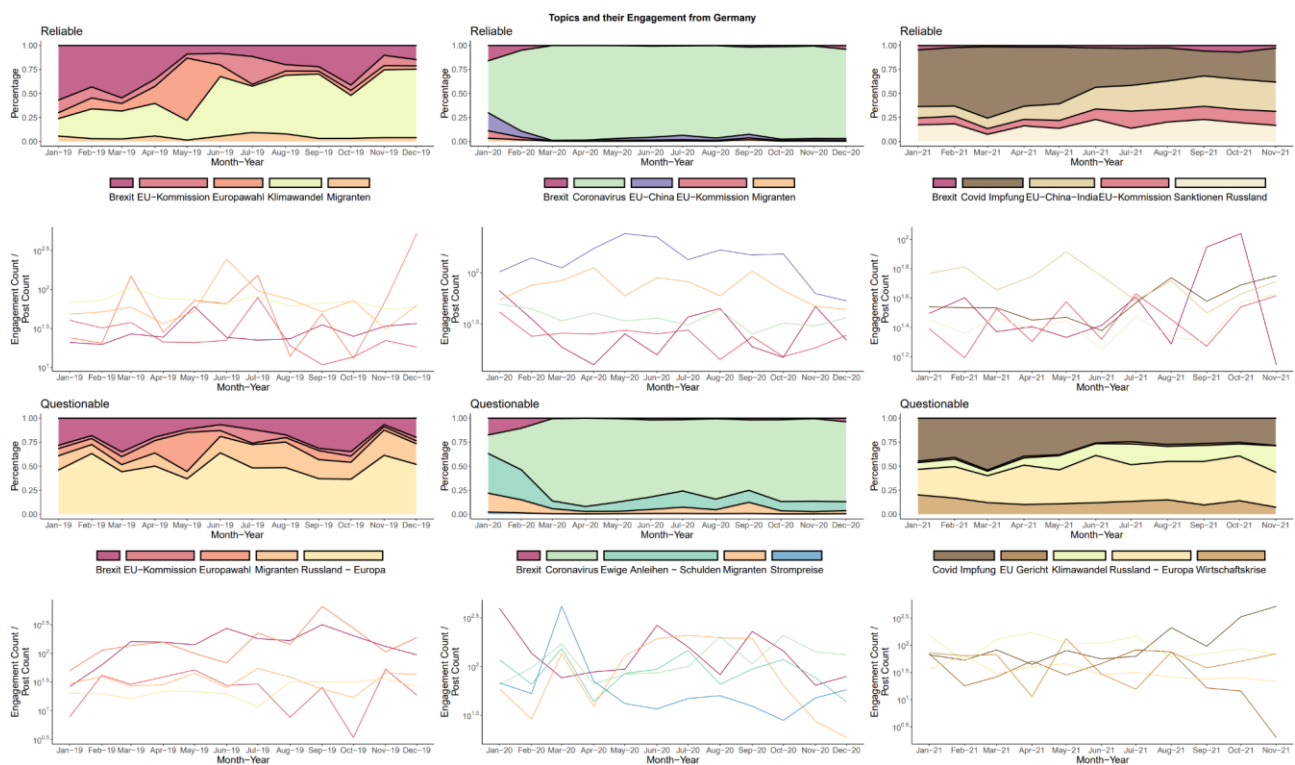

Figure 5. Most debated topics and their engagement in Germany.

Analogously, Figure 6 shows the most debated topics in Italy and the engagement received in terms of Likes, Replies, Quotes, and Retweets. We see the clear dominance of some topics such as "Milioni Euro", "Coronavirus" and "Vaccine covid" for Reliable sources in 2019, 2020, and 2021 respectively. "Coronavirus" and "Vaccine covid" are clearly dominant in the same years in Questionable sources as

well. However, we see "Migranti" as another dominating topic in 2019, sharing almost identical volume of tweets with "Milioni Euro". It can be noted that "Migranti" also emerged as the second most debated topic for Reliable sources in the same year and received a significantly higher engagement as compared to the rest of the topics until September 2019. However, in Questionable sources the topic was overshadowed in the following years by other issues. Interestingly, the topic "Green pass" emerged in Reliables sources towards the end of 2021.
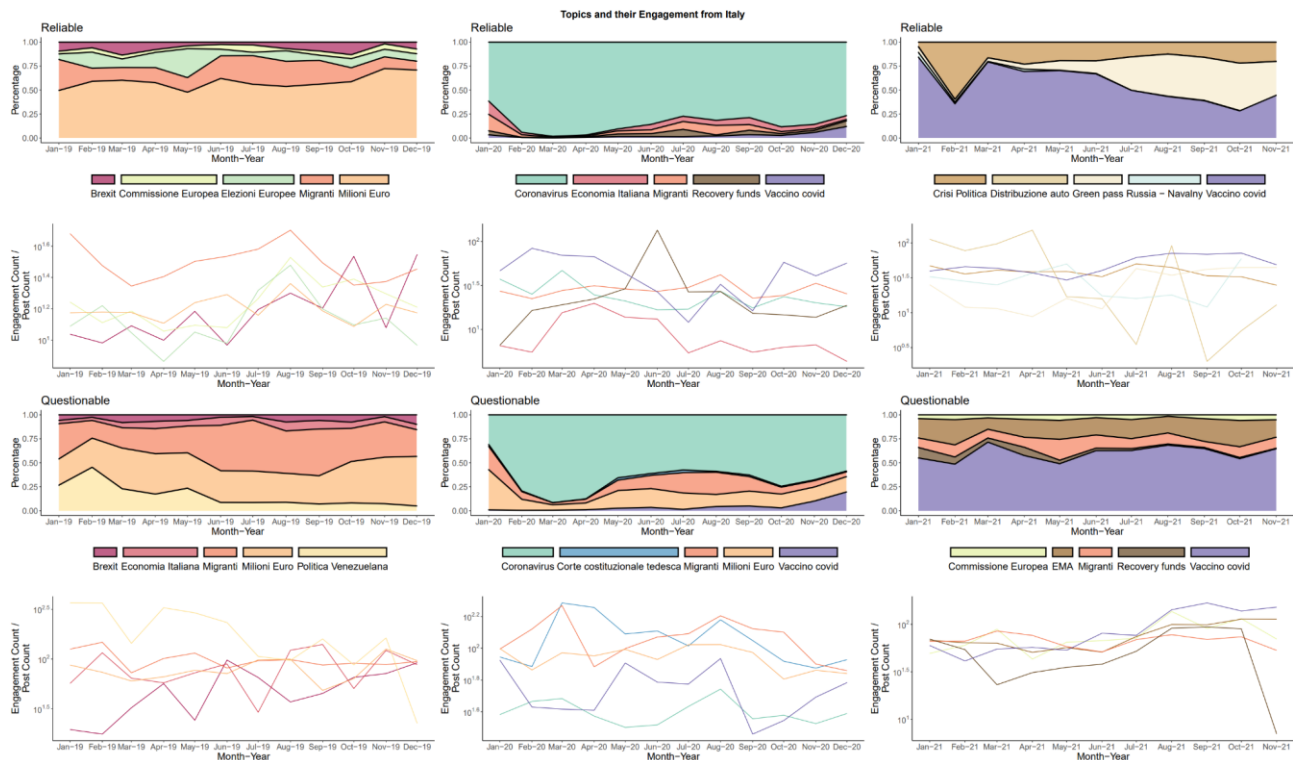

Figure 6. Most debated topics and their engagement in Italy.

Finally, Figure 7 shows the most debated topics in the UK and the engagement received in terms of Likes, Replies, Quotes, and Retweets. "Brexit" started as the most dominating topic in Reliable sources in 2019, and had its presence overshadowed by "Coronavirus" in the following year, which was clearly the most dominating topic. In 2021, there is no clear dominating issue, and the debate is focused on topics like "Brexit", "Climate Change", "Covid vaccine", "Nicola Sturgeon - Scotland", and "Refugees". "Brexit" is present in all three years of Questionable sources as well, although its presence decreases after 2019. Two topics are clearly dominant in Questionable sources in the UK in all years i.e., "Israel - Palestine" and "Sanctions on Iran", except for "Coronavirus" in 2020. In terms of engagement, we see that "Brexit" is the most engaging topic in 2019 and 2020 in Reliable sources. The same holds for Questionable sources

in 2019. Overall, discussions about "Covid vaccine" received high engagement, especially in Questionable sources.
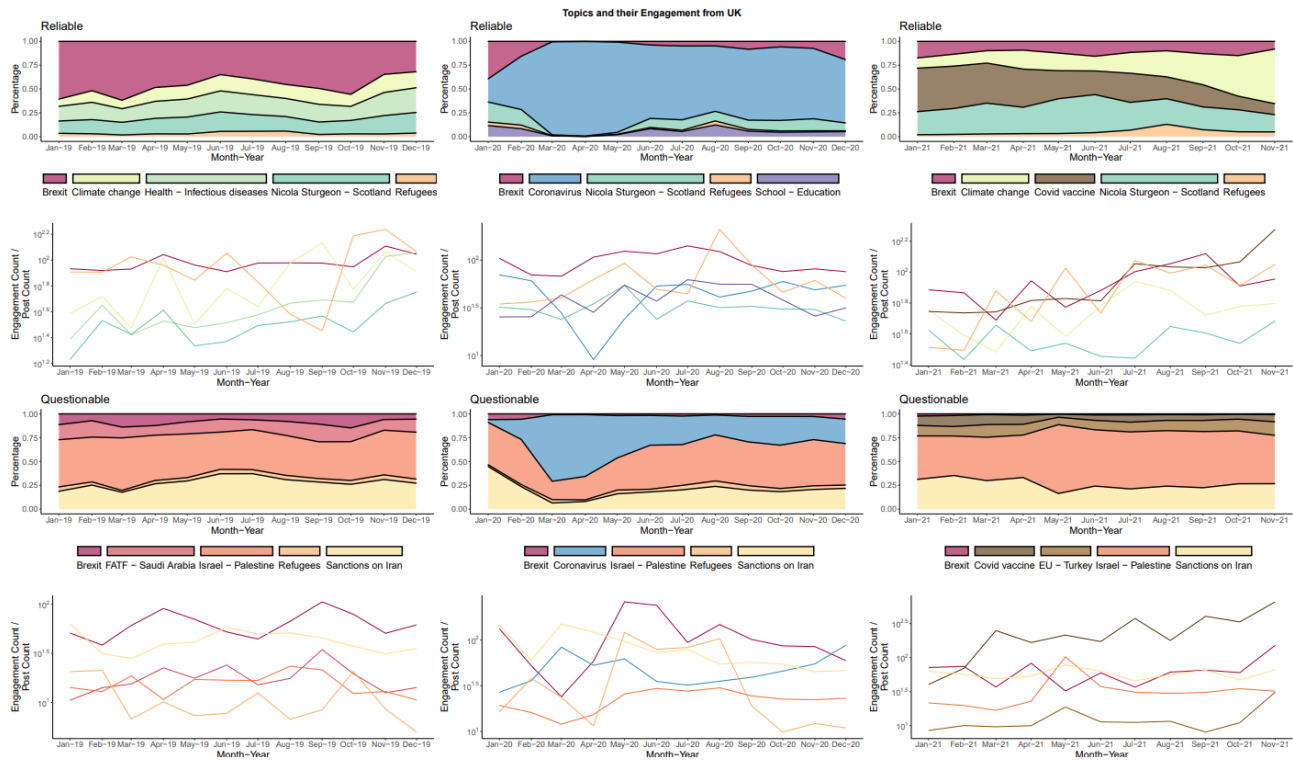

Figure 7. Most debated topics and their engagement in the UK.

Next, we analyse how the news sources in both categories shift across the most debated topics in each country over time. In Figure 8, the flow between two bars represents the proportion of news outlets moving among topics over years and the bars are proportional to the fraction of news outlets that produced content (i.e., tweets) about that topic. Flows connecting two bars of the same colour indicate the persistence of news sources in publishing content about that topic, while flows connecting two different colours indicate a shift in the news sources' focus. We notice some differences in the topics being debated, however issues that attracted high attention, such as Brexit, Coronavirus, and Covid Vaccine, respectively, in 2019, 2020, and 2021 are common to all countries. We may also observe the presence of some topics highly debated by only Questionable or Reliable news outlets, to the extent that they do not appear in all graphs for the same country, e.g., "EMA" in Italy. The latter is likely to be related

to the vaccine suspension performed by the European Medicines Agency which had a relevant impact on the narrative of Questionable news outlets pushing anti-vaccines views.
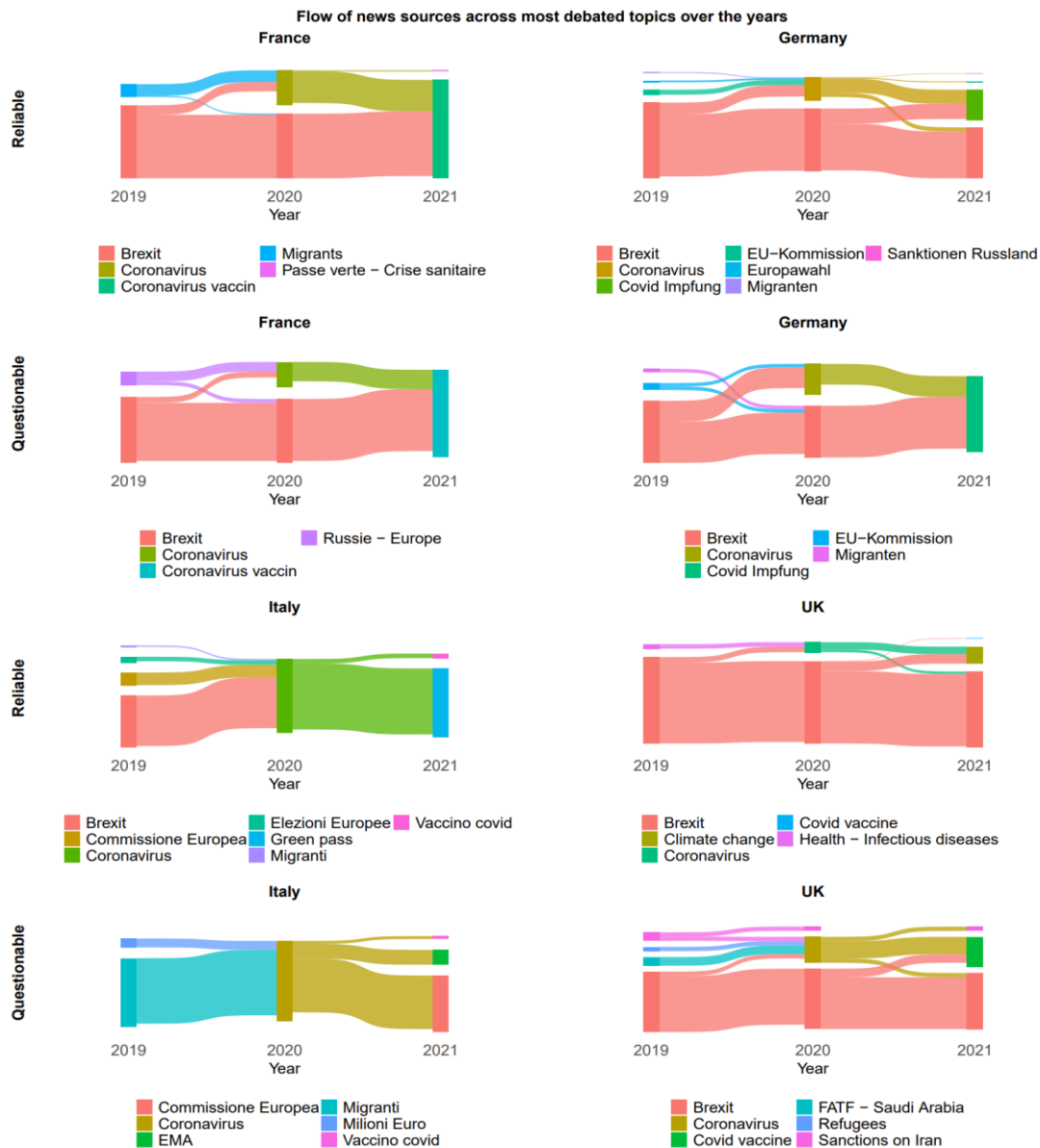

Figure 8. Change in news sources' publishing behaviour on the most debated topics over the years.

5. Anti-EU Misinformation: A case study

So far, we have studied the evolution of news production in four EU countries, highlighting the differences and similarities in the topics discussed by Questionable and Reliable news outlets of different countries. "Brexit", among all topics, fostered significant engagement and news production in all

countries during 2019, attracting the attention of both Reliable and Questionable news outlets as well as many politicians. Moreover, given the political background of the debate, the ideological stances are easily interpretable as "europeist" or "pro-EU" and "eurosceptics" or "anti-EU".

For these reasons, we deemed the debate around this controversial topic to be well-suited to study anti-EU misinformation. Thus, we exploit the interaction data which we collected about "Brexit" to reveal the presence of "anti-EU" news sources in the debate and then identify the ones more likely to spread misinformation. As discussed in Section 1, around debated topics users tend to cluster into ideologically homogeneous groups and get informed from similar information sources. Based on this assumption, we exploit the retweet information and use a latent ideology estimation technique to infer the opinion of the users in the Brexit debate. We validate our analysis by looking at the behaviour of politicians, who have an explicit opinion about Europe and europeism. The workflow of the analysis is illustrated in Figure 9.

First, we consider all the tweets that contain the keyword "Brexit" produced by the news outlets in our list. We filter this dataset keeping only original tweets that have been retweeted at least twice and download the list of all the retweeters (see Table 3 for more details on the data). Then, for each country, we create the matrix of interaction R, where rows represent users and columns represent news sources. The entry $r_{i,j}$ is equal to the number of times the user i retweeted content created by source j. Finally, we apply the latent ideology estimation technique (Barberá, 2015) to infer the opinion of users and news sources, which we refer to as "synthetic opinion". Intuitively, the latent ideology estimation technique decomposes the interaction matrix to assign a real value to each user. Two users that retweeted a similar set of news sources will have similar values, while two users with no common sources in their retweets will be assigned more diverse values. Notice that this technique is based solely on the interaction information, i.e., on the structure of the interaction matrix. Consequently, for each news source, we compute the median of its retweeters' values and use it as a proxy of its bias.

	UK	IT	FR	DE
Users	231,911	8,676	33,288	22,512
News outlets	167	89	129	127
Tweets	46,404	3,877	12,493	6,368
Retweets	1,385,023	24,856	97,909	53,352

Table 3 - Summary of the original tweets and retweets for the Brexit debate.

To validate the results of the latent ideology method, we look at politicians who shared the content produced by news outlets. To do so, we use the list of politicians from 26 countries collected between September 2017 and February 2021 available at TwitterPoliticians.org which has been explored in detail in previous work (van Vliet et al., 2020). For each politician, we assign a binary value, namely -1 or $+1$, depending on whether it is affiliated either with a europeist (-1) or eurosceptic $(+1)$ political party. We then compute the average of the binary values of all politicians that retweeted a news outlet and use this metric, which we call "political leaning", to confirm the solidity of the latent ideology analysis and identify the anti-EU group in the debate.

Results with only the news outlets for which we can estimate the political orientation based on politician retweeters are shown in Figures 10-13 for the United Kingdom, Italy, France, and Germany respectively. In general, our technique proved to be accurate in clustering news outlets depending on the opinion on the Brexit debate. Indeed, the news outlets on the left of the graphs are usually mostly retweeted by europeist politicians, while those on the right are shared by eurosceptics. Thus, based on this result, we identify the news outlets that have a synthetic value opinion greater than zero as anti-EU news sources. Nevertheless, we notice that countries exhibit some differences in the structure of the debate. While the UK has two distinct peaks in the distribution of the synthetic opinion, as shown in the top panel of Figure 10, France presents only one peak in the synthetic opinion distribution (top panel of Figure 12) and has no news outlets retweeted by politicians on the right side of the bottom panel of Figure 12. Germany and Italy present an intermediate situation, with quite unimodal synthetic opinion distributions and the presence of news outlets on the right side of the bottom panels. This can be explained by the different levels of attention created by the Brexit debate in the public sphere and thus the different volumes of news generated in each country, which is understandably dominated by the UK.
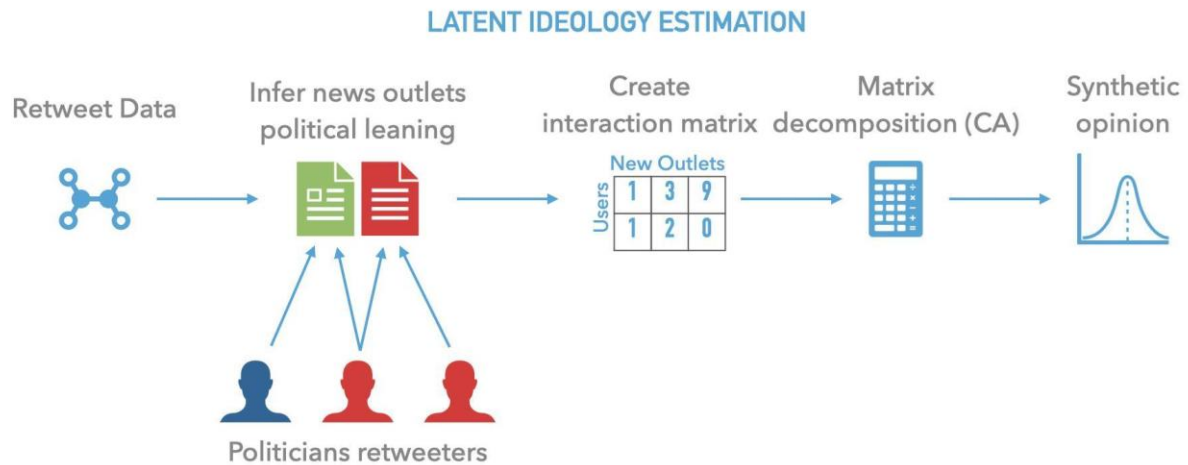
Figure 9: Workflow diagram of the latent ideology analysis.

Finally, we identify the news sources that are likely to produce anti-EU misinformation by cross-referencing the information obtained from latent ideology analysis with the reliability score in the NewsGuard dataset. To understand the specific topics created around Brexit, we perform topic modelling by using BERTopic on all the tweets produced by Questionable sources with anti-EU synthetic opinions. The results are shown in Figures 14, 15, and 16 for the UK, Italy, and Germany, respectively, and highlight the presence of topics associated with anti-EU populist arguments, such as "Brexit is a lie" or "Brexit betrayal", in which people are sceptical about the effective exit of the UK from the EU, as well as "Rising of nationalism", which highlights how nationalist are gaining momentum all over the world, and "UE policy against brexit", where the actions taken by the EU in response to Brexit are seen as an attempt to discourage other countries to leave EU. The corpus of France is not large enough to allow this type of analysis and is thus discarded here.
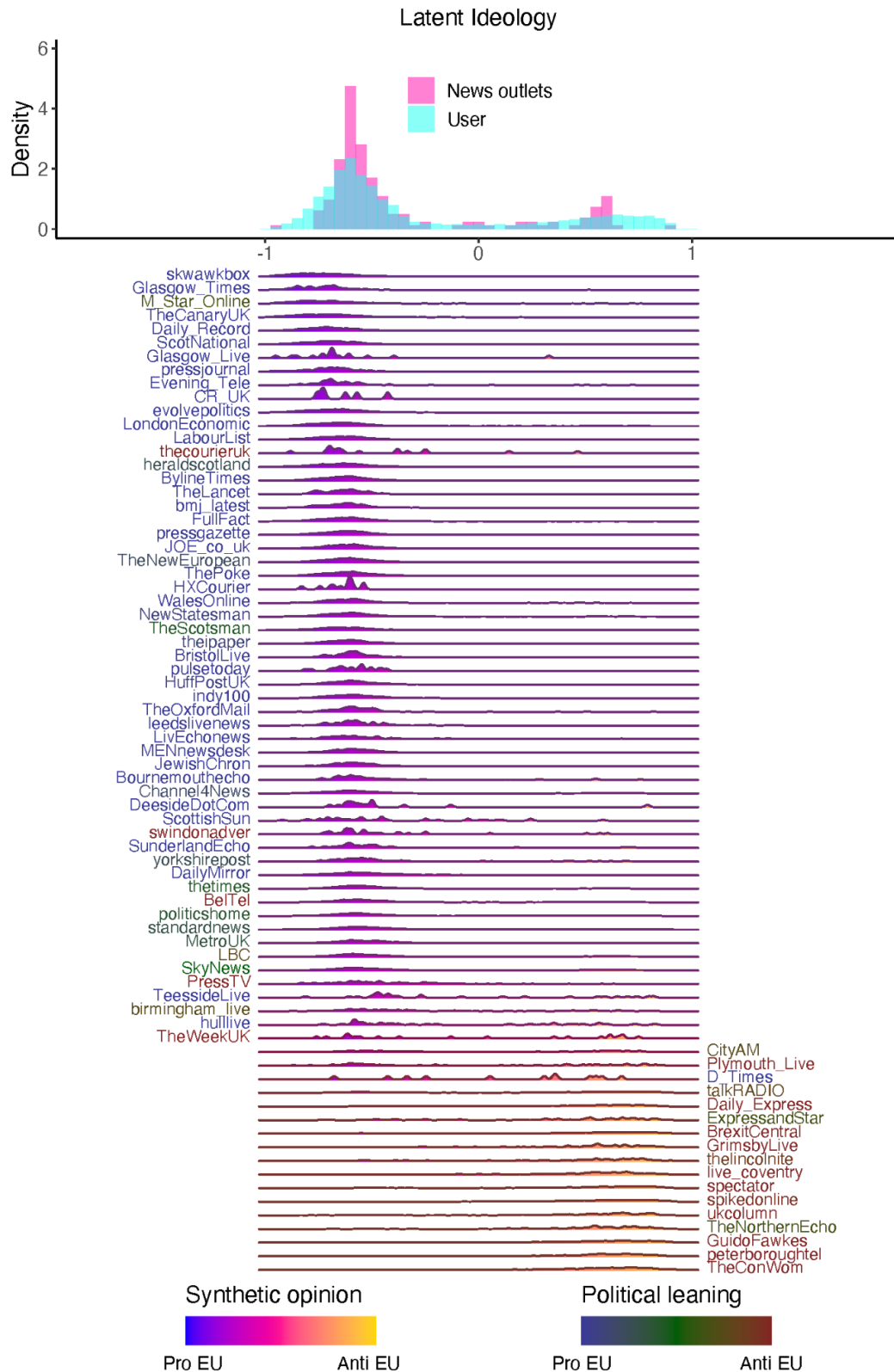
Figure 10. Latent ideology estimation for France news outlets with political leaning inferred from politician retweeters colour coded on news outlets' name.
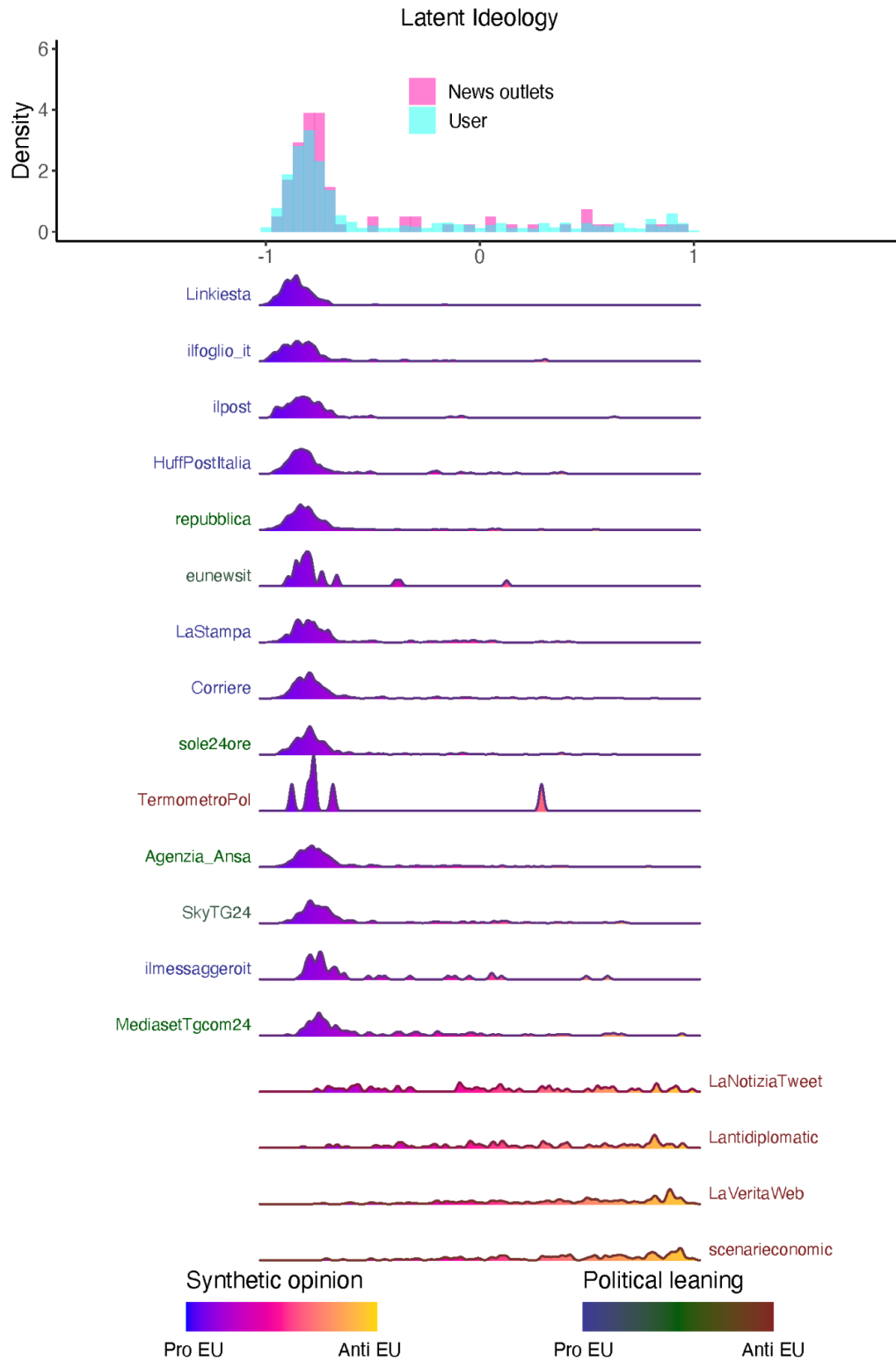
Figure 11. Latent ideology estimation for France news outlets with political leaning inferred from politician retweeters colour coded on news outlets' name.

Figure 12. Latent ideology estimation for France news outlets with political leaning inferred from politician retweeters colour coded on news outlets' name.
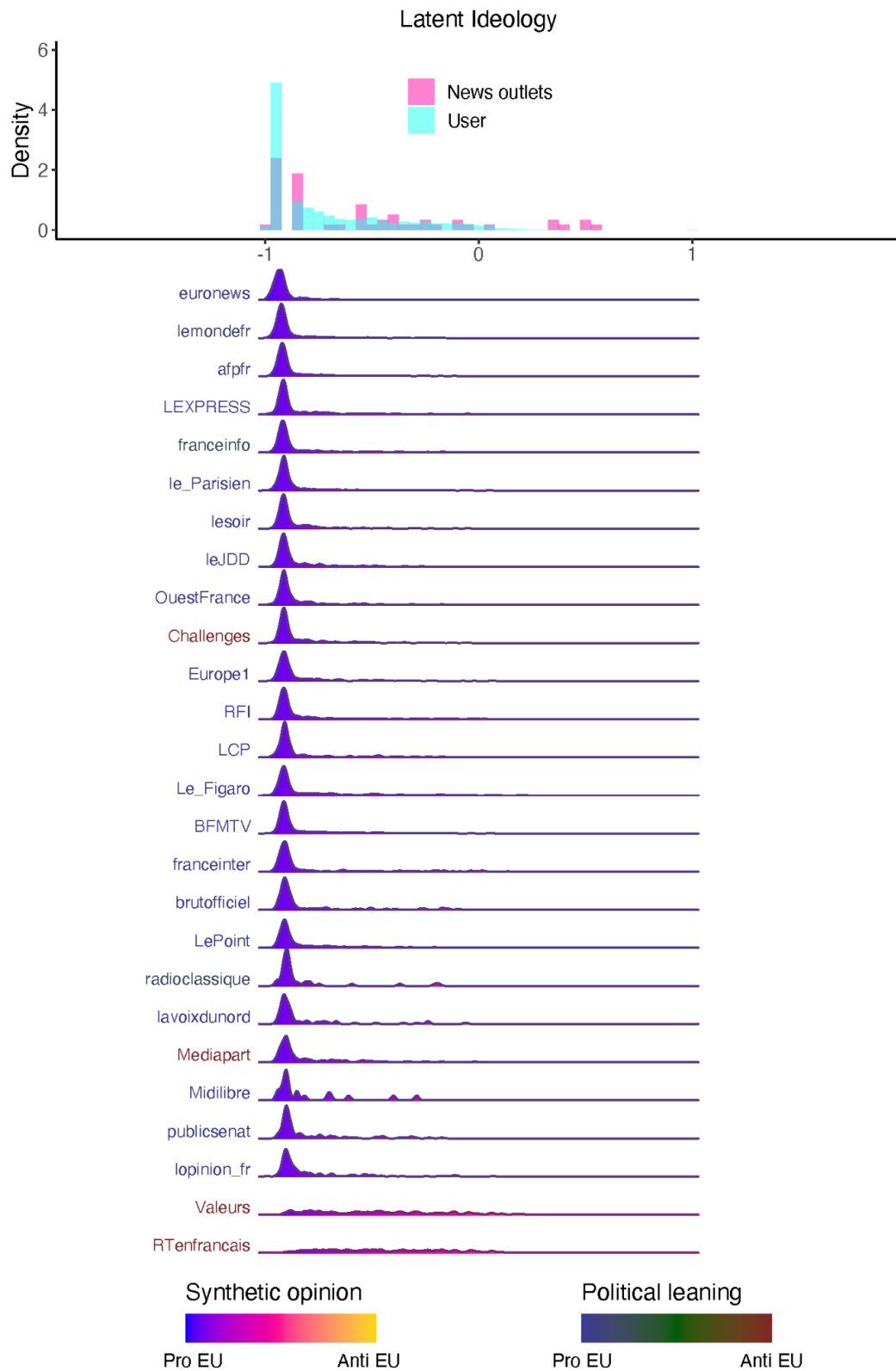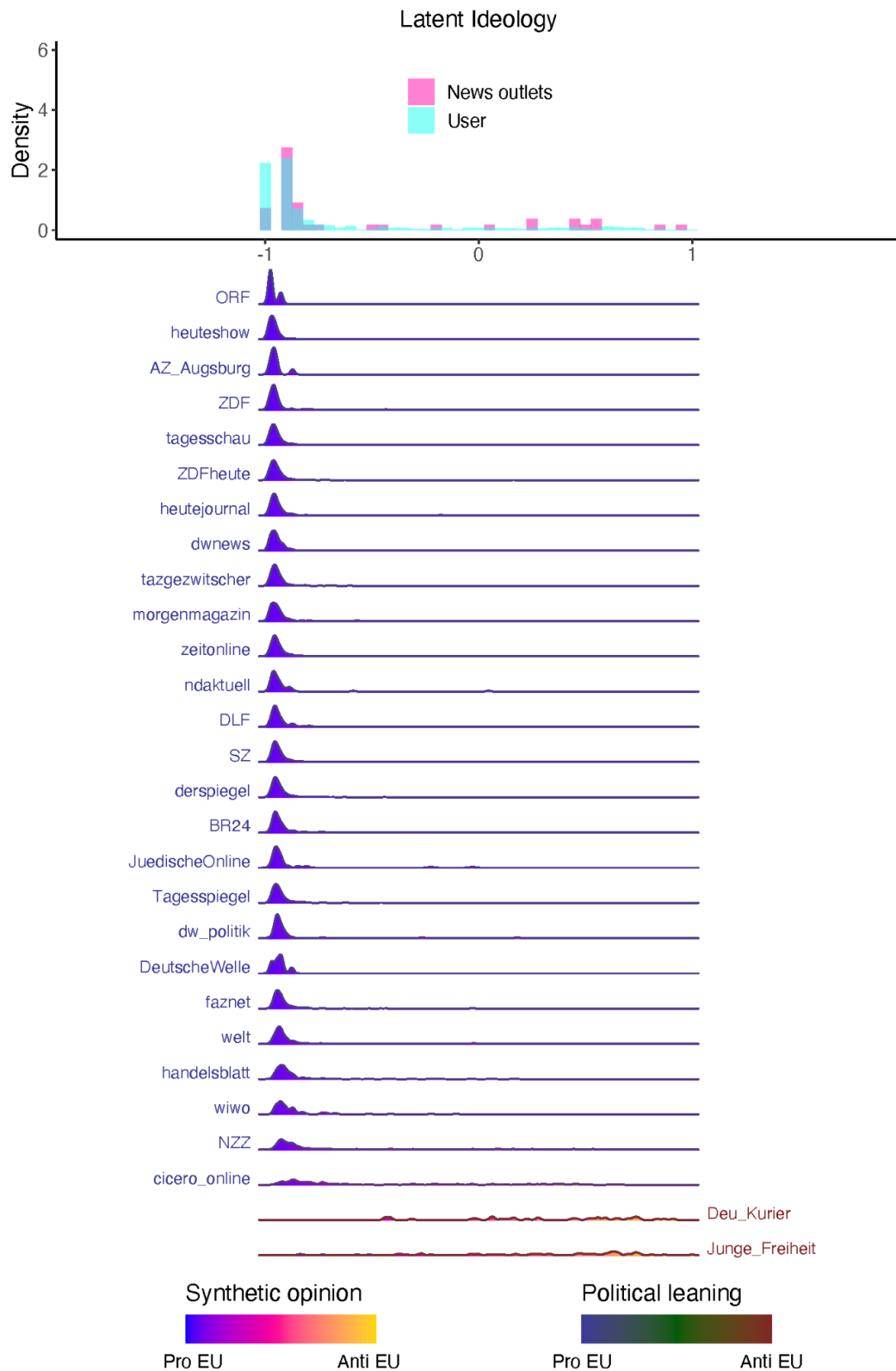
Figure 13. Latent ideology estimation for France news outlets with political leaning inferred from politician retweeters colour coded on news outlets' name.
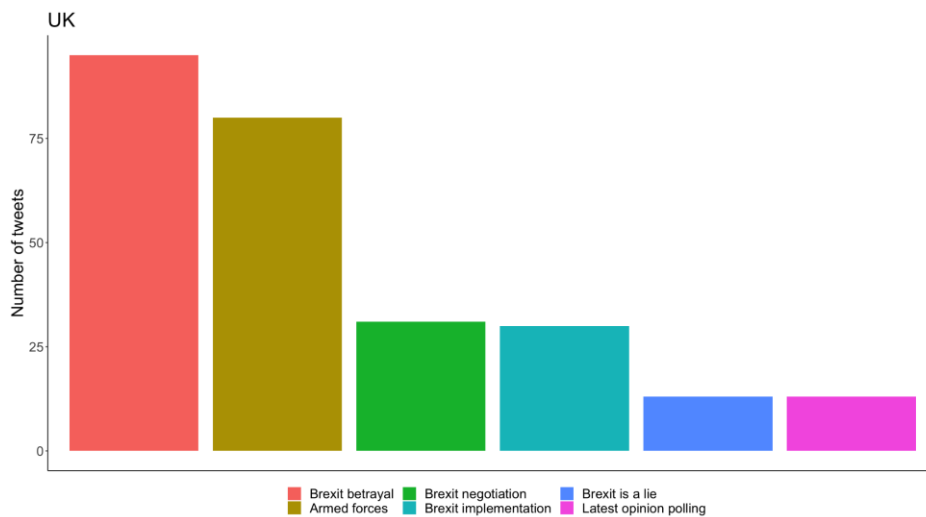
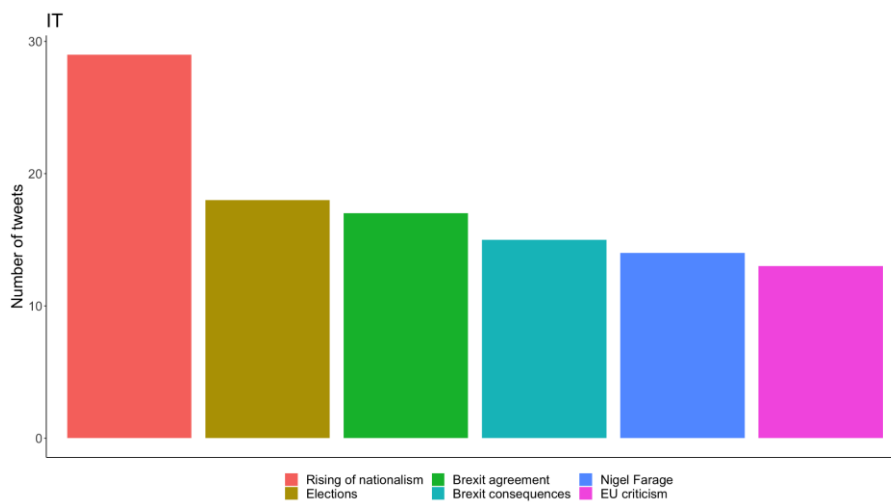*Figure 14. Most debated topics in the English Anti-EU misinformation corpus.*

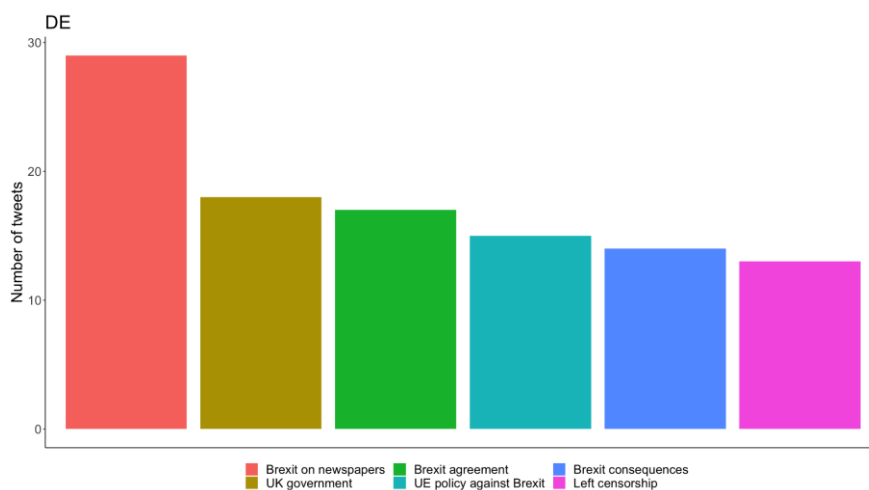*Figure 15. Most debated topics in the Italian Anti-EU misinformation corpus.*

Figure 16. Most debated topics in the Italian Anti-EU misinformation corpus.

6. How to fight misinformation

Debunkers strive to contrast misinformation spreading by providing fact-checked information on specific topics. One might assume that corrections (debunking activities) are likely to sound to target users such as information dissenting from their preferred narrative. Response to debunking attempts may result in the well-known backfire effect (Nyhan and Reifler, 2010), because corrections may be perceived as a further attempt to manipulate information, thus reinforcing users' original viewpoints. Empirical studies showed that debunking attempts are largely ineffective because they are not able to reach the target audience (Zollo et al., 2017). Furthermore, debunking has a high cost both in terms of human resources and time necessary for checking the information. Given the amount of content produced and circulating online, and the speed and scale at which information can spread on the Internet, relying solely on debunking to contrast misinformation is unfeasible. Moreover, it has been observed that the emergence of a certain topic in misinformation sources is confined to about 25 hours after its first appearance on official news (Del Vicario et al., 2019). Counteracting misinformation after its spreading is thus difficult and the time to act is limited.

Here we propose complementary strategies to counter the spreading of misinformation on social media. As we have discussed in Section 1, empirical evidence suggests that the increasing polarisation and segregation of users in echo chambers play a pivotal role in misinformation spreading on social media. This observation can be crucial for the definition of appropriate countermeasures. Smoothing polarisation and encouraging effective communication and dialogue between communities can instead become essential.

To this end, the analysis of users' behaviour and how they interact with information may prove useful to determine in advance the targets of hoaxes and false information (Del Vicario et al., 2019). In previous work, we presented a framework for the timely identification of polarising content that enables the "prediction" of future misinformation topics. Here, we advance that study by proposing the integration of the ideology metric presented in Section 5. Indeed, latent ideology estimation has proved to be a valid technique to measure the level of polarisation in online debates (Barberá, 2015; Falkenberg et al., 2022; Flamino et al., 2021). Since polarisation is one of the main drivers behind the formation of echo chambers and the proliferation of misinformation, its temporal evolution can be used as a predictor for the spreading of misinformation. If we observe a rapid increase in the debate polarisation around one particular topic, the environment is likely becoming more and more fragmented and thus prone to the spreading of inaccurate or false news. The increase of polarisation over time can be quantified on the synthetic opinion distribution using the Hartigans' dip test (Hartigan and Hartigan, 1985). For the case

study of Brexit presented in Section 5, we report here the values of the Hartigans' dip test for the distribution of users and news outlets in the four countries with the corresponding p-values (Table 4). The test detects a non-unimodal distribution only for the users' synthetic opinion distribution of the UK and Italy, which also have the two biggest anti-EU groups (see Figures 10 and 11). Thus, the significance of this value, together with its trends over time, can be used as a feature to be integrated into the framework for the early detection of misinformation-susceptible environments. Since most topics become subject to misinformation within a day from their first appearance on official news, an early warning strategy to counteract possible future misinformation is even more justified. Indeed, a timely identification of polarisation dynamics and potential misinformation targets may allow for the design of tailored counternarratives and appropriate strategies (Cinelli et al., 2022; Roozenbeek et al., 2022; Schmidt et al., 2020).

	UK	IT	FR	DE
Users	0.033***	0.0180**	0.0037	0.0046
News Outlets	0.0402	0.0245	0.0252	0.0223

*** $p<0.001$ **$p<0.01$ *$p<0.05$

Table 4 - Result of the hartigans' dip test of unimodality for users and news outlets synthetic opinion distribution

7. Discussion

As discussed so far, polarisation on social media platforms plays a crucial role in how information spreads –including false and unreliable content– and how the public debate unfolds. Recent findings have shown that polarisation is also linked to radicalisation phenomena such as online hate speech (Cinelli et al., 2021b). Therefore, investigating polarisation processes is key to design effective solutions and promote a civil space of public discussion.

Here, we have addressed such issues from a quantitative perspective, analysing more than 25M tweets and presenting a technique to infer the ideological separation of users (i.e. echo chambers) in debates around controversial issues. This enables us to 1) provide a complete overview of the public debate on European topics in the last three years, 2) study the evolution of misinformation at a large scale, and 3) develop a technical solution for combating misinformation. Despite the many benefits of this kind of approach, some limitations and shortcomings are worth discussing.

Misinformation classification. One of the main issues when working with huge amounts of data is how to classify misinformation. Clearly, examining every single piece of content in a dataset with millions of observations is unfeasible. Moreover, such a classification should be out of the researchers' competence.

For these reasons, in this work, we applied an approach based on the classification of the source domain provided by a third-party organization. This means that the reliability of a single tweet is derived from the reliability assigned to its source. Therefore, even when a specific piece of content does not contain false or unreliable information, it is classified as such if the source publishing the content is labeled as questionable. Viceversa, if a source classified as reliable disseminates false information, this is not labeled as such in our dataset. However, using NewsGuard to identify misinformation helps us in reducing the potential error derived from this approach. Indeed, since NewsGuard is considering a wide range of criteria –not limited to truthfulness– to rank news domains, and since this ranking is done by professional journalists taking into account the broader content production of such domains, it is reasonable to assume that the reliability score assigned to the domain holds for most of the content published by that domain. Nevertheless, we cannot exclude that some tweets in our dataset are misclassified.

Data bias. This work focuses on Twitter data for a variety of reasons: 1) the relevance of the platform in the online public debate around socio-political issues, involving the general public as well as news outlets, journalists, politicians, and policymakers; 2) the popularity of the platform in Europe; and 3) data availability. Clearly, Twitter does not represent the whole social media, nor the European population. Therefore, our findings are not necessarily generalizable to other media platforms. Future research to replicate this study in other social media and countries is thus desirable.

Polarisation. Here we measure polarisation in terms of multimodality in the users' ideology distribution. The ideology estimation requires the interaction network to be rich enough to ensure a robust analysis and correctly separate ideologically opposite factions. Usually, a large amount of data on users' interactions is thus necessary. While the issue is limited with Twitter data, it can become a barrier when focusing on social media platforms like Facebook or Instagram, where data access is more limited. To date, data policies on such platforms would prevent the ideology estimation to be performed. Moreover, there might be conflating factors influencing the ideology results. For example, geography may play a role in the online debate and the ideology estimation may reveal geographical segregation rather than ideological separation. This is why results have to be carefully analysed, as we have done here using external data (e.g. politicians) to validate clusters' ideological stance on Brexit.

Further research. Our analysis shows differences in the public debate around European issues across the selected countries. For example, in Germany, we can observe an increasing promotion of the topic "Russia – Europa" in Questionable since 2019, while in France the topic "EU – Islam" has been receiving a very high engagement since 2020. Possible explanations for such differences are out of the scope of

this work, but further qualitative investigations from a sociological and political science perspective are desirable.

8. References

Barberá, P., 2015. Birds of the Same Feather Tweet Together: Bayesian Ideal Point Estimation Using Twitter Data. Polit. Anal. 23, 76–91. https://doi.org/10.1093/pan/mpu011

Cinelli, M., De Francisci Morales, G., Galeazzi, A., Quattrociocchi, W., Starnini, M., 2021a. The echo chamber effect on social media. Proc. Natl. Acad. Sci. 118, e2023301118. https://doi.org/10.1073/pnas.2023301118

Cinelli, M., Pelicon, A., Mozetič, I., Quattrociocchi, W., Novak, P.K., Zollo, F., 2021b. Dynamics of online hate and misinformation. Sci. Rep. 11, 22083. https://doi.org/10.1038/s41598-021-01487-w

Cinelli, M., Peruzzi, A., Schmidt, A.L., Villa, R., Costa, E., Quattrociocchi, W., Zollo, F., 2022. Promoting engagement with quality communication in social media. PLOS ONE 17, e0275534. https://doi.org/10.1371/journal.pone.0275534

Del Vicario, M., Quattrociocchi, W., Scala, A., Zollo, F., 2019. Polarization and fake news: Early warning of potential misinformation targets. ACM Trans. Web. https://doi.org/10.1145/3316809

Falkenberg, M., Galeazzi, A., Torricelli, M., Di Marco, N., Larosa, F., Sas, M., Mekacher, A., Pearce, W., Zollo, F., Quattrociocchi, W., Baronchelli, A., 2022. Growing polarisation around climate change on social media (No. arXiv:2112.12137). arXiv. https://doi.org/10.48550/arXiv.2112.12137

Flamino, J., Galezzi, A., Feldman, S., Macy, M.W., Cross, B., Zhou, Z., Serafino, M., Bovet, A., Makse, H.A., Szymanski, B.K., 2021. Shifting Polarization and Twitter News Influencers between two U.S. Presidential Elections (No. arXiv:2111.02505). arXiv. https://doi.org/10.48550/arXiv.2111.02505

GCS, 2020. RESIST Counter-disinformation Toolkit.

Grootendorst, M., 2022. BERTopic: Neural topic modeling with a class-based TF-IDF procedure.

Hartigan, J.A., Hartigan, P.M., 1985. The Dip Test of Unimodality. Ann. Stat. 13, 70–84.

Joint Research Centre, Lewandowsky, S., Smillie, L., Garcia, D., Hertwig, R., Weatherall, J., Egidy, S., Robertson, R.E., O'connor, C., Kozyreva, A., Lorenz-Spreen, P., Blaschke, Y., Leiser, M., 2020. Technology and democracy: understanding the influence of online technologies on political behaviour and decision making. Publications Office of the European Union, LU.

Lewandowsky, S., Ecker, U.K.H., Seifert, C.M., Schwarz, N., Cook, J., 2012. Misinformation and Its Correction: Continued Influence and Successful Debiasing. Psychol. Sci. Public Interest 13, 106–131. https://doi.org/10.1177/1529100612451018

Marie, A., Altay, S., Strickland, B., 2020. The cognitive foundations of misinformation on science. EMBO Rep. 21, e50205. https://doi.org/10.15252/embr.202050205

Nyhan, B., Reifler, J., 2010. When Corrections Fail: The Persistence of Political Misperceptions. Polit. Behav. 32, 303–330. https://doi.org/10.1007/s11109-010-9112-2

Roozenbeek, J., van der Linden, S., Goldberg, B., Rathje, S., Lewandowsky, S., 2022. Psychological inoculation improves resilience against misinformation on social media. Sci. Adv. 8, eabo6254. https://doi.org/10.1126/sciadv.abo6254

Sammut, C., Webb, G.I. (Eds.), 2010. TF–IDF, in: Encyclopedia of Machine Learning. Springer US, Boston, MA, pp. 986–987. https://doi.org/10.1007/978-0-387-30164-8_832

Schaewitz, L., Kluck, J.P., Klösters, L., Krämer, N.C., 2020. When is Disinformation (In)Credible? Experimental Findings on Message Characteristics and Individual Differences. Mass Commun. Soc. 23, 484–509. https://doi.org/10.1080/15205436.2020.1716983

Schmidt, A.L., Peruzzi, A., Scala, A., Cinelli, M., Pomerantsev, P., Applebaum, A., Gaston, S., Fusi, N., Peterson, Z., Severgnini, G., De Cesco, A.F., Casati, D., Novak, P.K., Stanley, H.E., Zollo, F., Quattrociocchi, W., 2020. Measuring social response to different journalistic techniques on Facebook. Humanit. Soc. Sci. Commun. 7, 17. https://doi.org/10.1057/s41599-020-0507-3

Sunstein, C.R., 2017. A prison of our own design: divided democracy in the age of social media [WWW Document]. Democr. Audit UK. URL http://www.democraticaudit.com (accessed 10.28.22).

van Vliet, L., Törnberg, P., Uitermark, J., 2020. The Twitter parliamentarian database: Analyzing Twitter politics across 26 countries. PLOS ONE 15, e0237073.

Zollo, F., Bessi, A., Vicario, M.D., Scala, A., Caldarelli, G., Shekhtman, L., Havlin, S., Quattrociocchi, W., 2017. Debunking in a world of tribes. PLOS ONE 12, e0181821. https://doi.org/10.1371/journal.pone.0181821

Get in touch

✉ info@eumeplat.eu   </> www.eumeplat.eu