

Coordinated inauthentic behaviour on Facebook? A typology of manufactured attention

Platforms & Society
Volume 2: 1–13
© The Author(s) 2025
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/29768624251369784
journals.sagepub.com/home/pns



Richard Rogers¹  and Nicola Righetti² 

Abstract

This paper advances the discourse on coordinated inauthentic behaviour (CIB) on Facebook by extending its study beyond deceptive influence operations. Using techniques developed for its technical analysis, we undertake an empirical study of CIB that surfaces not only such operations but predominantly others, particularly media groups sharing news stories, political activists sharing memes, advertising networks promoting gambling and cyber scams as well as large public groups hijacked to spread ads. Based on these findings, we develop a typology of manufactured attention and make a series of observations concerning the analytical ambiguity of measuring CIB through (timed) link-sharing. The first is the question of when coordinating actors and coordination behaviours cross the line into inauthenticity, which is of methodological interest for platform policies and platform policy observers. The second aspect concerns the evolution of Facebook's (later Meta's) policies with respect to CIB, especially how the platform has defined and described it. We conclude with the observation that Meta's narrowing definition allows for once-erstwhile CIB activities to remain on-platform.

Keywords

Coordinated behaviour, attention hijacking, platform infrastructure, information operations, astroturfing

Introduction: situating our study of coordinated inauthentic behaviour

This paper contributes to the study of coordinated inauthentic behaviour (CIB) by extending it beyond explicit instances of influence operations, looking into the spectrum of coordination in the service of manufacturing attention (Goldhaber, 1997). By manufacturing attention, we mean to describe and capture shorthand the artificiality of audience fabrication through coordinated behaviour. We examine highly coordinated behaviours on Facebook, developing a typology through both coordination behaviour and actor analysis.

Using established techniques for CIB analysis, we first undertake an empirical analysis of near synchronous link-sharing that surfaces signature forms of content amplification through coordination (Giglietto et al., 2020b). Remarkably, apart from influence operations, we found from the actor analysis that a variety of collective online efforts—by media organisations, political activists and ad networks—coordinate their behaviour to very high degrees.

Subsequently, we address the question of the presence of these coordinating actors on Facebook. We do so in light of the evolving criteria of CIB put forward by Meta that, if

met, presumably would result in their removal from the platform (Gleicher, 2018). In conclusion, we find that Meta has winnowed its definition of CIB to the extent that most activities (once) falling under CIB are allowed or at least are not considered 'inauthentic'. Instead, Meta's focus is on coordinating 'adversarial actors', leaving others on-platform.

While coordination—extending beyond influence operations and including both authentic and inauthentic cases—has been addressed in recent literature (Mannocci et al., 2024) (see Figure 1), our contribution seeks to develop a typology comparing these operations with other forms of highly coordinated behaviour, examining similarities and differences. To give one poignant example at the outset, our empirical analyses of coordinating posts found nearly the same signature

¹Media Studies, University of Amsterdam, Amsterdam, The Netherlands

²Department of Communication Sciences, Humanities and International Studies, University of Urbino Carlo Bo, Urbino, Italy

Corresponding author:

Richard Rogers, Media Studies, University of Amsterdam,
Turfdraagsterpad 9, 1012 XT Amsterdam, The Netherlands.
Email: rogers@uva.nl



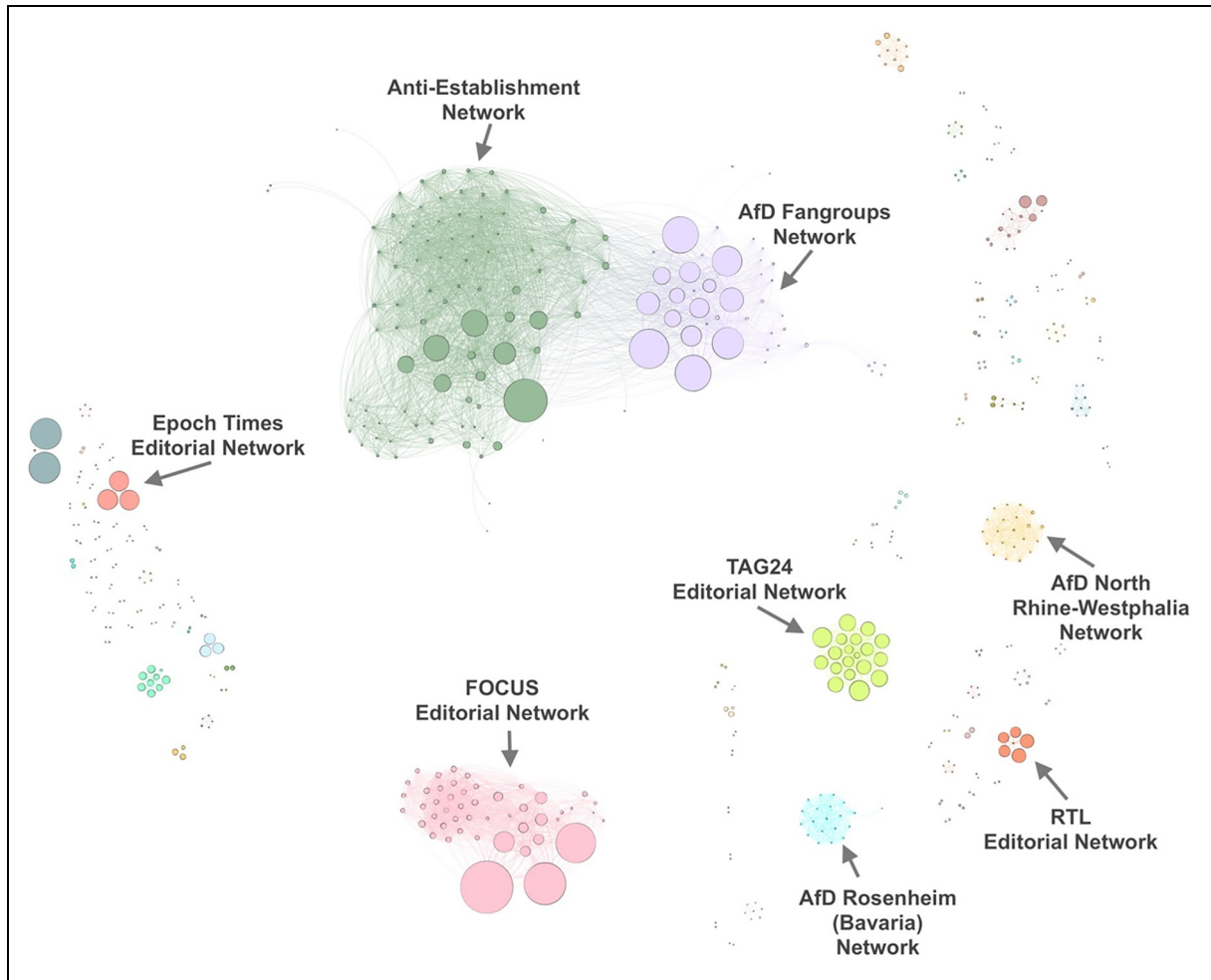


Figure 1. Coordinated link-sharing networks on Facebook during the 2021 German elections (Righetti et al., 2022).

forms of coordination between an influence operation and media groups sharing news stories. Another political messaging campaign, albeit more organic seeming, exhibits a form of coordination similar to an online gambling ad network. We additionally found distinctive forms of coordination between political activists sharing memes, advertising networks promoting cyber scams as well as large public groups hijacked to spread ads.

For each of these types, there are questions concerning the line to draw between non-offending and offending forms of attention manufacture (Brunton, 2013). By coordination tactics alone, we have not been able to distinguish between what one could call coordinated authentic and inauthentic behaviour without considering the actors or sources in question. Thus in our study of CIB we are occasioned to move beyond coordination only and additionally address actor types.

In the next parts, we first discuss research on forms of coordination relying on platform architecture, particularly how algorithms react to sharing behaviors and metrics (Burton et al., 2023). Thereafter we discuss the building

blocks of our typology (coordination behaviour and actor types) and put forward our findings. We conclude with a discussion of how the platform's (current) definition of CIB appears to allow coordination to thrive, where the exception is made for what Meta calls 'adversarial actors' which abuse the platform.

Research into coordinated inauthentic behaviour

As we have also found, coordinated behaviour on social media can be conducted by both covert and overt actors with a range of objectives. These objectives include mimicking organic engagement (Keller et al., 2020) and distributing content through multi-actor broadcasting and reposting (Giglietto et al., 2020b). Such coordinated campaigns are often driven by actors or bots synchronizing the promotion of the same or related content to achieve virality or a significant number of interactions and impressions, creating an appearance of popularity (Chan, 2024).

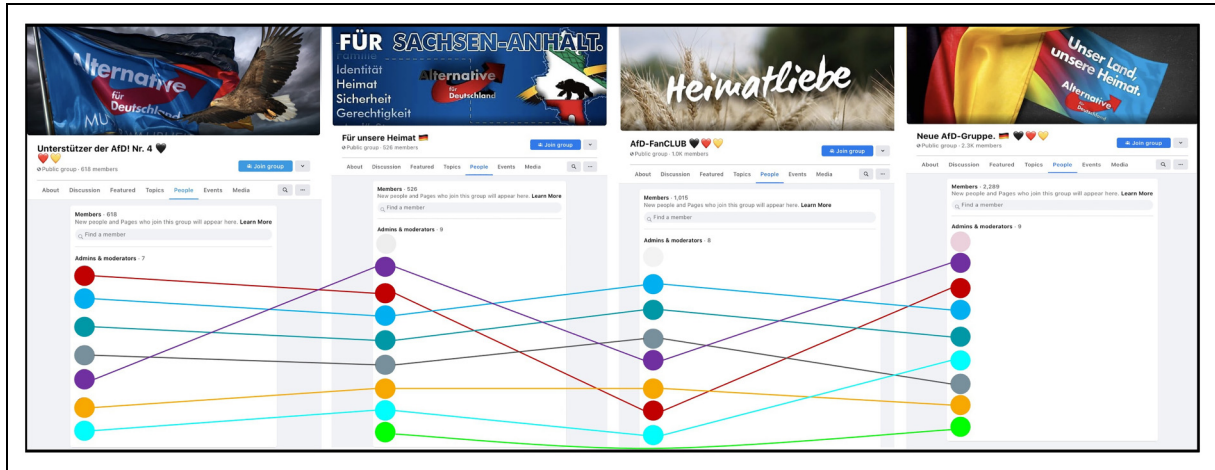


Figure 2. Recurring administrators of AfD fan groups on Facebook were identified in the analysis of coordinated behaviors. Names and profile pictures are anonymized. Each connected circle represents an individual, 2021. (Righetti et al., 2022).

Especially in the political arena, the primary objective of these campaigns is to ‘flood the space’, thereby exerting, or seeming to exert, influence and accruing symbolic power (McIntyre, 2018).

In empirical research in that arena, scholars have observed consistent, statistical associations between the spread of low-quality sources, disinformation, or misinformation and coordinated activity. For example, coordinated networks were found to be more likely to share news sources previously flagged by fact checkers both during the 2020 national Italian election and the 2019 European elections in Italy (Giglietto et al., 2020a). That also was the case leading up to the 2021 German elections, where coordinated entities have been found to be significantly more likely to share these domains (Righetti et al., 2022). Coordination research often finds misinformation networks in the run up to elections, which explains the emphasis in this area.

It should be noted, however, that this correlation has not been well accounted for. One reason for a higher degree of coordination of low-quality sources might be that the content is not just a part of a political campaign. It may be produced for profit and spread in a coordinated fashion to maximize circulation.

As we also found, misinformation and news link-sharing behaviours may be remarkably similar. Indeed, highly coordinated link-sharing is undertaken by legitimate editorial networks. For example, in the context of the 2021 German election, the more controversial Epoch Times (Roose, 2020) had sharing practices similar to the editorial networks for FOCUS, the German weekly news magazine and TAG24, the tabloid (see Figure 1).

A high prevalence of coordinated activity has also been observed within right-wing communities and parties, sometimes obscuring their origins (Benkler, 2018). In the UK general election of 2019, it was found that

conservatives coordinated on Twitter far more than those supporting the Labour party, with “higher degrees of automation and Twitter suspensions” (Nizzoli et al., 2021, p. 453). In the European context, among the largest coordinated networks active during the 2021 German elections were those organised by the right-wing AfD party (Figure 2) (Righetti et al., 2022). Besides official party pages, we also detected coordinated networks of AfD fan groups. These fan groups were found to be administered by the same set of individuals (see Figure 2). In turn, these fan groups were linked to Facebook groups that were not directly traceable to a specific political party but were still promoting a coherent right-wing ideology. Notably, one of the largest and most active groups, a populist faction that expresses distrust in the parliament by virtue of its very name, was discovered to have an AfD politician among its administrators. This example highlights the variety and ambiguity of the forms and purposes of coordination on Facebook.

There has been research on a range of other substantive areas where coordination takes place, highlighting its operational and geographical breadth. For example, it has been linked to coronavirus politics (Magelinski and Carley, 2020), protest in authoritarian regimes (Kulichkina et al., 2024), and cryptocurrency manipulation (Terenzi, 2023). Studies have documented coordinated networks on social media in diverse locations such as Australia (Graham et al., 2021), Nigeria (Giglietto et al., 2022), South Korea (Keller et al., 2020), the Philippines (Yu, 2022), Brazil and France (Gruzd et al., 2022). Although CIB is geographically widespread, it remains largely platform-specific, with most of the research orientated towards single-platform studies (Thiele et al., 2023). Most of the studies focus solely on platforms such as Twitter or Facebook, with cross-platform coordination being relatively rare, though not unheard of (DiResta et al., 2018; Howard et al., 2018).

Notably, most examples of coordinated behaviour under study are termed ‘inauthentic’. While misinformation and social media manipulation are often the focus of coordinated networks, other forms of coordination, such as movement campaigning, religious proselytization and fan support, can appear similar in their traces (Giglietto et al., 2023a). This is our point of departure.

We now turn to types of coordinated networks emerging from the empirical literature, specifically referring to coordinated link-sharing behaviour that may be flagged through CIB-style detection but are not political influence operations, misinformation or election-related campaigns. Social movements and activists can share the same message in a coordinated manner, which is a form of protest that would trigger the highly coordinated behaviour metric. For example, animal rights movements can mobilize groups of digital activists who, upon receiving instructions via a mailing list, are provided with directions to flood targeted social media pages or posts with coordinated messages (Righetti and Bertuzzi, 2020). Social movements can also employ coordinated campaigns to raise awareness about their cause (Canevez et al., 2024). Religious movements have also been found to engage in coordinated behavior to promote their beliefs and attract new followers, disguising themselves as mainstream religious organisations by playing on the similarities between faiths (Giglietto et al., 2023a). Similar coordinated networks appear when analyzing social movements organizing protests against authoritarian countries and authoritarian pro-government actors interested in fostering a certain narrative for strategic purposes like social control and protest suppression (Kulichkina et al., 2024).

Coordinated networks are increasingly utilized to promote scams (Terenzi, 2023). By leveraging such coordinated efforts, scammers can effectively disseminate deceptive schemes, amplifying their reach and impact on unsuspecting users. One prevalent form facilitated by coordinated networks involves deceptive cryptocurrency schemes. Scammers exploit the decentralized nature of cryptocurrencies and the anonymity offered by social media to lure victims into ‘get-rich-quick’ schemes. These include promises of free cryptocurrency through airdrops, rewards for completing simple tasks like sharing content or clicking on links, and opportunities to earn digital currency by watching advertisements or playing online games. Many of these schemes require users to provide personal data or invest initial sums, which may lead to financial losses when the promised rewards fail to materialize.

The impact of these coordinated scams is significant, as they manipulate public sentiment and exploit the lack of regulation in the cryptocurrency market. They can mislead many people by creating an illusion of widespread adoption and credibility. The use of coordinated networks amplifies the reach of these scams and also makes it difficult for authorities to track and shut down the operations.

Having briefly reviewed the various network types often identified by coordinated link-sharing detection techniques, we would like to turn to the study of coordination and the building of our typology.

Platform architectures and coordination

Coordination is generally defined in the dictionary as “making many different things work effectively as a whole” (Cambridge Dictionary, 2024). The coordination involved in our object of study, online attention manufacture, concerns collective effort to deliver messages to an audience (Keller et al., 2020). Following the interest in online misinformation in recent years, computational social scientists and other media scholars have paid increasing attention to coordination patterns on social media platforms from a trust-and-safety perspective, which is an umbrella term from industry concerning content moderation, or governing on-platform materials (Gruzd et al., 2023).

Coordination serves expressive and symbolic purposes, as it blends together multiple individuals through ritualistic purposes to reinforce internal solidarity and perform the idea of collectivity in front of bystanders (Durkheim, 2014). We can see this clearly when observing social movements and social protests. They not only coordinate the activities necessary to organize the protest movement but also coordinate at the symbolic level, managing the display of symbols and behaviours that constitute the entire choreography of the protest (Foster, 2003). These coordinated patterns unite the protest and lead observers to identify a multitude as a unit and attribute agency to it. Importantly for the purposes of our argument is that the numbers of those involved is also a proxy for their potential influence (Tilly, 1994). Indeed, it is often reported how many people are involved in a protest, and readers interpret it as such (Warner, 2002). The greater the number, the more serious the cause is taken.

The coordinated gathering at a place, the display of symbols of identity and the ritualistic choreography of the protest can also be rendered online (Gerbaudo, 2012). Here, other forms of communication allow for a similar impact as a coordinated mass of people protesting. Social media affordances favour specific types of coordinated display. The affordance of replicability, for instance, is one, where copying and pasting the same protest message lessens the burden of participation (boyd, 2010; Christensen, 2011). The same messages copied over and over, more or less intentionally, replicate online the feeling of a protest where many people display the same banners and messages. Here, the number of people involved is a proxy for their influence and the influence of their message.

Social media metrics, in fact, are commonly interpreted as measures of influence: the number of followers of influencers serves as an indicator for their popularity and influence, while likes, upvotes, shares on a message act as

proxies for social appreciation (Rogers, 2018). This, in turn, can influence bystanders' opinions and behaviour owing to a simple social inference process. Coordination can be staged to exploit this coordination property in an attempt to exert symbolic influence (Chan, 2024).

When we observe digital behaviours and study digital traces, we occasionally do not know who left the traces, which opens up the question of authenticity. Is one able to tell (and through which means) if the accounts whose coordinated actions we observe online are everyday users, bots, or a few people operating behind hundreds of accounts. Their motivations also may not be clear. This is where the recent specialized strand of research on coordinated inauthentic behaviour comes in. The goal is to identify organized efforts behind seemingly organic messages and engagements. The main problem and interest in this line of research is not so much the coordination itself but the question of deceptive behaviour which lies behind the notion of inauthenticity (Chan, 2024; Thiele et al., 2023).

Networks of users that function almost synchronously have been scrutinized for their critical role in manufacturing attention on social media, which is also called 'false amplification' (Giglietto et al., 2020b). The algorithmic architecture of social media platforms fosters coordination around specific digital objects. On Twitter, coordination often aims to push hashtags into trending topics, sometimes employing "weaponized bots" (Graham et al., 2021). On Facebook, URL links are typically shared in posts on Pages and Groups, aiming to elicit emotive reactions, long comment threads, and further sharing—key components for algorithmic amplification (Merrill and Oremus, 2021). These elements persuade the Facebook Feed to elevate the emotionally charged shared links. Especially for coordinated efforts on Facebook that seed social media posts with content from the web, URLs are often central to the study of online information ecologies and alternative infrastructures (De Maeyer, 2013; Rogers, 2017, 2023). Our approach, developed through methods built into software applications, concentrates on coordinated activities where URLs are shared.

A specific example of coordinated inauthentic behaviour is astroturfing (Keller et al., 2020). This term refers to coordinated campaigns involving the distribution of messages supporting a specific agenda. Now often undertaken on the internet, it mimics the actions of many, fabricating impressions and other social media metrics to suggest that a particular idea or opinion enjoys broad support (Chan, 2024).

Coordinated inauthentic activity can be used to increase the perceived importance of content online and spread it further, using networks of centrally controlled accounts and pages that synchronize the same content. This can have the additional objective of attempting to manipulate the social media algorithm that presides over sharing content on the platform (Giglietto et al., 2020b). The

underlying idea is that social media algorithms, which aim to identify potentially relevant content that is trending, could be tricked by receiving signals of popularity related to specific content. This could then lead the algorithms to further spread the content themselves (Zhang et al., 2016).

We wish to add that online environments have augmented coordinated behaviour, aiding coordination without central organization (Bennett and Segerberg, 2013). Messages that resonate with individuals' values can spread fast on social media networks and activate synchronous communicative behaviour that is then aggregated algorithmically into visible spaces on social media, where they can give the impression of an organized collectivity.

A typical example is the trending topic section of the former Twitter. This coordination is not centrally organized (or not supposed to be), as participants likely will not know each other. Coordination happens due to the networked and algorithmic infrastructure. The outcome can appear to observers similar to an organized effort, but the process is different. Additionally, given the multi-actor nature of social media, where multiple people participate in the communicative exchange, a whole spectrum of actions can be identified, ranging from organically to connective, passing through organizationally enabled connective actions (Kulichkina et al., 2024).

The methods developed to identify coordination rely on network analysis together with content and temporal signals (Giglietto et al., Forthcoming). Many studies (as ours) have employed the CooRnet software to surface coordinated link-sharing behaviour in a variety of countries and contexts (Giglietto et al., 2021). The software identifies Facebook accounts that repeatedly share the same URLs within a certain time frame, typically ranging from a few seconds to a minute. The foundational principle is that although a cluster of accounts may coincidentally share identical content, their recurrent sharing suggests an organised intent (Giglietto et al., 2020b). When such coordinated detection tools are applied, these surface accounts that are engaged in influence campaigns, such as state actors, front groups and other actors masking their identities. They also bring to light other forms of coordination for a variety of purposes, as we discuss.

Methodologies for surfacing coordinated behaviour

Notably, connective or centrally organized collective coordination emerges similarly when surfaced with current software tools for coordinated network analysis. Since 2020, three software tools for academic research have been developed: CooRnet, the Coordination Network Toolkit (Graham et al., 2024), and CooRTweet (Righetti and Balluff, 2025). These tools are similar yet distinct. CooRnet and CooRTweet are developed in R, while the

Coordination Network Toolkit is developed in Python. There are certain methodological differences. CooRnet specializes in detecting coordinated link-sharing networks on Facebook and focuses on common link sharing on Facebook. Because of its dependence on the CrowdTangle API, however, CooRnet is now in a state of suspension, awaiting an opening from its replacement, the Meta Content Library.

Although they differ in focus, specific features, and types of analysis they enable, all these tools share the same basic approach in operationalizing coordinated action. First, the tools are content-agnostic, meaning that the content shared by the social media accounts is irrelevant to detecting coordination. This is viewed as a point of strength in terms of generalizability and political neutrality (Giglietto et al., 2020), but it also leads to surfacing all manners of coordination activities, requiring further interpretation of the actors or sources, as we do below.

They also implement some established operationalizations of coordinated behaviour. Coordinated social media actors are conceptualized as accounts that perform at least r times the same action together within a time window t (Righetti and Balluff, 2025). The parameter r stands for repetition and is crucial to the methodology. In fact, the same action can be performed together by two social media accounts, even within a very short time frame, by chance only. Conversely, if two accounts repeatedly share the same content over time within a short time frame, this is likely not just by chance but due to some form of organization that links them together. The methodology has proved effective in a large empirical literature in detecting stable, coordinated networks over time. Therefore, the detection algorithm works like this: starting with a set of actors, it traces a connection between them if they perform the same action within a specified time frame (say, they share the same message within ten seconds of each other). It then measures how often this occurs and retains the accounts that have performed the same action at least r times within the time frame t . The resulting structure is defined as the coordinated network.

This research project relies on a dataset of Facebook accounts and URLs they shared in a coordinated manner that originates from an initial list of Facebook accounts sharing problematic content (304 pages and 2095 groups). Most of these accounts are from the Global Map of Coordinated Accounts 2023, a collection of URLs shared on Facebook containing aggregated engagement metrics, user feedback, and third-party fact-checker ratings (Messing et al., 2020). The URLs rated as problematic ('false', 'missing context', 'mixture of false headline', or 'missing context') by Facebook's third-party fact-checkers are used to collect all the Facebook public shares (within 7 days after publication) returned by the CrowdTangle API. Using CooRnet, we identify the coordinated accounts that performed coordinated link sharing behaviour rapidly

and repeatedly, sharing news stories rated as problematic by Facebook third-party fact-checkers.

Starting from this list, an original workflow developed for this research purpose detects new Facebook accounts sharing problematic content (Giglietto et al., 2023b). This data collection routine collects up to 100 overperforming posts shared by the monitored accounts every six hours. It extracts the shared URLs, identifies any other Facebook Pages or Groups that share the same URLs using the CrowdTangle link endpoint, and checks for coordination. Coordination is detected using the CooRnet methodology, which defines coordination based on two main parameters: the timing of co-sharing (or coordination interval) and the frequency of co-sharing by any two accounts (or percentile of edge weight). This workflow identifies accounts that co-share the same content within a time interval of up to 60 seconds more frequently than 95% of all other accounts co-sharing within the same interval. The new set of accounts that appear to share URLs on Facebook in a coordinated fashion, according to this operationalization, are then added to the initial list of monitored accounts and used in subsequent iterations of the data collection routine.

This data collection ran between October 2023 until CrowdTangle was shut down in August 2024. Our dataset was updated as of July 1, 2024, when our analysis began. In our case, we concentrate on the most coordinated accounts, defined as those that shared URLs more frequently than 99% of the accounts, corresponding to at least 23 URLs, each posted no more than 60 s apart. We chose this higher cut-off to minimize false positives, focusing on highly coordinated accounts. At the same time, we recognize that other approaches could capture less explicit forms of coordination. This set of highly coordinated accounts includes 743 nodes—59 Pages and 684 Groups—connected by 16,670 edges, corresponding to 6,200 unique URLs. The network was further divided into 'communities' of accounts sharing the same content using the Louvain algorithm, a popular social network analysis method for community detection. It resulted in 19 coordinated communities (or networks as we also call them), which, along with the URLs they shared, form the main focus of this analysis (see Figure 3).

Regarding reproducibility, it is worth noting that despite the discontinuation of CrowdTangle and CooRnet in August 2024, coordination analysis remains possible with the new R package CooRTweet (Righetti and Balluff, 2025), which allows researchers to analyze coordination using any social media data source, including the Meta Content Library.

Coordination typology-making

Our approach to typology-building provides a methodological foundation for the analysis of the concept of coordinated inauthentic behaviour on social media. As said,

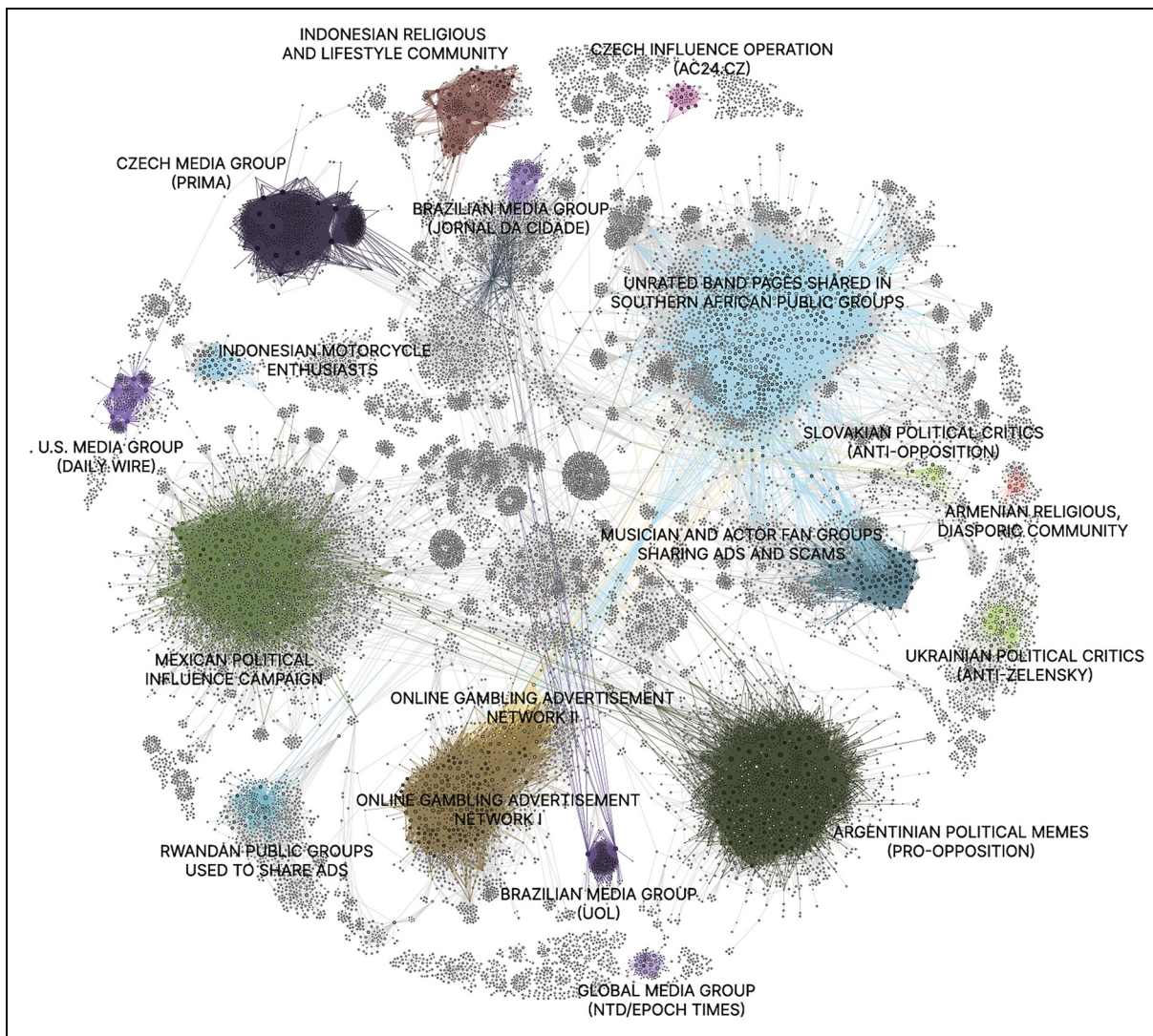


Figure 3. Coordinated link-sharing networks on Facebook, July 2024. Graphic by Alessandra Facchin.

the point of departure is a list of networks exhibiting high degrees of coordination of link-sharing on Facebook. We subsequently examine the actors sharing the content. We generally follow Chaffee's framework of concept explication (1991), which starts with a list of relevant items within a higher-level concept without implying relationships or hierarchy and subsequently organises these items into distinct categories based on specific criteria, highlighting relationships and differences and offering a structured framework for further analysis. For Chaffee this process is the foundation for concept explication, or defining and clarifying a concept for its reuse in research.

Following the procedure, we analyse the degree of centralised coordination as well as the actors undertaking such coordination. Then, for our typology building, we describe the actors and compare the degree of centralised coordination. It allows us to make findings concerning how media

groups and influence operations have the same coordination signature.

One cautionary note is that our typology has the limitation that it is of one particular time frame and platform and composed of only those networks exhibiting a very high degree of coordination. As with others, it may be a time-bound typology (Chaffee, 1991), especially given it comes from a platform where content moderation is particularly active. It is also the very first empirically-based effort at sorting coordinated behaviour types beyond disinformation and thereby could be considered preliminary.

To create a typology of coordinated behaviour, qualitatively, we characterise the entities sharing media in each network. Quantitatively, we study the degree of coordination through sharing.

Examining 19 communities, we first create thick descriptions of them based on the actor types and the media sharing, including the contents. Concentrating on the actors and

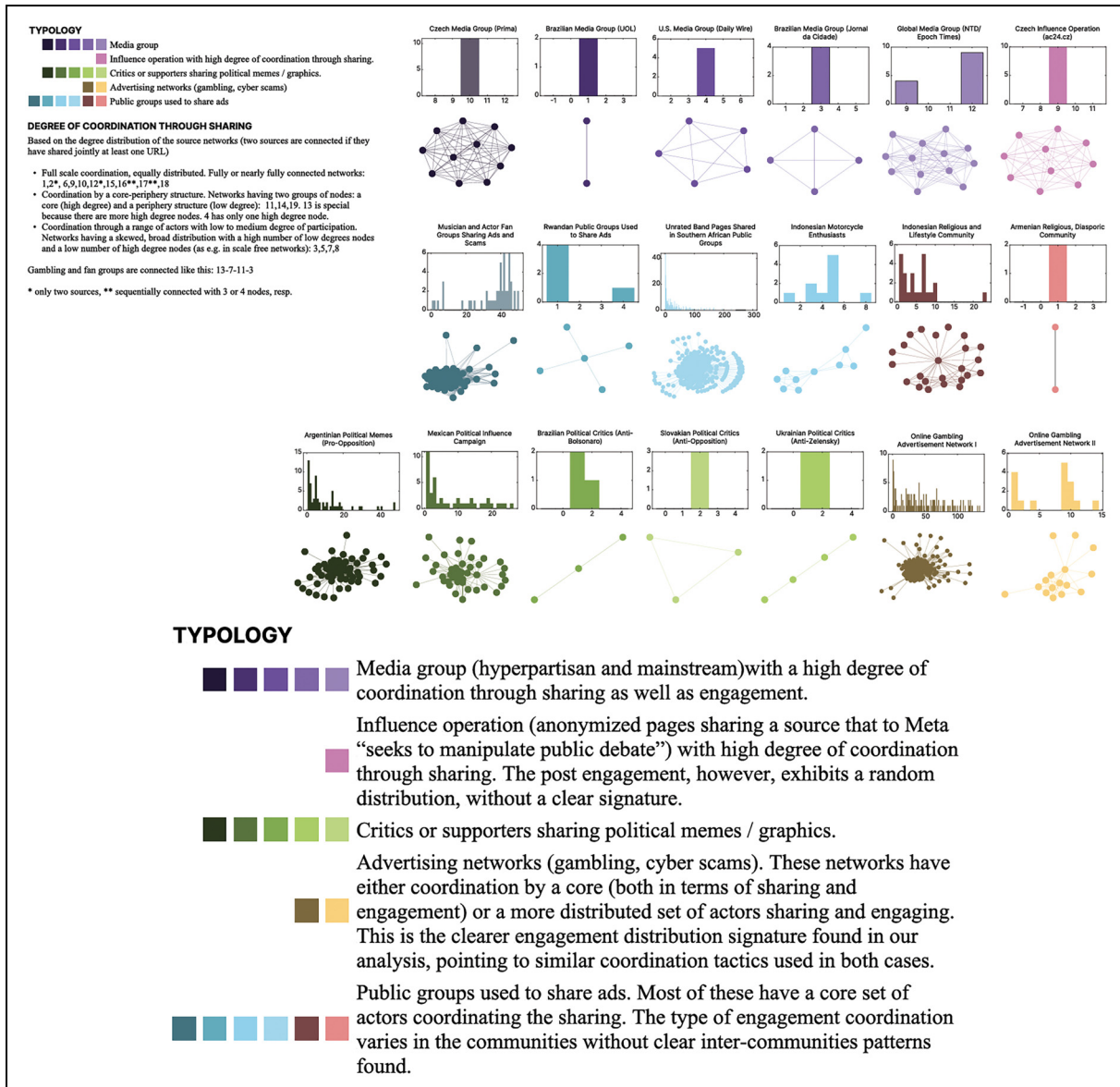


Figure 4. Typology of coordinated link-sharing networks on Facebook, July 2024. Graphic by Alessandra Facchin.

the media being shared, we subsequently characterise six types: media groups (from the tabloid to the mainstream sharing stories), advertising networks (sharing online gambling and cyber scams), large public groups (such as TV channels) used to share various kinds of ads, critics or supporters of politicians (sharing memes / graphics) as well as an influence operation, where a set of anonymous pages all share content from a source classified as disinformation (see Figure 4).

The quantitative analysis yielded degrees of coordination, from full scale coordination (all sharing all) to coordination by a core to a broad distribution of sharing participants. Certain attributes are common to communities, where for example the media groups as well as the influence operation have a high degree of coordination. The type,

critics or supporters of politicians sharing memes and graphics, has a lower degree of distribution; in that case, the Brazilian, Ukrainian and Slovakian communities have a core set of participants, while the other two have different signatures. One stands out. The pro-Amló Mexican community has a broad distribution of identifiable contributors and might be a case of ‘participatory propaganda’ (Wanless and Berk, 2020), where there is bottom-up, grassroots-like coordination. The online gambling communities are among the largest communities, where one is driven by a core group and another by a broader distribution.

With respect to the question of what constitutes inauthentic behaviour, we have found a continuum. One network has a set of actors who do not identify themselves and coordinate sharing (the Czech influence operation).

Another network operates Pages with seemingly grassroots names, yet all share the posts of its source (the U.S. Daily Wire Media Group). Given the lower degree of centralisation, the pro-Amlo community's sharing appears to be more organic.

Whether the analysis of the actors and what they are sharing or the measures with respect to the degree of coordination should form the basis of the typology-making is one starting point of the discussion, for they do not completely overlap, with certain exceptions. What we characterised as an influence operation and what we characterised as the operations of a media group have the same 'signature'; they both have full-scale, equally distributed coordination through sharing. That newsrooms and fake newsrooms coordinate sharing in the same manner is one significant finding. Another is the continuum of inauthentic behaviour where sharing is performed by accounts which have varying degrees of anonymity and disclosure. The influence operation has the greatest anonymity and least disclosure, for it is anonymised; next on the spectrum is the U.S. media group, which identifies itself but does not disclose its link to the master source. Identifying oneself without disclosing the link to the campaign or group behind the coordination is common to many of the advertising networks. On the other end of the spectrum are the organic sharers of pro-Amlo material who identify themselves and openly disclose their support for the political cause.

Facebook's definitional turns

As one journalistic account put it, coordinated inauthentic behaviour (CIB) may "sound technical and objective" (Douek, 2020), but as we have found it is contingent on two determinations, which have been subject to changing emphases by Facebook (later Meta), which coined the term (Gleicher, 2018). One determination is the authenticity of coordination and the other is that of the users. In the following we discuss these aspects of the notion from Meta, where the behaviour may result in user removal from its platform.

As we related, we found a range or continuum of what could be considered inauthentic coordination and actors. Having operationalised the study of CIB using Facebook data collected for its detection, we found some nineteen networks of actors coordinating their activities to a very high degree, which we then developed into a typology of coordinating behaviour. Moreover, one type of network (influence operation) has accounts that do not identify themselves, and another does not disclose their relationship to a master source they amplify (Daily Wire). As we discuss, the change in emphasis of Meta's definition of CIB would have the implication that the Daily Wire would not be identified or targeted by the platform. Put differently, according to the definitions put forward by Facebook, certain of these networks would have been removed some

years ago, but given Meta's most recent definition, however, fewer would qualify for a takedown. We discuss the implications of the narrowing of the definition by way of conclusion.

Facebook's discussion of coordination as well as inauthentic activities that abuse its platform has evolved over the years. The changes have been subtle. Initially the efforts connected coordination with the spreading of fake or 'false news', as Facebook termed it. Later, it would refer to a separate activity regardless of content. Especially during election seasons, Meta emphasises in its more recent policy rationale documents that these "standards apply agnostic of content, political or otherwise" (Meta, 2022).

The breadth of those committing the offending acts of coordination also narrowed. For example 'spam' networks were once counted (and taken down), fitting the definition of CIB as "people or organisations working together to create networks of accounts and Pages to mislead others" (Gleicher, 2018). Some five years later the emphasis would rest with 'adversarial' actors.

In a sense this simplification (or further specification) of coordinated inauthentic behaviour has allowed Meta to absorb criticism that their takedown policies are politically imbalanced; it also enables certain pages and accounts to remain online, such as borderline commercial activities and hyper-partisan networks of sources, which receives reportedly high levels of engagement won in part through coordination (Christin et al., 2024). Daily Wire is a case in point. In 2019 it was accused of running a network of Facebook Pages to "exclusively promote content from the Daily Wire" without revealing these Pages' affiliation with it (Legum, 2019). The journalistic piece demonstrated the effect of the amplification work, showing how it significantly outperforms in engagement measures other major media outlets.

From 'fake' to 'adversarial'

While there are earlier mentions of CIB-like terminology, Facebook's 2017 White Paper, entitled Information Operations and Facebook, is the first to discuss it in any depth (Weedon et al., 2017). It was published on the heels of the uproar about the platform's 'fake news' problem, discovered six months earlier by data journalists at BuzzFeed News (Silverman, 2016). In that widely reported work, imposter and hyper-partisan news sites (as their 'fake news' definition had it) outperformed mainstream news on Facebook just prior to the U.S. presidential elections. The fake news received more interactions (likes, shares and comments) than mainstream sources for election-related posts.

The Facebook company White Paper was the first and arguably its most elaborate reaction to the problem, for it sought to address concerns that Facebook, as argued in governmental circles in the US and Europe at the time, 'was not

doing enough' (Ahmed, 2017; Clifford, 2017; Eddy and Scott, 2017; Lomas, 2017). In the paper, it first dismisses 'fake news' as a term 'overused and misused', introducing 'false news' in its place. It defines it as news that contains "intentional misstatements of fact" intended to "arouse passions, attract viewership, or deceive" (Weedon et al., 2017, p. 5). After redefining 'fake news' as 'false news', the paper subsequently introduces related terminology, 'false amplifiers' that the authors from the Facebook Security team explain expands how they conceive of 'abusive behaviour' on the platform from "account hacking, malware, spam and financial scams" to "attempts to manipulate civic discourse and deceive people" (Weedon et al., 2017, p. 3).

As with the 2016 fake news problem, when the Russian Internet Research Agency infiltrated Facebook and ran a covert campaign, these manipulations are considered information (or influence) operations, or what the authors point out are the kinds of activities at the "heart of the [white] paper" (Weedon et al., 2017, p. 4). These are "actions taken by organised actors (governments or non-state actors) to distort domestic or foreign political sentiment" (Weedon et al., 2017, p. 5).

But false amplifiers and amplification are not only influence operations of the kind that took place in 2016 and beyond. Some are broader than that. False amplifiers the authors define as "coordinated activity by inauthentic accounts with the intent of manipulating political discussion" (Weedon et al., 2017, p. 5). In a section about what 'false amplification looks like' the authors describe coordinated "sharing of content and repeated, rapid posts"; "repeated comments", or "likes"; "astroturf groups"; "Pages with the specific intent to spread sensationalistic or heavily biased news or headlines"; and "inflammatory and sometimes racist memes, or manipulated photos and video content" (Weedon et al., 2017, p. 9).

Terminologically speaking, these "false amplifiers" became "coordinated inauthentic behaviour" a year later (Gleicher, 2018). There is a Facebook video from 2018, part of Facebook's community standards webpages, where the head of security explains that coordinated inauthentic behaviour 'is when groups of pages or people work together to mislead others about who they are or what they're doing' (Gleicher, 2018). An important point is that it concerns 'behaviour' rather than the content itself, separating it from 'false news'. This makes such behaviour content-agnostic.

Between 2018 and early 2023 (when Meta archived its coordinated inauthentic behaviour blog posts), the platform repeatedly reported about CIB-related takedowns, eventually bundling these into quarterly reports (still published at the time of writing in late 2024) about what they call 'adversarial threat actors'. Since 2022 CIB, principally, occurs "[when] adversarial threat actors use fake accounts to engage in sophisticated inauthentic tactics in order to influence public debate" (Meta, 2022). Here the emphasis, in other words, concerning the actor types shifts from

"people and organisations" more generally to 'adversarial' or hostile groups, narrowing the actor types undertaking inauthentic coordination.

The narrowing of the definition of actor engaged in coordinated inauthentic behaviour, coupled with the removal of types of content considered as typical or illustrative of it, reduces the term's coverage. It thereby becomes what one scholar has dubbed a 'strategically ambiguous concept' whose evolution and changing application enables coordinated behaviour to remain on-platform (Graham, 2024).

Conclusions: 'authentic enough'

Drawing on an empirical study of coordination behaviours sourced on Facebook with the analytical technique, we identify examples of the purposes and develop a typology of coordination. In doing so, we are interested in the extent to which current techniques of attention manufacture are similar across coordination types. More quantitatively, the manner in which we discuss similarity is through degrees of centralisation of coordination. Do all the actors pushing out links within a short time frame share them all or only some of them? Could high degrees of centralised coordination be considered a marker of inauthentic coordinated behaviour, or is it rather (or additionally) a sign of optimised platform marketing?

More qualitatively, we examine the entities or actors sharing the URLs. We found some nineteen networks of actors exhibiting high degrees of coordination sharing a variety of material from news stories, political content and a wide range of ads (from religious products and services to gambling). Does the actor type or content shared provide a strong marker for CIB? Are there particular combinations of centralised coordination and actor as well as content types indicative of CIB? Finally, we conjoin the two outlooks in the typology, identifying five types of coordination: influence operations, media groups sharing news stories, political activists sharing memes, advertising networks promoting gambling and cyber scams as well as large public groups hijacked to spread ads. In the discussion of authenticity, we draw on Meta's (evolving) understandings, where questions of the intention of the coordination as well as the identity of the actors have been central.

By way of conclusion, the analysis of CIB by the EU Disinfo Lab, the non-governmental organisation, is helpful, for it describes how EU legislation and platform companies discuss it (EU Disinfo Lab, 2024). The term, coordinated inauthentic behaviour, is mentioned in key European legislation, including the Digital Services Act. Platforms may be abused when actors manipulate their algorithms through artificial amplification, which should be curtailed. EU legislation largely leaves the operationalisation of CIB detection to the platforms. With the exception of Facebook (later Meta), CIB is not well described or defined by other social media platforms under consideration, including Google

(Web Search), YouTube and X/Twitter. As we have discussed above, Meta has made great efforts to both define and provide examples of CIB, most recently emphasising ‘adversarial actors’ as those behind inauthentic behaviour. It thereby associates CIB mainly with influence operations, where actor identities are faked or obscured.

Thus, those campaigns driven by coordination efforts undertaken by accounts deemed non-adversarial could be considered ‘authentic enough’ (Lindquist, 2021; Lindquist and Weltevrede, 2024). This is the main implication of the finding that a spectrum of actors employing boosting strategies for their content remain on-platform.

A second implication of that finding concerns how to address high-degree coordination. Is it to be viewed as false amplification or a more everyday practice to circulate content? It is a pertinent question, given the platform architecture, especially how success is measured by metrics that invite engagement management of one form or another. Put more forcefully, the manner in which the platform is designed for highlighting content could be said to invite coordination, as the variety of actors practicing it testifies.

This work has been funded by the European Union under the Horizon Europe vera.ai project, Grant Agreement number 101070093.

Acknowledgements

The authors would like to thank the participants in the project, Coordinated Inauthentic Behaviour on Facebook? at the 2024 Digital Methods Summer School, Media Studies, University of Amsterdam.


Funding


The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the HORIZON EUROPE Digital, Industry and Space, grant number vera.ai, 101070093.

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

ORCID iDs

Richard Rogers  <https://orcid.org/0000-0002-9897-6559>

Nicola Righetti  <https://orcid.org/0000-0002-9257-5113>

References

- Ahmed K (2017) *Zuckerberg: My Facebook Manifesto to Re-boot Globalisation*. BBC News.
- Benkler Y (2018) *Network Propaganda: Manipulation, Disinformation, and Radicalization in American Politics*. Oxford: Oxford University Press.
- Bennett WL and Segerberg A (2013) *The Logic of Connective Action: Digital Media and the Personalization of Contentious Politics*. Cambridge: Cambridge University Press.
- boyd d (2010) Social network sites as networked publics: Affordances, dynamics, and implications. In: *A Networked Self*. London: Routledge, 47–66.
- Brunton F (2013) *Spam: A Shadow History of the Internet*. Cambridge, MA: MIT Press.
- Burton AG, Chun WHK, Bounegru L, et al. (2023) *Algorithmic Authenticity: An Overview*. Leuphana: Meson Press.
- Cambridge Dictionary. (2024) *Coordinate*.
- Canevez RN, Maikovska K and Zwarun L (2024) Tactics and affordances in the mediatization of war: Pro-Ukrainian cyber resistance on telegram. *Digital War 5*: 1–14.
- Chaffee SH (1991) *Explication, Communication Concepts*. London: Sage Publications.
- Chan J (2024) Online astroturfing: A problem beyond disinformation. *Philosophy & Social Criticism* 50(3): 507–528.
- Christensen HS (2011) Political activities on the Internet: Slacktivism or political participation by other means? *First Monday* 16(2).
- Christin A, Bernstein MS, Hancock JT, et al. (2024) Internal Fractures: The Competing Logics of Social Media Platforms. *Social Media + Society* 10(3). <https://doi.org/10.1177/20563051241274668>.
- Clifford C (2017) *Mark Zuckerberg shares his strategy for leading Facebook through controversy*.
- De Maeyer J (2013) Towards a hyperlinked society: A critical review of link studies. *New Media & Society* 15(5): 737–751.
- DiResta R, Shaffer K, Ruppel B, et al. (2018) *The Tactics and Tropes of the Internet Research Agency* [Report prepared for the United States Senate Select Committee on Russian Interference in the 2016 Election]. *New Knowledge*.
- Disinfo Lab EU (2024) *Coordinated Inauthentic Behaviour detection tree*.
- Douek E (2020) *What Does “Coordinated Inauthentic Behavior” Actually Mean? There’s no clear definition, and that’s worrisome*. Slate.
- Durkheim É (2014) *Les formes élémentaires de la vie religieuse: Le système totémique en Australie*. Paris: Puf.
- Eddy M and Scott M (2017) *Facebook and Twitter Could Face Fines in Germany Over Hate Speech Posts*.
- Foster SL (2003) Choreographies of protest. *Theatre Journal* 55(3): 395–412.
- Gerbaudo P (2012) *Tweets and the Streets: Social Media and Contemporary Activism*. London: Pluto Press.
- Giglietto F, Graham T and Righetti N (Forthcoming) Navigating coordination and inauthentic behaviour: Challenges and innovations in social media detection. In *Routledge Companion to Social Media and Politics*. London: Routledge Taylor & Francis Group.
- Giglietto F, Marino G, Mincigrucci R, et al. (2023a) A workflow to detect, monitor, and update lists of coordinated social Media accounts across time: The case of the 2022 Italian election. *Social Media+ Society* 9(3): 20563051231196866.
- Giglietto F, Mincigrucci R and Marino G (2023b) *Detecting Coordinated Link Sharing Behaviour with CooRnet*. *Digital*

- Methods Winter School*. Digital Methods Winter School, Amsterdam. https://github.com/fabiogiglietto/CooRnet_at_DMI_WS_2023
- Giglietto F, Olaniran S, Mincigrucchi R, et al. (2022) Blowing on the fire: An analysis of low quality and hyper partisan news sources circulated by coordinated link sharing networks in Nigeria. Available at SSRN 4162030. <https://doi.org/10.2139/ssrn.4162030>.
- Giglietto F, Righetti N and Rossi L (2020a) *CooRnet. Detect Coordinated Link Sharing Behavior on Social Media* [Computer software]. <https://github.com/fabiogiglietto/CooRnet>
- Giglietto F, Righetti N, Rossi L, et al. (2020b) It takes a village to manipulate the Media: Coordinated link sharing behavior during 2018 and 2019 Italian elections. *Information, Communication & Society* 23(6): 867–891.
- Giglietto F, Righetti N, Rossi L, et al. (2021) CooRnet: An Integrated Approach to Surface Problematic Content, Malicious Actors, and Coordinated Networks. *AoIR Selected Papers of Internet Research*. <https://doi.org/10.5210/spir.v2021i0.12170>.
- Gleicher N (2018) *Coordinated Inauthentic Behavior Explained*. Facebook Newsroom. Facebook. <https://about.fb.com/news/2018/12/inside-feed-coordinated-inauthentic-behavior/>.
- Goldhaber MH (1997) The attention economy and the net. *First Monday* 2(4).
- Graham T (2024) *The 'Inauthenticity' Paradox: How Platforms Profit From And Shape Coordinated Inauthentic Behaviour*.
- Graham T, Bruns A, Angus D, et al. (2021) #IStandWithDan versus #DictatorDan: The polarised dynamics of twitter discussions about Victoria's COVID-19 restrictions. *Media International Australia* 179(1): 127–148.
- Graham T, Hames S and Alpert E (2024) The coordination network toolkit: A framework for detecting and analysing coordinated behaviour on social media. *Journal of Computational Social Science* 7: 1–22.
- Gruzd A, Mai P and Soares FB (2022) How coordinated link sharing behavior and partisans' narrative framing fan the spread of COVID-19 misinformation and conspiracy theories. *Soc Netw Anal Min* 12(1): 18.
- Gruzd A, Soares FB and Mai P (2023) Trust and safety on social Media: Understanding the impact of anti-social behavior and misinformation on content moderation and platform governance. *Social Media+ Society* 9(3): 20563051231196878.
- Howard PN, Ganesh B, Liotsiou D, et al. (2018) *The IRA, social media and political polarization in the United States, 2012-2018*.
- Keller FB, Schoch D, Stier S, et al. (2020) Political astroturfing on twitter: How to coordinate a disinformation campaign. *Political Communication* 37(2): 256–280.
- Kulichkina A, Righetti N and Waldherr A (2024) Protest and repression on social media: Pro-Navalny and pro-government mobilization dynamics and coordination patterns on Russian Twitter. *New Media & Society*.
- Legum J (2019) *Facebook Allows Prominent Right-wing Website to Break the Rules*. Popular Information.
- Lindquist J (2021) Good enough imposters: The market for Instagram followers in Indonesia and beyond. In: Woolgar S, Vogel E, Moats D, et al. (eds) *The Imposter as Social Theory: Thinking with Gatecrashers, Cheats and Charlatans*. Bristol: Bristol University Press, 269–291.
- Lindquist J and Weltevrede E (2024) Authenticity Governance and the Market for Social Media Engagements: The Shaping of Disinformation at the Peripheries of Platform Ecosystems. *Social Media + Society* 10(1). <https://doi.org/10.1177/20563051231224721>.
- Lomas N (2017) *Facebook's content moderation rules dubbed 'alarming' by child safety charity*.
- Magelinski T and Carley KM (2020) Detecting coordinated behavior in the Twitter campaign to Reopen America. *Center for Informed Democracy and Social-Cybersecurity Annual Conference, IDeaS*.
- Mannocci L, Mazza M, Monreale A, et al. (2024) Detection and characterization of coordinated online behavior: A survey. *arXiv preprint arXiv:2408.01257*.
- McIntyre L (2018) *Post-truth*. Cambridge, MA: MIT Press.
- Merrill JB and Oremus W (2021) Five points for anger, one for a 'like': How Facebook's formula fostered rage and misinformation. *The Washington Post*, 26.
- Messing S, DeGregorio C, Hillenbrand B, et al. (2020) *Facebook Privacy-Protected Full URLs Data Set*.
- Meta (2022) *Inauthentic Behaviour*.
- Nizzoli L, Tardelli S, Avvenuti M, et al. (2021) Coordinated behavior on social Media in 2019 UK general election. *Proceedings of the International AAAI Conference on Web and Social Media* 15: 443–454.
- Righetti N and Balluff P (2025) CooRTweet: A generalized R software for coordinated network detection. *Computational Communication Research* 7(1): 1.
- Righetti N and Bertuzzi N (2020) Digital animal advocacy: A study on Facebook communication styles of Italian animal rights organizations and their Followers' reactions. *Mediascapes Journal* 16: 128–150.
- Righetti N, Giglietto F, Kakavand AE, et al. (2022) *Political Advertisement and Coordinated Behavior on Social Media in the Lead-Up to the 2021 German Federal Elections*. https://www.medienanstalt-nrw.de/fileadmin/user_upload/NeueWebsite_0120/Zum_Nachlesen/BTW22_Political_Advertisement.PDF
- Rogers R (2017) Digital methods for cross-platform analysis. In: *The SAGE Handbook of Social Media*. London: Sage, 91–110.
- Rogers R (2018) Otherwise engaged: Social media from vanity metrics to critical analytics. *International Journal of Communication* 12: 23.
- Rogers R (2023) *The Propagation of Misinformation in Social Media: A Cross-platform Analysis*. Amsterdam: Amsterdam University Press.
- Roose K (2020) How The Epoch Times Created a Giant Influence Machine. *International New York Times*, NA-NA.

- Silverman C (2016) *This analysis shows how viral fake election news stories outperformed real news on Facebook*. BuzzFeed News. <https://www.buzzfeednews.com/article/craigsilverman/viral-fake-election-newsoutperformed-real-news-on-facebook>
- Terenzi M (2023, October 18) Exploring the Dark Side of Cryptocurrencies on Facebook and Telegram: Uncovering Media Manipulation and “Get-Rich-Quick” Deceptive Schemes. In: *Selected Papers of #AoIR2023: The 24th Annual Conference of the Association of Internet Researchers*.
- Thiele D, Milzner M, Heft A, et al. (2023) 50 Shades of Astroturf and How to Choose Them: A Theoretical Model and Review of Coordinated Social Media Manipulation. In: *73th Annual ICA Conference*.
- Tilly C (1994) Social movements as historically specific clusters of political performances. *Berkeley J Sociol* 38: 1–30.
- Wanless A and Berk M (2020) The audience is the amplifier: Participatory propaganda. In: *The SAGE Handbook of Propaganda*. London: Sage, 85–104.
- Warner M (2002) Publics and counterpublics (abbreviated version). *Quarterly Journal of Speech* 88(4): 413–425.
- Weedon J, Nuland W and Stamos A (2017) *Information Operations and Facebook*. Facebook.
- Yu WES (2022) A framework for studying coordinated behaviour applied to the 2019 philippine midterm elections. Proceedings of Sixth International Congress on Information and Communication Technology: ICICT 2021, London, 2, 721–731.
- Zhang Y, Ruan X, Wang H, et al. (2016) Twitter trends manipulation: A first look inside the security of twitter trending. *IEEE Trans Inf Forensics Secur* 12(1): 144–156.