Journal of Sociocybernetics

Journal of Sociocybernetics ISSN 1607-8667

ORIGINAL ARTICLE

Modeling persuasion in social media: a theoretical approach to algorithmic content distribution and manipulation

Massimo Terenzi
University of Urbino, Dipartimento di Scienze della Comunicazione

This paper proposes a theoretical model integrating sociocybernetics and persuasive technology to analyze the systemic dynamics of content dissemination and manipulation on social media platforms. Recommendation algorithms, conceptualized as active agents within digital communication systems, mediate interactions between users but are also susceptible to exploitation by external actors seeking to amplify deceptive or harmful content. Using a case study on coordinated cryptocurrency-related activities across Facebook and Telegram, this study demonstrates how manipulators exploit algorithmic vulnerabilities to simulate engagement and disrupt platform mechanisms, thereby amplifying manipulative strategies. The study employs computational techniques to detect Coordinated Link-Sharing Behavior (CLSB), revealing how manipulators leverage these dynamics to amplify fraudulent content. The findings highlight the dual role of algorithms as both selectors of relevant information and targets of manipulation, illustrating their broader influence on digital social interactions. This research contributes to the theoretical understanding of algorithms as cybernetic operators. It underscores the importance of implementing robust mechanisms to mitigate manipulative behaviors, aligning with regulatory efforts such as the regulatory framework of the European Union's Digital Services Act (DSA).

url: 10.26754/ojs_jos/jos.2024110992

1. Introduction

When navigating social media, one might easily overlook the underlying mechanisms at work—algorithms that determine the content displayed on your feed. The posts we view, share, and engage with are not random; they are selectively curated by these digital gatekeepers, which profoundly shape our experience and the information we consume. Recommendation algorithms play a significant role in shaping content visibility, determining which stories, ideas, and discussions are brought to the forefront and which are pushed to the periphery of users' feeds (Narayanan, 2023). Yet, with this power comes vulnerability. These algorithms are not foolproof: they can be susceptible to manipulation tactics that exploit their underlying logic, often resulting in the amplification of misleading or deceptive content (Marwick and Lewis, 2017). The expansive nature of the Internet, particularly social media, allows it to host an almost limitless array of topics. This includes everything from everyday social interactions to niche interests, controversial debates, and even misinformation. The absence of a clear, universally applied editorial filter or "selective code" means that virtually any type of content—whether accurate, relevant, or misleading—can find an audience (Gillespie, 2014).

Traditional media rely on editorial standards to guide content selection, ensuring a level of quality and accuracy. In contrast, social media platforms often base content visibility on user engagement metrics and algorithmic decisions. These mechanisms do not inherently distinguish between credible information and falsehoods, creating an environment where misleading or sensational content can thrive. The algorithms, designed to maximize user engagement, typically amplify content that captures attention, regardless of its veracity. This process blurs the distinction between reliable information and manipulation, making it increasingly difficult for users to discern the credibility of what they encounter (Pariser, 2011). As a result, the prioritization of highly engaging content inadvertently increases the visibility of misleading or harmful narratives, allowing bad actors to exploit these systems and push deceptive content into mainstream visibility (Acker, 2018).

Given this complex landscape, this paper will not focus on how individual users experience recommendations from their perspective, which might seem like an exchange between the user and the algorithm. Instead, this work will adopt a broader perspective, concentrating on the macro-level dynamics of content dissemination within the platform. This approach is essential to understanding the full scope of how information circulates on social media, beyond the limited viewpoint of personalized recommendations.

By analyzing systemic content dissemination, the study explores how recommendation algorithms mediate communication, amplifying both genuine and manipulative narratives. A case study on cryptocurrency scams on digital platforms such as Facebook and Telegram will be presented, to demonstrate how such theories provide new insights into the dynamics of online manipulation. This approach highlights the inherent conflict between the goals of social platforms and those of malicious actors who seek to manipulate these platforms for their own purposes, creating a competitive struggle for control over the flow of information.

Finally, the paper will discuss the theoretical and practical implications of these findings, with a particular focus on the European regulatory landscape, as framed by the Digital Services Act (DSA).

2. Ties, networks, and algorithmic systems

Understanding the dynamics of human interaction within complex networks has long been a central focus in the study of social systems, providing a foundation for the evolution of modern social networking platforms. Early work in mathematical sociology, such as White's structural analysis of networks, emphasized that social outcomes are shaped not by individual attributes but by relationships and their structural configurations (White, 1970). Granovetter's "strength of weak ties" further demonstrated how loosely connected individuals serve as crucial bridges between otherwise disconnected communities, enabling the rapid diffusion of information (Granovetter, 1973). These foundational insights from network theory have informed the design of digital platforms, which amplify relational dynamics through recommendation algorithms. These algorithms, originally developed for e-commerce, have been adapted to social media environments, where they now mediate not just individual content preferences but broader patterns of information circulation (Sarwar et al., 2000).

Niklas Luhmann's systems theory complements these frameworks by emphasizing the role of communication in shaping social realities. For Luhmann, the media act not as passive mirrors of society but as active constructors, selectively highlighting certain topics while excluding others (Luhmann, 2000). This selective process, while unavoidable, creates a mediated version of reality, influencing public perception and setting the agenda for societal discourse.

With the advent of the Internet, this mediation process has been disrupted and transformed. Initially envisioned as a "medium of media," the Internet promised a democratically structured space free from traditional gatekeeping, allowing multiple realities to coexist. However, this inclusivity has given rise to what Webster (2014) terms the marketplace of attention, where competing narratives vie for visibility, often at the expense of credibility. As traditional media selectively shaped societal discourse, the advent of algorithms has intensified this process by automating content curation based on user engagement metrics and embedding selective mechanisms into digital communication systems (Gillespie, 2018).

Sociocybernetics offers a valuable lens for understanding these dynamics. For example, Luhmann's concept of 'operational closure' is particularly relevant here: algorithms process external inputs (user interactions) but produce outputs (content recommendations) based on their internal logic and goals. This self-referential process creates a selective mechanism that amplifies certain topics while excluding others, effectively constructing a mediated version of digital reality (Beer, 2017).

As discussed in the following sections, the literature identifies at least two ways to interpret the relationship between users and content on social media. The first perspective positions the user as a direct interlocutor of the algorithm, where recommendations are tailored to individual preferences and behaviors. From this vantage point, the algorithmic system addresses the user personally, curating content that aligns with their tastes, interests, and specific digital footprints.

However, a second perspective shifts the focus to the macro level, examining recommendations as platform-wide processes that shape the visibility and dissemination of content across broader networks. In this context, content is not merely consumed but also reinterpreted and recontextualized by users through collective engagement. This dual dynamic—personalized consumption and collaborative reinterpretation—highlights the participatory role of users within algorithmically mediated environments, where individual agency interacts with systemic patterns of information circulation.

User-centric recommendations

When focusing on the personal user experience, algorithms operate on the basis of detailed profiling of user preferences and behavior. Algorithms collect data from user interactions, such as likes, comments, and time spent on certain content, to build a detailed profile that allows the feed to be personalized. For example, a user who frequently interacts with cooking-related content will see a majority of posts on recipes, cooking utensils, and cooking videos in their feed. This type of personalization aims to keep the user on the platform longer, maximizing engagement (Solsman, 2018; Hern, 2022). In other words, the algorithm suggests content that is designed to resonate with the user's individual preferences, creating a highly personalized and tailored experience. Scholars seem to agree that at least two main forms of recommendation can be distinguished: the first related to collaborative filtering and the second to content-based recommendations (e.g., Bobadilla et al. 2013; Bozdag 2013). Between the two poles of this axis is a continuum of hybrid approaches that integrate both filtering methods. The content-based filtering method is based on the popularity of the content among users and is the simplest approach. In this model, what is deemed popular tends to become increasingly popular, as the visibility of a piece of content further increases its popularity among other users, triggering a cycle of exponential growth. This mechanism favors a rapid growth in the popularity of a piece of content, leading it to become relevant in a short period of time. Conversely, content that loses popularity is destined to gradually disappear as users find it less and less interesting or worthy of attention. At the opposite end of this continuum is so-called collaborative filtering, a process that works similarly to word-of-mouth. This recommendation mechanism flags and suggests content based on the values assigned to that content by networks of people, communities, or users with similar profiles and interests. The algorithm, therefore, divides users into communities of interest and assigns content that is liked by a member of the community, or several members, extending it to other users of the same community, with the assumption that such content may also be of interest to them. In this way, collaborative filtering exploits the dynamics of affinities between users to optimize content distribution, creating a recommendation system based on shared collective preferences (Batmaz et al. 2019).

Platform-wide recommendations

On the other hand, by broadening the point of observation and placing the focus on content, the algorithm's logic shifts towards a global recommendation of content across the entire platform. This shift underscores the importance of analyzing content circulation at the macro level rather than just focusing on individual user interactions. By adopting this broader perspective, the analysis moves beyond the micro-level recommendation mechanisms, such as collaborative filtering and content-based recommendations, and instead captures the systemic dynamics that influence how content is disseminated across wider audiences. Building on this, Webster (2014, p. 12) emphasizes that understanding these broader audience dynamics is crucial for analyzing how attention is distributed across media platforms. Interactions between users can produce mass behavior that influences content flow and visibility, which cannot be fully explained by focusing solely on individual user actions (Webster 2014, p. 12). Irrespective of individual user preferences, algorithms are designed to identify which content is rapidly gaining resonance by analyzing patterns of growth of interactions on specific topics. For example, a rapid growth of interest in a certain topic may lead algorithms to further promote that content, exploiting its potential virality. The novelty of the content may be another significant selection criterion: fresh and up-to-date content usually tends to receive more attention than older content.

This macro-level perspective aligns with the broader dynamics of mediated communication as discussed by Tilak and Glassman (2020). They highlight the Internet's dual role as a space for both hierarchical control and participatory engagement, framing it as comprising a "first-order Internet" dominated by top-down structures and a "second-order Internet" driven by

non-hierarchical, collaborative interactions. Within platform-wide recommendation systems, this duality becomes particularly evident. Algorithms, in their attempt to maximize engagement, often trivialize interactions by prioritizing superficial connections, effectively creating what Tilak and Glassman describe as "surface communities." These are networks bound by shared consumption rather than meaningful dialogue or critical discourse. However, in the framework of the second-order Internet, they also note the potential for creating alternative "lifeworlds," where users collaboratively reinterpret and disseminate content, often challenging the dominant narratives imposed by algorithmic selection mechanisms (Tilak & Glassman, 2020).

In this sense, platform algorithms can simultaneously reinforce hierarchical control—through the commodification of interactions and the prioritization of trivial ties—and open pathways for bottom-up reinterpretation of content. Interactions at scale, beyond individual recommendations, shape systemic patterns of content circulation. For example, as algorithms detect and amplify topics gaining rapid traction, they reinforce emergent trends while simultaneously creating opportunities for alternative narratives to gain visibility, albeit within the constraints of platform logic. Extending the analytical focus from individual user behavior to these macro-level interactions allows for a critical assessment of how platforms mediate both the flow of information and the conditions that either foster participatory engagement or enable its appropriation by dominant structures.

3. The problem of content relevance on social media

The assessment of relevance—that is, the process by which algorithms must identify, among trillions of particles of information, the content that meets the criteria of a set of users and platform targets—is by no means simple. These calculations, although they have evolved and complexified over time, have never been trivial. This is a particularly contentious issue for sociologists, as relevance is a fluid concept, open to interpretation and closely linked to what is considered newsworthy or popular (Couldry & Hepp, 2017; Bucher, 2018; Van Dijck, 2013). There is no independent and neutral metric to determine what is most relevant, so it is up to the platform builders to define these criteria and modify the algorithms so that they distribute results in accordance with these principles. As Gillespie notes, there is no such thing as a bias-free algorithm, as this would presuppose the existence of a completely unbiased algorithm, which is clearly impossible (Gillespie, 2014). The main problem in defining relevance in platform algorithms lies in the fact that the criteria and assumptions on which they are based generally remain hidden. How these criteria are defined and weighted by engineers remains a secret of the platform, a secret that may incorporate commercial or political benefits, operating below the threshold of user awareness. Making these criteria visible would allow competitors to copy, surpass, or improve the service offered while requiring technical explanations that users would find difficult to understand due to the advanced skills required. More importantly, this would provide tools to those seeking to manipulate the system and to hack it more effectively (Gillespie 2014, p. 10). From a purely commercial point of view, it is evident that algorithms are designed to meet the needs and objectives of all those institutions seeking to capitalize and generate revenue through their use. This reality, however, becomes extremely complex, making it difficult to accurately discern the economic and social implications of this. Each algorithm is structured based on organizational and, in many cases, political principles that govern its operation. Algorithms take different approaches to content, determining which of them should become popular. In the past, it was considered that content that was already popular was deserving of further visibility, as was evident in early search engine studies. However, the criteria used by algorithms have multiplied over time, becoming increasingly complex and articulated. Nevertheless, the fact remains that what is considered popular by the algorithms is decided upstream by those who design them. Users generally do not question how algorithms work, as they perceive them as neutral tools, serving the platform experience. As a result, algorithms are treated as ‘black boxes’, i.e. systems whose inner workings are opaque and unclear (Gillespie 2014, p. 12).

Furthermore, it is important to consider that although the workings of algorithms remain inscrutable, they are constantly and invisibly changing over time. These changes occur gradually and subtly, thus avoiding too radical a change that could be felt by users. Algorithms, therefore, are extremely malleable entities, constantly evolving through a series of A/B tests, in which different results are presented to users to understand which one works best in terms of engagement and interaction (Gillespie 2014, pp. 12-13). This is especially true for new deep learning techniques, which allow algorithms to adapt over time based on interaction with content and users, making them dynamic and constantly evolving (Batmaz et al., 2019). In algorithm design, the issue of diversity becomes relevant, as many recommendation systems include elements of surprise and serendipity (Möller et al., 2018). Although some scholars believe that serendipity is a fundamental element for all recommender systems, its functioning within algorithms represents a rather complex mechanism to study and implement. As a rule, serendipity is based on the inclusion of elements that are not chosen in the manner described above but are selected randomly. This happens for several reasons. Firstly, to introduce elements of diversity and ‘disturbance’ into the flow of similar content, interrupting a steady flow that might otherwise cause the user’s attention to wane. Secondly, these elements serve a fundamental principle of algorithms: allowing the algorithms themselves to learn. By introducing novel elements, algorithms can better identify users’ preferences, refine their profiles, and improve their ability to understand which content is most engaging. However, integrating and operationalizing the concept of serendipity within algorithms is far from simple. The reason lies in the difficulty of reaching a consensus on what should constitute serendipity and how it should be balanced within the system. A constant balance has to be struck between the accuracy of recommendations and the introduction of novelty to maintain the effectiveness of the system without compromising the user experience (Möller et al., 2018, p. 4).

The problem of recommendation accuracy is often confronted with the promise of algorithmic objectivity, as algorithms are supposed to act as truth stabilizers, ensuring accurate evaluations free of error, bias, or subjectivity. However, automatic as they are, algorithms cannot completely disregard human intervention, particularly the platforms that run them, especially in the context of social media (Gillespie, 2014). Indeed, there is content that algorithms could restrict or prevent from being distributed, such as potentially harmful content, such as problematic information that incites violence, bullying, or other illegal practices. During significant events, such as the COVID-19 pandemic or political elections, many platforms have taken specific measures, such as tagging content with health or political information, and directing users to verified sources. It is the task of the algorithm managers to define a set of values and standards that give legitimacy to the algorithms, ensuring that they meet criteria of accuracy and truthfulness, without neglecting the needs of advertisers, who need to promote their content through advertising. It is also crucial that the algorithm is clearly explained to the public so that the goodness of its results is legitimized (Giglietto et al., 2022a). Ultimately, providers must ensure that their algorithms are impartial and objective in delivering results. In the past, the selection of content and topics was the prerogative of journalists, whose objectivity depended on a set of journalistic standards learned during their careers. This professional ethic was supposed to limit the influence of personal bias or political convictions while respecting a well-defined editorial line. In the case of algorithms, on the other hand, the promise of objectivity is theoretically infused in a mechanical, computer-based neutrality embedded in the circuits of the machine itself. Neutrality, in this context, should be codified within the algorithms. However, as we know, this is not always the case (Tufekci, 2015). It is not always clear, and often impossible to explain to users, how machines and algorithms in particular advertise, disseminate, or promote content.

In some cases, platforms have been accused of influencing their algorithms by manually altering the organic reach and circulation of certain content to the detriment of others. This manual intervention was not codified within the algorithm and was not based on criteria of algorithmic relevance or content value according to standardized measures of ‘newsworthiness’, but rather responded to the logic decided by the guardians of the platforms. An emblematic example is the case of TikTok's ‘heating button’, in which human operators gave a popularity boost to specific content, following criteria that were not embedded in TikTok’s algorithm, but reflected the personal decisions of those who ran the platform (Baker-White, 2023). Another example concerns the controversy over Twitter (now X), where the content of Elon Musk, the new owner, was promoted more visibly than others (Schiffer & Newton, 2023). The intervention raised ethical concerns among employees and users, highlighting how the power to control algorithms can be used to manipulate the visibility of content, challenging the promise of neutrality and objectivity that algorithms are supposed to guarantee. In such cases, potentially problematic content can be promoted, fuelling human bias and leading to the dissemination of ideologically driven content. These manual interventions undermine the trust in neutrality and impartiality that algorithms should guarantee, highlighting how, despite the promise of objectivity, algorithmic processes are often influenced by subjective human decisions.

4. How is information disseminated on social media?

The decision to focus on analyzing audiences rather than individual user behavior stems from the need to understand the broader mechanisms of information propagation on social media. While examining how algorithms recommend content to individual users offers valuable insights, it provides only a partial view of how content circulates on platforms. Focusing solely on the micro level risks reducing the interaction to a one-to-one dialogue between the user and the algorithm, obscuring the larger dynamics at play. To capture the "big picture" of how content moves through social media ecosystems, it is essential to examine how algorithms function at the macro level. This approach allows for a comprehensive understanding of virality, engagement, and the systemic influence of platforms, which align with their businesses' goals, such as maximizing user engagement and ad revenue. In the following section, we will explain how these processes work and how they contribute to the amplification of content across wider audiences.

Traditionally, three different forms of information propagation mechanisms have emerged in the history of algorithms, which were initially designed to distribute content during the early days of social media. The three models are the subscription model, the network model, and the algorithmic model (Narayanan 2023, p. 9). In the subscription model users follow a set of accounts and receive updates from those accounts in their feeds. The second model, the network model, expands on this by showing users not only the content produced by the accounts they follow but also the posts that those users have amplified by engaging with them, for example by sharing them or liking them. The third model, the algorithmic model, represents pure recommendation. In this model, users see only the content that the algorithm predicts will generate the highest engagement, or content that is most likely to elicit interaction from the user. In this form of information propagation, the social network itself plays a less important and often less visible role (Narayanan, 2023). Consider, for example, TikTok and its "For You" page. Here, users spend the majority of their time, and the content they see is simply recommended by an algorithm, largely independent of the people they follow. It is important to note that there isn't just one algorithm that regulates content distribution across platforms. Often, platforms use a combination of all these models. The role of the algorithmic model is becoming increasingly significant across platforms, as users are more frequently exposed to new content through pure recommendation systems. A prime example of this is TikTok, where users are continually presented with fresh content based solely on algorithmic predictions, rather than on their existing social connections or subscriptions. This model also plays a key role in understanding how content virality works. Virality is something that is not directly predictable. It is an emergent property in a complex system, such as a social network site. Virality follows probabilistic models, meaning that some degree of unpredictability is inherent. Naturally, when virality is reduced or suppressed through mechanisms like demotion—such as in cases of shadow-banning[1]—a small intervention by the platform can lead to a drastic drop in the visibility of certain posts. Therefore, it can be deduced that predicting the virality of content is extremely difficult (Narayanan, 2023, p. 16). The only tool that platforms have at their disposal to predict or at least attempt to control virality is the prediction of engagement.

All platforms have various types of goals, which are often related to the revenue they generate from advertisements, keeping users engaged with their screens, and perhaps even some more noble goals. However, none of these objectives can be achieved simply by making decisions based on individual users and the specific content they are shown at a particular moment in time.

This is because there is no way to measure the long-term impact of a single post on a single user. To address this challenge, platforms rely on engagement. Engagement serves as a kind of score, determined by the actions users take on each post that appears in their feed. The way users interact with content influences how recommendation algorithms rank the available content, predicting how likely users are to engage with it. In this sense, «engagement acts as a proxy for higher-level goals (Narayanan 2023, p. 23). More generally, the selection of this content is performed according to various metrics and criteria, which vary from platform to platform. Algorithms constantly monitor these interactions to assess the level of engagement of each piece of content. Generally, a high number of interactions signals to the algorithm that the content is interesting to users, leading to its greater amplification.

5. From persuasion to manipulation: the dual role of algorithms in content recommendation

1 Shadow-banning is defined as a practice where a user's content or account is restricted or hidden from visibility on a social media platform without their explicit knowledge. While the user can still post and interact normally, their content may not appear in feeds, search results, or hashtags, limiting its reach and engagement. This method is often used by platforms to reduce the visibility of content violating the platforms policies or is deemed problematic without outright banning the user (Savolainen, 2022).

AI recommendation algorithms have been described for their manipulative capacity because, unlike traditional mass media, they can employ a range of persuasive techniques that differ in the large amount of information that can be obtained about users. This information ranges from demographic data to actual cognitive profiles of individuals, collecting habits, places visited, relationships, and preferences. Consequently, an algorithm can capture and translate this individual information in a very refined way, adapting the content presented to individuals according to their preferences and cognitive profiles (Bozdag, 2013). Users can take much more risk with this algorithm-induced manipulation because they can lose their capacity for judgment. This manipulation is characterized by two fundamental aspects: it is covert and exploits cognitive vulnerabilities in individuals' decision-making processes. This type of manipulation uses the so-called 'nudges', which are structural choices designed to change the behavior of individuals predictably (Faraoni, 2023). It is very effective because it creates a structure for the users who are targeted by these nudges, influencing their decisions towards specific directions chosen by the nudge designers themselves. Modern algorithmic systems allow these nudges to operate in real-time, changing outputs according to the user's actions. This mechanism illustrates how recommendation algorithms work to select and amplify the most relevant content, but also how they can be exploited by those seeking to manipulate the system (Faraoni, 2023).

In fact, it is not just platform operators who influence the algorithm to promote specific content. External actors, known as manipulators, may also exploit the same algorithmic dynamics to amplify deceptive and problematic content. While direct control of the algorithm by platform owners can lead to questionable decisions, there are also more subtle and widespread forms of manipulation orchestrated by external parties who use sophisticated techniques to influence algorithms from the outside (Acker, 2018). These manipulators operate in a context where user attention is an increasingly scarce and valuable resource, making the ability to manipulate attention metrics extremely profitable (Zhang et al., 2018). Digital media manipulation, which can be understood as the set of practices aimed at the systematic creation and dissemination of disinformation, fake news, or propaganda, is often motivated by ideological, economic, or status and attention-seeking purposes (Marwick and Lewis 2017, p. 27). Media manipulators not only exploit inherent vulnerabilities in algorithmic systems to spread disinformation and influence public discourse. As a matter of fact, these practices also exploit a media ecosystem characterized by internet subcultures and participatory culture, as theorized by Henry Jenkins (Jenkins, 2006). This type of sociotechnical exploitation shows how algorithms and recommender systems can be vulnerable to manipulation techniques, creating an ecosystem in which misinformation can flourish. Specifically, manipulation occurs not only through the creation of fake content but also by exploiting the very design of digital infrastructures (Golebiewski and boyd, 2019; boyd, 2024). Manipulators attempt to influence these metrics to artificially increase the visibility of their content through techniques such as coordinated link sharing (Giglietto et al., 2020b). Here, manipulators use multiple accounts to post the same content in a short period of time, thereby fooling the platform's algorithm. These tactics exploit the workings of the algorithm to make manipulative content reach a wider audience, increasing its impact and spread. Building on the foundational concepts of content selection and algorithmic dynamics, the following section provides a detailed examination of how these principles manifest in practice. The proposed model integrates established theories while refining and expanding them to capture the dual role of recommendation systems as both selectors of content and dynamic agents shaping communication patterns. Here, recommendation systems are conceptualized not only as tools for selecting content but as dynamic agents that actively shape communication patterns and influence user behavior and information flows across the platform.

6. The persuasive function of the recommendation: a theoretical model

As outlined in the preceding sections, algorithms on social media platforms function as critical components within the broader system of communication. They do not merely act as passive channels, but actively participate in the regulation and circulation of information, shaping the system's overall dynamics. Within this framework, the platform's algorithms are seen as key operators that influence the equilibrium of the system, determining which content gains prominence and how user behavior is collectively shaped. The theoretical model proposed in this paper integrates sociocybernetic concepts to frame recommendation algorithms as self-referential systems operating within a social media ecosystem. These algorithms embody key features of sociocybernetic systems: they observe user interactions, process feedback, and adjust their operations dynamically to optimize engagement. By doing so, they create a form of recursive communication between users and the platform, aligning with Luhmann's notion of system-environment interaction (Luhmann, 1995). This perspective highlights how algorithms mediate between individual preferences and systemic content flows, thereby acting as regulators of digital social realities.

One of the key ideas that emerges from this perspective is that algorithms function as important nodes or gateways through which communication between a generic 'Alter' and 'Ego' takes place on social media. Traditionally, it is assumed that users primarily interact with the algorithm itself, with little thought given to how their preferences affect others (Bucher, 2018). However, the engagement expressed by a generic 'Ego' user is not confined to their experience alone. Instead, it is captured by the algorithm, which amplifies this information and disseminates it to multiple 'Alter' users' feeds, effectively making the algorithm a channel of communication that enhances and extends the interaction. The shift from viewing algorithms as passive intermediaries to recognizing them as active agents in the communication process offers a nuanced understanding of their role in content dissemination. This perspective complements existing models by highlighting how algorithms amplify and mediate interactions between users, shaping the overall communicative environment on social media platforms. In essence, the platform's algorithm acts as a megaphone that determines what should be considered relevant, giving these contents a kind of algorithmic relevance. In selecting content to be amplified, the algorithm inevitably excludes other content, which will not receive the same relevance or visibility. Thus, the algorithm not only distributes information but also creates a hierarchy of what is considered important and worthy of attention, ranking the available content and predicting how likely users are to engage with it (Narayanan, 2023).

This is where the concept of persuasive technology becomes particularly relevant. Persuasive technology, outlined by Fogg (2002) is based on the idea that technological devices, such as computers and algorithms, are not merely passive tools, but can act as real social actors, capable of influencing users' attitudes and behavior. In the context of this study, persuasive technology specifically refers to recommendation algorithms implemented by social media platforms. These systems, designed to maximize user engagement, utilize feedback loops and predictive models to personalize content for individual users. While the broader concept of persuasive technology can encompass diverse domains—including propaganda or marketing strategies—we focus exclusively on the role of recommendation algorithms as active agents in shaping

communication patterns and influencing user attention.

According to Fogg, digital technologies can act as 'persuasive social actors' through the use of social signals, such as interactive language, positive feedback, and the assumption of social roles (Fogg, 2002). These signals stimulate automatic responses in users, shaping their interactions with technology in ways that reflect real social dynamics. This framework is particularly relevant in the realm of social media, where recommendation algorithms represent an advanced form of persuasive technology (Floridi, 2024). Building on Fogg's ideas, these algorithms do not merely suggest content—they actively interact with social systems, influencing user behavior and shaping social dynamics by leveraging persuasive techniques to maximize engagement. In other words, persuasive technology can be seen as a cybernetic control system that uses feedback strategies to influence and direct user behavior toward specific goals.

In a generic representation of persuasive technology, there is a source (S), the persuader, who uses the technology to formulate a message (M) with the aim of persuading a target (D) to achieve a certain purpose (G), as highlighted in Fig. 1.

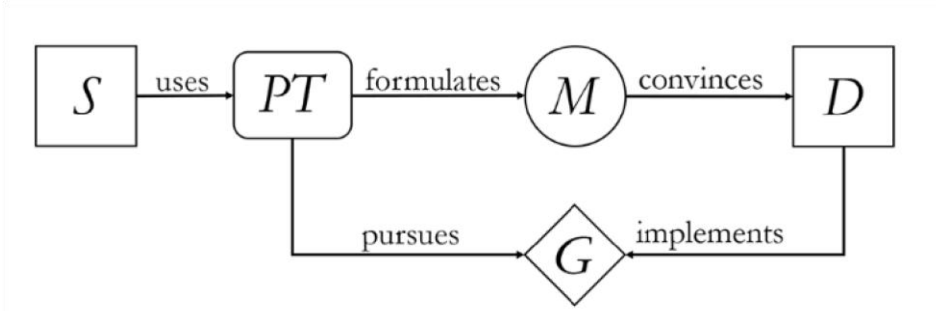

Fig. 1 – Diagram of a general persuasive system (Floridi, 2024)

In this context, a scenario can be imagined involving two users: a sender and a receiver. However, on social media, these users do not necessarily represent specific individuals, as we have already noted. Instead, they can be conceptualized as 'Alters' and 'Egos,' emphasizing that these users serve as generic representations of any two participants in the communication process on the platform. This approach highlights the importance of examining not only individual user experiences but also the broader dynamics of audience behavior and interaction, which gain particular relevance when adopting this specific lens of observation. While individual user interactions remain foundational, this study emphasizes how collective engagement patterns provide valuable insights into the systemic functions of platform algorithms. With these premises, it is now viable to explore how persuasive technology operates within social media platforms to shape these interactions. We will now examine how this model operates in practice, as portrayed in Figure 2.

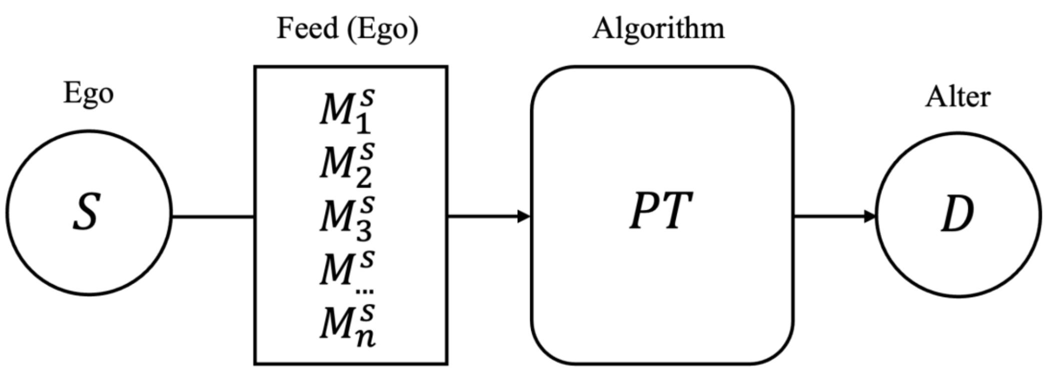

Fig. 2 – Outline of persuasive technology in Social Media

Consider that, after seeing a set of posts in their feed (denoted as MS in Fig. 2), user S interacts with them, either by liking, commenting, sharing, or even creating a new post inspired by the content. The algorithm captures the result of S's action and, consequently, shows user D the post with which user S interacted, or a derivative post created by S (see Fig. 2). This dynamic is driven by the algorithm's ability to capture engagement signals that user 'Ego' sends by interacting with or generating posts in their feed. This iterative process creates a dynamic feedback loop, where the interactions of users like S and D continuously shape the visibility and circulation of content on the platform. The algorithm prioritizes posts that have garnered the most engagement, increasing their likelihood of being shown to broader audiences, including user D. Notably, this loop also includes the possibility for users like S to create new posts inspired by or derived from previous interactions. These new posts, once generated, enter the same evaluative cycle, where the algorithm measures their engagement potential and determines their prominence within user feeds. This dynamic underscores the dual role of users as both consumers and producers of content, while the algorithm acts as a selective mediator, amplifying high-engagement material and perpetuating the cycle of content dissemination.

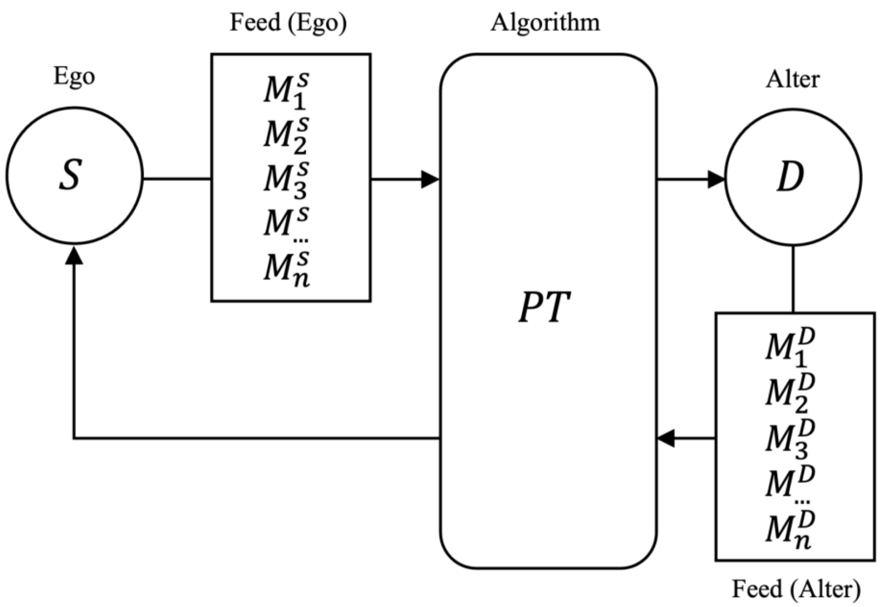

Fig. 3 – Feedback mechanism in social media persuasive technology

One of the most intriguing aspects of this form of social media communication lies in its ability to establish a continuous communication cycle between two generic users (S and D) mediated by the persuasive technology embedded in the platform's algorithm. In this dynamic, the algorithm serves as an active intermediary, shaping and directing the messages (M) exchanged between users. This mechanism generates a feedback loop fueled by user interactions, where the content users engage with gains increasing visibility and, in some cases, achieves virality. Unlike traditional forms of persuasion, where the sender's influence on the receiver is direct, this system creates a more intricate dynamic. Both users are actively engaged in a form of double contingency (Luhmann, 1995), where their interactions are mutually conditioned yet mediated by the algorithm. On the one hand, individual user preferences guide the algorithm's recommendations. On the other, the algorithm exercises selective control, amplifying specific content and determining its broader dissemination. This dual process ultimately shapes the communicative environment of the platform, blending user agency with algorithmic mediation. This continuous feedback loop not only amplifies content based on user interest but also reinforces user behavior, encouraging repeated patterns of viewing and interaction. The result is a system where user engagement and algorithmic mediation work together to create a dynamic and ever-evolving communicative space between users, as Figure 4 outlines.

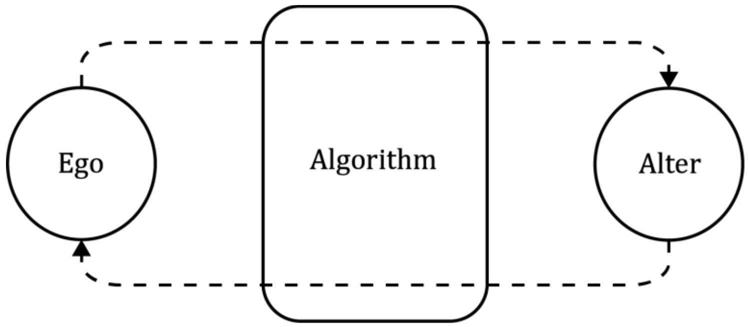

Fig. 4 – Mutual conditioning through persuasive technology

From this perspective, the algorithm functions as an active operator that shapes the formation and distribution of information within the social media system, acting as a structural element in the construction of digital social reality. The algorithm not only reflects existing social dynamics, but actively shapes perceptions and relationships between users by selecting and prioritizing content. This process scales seamlessly from individual exchanges to collective dynamics, where content becomes a focal point for broader patterns of reinterpretation and dissemination. Moreover, while this model illustrates one-to-one, Alter-Ego interactions, it also serves as a foundation for understanding many-to-many dynamics, where 'Alter' and 'Ego' do not represent fixed individuals but rather stand in for groups or communities engaging collectively. Indeed, the algorithm mediates direct exchanges between individuals and broader patterns of interaction that emerge when content is consumed, discussed, and reinterpreted within larger audiences.

This interpretation emphasizes the scalability of the feedback loop, which adapts seamlessly from individual interactions to collective dynamics, reinforcing the platform's role in shaping personal experiences and narratives.

However, the very dynamics that allow algorithms to mediate interactions and shape digital communication also expose significant vulnerabilities, particularly their susceptibility to manipulation. External actors can exploit these systems, intentionally distorting the algorithm's selection processes to serve their own agendas. This form of manipulation challenges the integrity of the system and creates a tension between the platform's intended goals—what we earlier referred to as *G* in the persuasive technology model (see

Fig. 1)—and the objectives pursued by manipulators. While platforms' high-level goals (G) are primarily aimed at maximizing user engagement and monetizing attention—objectives deeply embedded in their business models (Webster, 2014; Narayanan, 2023)—manipulators operate with divergent objectives, such as achieving virality or ensuring widespread coverage of specific content. In the following section, we examine how such actors compete with and undermine platform goals, introducing significant implications for the dynamics of content dissemination.

7. A practical example of persuasive amplifier manipulation: the case of crypto-communities on Facebook and Telegram

From a sociocybernetic perspective, this manipulation represents a deliberate distortion of the double contingency dynamic. By strategically injecting signals into the system, these actors interfere with the algorithm's selection process, effectively bypassing its intended purpose and redirecting outputs toward their objectives. In doing so, they transform the algorithmic feedback loop into a tool for their own ends, distorting its role as a mediator of genuine engagement. To further clarify this phenomenon, we present a specific case study that illustrates these dynamics of content manipulation in greater detail, offering a concrete context to apply and explore the theoretical concepts discussed so far.

Methods

The present analysis addresses the coordinated dissemination of cryptocurrency-related content on Facebook and Telegram and demonstrates how coordinated posting exploits algorithm vulnerabilities to achieve manipulators' goals. This investigation builds on prior research into Coordinated Inauthentic Behavior (CIB) in the context of disinformation on Facebook, as defined by Gleicher (2018), which served as the methodological foundation for identifying problematic information flows. The computational technique used to detect this coordinated behavior identifies the spread of the same link on Facebook in a very short time span, suggesting with good probability that the content is spread by a central entity or shared strategy. In other words, all Facebook entities[2] that share the same content in such short time intervals could be nothing more than a single agglomeration of pages governed by the same entity. This coordinated behavior fools Facebook's algorithm into believing that the links shared are particularly relevant. Since these links are spread simultaneously by many entities, the algorithm interprets this as a signal of popularity, promoting the content even though the interest of the actual users is not genuine (Giglietto et al., 2020a). The approach based on the detection of Coordinated Link Sharing Behavior (CLSB) has proven particularly effective in addressing the challenges posed by the blurry boundaries of

2 The term 'entities' refers to individual verified public profiles, Facebook groups, or Pages.

manipulation, especially when empirically distinguishing between problematic and non-problematic information (Jack, 2019; Lazer et al., 2018). Recent studies have further emphasized the value of shifting the analytical lens from content to the dynamics of information dissemination within online networks, highlighting how coordinated activities can serve as key indicators of manipulation (Hui et al., 2019; Keller et al., 2020). This inauthentic coordinated behavior has been observed in several other studies on topics such as politics or health, demonstrating how coordinated actions can hack and bend the operating logic of Facebook's algorithm (e.g. Ayers et al., 2021; Broniatowski, 2021; DFRlab, 2020; Giglietto et al., 2019b; Giglietto et al., 2020b).

In addition to analyzing coordinated dissemination patterns, this study incorporates content decay as a complementary indicator of problematic activity. Decay, understood as the inaccessibility or obsolescence of shared links, highlights the ephemeral nature of manipulated content (Bastos, 2021). Specifically, for the posts identified through coordinated link-sharing behavior, we aim to measure the proportion of links that have decayed—becoming inactive or inaccessible—compared to those still active. This ratio serves as a concrete indicator of the content's problematic nature, providing insight into the extent to which coordinated posts are tied to transient, potentially exploitative strategies.

This methodological approach enables us to address the following research question (RQ): to what extent does the dissemination of cryptocurrency-related content exhibit signs of problematic behavior, specifically through inauthentic coordination and content decay over time? By analyzing both the structure of coordinated dissemination and the status of shared links, this study aims to provide a nuanced understanding of manipulative strategies operating within these systems.

Data

The earlier work by Giglietto et al. (2022b), which we employ as the foundation for this case study, examined a wide range of Facebook communities operating within the Nigerian digital landscape, uncovering numerous clusters of accounts engaged in the coordinated dissemination of content. Using the CooRnet library (Giglietto et al., 2020a), which is an R package leveraging data from the CrowdTangle[3] API, the study identified coordinated link-sharing behavior (CLSB) among Nigerian actors. CLSB refers to the repeated sharing of the same URLs by multiple Facebook entities—such as groups, pages, and verified profiles—

3 A public insights tool previously owned by Meta used to track, analyze, and compare content performance across social media platforms. It was widely utilized by researchers, journalists, and organizations to monitor trends, identify viral content, and study engagement metrics on platforms like Facebook, Instagram, and Reddit. Despite its popularity, CrowdTangle was discontinued in 2024

within a very short time frame. This behavior is detected through an algorithm that estimates a time threshold for coordination and identifies entities repeatedly sharing the same content within that interval[4]. The CooRnet library facilitated the extraction of various results, including spreadsheets detailing the coordinated entities and graph files specifically prepared for conducting social network analyses of these entities.

Among the many components identified through Social Network Analysis performed with Gephi, one was selected for closer examination due to its thematic focus on cryptocurrencies, as suggested by the names of the entities it encompassed. This component, consisting of 1,212 nodes, exhibited notable structural properties that indicated high internal cohesion and organization. Specifically, a modularity score of 0.456 revealed 32 distinct communities, while an average clustering coefficient of 0.737 highlighted a highly interconnected structure, reinforced by the presence of 388,267 closed triangles.

Within this component, a particularly intriguing cluster comprising 151 Facebook groups was identified for further exploration (see Fig. 5). This cluster stood out for its strong internal connectivity, thematic alignment with cryptocurrency-related activities, and the presence of highly influential nodes actively disseminating information across the network. The analysis of node-level metrics revealed a dynamic structure: some nodes exhibited significant connectivity, with the highest degree reaching 245, positioning them as central hubs within the cluster.

Additionally, strength values, which reflect the intensity of weighted interactions, peaked at 2,765, indicating concentrated activity and engagement around specific nodes.

Prominent examples include nodes such as “Airdrop,” with 871 coordinated shares, and “BCPAY FINTECH AND OUTRACE PLAY2EARN COMMUNITY,” with 496 coordinated

shares. These nodes played a critical role in driving the coordinated dissemination of content. Several accounts within the cluster also demonstrated significant reach, with top-performing groups like “Airdrop - Token Crypto” attracting tens of thousands of participants.

This alignment of structural coherence and thematic focus aligns with Wasserman and Faust’s (1994) observation that clusters demonstrating such characteristics are particularly effective for representing broader phenomena. Consequently, this cluster was chosen for its potential to provide meaningful insights into coordinated behavior within the realm ofcryptocurrency-related activities.

4 For further information regarding the CooRnet tool, please refer to: https://coornet.org

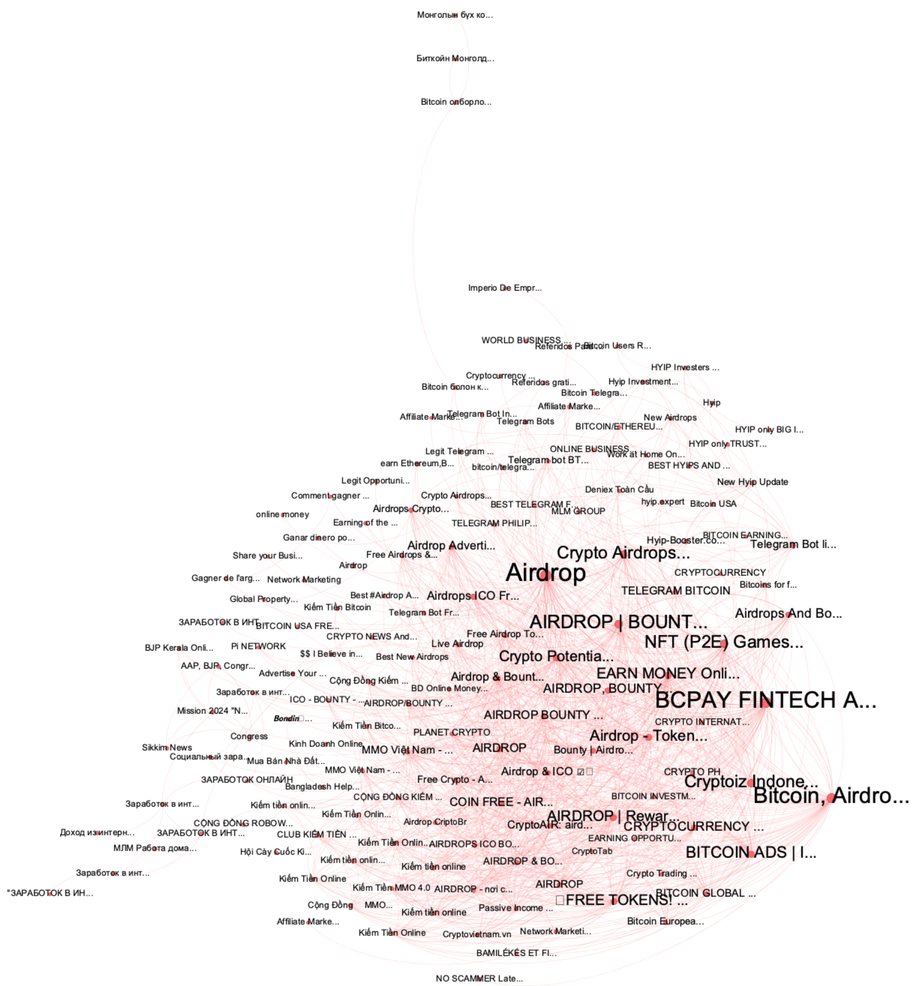

Fig. 5 - Visualisation of the cluster of crypto-related Facebook groups[5]

Starting from the list of groups included in the selected cluster, a new round of data collection was initiated via the CrowdTangle API. All posts from the 151 groups were retrieved over a four-month period, spanning from April 1, 2020, to September 2, 2020. This dataset comprised approximately 100,000 posts.

Data analysis

---

[5] The image above depicts a network visualization of the 151 coordinated Facebook groups. Each node in the network represents a group, and the nodes size represents their degree centrality.

From these posts, all embedded links (URLs) were extracted, resulting in a preliminary dataset of 12,033 unique URLs. To refine this dataset for further analysis, we prioritized the links based on their engagement metrics, including the total number of likes, comments, and shares associated with posts containing each URL.

To achieve this, an automated examination of the top 1,052 URLs—selected based on their high interaction volumes—was conducted. HTTP status codes were analyzed using a bulk-checking tool capable of processing multiple URLs simultaneously[6]. Links returning a “200 OK” status code were categorized as active, while those with a “404 Not Found” status were classified as inactive. Additionally, links with “301” or “302” status codes were recorded as redirects, offering insights into whether content decay or reorganization was occurring within the cluster. The analysis revealed that a significant 26% of those links were inactive. Among the active links, 21.1% pointed to Telegram domains (e.g., t.me), a platform known for its privacy-focused features and decentralized structure. Unlike Facebook, which has progressively tightened its content moderation policies, Telegram operates with minimal regulation, offering a more permissive environment for actors aiming to avoid detection (Schulze et al., 2022).

Given the lack of automated tools for analyzing content on Telegram, a manual review was conducted to examine the content associated with these links. This review focused on assessing the accessibility of Telegram domains was evaluated. As a result, a significant portion of the links—43.8%—were found to be either inactive or flagged as scams by the platform itself (as shown in the example in Fig. 6), with 24.4% explicitly marked as fraudulent by Telegram.

6 The tool used for this task can be found at: https://httpstatus.io

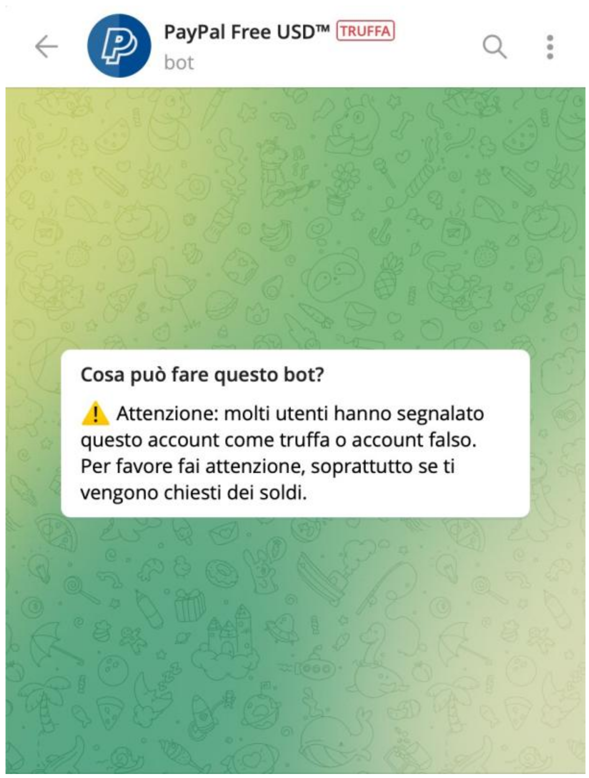

Fig. 6 – Example of Telegram bot labeled as a scam

These findings underscore the ephemeral nature of problematic content on Telegram and its potential use as a platform for coordinated manipulative activities. While the combination of automated and manual methods provided a robust approach to analyzing the dataset, a few limitations must be acknowledged. A significant challenge was the high proportion of inactive links (26%), which, while indicative of problematic posting activity, limits the scope of retrospective analysis, creating a gap in our ability to reconstruct the full context of manipulative activities (Bastos, 2021).

The significant proportion of active Telegram links flagged as scams (24.4%) provides a strong indicator of broader fraudulent activity within the platform. While these flagged links represent only a portion of the observed ecosystem, their recurrence suggests systemic patterns of manipulation. A closer, manual examination of the active links revealed that many relied on well-established marketing techniques—including cryptocurrency rewards, giveaways, and referral-based incentives—to encourage users to engage in social activities, such as sharing content, joining groups, or inviting others to participate. These strategies, while not inherently fraudulent, are often used to build engagement and amplify content visibility in ways that align with manipulative objectives (Faraoni 2023). A deeper qualitative analysis of these strategies, particularly focusing on the activities and roles of

Telegram entities such as groups and bots, has been conducted in a separate study (Author, 2023).

Importantly, although we cannot definitively classify all active spaces as scams, their methods reveal a clear reliance on mechanisms that incentivize user activity under the guise of potential financial gain. This observation strengthens the hypothesis that the now-inactive links likely contained similar initiatives. Their elimination could reflect deliberate efforts to abandon campaigns after short-term goals were achieved, or after detection risks increased.

This combination of manipulative marketing strategies, the prevalence of flagged content, and the high proportion of inactive links underscores a broader trend of attention and engagement manipulation within Telegram spaces. These findings point to a recurring use of transient content designed to exploit platform vulnerabilities and user behavior for coordinated, potentially exploitative purposes.

8. The tug-of-war between manipulators and platforms: implications for algorithmic integrity and content quality

This study reveals a particular strategy used by manipulators to spread cryptocurrency-related initiatives and projects. These manipulators try to obtain as many adhesions as possible from a mainstream social media platform such as Facebook, exploiting the relatively loose moderation controls that are operated on external links, which are not directly controllable by the platform. Consequently, these malicious initiatives developed and took shape mainly on Telegram, a platform considered more fringe and with less strict moderation (Schultze et al., 2022). In this context, manipulators use the visibility gained on Facebook through the coordinated sharing of links to deceive the algorithm, making their content appear more relevant than it really is.

To make their dragnet strategy of capturing users on Facebook more effective, manipulators actively influence the feedback mechanism depicted in Figure 3. By inserting themselves into the platform’s communicative loop, they manipulate the reciprocal interactions between users and the algorithm. Manipulators simulate patterns of engagement through a coordinated activity that the platform interprets as genuine interest from its ‘Alter’ and ‘Ego’ users. By doing so, these actors insert themselves into the cycle of double contingency that characterizes communication between users on platforms like Facebook. In this many-to-many dynamic, manipulators leverage a network of accounts under their control to distort what the algorithm perceives as genuine interest. These accounts, whether real users or bots, engage in highly coordinated posting activity that amplifies specific content within extremely short time intervals. This pattern of posting, as observed in this study, strongly suggests the involvement of sockpuppet accounts—fictitious identities operated by the same entity.

By orchestrating such behavior, manipulators effectively simulate organic engagement, tricking the algorithm into prioritizing manipulated content while undermining the integrity of the platform's recommendation system. In essence, instead of allowing naturally engaging content to rise organically, they falsify engagement metrics, ensuring that posts they aim to promote—those containing links to Telegram, for example—achieve higher visibility.

Once the manipulated content gains traction within the algorithm, the feedback loop takes over, amplifying its reach further. As a result, the manipulated content continues to circulate widely and without cost (i.e., without the need to pay for advertising space), exploiting Facebook's recommendation system to reach users predisposed to similar activities.

9. Conclusions

This article emphasizes the importance of integrating sociocybernetics and persuasive technology theory to understand the dynamics of content manipulation on social media. By framing recommendation algorithms as cybernetic operators within a communication system, this study has shown how these technologies mediate interactions, shape content flows, and create feedback loops that influence user behavior and societal discourse.

The balance between platforms' recommendation algorithms and the manipulative strategies employed by external actors emerges as a complex dynamic. Platforms must reconcile their economic objectives with the need to maintain quality information, while manipulators exploit these same algorithms to amplify content that serves their interests. The consequences of this competition between manipulators and platforms are significant. If the goals of the manipulators prevail, the virality of low-quality, uninformative, or even fraudulent content increases, distorting the information landscape accessible to users in their feeds. This also aligns with broader concerns raised by theories such as the Dead Internet Theory (Walter, 2024), which suggests that an increasing proportion of online content is artificially generated or amplified by non-human actors. Such dynamics risk creating a digital ecosystem where genuine human interaction is overshadowed by synthetic activity, eroding user trust and undermining the authenticity of online communication. As noted by Gillespie (2014), recommendation algorithms must maintain a certain standard of perceived quality to sustain user trust, while persistent exposure to misleading or irrelevant content—such as feeds heavily populated by cryptocurrency-related scams—risks eroding users' confidence in the platform.

On a theoretical level, this study contributes to understanding how persuasive technologies can be appropriated not only by platforms to influence user behavior but also by manipulators to bypass system rules. The concept of 'hyper-persuasion' (Floridi, 2024), reflecting the advanced capabilities of artificial intelligence and machine learning, highlights the dual-use nature of these technologies. While they can enable positive outcomes, they also pose significant ethical challenges when exploited for manipulative purposes, including disinformation or economic

fraud.

On a practical level, this work underscores the urgency for platforms to develop transparent and robust systems to counter manipulative behaviors. Implementing advanced detection mechanisms for fake user behavior, alongside greater transparency in recommendation algorithms, is essential. Users should be empowered to understand how content is selected and distributed, fostering trust and accountability.

From a regulatory perspective, the findings of this study are closely aligned with the objectives of the European Union's Digital Services Act (DSA). Recital 84 of the DSA (European Union, 2022) explicitly emphasizes the systemic risks posed by algorithmic systems, including recommendation and advertising algorithms. This regulation shifts the focus from the nature of the content itself to the behaviors that exploit platforms to disseminate and amplify misleading or deceptive information. Practices such as the intentional and coordinated use of fake accounts, bots, and other inauthentic behaviors are explicitly addressed, underscoring the need for platforms to mitigate these risks through better algorithmic design and moderation.

Bibliography

Acker, A. (2018). *Data craft: The manipulation of social media metadata.*

Ayers, J. W., Chu, B., Zhu, Z., et al. (2021). Spread of misinformation about face masks and COVID-19 by automated software on Facebook. *JAMA Internal Medicine*. https://doi.org/10.1001/jamainternmed.2021.2498

Bastos, M. (2021). This account doesn't exist: Tweet decay and the politics of deletion in the Brexit debate. *American Behavioral Scientist, 65*(5), 757–773. https://doi.org/10.1177/0002764221989772

Baker-White, E. (2023, January 20). TikTok's secret "heating" button can make anyone go viral, says employee. *Forbes*. https://www.forbes.com/sites/emilybaker-white/2023/01/20/tiktoks-secret-heating-button-can-m ake-anyone-go-viral/

Batmaz, Z., Yurekli, A., Bilge, A., & Kaleli, C. (2019). A review on deep learning for recommender systems: Challenges and remedies. *Artificial Intelligence Review, 52*(1), 1–37.

Beer, D. (2016). The social power of algorithms. *Information, Communication & Society, 20*(1), 1–13. https://doi.org/10.1080/1369118X.2016.1216147

Bobadilla, J., Ortega, F., Hernando, A., & Gutiérrez, A. (2013). Recommender systems survey. *Knowledge-Based Systems, 46,* 109–132. https://doi.org/10.1016/j.knosys.2013.03.012

Boyd, d. (2024). Making sense of data voids. *Data & Society*. In *Reckoning with Mis/Disinformation in 2024*. https://datasociety.net/points/making-sense-of-data-voids/

Bozdag, E. (2013). Bias in algorithmic filtering and personalization. *Ethics and Information Technology, 15,* 209–227.

Broniatowski, D. A. (2021). Towards statistical foundations for detecting coordinated inauthentic behavior on Facebook. https://iddp.gwu.edu/sites/g/files/zaxdzs3576/f/downloads/Coordinated%20Behavior%20On%20Facebook;%20Broniatowski.pdf

Bucher, T. (2018). *If… then: Algorithmic power and politics*. Oxford University Press.

Couldry, N., & Hepp, A. (2018). *The mediated construction of reality*. John Wiley & Sons.

DFRlab. (2020). Why the debunked COVID-19 conspiracy video "Plandemic" won't go away. https://medium.com/dfrlab/why-the-debunked-covid-19-conspiracy-video-plandemic-wont-go-a way-c9dd36c2037c

European Union. (2022). Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market for Digital Services (Digital Services Act) and amending Directive 2000/31/EC. *Official Journal of the European Union*. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32022R2065

Faraoni, S. (2023). Persuasive technology and computational manipulation: Hypernudging out of mental self-determination. *Frontiers in Artificial Intelligence, 6,* Article 1216340.

Floridi, L. (2024). Hypersuasion: On AI's persuasive power and how to deal with it. *Philosophy & Technology*.

Fogg, B. J. (2002). Persuasive technology: Using computers to change what we think and do. *Ubiquity, 2002*(December), Article 2.

Giglietto, F., Righetti, N., & Marino, G. (2019a). Understanding coordinated and inauthentic link-sharing behavior on Facebook in the run-up to 2018 general election and 2019 European election in Italy. https://doi.org/10.31235/osf.io/3jteh

Giglietto, F., Righetti, N., Marino, G., & Rossi, L. (2019b). Multi-party media partisanship attention score: Estimating partisan attention of news media sources using Twitter data in the lead-up to 2018 Italian election. *Comunicazione Politica, 1,* 85–108. https://doi.org/10.3270/93030

Giglietto, F., Righetti, N., & Rossi, L. (2020a). CooRnet: Detect coordinated link-sharing behavior on social media. http://coornet.org

Giglietto, F., Righetti, N., Rossi, L., & Marino, G. (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. https://doi.org/10.1080/1369118X.2020.1739732

Giglietto, F., Farci, M., Marino, G., Mottola, S., Radicioni, T., & Terenzi, M. (2022a). Mapping nefarious social media actors to speed-up COVID-19 fact-checking. https://doi.org/10.31235/osf.io/6umqs

Giglietto, F., Olaniran, S., Mincigrucci, R., Marino, G., Mottola, S., & Terenzi, M. (2022b).
Blowing on the fire: An analysis of low-quality and hyper-partisan news sources circulated by coordinated link sharing networks in Nigeria. *SSRN*. https://doi.org/10.2139/ssrn.4162030

Gillespie, T. (2014). The relevance of algorithms. In *Media technologies: Essays on communication, materiality, and society*. The MIT Press. https://doi.org/10.7551/mitpress/9780262525374.003.0009

Gillespie, T. (2018). *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions that Shape Social Media*. Yale University Press.

Gleicher, N. (2018, December 6). Coordinated inauthentic behavior explained [Interview]. *Meta Newsroom*. https://about.fb.com/news/2018/12/inside-feed-coordinated-inauthentic-behavior/

Golebiewski, M., & boyd, d. (2019). Data voids: When missing data can be easily exploited. *Data & Society*. https://datasociety.net/wp-content/uploads/2019/11/Data-Voids-2.0-Final.pdf

Granovetter, M. (1973). The strength of weak ties. *American Journal of Sociology, 78*(6), 1360–1380. https://doi.org/10.1086/225469

Hern, A. (2022). How TikTok's algorithm made it a success: 'It pushes the boundaries.' *The Guardian*. https://www.theguardian.com/technology/2022/oct/23/tiktok-rise-algorithm-popularity

Hui, P.-M., Yang, K.-C., Torres-Lugo, C., Monroe, Z., McCarty, M., Serrette, B., Pentchev, V., & Menczer, F. (2019). BotSlayer: Real-time detection of bot amplification on Twitter. *Journal of Open Source Software, 4*(42), 1706. https://doi.org/10.21105/joss.01706

Jack, C. (2019). Wicked content. *Communication, Culture and Critique, 12*(4), 435–454. https://doi.org/10.1093/ccc/tcz043

Jenkins, H. (2006). *Fans, bloggers, and gamers: Exploring participatory culture*. NYU Press.

Keller, F. B., Schoch, D., Stier, S., & Yang, J. (2020). Political astroturfing on Twitter: How to coordinate a disinformation campaign. *Political Communication, 37*(2), 256–280. https://doi.org/10.1080/10584609.2019.1661888

Lazer, D. M. J., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F.,
Metzger, M. J., Nyhan, B., Pennycook, G., Rothschild, D., Schudson, M., Sloman, S. A., Sunstein, C. R., Thorson, E. A., Watts, D. J., & Zittrain, J. L. (2018). The science of fake news. *Science, 359*(6380), 1094–1096. https://doi.org/10.1126/science.aao2998

Longo, M. (2005). Doppia contingenza on line: Considerazioni teoriche sulla comunicazione in rete. In A. Milanaccio (Ed.), *Immaginario e società globale* (pp. 71–91). Manni.

Luhmann, N. (1995). Social systems. Stanford University Press.

Luhmann, N. (2000). *The reality of the mass media* (K. Cross, Trans.). Stanford University Press. (Original work published 1996).

Marwick, A., & Lewis, R. (2017). Media manipulation and disinformation online. *Data & Society*. https://datasociety.net/pubs/oh/DataAndSociety_MediaManipulationAndDisinformationOnline. pdf

Möller, J., Trilling, D., Helberger, N., & van Es, B. (2018). Do not blame it on the algorithm: An empirical assessment of multiple recommender systems and their impact on content diversity.

Information, Communication & Society, 21(7), 959–977. https://doi.org/10.1080/1369118X.2018.1444076

Narayanan, A. (2023). Understanding social media recommendation algorithms. *Knight First Amendment Institute at Columbia University*. https://doi.org/10.7916/khdk-m460

Pariser, E. (2011). *The filter bubble: How the new personalized web is changing what we read and how we think*. Penguin Press.

Sarwar, B., Karypis, G., Konstan, J., & Riedl, J. (2000, October). Analysis of recommendation algorithms for e-commerce. In *Proceedings of the 2nd ACM Conference on Electronic Commerce* (pp. 158–167).

Savolainen, L. (2022). The shadow banning controversy: Perceived governance and algorithmic folklore. *Media, Culture & Society, 44*(6), 1091–1109. https://doi.org/10.1177/01634437221077174

Schiffer, Z., & Newton, C. (2023, February 14). Yes, Elon Musk created a special system for showing you all his tweets first. *Platformer*. https://www.platformer.news/yes-elon-musk-created-a-special-system/

Schulze, H., Hohner, J., Greipl, S., Girgnhuber, M., Desta, I., & Rieger, D. (2022). Far-right conspiracy groups on fringe platforms: A longitudinal analysis of radicalization dynamics on Telegram. *Convergence: The International Journal of Research into New Media Technologies, 28*(4), 1103–1126.

Solsman, J. (2018). YouTube's AI is the puppet master over most of what you watch. *CNET*.

Tilak, S., & Glassman, M. (2020). Alternative lifeworlds on the Internet: Habermas and democratic distance education. *Distance Education*, *41*(3), 326-344.

Tufekci, Z. (2015). Algorithmic harms beyond Facebook and Google: Emergent challenges of computational agency. *Colorado Technology Law Journal, 13,* 203–217.

Van Dijck, J. (2013). *The culture of connectivity: A critical history of social media*. Oxford University Press.

Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications*. Cambridge University Press.

Walter, Y. Artificial influencers and the dead internet theory. *AI & Soc* (2024). https://doi.org/10.1007/s00146-023-01857-0

Webster, J. G. (2014). *The marketplace of attention: How audiences take shape in a digital age*. MIT Press.

White, H. C. (1970). *Chains of opportunity: System models of mobility in organizations*. Harvard University Press.

Zhang, Y., Wells, C., Wang, S., & Rohe, K. (2018). Attention and amplification in the hybrid media system: The composition and activity of Donald Trump's Twitter following during the 2016 presidential election. *New Media & Society, 20*(9), 3161–3182.