



---

## Exploring the Zombie Internet: Anatomy of Three Deceptive Information Operations on Facebook

Journal:	<i>Journal of Information, Communication &amp; Ethics in Society</i>
Manuscript ID:	JICES-05-2025-0109.R1
Manuscript Type:	Journal Paper
Keywords:	facebook, information operation, deception, coordinated behaviour, gambling, propaganda

SCHOLARONE™  
Manuscripts

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

## Introduction

In recent years, theories about the toxicity of social media ecosystems have proliferated in academia due to a surge in deceptive content, including AI-generated imagery, phishing schemes, and manipulative visuals designed to provoke emotional reactions or sway audiences (DiResta & Goldstein, 2024; Donovan et al., 2020). Although taxonomies of problematic information are well-established (Jack, 2017; Wardle & Derakhshan, 2017), these frameworks often fail to capture the full scope and complexity of contemporary information operations (Starbird et al., 2019). These operations extend beyond merely spreading falsehoods; they deliberately pollute the digital information ecosystem by flooding social media with low-credibility content – such as clickbait, divisive memes, and emotionally charged posts – thereby eroding public trust and fostering cynicism.

Previous studies, journalistic investigations and interviews with political strategists have unveiled this ecosystem's intentional pollution. King's "spaghetti approach" (2024) and Bannon's "flood the zone" strategy (Broadwater, 2025) demonstrate methods of overwhelming social media with deceptive content to test effectiveness of online communication and drown out opposition. These tactics don't merely mislead but create confusion by blurring boundaries between false and legitimate information, ultimately undermining epistemic foundations of public discourse (Molina, 2023).

These strategies transform platforms like Facebook into fragmented digital ecosystems where the boundaries between human and artificial agents blur, a development Koebler (2024) refers to as the Zombie Internet. These environments are characterized by automated conversations, manufactured controversies, and deceptive identities that dissolve what was once a shared consensus about reality. This represents a fundamental shift from earlier concerns about isolated instances of misinformation to a more pervasive challenge to the integrity of online information itself.

While current efforts to preserve ethical digital information environments often focus narrowly on false news, our research addresses more pervasive deceptive communication activities that carry significant yet understudied impacts. Unlike false news campaigns, these information operations rely on iterative testing and adaptation of messages, often deploying a mix of benign and harmful content to obscure intent and evade detection (Chadwick & Stanyer, 2022). These tactics function as gaslighting strategies (Jack, 2017) that systematically manipulate targets into questioning their own perception of reality and judgment of events.

To build on this discussion, we introduce the concept of deceptive information operation to better capture these chaotic, boundary-testing communication activities. We draw on two complementary bodies of theory: Starbird et al.'s (2019) work on strategic information operations and Chadwick and Stanyer's (2022) concept of deception in mis/disinformation. This integrated approach expands the focus beyond disinformation as false content, encompassing intentional communicative acts carried out with the "intention to mislead, resulting in attitudinal or behavioural outcomes that correspond with the prior intention" (Chadwick & Stanyer 2022, p. 2). By articulating deceptive information operations as a broader category, we offer conceptual tools capable of identifying harmful communicative activity that platform labels such as Coordinated Inauthentic Behaviour (CIB) (Gleicher, 2018) address only partially.

To explore this concept empirically, we analyse three deceptive information operations active on Facebook. Despite its declining popularity among younger audiences, Facebook remains widely used globally, particularly among older and less digitally confident users who may be more exposed to confusing or deceptive content (Guess et al., 2019). Our study draws on a monitoring tool developed within the [Anonymized Research Project] to track accounts that repeatedly share content flagged as problematic by Meta's third-

party fact-checkers. We also rely on the Meta Content Library (Meta Platforms, Inc., 2025), making this one of the first studies to employ this dataset at scale. Each case is examined through three dimensions: the operation's deceptive features, the exploitation of platform affordances, and the participatory behaviours that allow such activities to persist and evolve.

This research significantly advances our understanding of deceptive information operations by examining them through these three dimensions, offering a more comprehensive framework than existing taxonomies. Our methodological approach using the Meta Content Library and the [Anonymized Research Project] monitoring tool establishes new techniques for detection and analysis that transcend conventional false news narratives, addressing critical gaps in approaches to safeguarding digital information ecosystems.

The paper is structured as follows: We begin by defining deceptive information operations and outlining their key characteristics, followed by a presentation of the [Anonymized Workflow] and analytical framework used to detect and study these operations. We then examine three case studies, each illustrating different types of networks and distinct deceptive intentions: a network disseminating pro-Putin propaganda, a network promoting online casinos and gambling, and a network of poorly moderated public groups spamming adult content with fraudulent links. We conclude by discussing the broader implications of our findings for platform governance and future research.

## Theoretical Background

This paper coins the definition of *deceptive information operations* by combining the theoretical framework of "strategic information operations" proposed by Starbird et al. (2019) and that of deception proposed by Chadwick and Stanyer (2022). This concept

extends beyond the traditional focus on dis/misinformation – which typically concentrates on wholly or partially false information – by encompassing organised, communicative activities on social media designed to circulate information with the intent to cause harm (Weedon et al., 2017). The theoretical framework supporting this study is organised along three lines: [1] the aim to deceive users online, [2] the exploitation of digital media platforms' affordances to amplify deception, and [3] the involvement of multiple user accounts in deception. The research questions are thus formulated in alignment with these three dimensions.

### **Information Aimed at Deceiving**

Deceptive content spread through information operations has been examined from multiple perspectives. Meta and other platforms initially framed these operations through the lens of authenticity versus inauthenticity, defining Coordinated Inauthentic Behaviour (CIB) as a phenomenon in which groups of individuals work together to mislead others (Gleicher, 2018; Graham, 2024). In addition to CIB, Meta coined the definition of Coordinated Social Harm (CSH) that captures networks of users who systematically violate platform policies to inflict harm (Gleicher, 2018; Rogers & Righetti, 2024).

The notion of “causing harm” is a recurring theme in discussions on disinformation operations. According to EU DisinfoLab, a think tank working on the topic, these operations are deliberate actions intended to inflict harm in pursuit of personal, corporate, or institutional gains (2020). Such operations generally fall into three primary categories. The first is political influence, both domestic (Vargo et al., 2018) and foreign (Parham, 2018; Select Committee on Intelligence, 2019), where actors seek to shape public opinion, destabilise societies, or undermine adversaries to advance specific agendas. The second category, issue-based campaigns, involves coordinated efforts to promote

particular causes, such as conspiracy theories or no-vax sentiments, to give two examples.

While often grassroots, these campaigns share strategic similarities with politically motivated operations (Huth et al., 2020). Finally, lucrative information operations are financially motivated, leveraging clickbait, traffic redirection to monetised websites, and exploiting online advertising systems (Chergarova et al., 2022; Silverman & Alexander, 2016).

Chadwick and Stanyer (2022) propose deception as a concept for understanding the harm malicious actors cause. They see deception occurring when an actor's intent to mislead others results in attitudinal or behavioural changes that align with their objectives (2022, p.6). Previously, Buller and Burgoon (1996) argued that deception depends on successfully executing misleading intentions during communication. This perspective emphasises the interaction between deceivers and the deceived, framing deception as a communicative act to foster false beliefs or misunderstandings (Chadwick & Stanyer 2022, p.3). By focusing on the theoretical links between intent, process, and outcome (Buller & Burgoon, 1996; Rubin, 2016), this approach offers an empirical framework for analysing information operations from the perspective of "attributes and actions of deceptive entities" (Chadwick & Stanyer, 2022, p.14).

In this study, we examine if these networks of actors are involved in information operations to spread deceptive content. Anchoring the study of information operations within this framework clarifies what may be delineated as a deceptive information operation, and qualitative case studies' analysis allows us to uncover this interplay. Accordingly, we posed the following research question:

RQ1. How are the information operations surfaced by the [Anonymized Workflow] deceptive?

### Exploiting Digital Media Platforms' Affordances To Amplify Deceptive Information

Strategic information operations leverage algorithm-driven information flows to maximise their resilience and reach (Starbird et al., 2019). Through this technique, these operations achieve rapid dissemination and show adaptability to different technological contexts and audiences (Bradshaw & Howard, 2018). Social media platforms, in particular, play a central role in facilitating their spread. They provide tools for advertisers, such as impression management and targeting, which malicious actors can exploit to deliver content to carefully targeted audiences. Combined with algorithms that prioritise engagement, these tools enable the rapid and widespread sharing of content, creating self-reinforcing loops that amplify the visibility and impact of these operations (Gillespie, 2021). Moreover, social media platforms blur the lines between news, entertainment, and opinion within a single feed, collapsing contextual boundaries (Newman et al., 2021). This genre collapse makes it harder for users to discern the intent behind a given piece of content, especially when assessing the reliability of information is hindered by the multiplication and fragmentation of channels that characterise the post-broadcast information environment (Prior, 2014). Exploiting this ambiguity, malicious actors often mix accurate and misleading information or adopt fabricated identities to disguise their true motives (Wardle & Derakhshan, 2017).

Algorithms primarily govern the dissemination and consumption of information on social media. Platforms use complex algorithms to sort, filter, and prioritise content, optimising for user engagement and time spent on the platform (Fowler, 2016). Scholars have argued that by personalising content based on users' interests and behaviours, these algorithms may create "filter bubbles" that reinforce biases and limit exposure to diverse viewpoints (Gillespie, 2014). Algorithms tailor content and search results based on users' interests, past behaviours, and geographic location. Although the internet theoretically expands

access to information, algorithmic curation often undermines the free flow of ideas necessary for a healthy democracy. Content that sparks outrage, confirms biases, or drives engagement tends to be prioritised, facilitating rapid dissemination regardless of accuracy (Al-Rawi, 2019). As a result, while information access has increased in the current ecosystem, algorithms have made encountering critical or diverse perspectives harder.

We thus asked:

RQ2: How do malicious actors exploit Facebook's platform features to conduct information operations?

### **Participatory Behaviour In Proliferating Deceptive Information**

On social media, users with similar goals are able to coordinate their activities without centralised control (Bennett & Segerberg, 2013; Jenkins et al., 2015). In the context of information operations, this coordination often serves malicious purposes (Authors, 2020a). Scholars and social media platforms have described these collaborative efforts in various ways, including terms like dark participation (Quandt, 2018), CIB (Gleicher, 2018), and coordinated influence operations (O'Donovan, 2024). The most widely recognized definition is Meta's CIB, which refers to networks of inauthentic accounts working together to deceive users, platforms, or systems to influence public discussion, fraud, or evade platform policy enforcement (Romero-Vicente et al., 2024).

Coordination plays a crucial role in the success of strategic information operations. It enables the systematic creation of misleading content and amplifies its dissemination by exploiting platform algorithms (Gleicher, 2018; Graham & QUT Digital Observatory, 2020). Because every action on social media is visible and measurable – and these metrics influence how content is prioritised by indexing algorithms – coordinated behaviour can significantly boost the perceived impact of deceptive content. This visibility is a proxy

for influence, making coordinated activities especially effective in manipulating public discourse (Allen, 2022). Using this backdrop, the third research question examines the implications of users' participatory endeavours in strategic information operations' creation and circulation.

RQ3. How do users' participatory endeavours influence the circulation of information operations?

## Method

This study investigates three cases of deceptive information operations on Facebook discovered by a tool implementing a workflow designed for monitoring lists of problematic actors and their activities (Authors, 2023). The workflow builds upon existing methodologies for detecting information operations by integrating automated content collection, behavioural pattern analysis, and iterative account tracking. The workflow and analytical approach is displayed in Figure 1.

### Anonymized Workflow: Design and Implementation

The **Anonymized Workflow<sup>1</sup>** operates through a multi-step process that systematically detects, analyses, and updates lists of coordinated social media accounts. Unlike static network analysis approaches that rely on one-time data collection, this workflow utilises an adaptive methodology to track evolving coordination patterns over time (Authors, 2023).

The process begins with the identification of a curated seed list of social media accounts flagged as potentially problematic. These accounts are selected based on entries in verified

---

<sup>1</sup> Results of the workflow detection based on the Italian elections 2022 are available in a scientific journal article by Authors (2023).

fact-checking databases, investigative reports as well as academic research. Starting with this high-confidence baseline ensures that the initial sample is both credible and analytically robust. Once the seed list is established, the system initiates automated data collection. Scheduled queries are employed to systematically retrieve posts from the seed accounts at predetermined intervals. The system prioritises content that exhibits high engagement – measured through shares, reactions, and comments – as these indicators often signal broader influence or coordinated activity.

To detect patterns of coordinated behaviour, the workflow utilises the [Anonymized] Tool, which is specifically designed to identify multiple forms of synchronous sharing. These include the simultaneous posting of identical URLs across multiple accounts, similarities in image text and post blurbs, and temporal clustering in content distribution. Such behaviours often point to coordinated efforts to amplify specific narratives and distort public discourse.

Following detection, the workflow moves into the feature extraction and network expansion phase. Here, elements like URLs, images, and text snippets are analysed across the platform to identify other accounts that exhibit similar patterns of dissemination. Accounts that consistently engage in such activities are incorporated into the monitoring pool, allowing the system to adapt, in real time, to shifts in coordination strategies.

Between October 2023 and August 2024, a CrowdTangle-based implementation of this dynamic tracking approach demonstrated substantial scalability in detecting coordinated information operations. The system initiated with a seed dataset of 1,225 Facebook accounts – comprising both Pages and public groups – each strategically selected based on documented dissemination of misinformation, specifically having coordinatedly shared a minimum of four URLs from a corpus of 36,091 web pages identified as false by Meta's third-party fact-checking partners between 2017 and 2022 (Messing et al. 2020). During

the 10-month operational period, the system identified 10,681 unique coordinated links, revealing the underlying content infrastructure of influence operations, and discovered 2,126 previously undetected accounts beyond the original seed list. The incorporation of these newly identified accounts into the monitoring framework established a self-reinforcing detection mechanism that ultimately captured 7,068 coordinated posts displaying synchronized sharing patterns. Analysis revealed 17 distinct networks pursuing varied operational objectives across multiple geographic regions. These networks encompassed anti-vaccine discourse, migration-related narratives, and regionally targeted campaigns, with notable activity clusters in Southeast Asia, Eastern Europe, and Latin America. System operations ceased on August 14, 2024, following Meta's deprecation of the CrowdTangle platform.

### Case Studies Selection

We selected a convenience sample from the 17 identified networks to examine a representative cross-section of deceptive information operations. This approach enabled us to examine multiple dimensions of coordinated behaviour, including the rapid, synchronous dissemination of identical or near-identical links and images within a single group, across multiple groups by networked actors, or by users exploiting weak moderation to amplify illicit content.

Our sample encompasses a variety of actors engaged in coordinated sharing, such as individual users, automated accounts, organised groups, and pages operated by vested interests. This diversity illustrates that coordination is not limited to a specific type of actor but involves participants with different motivations and strategies. The cases also reflect the wide array of misleading objectives pursued through coordinated sharing, including

spreading misinformation, promoting engagement bait, conducting political propaganda, amplifying scams, and distributing spam and pornographic content.

After identifying these cases, we imported the lists of accounts into the Meta Content Library (the new tool designed by Meta in lieu of CrowdTangle) to streamline data retrieval and analysis. This step enabled the systematic observation of relevant posts, engagement metrics, and associated dissemination patterns of these accounts. However, the adaptation to MCL also presented challenges, such as the severe limitation on the number of downloadable posts, which necessitated conducting observations within the platform.

### The Analytical Approach

Our analytical approach followed an adaptation of grounded theory (Charmaz, 2014; Starbird et al., 2023), allowing us to derive conceptual categories inductively from qualitative data rather than imposing predefined typologies. This approach was well suited to examining how deceptive information operations manifest on Facebook, given their dynamic and context-dependent nature. The analysis scheme is available in the Appendix. The online dimension of deceptive information operations enables the study of motives, tactics, and narratives through the digital traces they leave on platforms – traces that, while incomplete (Starbird et al., 2019), are nonetheless observable and measurable. We analysed a dataset combining account-level metadata and behavioural patterns with the top 500 most-viewed posts for each case study within the time period between October 2023 and August 2024, tracked on the Meta Content Library (Meta Platforms, Inc., 2025) between November 2024 and January 2025.

Our analysis involved an iterative movement between micro – and macro-level perspectives: systematic open coding of individual posts and close reading of emerging patterns. Through constant comparison, analytical categories emerged organically from the

1  
2  
3 data and were subsequently interpreted through the theoretical frameworks of Starbird et  
4  
5 al. (2019) and Chadwick and Stanyer (2022), as discussed in the previous section.  
6  
7 Several distinctive patterns were identified:  
8  
9

- 10  
11 • Identity management, including account name changes, thematic incoherence, and  
12 indicators of administrator identity;  
13  
14 • Behavioural coordination, marked by bursts of identical posting, cross-posting  
15 across groups, and the use of cloaked or masked links;  
16  
17 • Content manipulation, involving AI-generated visuals, brandjacking, and the  
18 blending of benign and harmful narratives.  
19  
20  
21

22 Each case study was assigned to an author for primary coding. To ensure consistency, the  
23 lead author cross-coded 30% of the dataset, and all coding decisions were reviewed  
24 collaboratively to refine and validate the emerging categories.  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36

37 Figure 1. Workflow and analytical approach.  
38  
39

## Findings

40 This section illustrates the three case studies selected from a map of account networks  
41 exhibiting coordinated sharing behaviour identified by the [Anonymized Workflow]. The  
42 selected cases illustrate distinct types of deceptive information operations: political  
43 propaganda, online gambling promotion, and the dissemination of a mix of  
44 misinformation, memes, and adult content.  
45  
46

47 The subsequent sections will analyse the case studies in depth. Following our theoretical  
48 frame, the analysis of each deceptive information operation has a three-fold structure  
49 following the three research questions: [1] describing deceptive content, [2] investigating  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3 how Facebook affordances are exploited, and [3] observing users' participatory  
4 behaviours.  
5  
6  
7  
8

## 9 Pro-Putin Propaganda 10

### 11 *Spreading Deceptive Content for Political Influence* 12

13 The first case study identified 27 public groups on Facebook that exhibit strong pro-  
14 Putin  
15 affiliations<sup>2</sup>. This is the most traditional form of information operation to promote political  
16 propaganda, resonating with Chadwick and Stanyer's (2022) framework of deception  
17 driven by political motivations. The deliberate integration of apparently benign material,  
18 such as posts celebrating Moscow's urban development, alongside more explicitly  
19 politically polarising content, such as mocking the Ukrainian President, exemplifies the  
20 gaslighting nature of these information operations, fostering confusion and advancing a  
21 specific agenda (Jack, 2017). These groups display notable structural similarities, including  
22 the frequent replication of naming conventions, identical group descriptions, and  
23 synchronised name changes preceding key geopolitical events. For instance, the group  
24 *THE BEST PRESIDENT, VLADIMIR VLADIMIROVICH PUTIN*, which has  
25 approximately 35,000 members, includes a profile description outlining community rules  
26 that explicitly prohibit abusive language and spam while also not overtly expressing a pro-  
27 Putin stance. Other groups with nearly identical names and descriptions, which have twice  
28 as many members, also exhibit patterns of coordinated behaviour. Notably, several groups  
29 underwent name changes in early 2022, just weeks before Russia invaded Ukraine,  
30 suggesting strategic adaptation in response to geopolitical developments.

31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
<sup>2</sup> The investigation focused on accounts available in the Meta Content Library at the time of the analysis between November 2024 and January 2025 (N=15). The complete list of accounts is available at this link: [Anonymized GitHub Repository].

The content disseminated within these groups can be grouped into two primary thematic categories: first, posts that disparage Western leaders and Ukraine, and second, content that seeks to bolster the image of Putin and Russia. Among the most widely circulated posts in the observed period was a meme targeting Volodymyr Zelenskyy, shared in coordination across multiple pro-Putin groups within a tightly constrained time frame. On February 12, 2024, between 9:12 and 9:18 CEST, six groups posted an identical image portraying the then U.S. President Joe Biden in a Soviet military uniform, sternly holding a cat with Zelenskyy's face, accompanied by the caption: "Dad, don't leave me... Russia has won, I don't need you anymore" (See Fig. 2). The timing of this coordinated posting coincided with significant battlefield developments, including a Russian drone strike in Kharkiv that resulted in civilian casualties<sup>3</sup>. This suggests that the disinformation network actively leveraged real-world events to reinforce the narrative of Russian military superiority while portraying Western support for Ukraine as diminishing.

In contrast to the ridicule-oriented posts, another subset of content was designed to project a positive image of Putin and Russia. Several widely circulated posts sought to depict Moscow as a superior urban centre compared to Western cities. In one instance, two posts prominently featured American journalist Tucker Carlson, who apparently described Moscow as "a more pleasant city than anywhere in the U.S." during his speech at the World Government Summit in Dubai. Another post claimed that a Canadian farming family had relocated to Russia to preserve traditional values, reinforcing a broader narrative of Russia as a bastion of moral integrity in contrast to perceived Western decadence. Additionally, several AI-generated images depicted Putin in various heroic poses, including one surrounded by young women, seemingly reinforcing his appeal to younger generations.

<sup>3</sup> <https://www.aljazeera.com/news/2024/2/10/overnight-russian-drone-attack-kills-at-least-seven-in-ukraines-kharkiv>

Other images drew on nationalistic symbolism, such as a Russian bear wielding a Kalashnikov rifle in front of a national flag or a lion juxtaposed against hyenas bearing the flags of Ukraine, the EU, and the US. These visual representations functioned as implicit ideological statements, reinforcing a binary opposition between Russia and the West.

#### *Exploiting Facebook's Affordances for Algorithmic Amplification*

The systematic amplification of content within these groups reveals a deliberate exploitation of Facebook's platform affordances to maximise visibility and engagement. In particular, coordinated sharing behaviour is employed to game Facebook's ranking algorithms, prioritising content that generates engagement. By ensuring that identical posts appear in multiple locations within short intervals, the actors behind these operations effectively created an illusion of organic popularity, increasing the likelihood that the content would be further highlighted by the platform's recommendation algorithms. A good example demonstrating a high degree of coordination in its dissemination of content is WorldRusWorld<sup>4</sup>, managed by Bulgarian and Russian moderators and one of the most active content producers within this network. On October 1, 2023, for instance, WorldRusWorld posted an identical pro-Putin image in ten different groups within a three-minute window (16:21–16:24 CEST). The near-simultaneous posting pattern suggests an organised effort to amplify reach and visibility, indicative of a coordinated campaign.

The structured timing of content dissemination aligns with previous studies on coordinated influence operations, where automated or semi-automated systems play a key role in content propagation (O'Donovan, 2024). Furthermore, while most of the content was in Russian, some of the most widely viewed posts appeared in Arabic and English, indicating an effort to target non-Russian audiences (see Figure 2 for an example).

---

<sup>4</sup> This Page disappeared at the time of writing (March 2025) .

### Coordinated Participation and Broader Implications

The engagement metrics associated with these posts reveal significant variation in effectiveness. Some failed to attract substantial engagement, with posts generating only a few hundred views and minimal interaction. However, others demonstrated considerable reach, with the most successful posts exceeding 100,000 views and accumulating thousands of reactions, comments, and shares. For instance, the Zelenskyy meme campaign in Figure 2, which displayed clear signs of coordination, accumulated over 1.1 million views, underscoring the extensive reach and potential impact of these efforts.

While the initial coordination of content is orchestrated by a relatively small group of actors, the amplification of these messages often depends on the actions of ordinary users who unwittingly participate in spreading misleading content. This interplay between orchestrated coordination and organic participation reflects the broader challenges posed by contemporary disinformation campaigns, where the distinction between authentic and inauthentic engagement becomes increasingly blurred.

Overall, this first case study highlights a state-aligned operation promoting pro-Putin propaganda. Coordinated Facebook pages and groups create the illusion of widespread grassroots support for Russia's geopolitical goals. Automated tools synchronise content dissemination, exploiting Facebook's features to amplify narratives supportive of the Russian government. This behaviour aligns with previous findings on using social media affordances for propagandistic purposes (Authors 2020b).

Figure 2. On the left: Meme mocking Zelenskyy with the satirical caption shared simultaneously across multiple pro-Putin accounts; on the right: Post in Arabic in poetic language to express longing and deep emotions for Putin.

### Online Casino Engagement Bait

#### *Spreading Deceptive Content for Gambling Promotions*

The second case study focuses on a coordinated campaign promoting online casinos and games through 260 Facebook groups<sup>5</sup>. These groups, with memberships ranging from 30,000 to nearly 600,000, use automation and generative AI to produce and share content promoting a set of digital casino websites. The consistent thematic branding in group names reflects a deliberate strategy to attract users with gambling incentives. Notably, among the 500 most-viewed posts from the network, 368 were no longer accessible at the time of the analysis, likely due to systematic policy violations leading to content takedowns. This suggests a strategy designed to maximise exposure before removal.

These groups signal deception through branding tactics, prominently featuring gambling brands like Orion Stars and Juwa – appearing 111 and 90 times, respectively – without verified affiliations. For example, the third-largest group, *FREE PLAY ALL GAMES (JUWA CITY ORION STARS / FIRE KIRIN / GAME VAULT)* (429,332 members), illustrates how these brands gain visibility within the network. By leveraging well-known names, these groups promote lesser-known casino apps through brandjacking (Du Plessis, 2018).

Beyond misleading branding, these operations cause harm by luring users into scams under the guise of legitimate gambling platforms, exploiting vulnerabilities, and fostering economic harm. They also employ synthetic content, such as AI-generated visuals and comments, further polluting the digital environment (DiResta & Goldstein, 2024) and fueling distrust. The operation's ability to test and refine policy circumvention strategies

<sup>5</sup> The investigation focused on accounts available in the Meta Content Library at the time of the analysis (N=194). The complete list of accounts is available at this link: [Anonymized GitHub Repository].

amplifies its impact and lays the groundwork for future campaigns. The accumulation of social media assets and networked groups increases their potential for repurposing in other deceptive operations (Donovan et al., 2020), including disinformation or fraud, posing a broader threat to digital integrity.

#### *Exploiting Facebook's Affordances for Algorithmic Amplification*

Automation plays a key role in these groups for maintaining high posting frequency and consistency. The groups produced about 10,000 posts in a single month (Fig. 3), a difficult output manually. The near-simultaneous distribution of identical promotional content across multiple groups suggests the use of scheduling tools and bots, ensuring continuous and widespread publication. This tactic amplifies visibility within Facebook's algorithmic feed, increasing user exposure.

The dataset's most-viewed posts show a stark imbalance, with disproportionately high comment count relative to reactions and shares. One post, for instance, amassed 169.1K views and 44.1K comments but only 7 reactions and 0 shares – an anomaly suggesting artificial comment inflation. These comments often contain generic praise (“great platform,” “awesome gameroom”), random alphanumeric strings (“k4,” “oyy,” “flq”), or unspaced words (“Awesomekeepuptheamazingworkandkeepdoingyourbest”), aligning with synthetic engagement strategies (Cresci et al., 2016). By fabricating popularity, these campaigns mislead users into perceiving widespread endorsement, a tactic consistent with broader concerns about increasing automation in digital interactions as within the broader discussions on the ‘Dead Internet Theory’ (Muzumdar et al., 2025) positing that much of the online content and interactions are generated by bots rather than humans.

Group administrators play a crucial role in sustaining this automation-driven strategy. Many administrator accounts show minimal non-automated activity, resembling

compromised profiles used in spam networks. The analysis suggests these accounts are centrally managed, facilitating seamless gambling content distribution.

For instance, the largest group, *ORION STARS, JUWA, GAME VAULT, FIRE KIRIN, MOOLAH - FREEPLAY* (595,839 members), is managed by 6 administrators and 5 moderators. Except for the page “OrionStars Freeplay 24h/7,” the remaining admins have low engagement, only a few followers, and very little content. Many label themselves as “digital creators,” yet their profiles lack meaningful activity. Some administrators, like “Dm me 25\$ Freeplay” and “Message me for 30Freeplay,” exist solely to promote entrance bonuses, lacking original content or followership.

Shared content analysis reveals frequent low-engagement posts, emphasising promotion. Posts often advertise external casino pages offering bonuses or invite users to join private chats that include reward mentions in their titles.

### *[Synthetic] Coordinated Participation and Broader Implications*

Although user participation in these gambling groups appears extensive, much of it is artificially generated through AI visuals and engagement bait tactics that simulate activity. This deceptive model blends synthetic and real user interactions, making genuine users unknowingly complicit in amplifying misleading content.

Our analysis of the 500 most-viewed posts reveals centralised coordination. AI-generated imagery plays a key role, enhancing visual appeal while streamlining content production. These graphics – often depicting slot machines, lions, sharks, and regal female figures – serve to attract attention while enabling rapid, cost-effective content creation. A pattern of image recycling is evident, with the same promotional graphics reappearing across multiple groups, such as the AI-generated shark image (Fig. 4) repeatedly disseminated

1  
2  
3 across the network on the same day, often by anonymous Facebook profiles posting in  
4  
5 near-simultaneous intervals (e.g., multiple posts on August 6, 2024, at 01:40 AM).  
6  
7  
8  
9

10 This follows an established pattern where content from central Facebook pages is shared  
11 by private or public profiles in these groups. Public profiles engaging in this behaviour  
12 typically describe themselves as digital creators, have at least 1,000 followers, and blend  
13 casino promotions with investment and life-coaching content. One such profile, for  
14 example, mirrors behaviours observed in the following case study on exploiting  
15 unmoderated groups for deceptive content. These accounts remain largely inactive over  
16 time, apart from periodic profile image changes. Engagement incentives, such as free play  
17 bonuses for likes and comments, encourage real users to interact with manipulated posts.  
18 This boosts visibility, attracting further engagement and reinforcing the content's presence  
19 in the platform's algorithmic ranking.  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34

35 **Figure 3.** Boxes showing the activity of some of the accounts in the network.  
36  
37  
38  
39

40 **Figure 4.** Four examples of posts shared from the network accounts displaying AI-  
41  
42 generated visuals.  
43  
44  
45

## 46 **Unmoderated Groups Flooded with Explicit Imagery** 47 48

### 49 *Spreading Deceptive Content for Circulating Explicit and Illicit Material* 50 51

52 The third case study examined 222 public Facebook groups, which formed two distinct  
53 communities: one with 187 accounts and another with 35<sup>6</sup>. The findings reveal deceptive  
54  
55  
56  
57  
58

59 <sup>6</sup> The investigation focused on accounts available in the Meta Content Library at the time of the analysis  
60 (N=111). The complete list of accounts is available at this link: [Anonymized GitHub Repository].

practices exploiting large, poorly moderated spaces. The deceptive nature of these operations manifests itself in two distinct ways. First, some groups are misleadingly renamed, shifting between political and entertainment themes. For example, a group originally named *EFF Juju Nation* was later renamed *DSTV Premier League Soccerzela* and changed names 13 times. Before the South African elections in May 2024, it was temporarily rebranded as *2024 Elections South Africa*, then reverted to a soccer-related name post-election. Despite these shifts, its core content remained apolitical, suggesting a strategy to manipulate public perception while evading detection. Similarly, *The Golden Girls Fan Club* was previously linked to the *Pakistan Peoples Party (PPP)*, posting pro-PPP content before going dormant for nearly a year. It later rebranded as a fandom group for The Golden Girls while mixing sitcom-related posts with spam and celebrity death hoaxes, exposing users to misinformation when they are most susceptible – while consuming entertainment content.

Second, the lack of moderation allows misinformation on wide-ranging topics, adult content with scammy links, fabricated job ads, and illegal services like follower sales to spread unchecked. A notable example of misinformation is the false claim that Coca-Cola contains alcohol. Despite being debunked by fact-checkers<sup>7</sup>, it circulated widely, accumulating 3.5 million views and 37,300 shares at the time of analysis (Fig. 5).

#### *Exploiting Facebook's Affordances for Algorithmic Amplification*

The [Anonymized Workflow] detected a sophisticated deception technique within this information operation on Facebook. The operation begins when private users share a “listening to” activity status update inside these groups. This status includes a link to a Facebook page categorized as Musician/Band. However, these pages are typically empty,

<sup>7</sup> See <https://factcheck.afp.com/doc.afp.com.32B23RF>.

with profile images that change periodically to clickbait visuals—riddles, pornography, or emotionally charged images of deceased individuals or people with visible disabilities. We hypothesise that this tactic helps evade Meta's detection of CIB.

This is not the only circumvention method observed. For instance, adult content is often disguised under unrelated blurbs narrating the history of BMW, the automobile company, all posted with a set of hashtags, including *#carporn* (see Fig. 5). These posts are strategically coordinated to maximise visibility and mask the real link destination under seemingly safe domains, including, ironically, *facebook.com*. An anonymous participant shares a post in a public group originally created by another user. It structures as a “link” status update (e.g., “reading,” “listening”, “attending”) and appears automotive-related, referencing BMW’s history. However, its preview image contains explicit content. The post links to an external website through *tinyurl.com* link shortener.

Many of these posts follow a formula that maximises engagement by leveraging sensationalism, humour, or controversy. Engagement bait – such as provocative memes and misleading headlines – ensures sustained visibility. A particularly viral example featured a female teacher in a classroom with the caption “Teachers nowadays are 

#### *Coordinated Participation and Broader Implications*

The user base of these accounts plays a key role in orchestrating information operations, often with or without administrators' knowledge. The lack of moderation in these groups allows deceptive content to proliferate. Among the 500 most-viewed posts, 38 surpassed one million views, with the top three exceeding 3 million. One widely circulated example was a manipulated meme of Phineas from Phineas and Ferb, humorously questioning his shirt-wearing habits. Though seemingly harmless, its viral spread across X, YouTube,

Reddit, and TikTok (confirmed via Google Lens) illustrates how user engagement drives visibility, regardless of accuracy or intent.

Some groups adopt misleading appearances in order to attract more members. For example, the group *Sons and Daughters of Apostle and Dr. Lizzy Johnson Suleiman* presents itself as a religious community, yet users there primarily share adult content (DiResta & Goldstein, 2024).

Broadly, this case study underscores how deception is deeply embedded in Facebook's information operations and CIB. By exploiting platform affordances – such as name changes, AI-generated content, and algorithmic amplification – these networks manipulate public discourse while avoiding detection. High user engagement further fuels their spread, ensuring the continued circulation of deceptive content.

Figure 5. On the left: A widely circulated image of a female teacher in a classroom, shared with a provocative caption; In the center: A viral piece of misinformation claiming that Coca-Cola contains alcohol. On the right: An anonymous participant shares a post in a public group originally created by another user. The post blurb briefly tells BMW's history. However, its preview image contains explicit content.

## Discussion and Implications of Findings in Light of the Recent Platforms Lifting Restrictions on Deceptive Content

Previous research shows that information operations exploit social media affordances and mobilise distributed online crowds to pursue malicious objectives (Starbird et al., 2019). According to Chadwick and Stanyer's (2022) framework, malicious objectives range from financial gain to political influence. This study integrates insights from both to advance the concept of *deceptive information operations*—a broader analytical lens capturing

misleading and manipulative activities often overlooked in traditional accounts of disinformation.

Across all three case studies analyzed using grounded theory, we observe common patterns. Deceptive content – anti-democratic political propaganda, gambling promotion (which falls outside Meta's policies), and adult material – are amplified by actors who exploit two interconnected platform vulnerabilities central to strategic information operations (Starbird et al., 2019): user participatory behaviour and biases in Facebook's algorithmic feed. First, Facebook's algorithmic ranking system amplifies engagement-driven content, a process often intensified through participatory or simulated-participatory behaviours such as coordinated sharing. Second, moderation is inconsistently enforced. Together, these vulnerabilities create information asymmetries between operators and audiences, blurring the line between authentic and manipulated activity.

The state-backed political operation created coordinated pages and groups, using automated sharing and AI-generated images to trigger the algorithmic amplification of pro-Kremlin narratives. It also blended in seemingly harmless content to reduce the risk of detection and removal. The gambling operation mass-produced promotional materials with generative AI and bot-driven engagement, overwhelming the platform more quickly than moderation systems could respond. The operation distributing adult material repurposed large, seemingly benign Facebook groups—such as fandom or meme communities—for pornographic scams, exploiting spaces with minimal or absent moderation and embedding harmful content alongside unrelated posts to avoid bans. Across all three cases, the same dual mechanism enabled large-scale circulation of harmful content: algorithms incentivised engagement, while Facebook's enforcement framework failed to reliably distinguish automated from genuine participation, or deceptive material from legitimate content.

Taken together, these dynamics evoke Koebler's (2024) notion of the "Zombie Internet": an information ecosystem in which the distinction between human and automated behavior becomes increasingly difficult to maintain. As algorithms prioritize engagement over considerations of authenticity, and moderation systems struggle to reliably detect coordinated activity, the boundary between genuine participation and manipulation progressively erodes. The result is a feedback process sustained by manufactured interactions and synthetic identities, enabling harmful content to circulate across platforms with limited friction.

While our study identifies recurring patterns, its descriptive design does not allow us to determine whether Facebook is itself undergoing "zombification." Addressing that question would require more systematic, longitudinal research. Nonetheless, the findings reveal structural weaknesses in platform governance. Moderation on Facebook appears uneven: deceptive information operations can evade detection, amplify visibility before intervention, or persist in poorly moderated spaces. This inconsistency reflects not only enforcement gaps but also the fundamental difficulty in distinguishing authentic from "inauthentic", to use a Meta's concept (Gleicher, 2018), activity at scale. Importantly, these vulnerabilities are not unique to Facebook. The same affordances that sustain online communities – algorithmic curation, engagement-driven visibility, and group autonomy – also complicate governance efforts. This reveals a deeper tension: the very features that underpin social media's commercial success are those that make it susceptible to manipulation. The "Zombie Internet," in this sense, does not emerge from deliberate neglect or conspiracy, but from the structural collision between engagement-oriented design and the limits of platform governance.

Against this backdrop, the need for further research becomes increasingly urgent, particularly as social media companies appear to be scaling back safety investments

(Kaplan, 2025). Our reporting of these findings to Meta in October 2023 and November 2024 was met with acknowledgement but limited follow-up action. These exchanges also reveal a deeper problem: platform categories such as CIB rely on a narrow and ultimately bland definition of harm, overlooking a wider spectrum of practices that constitute deceptive information operations. This study contributes by advancing a clearer conceptualisation of these operations and by grounding it in concrete empirical cases – an essential step, since effective governance is difficult without clarity about what counts as harmful activity.

As social media continues to shape public discourse, the patterns identified here underscore the need for stronger platform accountability and sustained collaboration between independent researchers and technology companies. Without structured partnerships that support systematic investigation and timely responses to emerging threats, the integrity of the information ecosystem will remain precarious.

## References

- Allen, J., 2022. Misinformation Amplification Analysis and Tracking Dashboard. Integrity Institute. Available at: <https://integrityinstitute.org/blog/misinformation-amplification-tracking-dashboard>.
- Al-Rawi, A., 2019. Viral news on social media. *Digital journalism*, 7(1), pp.63–79. Available at: <https://www.tandfonline.com/doi/abs/10.1080/21670811.2017.1387062>.
- Bennett, L.W. & Segerberg, A., 2013. *The Logic of Connective Action: Digital Media and the Personalization of Contentious Politics*. Cambridge University Press.
- Bradshaw, S. & Howard, P.N., 2018. Challenging truth and trust: A global inventory of organized social media manipulation, Oxford Internet Institute. Available at: <https://demtech.oi.i.ox.ac.uk/research/posts/challenging-truth-and-trust-a-global-inventory-of-organized-social-media-manipulation/>.
- Broadwater, L., 2025. Trump’s “flood the zone” strategy leaves opponents gasping in outrage. *The New York Times*. Available at: [https://www.nytimes.com/2025/01/28/us/politics/trump-policy-blitz.html?unlocked\\_article\\_code=1.904.JWp\\_.6qWEPiGSsp2o&smid=url-share](https://www.nytimes.com/2025/01/28/us/politics/trump-policy-blitz.html?unlocked_article_code=1.904.JWp_.6qWEPiGSsp2o&smid=url-share).
- Buller, D.B. & Burgoon, J.K., 1996. Interpersonal deception theory. *Communication theory: CT: a journal of the International Communication Association*, 6(3), pp.203–242. Available at: <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1468-2885.1996.tb00127.x>.
- Chadwick, A. & Stanyer, J., 2022. Deception as a bridging concept in the study of disinformation, misinformation, and misperceptions: Toward a holistic framework. *Communication theory: CT: a journal of the International Communication Association*, 32(1), pp.1–24. Available at: <http://dx.doi.org/10.1093/ct/qtab019>.
- Charmaz, K. (2014). Grounded theory in global perspective: Reviews by international researchers. *Qualitative inquiry*, 20(9), 1074-1084.
- Chergarova, V. et al., 2022. Cryptocurrency fraud: A study on the characteristics of criminals who are using fake profiles on a social media platform to persuade individuals to invest into cryptocurrency. *Issues In Information Systems*, 23(3), pp.242–252. Available at: [https://iacis.org/iis/2022/3\\_iis\\_2022\\_242-252.pdf](https://iacis.org/iis/2022/3_iis_2022_242-252.pdf).
- Cresci, S. et al., 2016. DNA-Inspired Online Behavioral Modeling and Its Application to Spambot Detection. *IEEE intelligent systems*, 31(5), pp.58–64. Available at: <http://dx.doi.org/10.1109/MIS.2016.29>.
- DiResta, R. & Goldstein, J.A., 2024. How Spammers and Scammers Leverage AI-Generated Images on Facebook for Audience Growth. *arXiv [cs.CY]*. Available at: <http://arxiv.org/abs/2403.12838>.
- Donovan, J. et al., 2020. The Media Manipulation Definitions. *The Media Manipulation Casebook*. Available at: <https://mediamanipulation.org/definitions>.
- Du Plessis, C., 2018. Examining the brandjacking phenomenon in the digital age. In *Communication, Mass Media & Society Conference (COMASS18)*, July. pp. 19–20. Available at: <https://scholar.google.com/citations?user=4BGeXw0AAAAJ&hl=en&oi=sra>.
- Fowler, G.A., 2016. What if Facebook Gave Us an Opposing-Viewpoints Button. *Wall Street Journal*.
- Authors, 2023. <https://doi.org/10.1177/20563051231196866>

- 1  
2  
3 Authors, 2020a. <https://doi.org/10.1145/3400806.3400817>.
- 4  
5 Authors, 2020b. <http://dx.doi.org/10.1080/1369118X.2020.1739732>.
- 6  
7 Authors, 2020. <https://ora.uniurb.it/handle/11576/2675493>.
- 8  
9 Gillespie, T., 2021. Custodians of the internet, New Haven, CT: Yale University Press.
- 10  
11 Gillespie, T., 2014. The relevance of algorithms. In Media Technologies. The MIT Press, pp. 167–194.  
12 Available at: [https://direct.mit.edu/books/edited-volume/chapter-pdf/2273545/9780262319461\\_cai.pdf](https://direct.mit.edu/books/edited-volume/chapter-pdf/2273545/9780262319461_cai.pdf).
- 13  
14 Gleicher, N., 2018. Coordinated inauthentic behavior explained. Retrieved August, 19, p.2019.
- 15  
16 Graham, T., 2024. The “inauthenticity” paradox: How platforms profit from and shape coordinated  
17 inauthentic behaviour. Available at:  
18 <https://sites.google.com/uniurb.it/csbDetectionconf/home#h.te0xvcsx5g4q>.
- 19  
20 Graham, T. & QUT Digital Observatory, 2020. Coordination Network Toolkit, Queensland University  
21 of Technology. Available at: [https://doi.org/10.25912/RDF\\_1632782596538](https://doi.org/10.25912/RDF_1632782596538).
- 22  
23 Guess, A., Nagler, J. & Tucker, J., 2019. Less than you think: Prevalence and predictors of fake news  
24 dissemination on Facebook. Science advances, 5(1), p.eaau4586. Available at:  
25 <http://dx.doi.org/10.1126/sciadv.aau4586>.
- 26  
27 Huth, K., Peters, J. & Seufert, J., 2020. The heartland lobby. [correctiv.org](https://correctiv.org/en/top-stories/2020/02/11/the-heartland-lobby/). Available at:  
28 <https://correctiv.org/en/top-stories/2020/02/11/the-heartland-lobby/>.
- 29  
30 Jack, C., 2017. Lexicon of lies: Terms for problematic information. Data & Society, 3. Available at:  
31 <https://apo.org.au/sites/default/files/resource-files/2017/08/apo-nid183786-1180516.pdf>.
- 32  
33 Jenkins, H., Ito, M. & Boyd, D., 2015. Participatory Culture in a Networked Era: A Conversation on  
34 Youth, Learning, Commerce, and Politics, John Wiley & Sons.
- 35  
36 Kaplan, J., 2025. More speech and fewer mistakes. Meta. Available at:  
37 <https://about.fb.com/news/2025/01/meta-more-speech-fewer-mistakes/>.
- 38  
39 King, M., 2024. The Spaghetti Approach: How Trivial Misinformation Can Cause Significant  
40 Damage. Available at: [https://www.linkedin.com/pulse/spaghetti-approach-mark-king-onnvc?utm\\_source=share&utm\\_medium=member\\_android&utm\\_campaign=share\\_via](https://www.linkedin.com/pulse/spaghetti-approach-mark-king-onnvc?utm_source=share&utm_medium=member_android&utm_campaign=share_via).
- 41  
42 Koebler, J., 2024. Facebook’s AI spam isn’t the “dead internet”: It’s the zombie internet. 404 Media.  
43 Available at: <https://www.404media.co/facebook-ai-spam-isnt-the-dead-internet-its-the-zombie-internet/>.
- 44  
45 Authors, 2023. DOI: 10.1445/106772
- 46  
47 Messing, S., DeGregorio, C., Hillenbrand, B., King, G., Mahanti, S., Mukerjee, Z., Nayak, C., Persily,  
48 N., State, Bogdan, & Wilkins, A. (2020). Facebook Privacy-Protected Full URLs Data Set  
49 (Version 10). Harvard Dataverse. <https://doi.org/10.7910/DVN/TDOAPG>
- 50  
51 Meta Platforms, Inc. (2025). Meta Content Library API version v5.0.  
52 Available at: <https://developers.facebook.com/docs/content-library-and-api/citations>.
- 53  
54 Molina, M.D., 2023. Do people believe in misleading information disseminated via memes? The role  
55 of identity and anger. New media & society. Available at:  
56 <https://journals.sagepub.com/doi/abs/10.1177/14614448231186061>.

Muzumdar, P. et al., 2025. The Dead Internet Theory: A survey on artificial interactions and the future of social media. arXiv [cs.CY]. Available at: <http://arxiv.org/abs/2502.00007>.

Newman, N. et al., 2021. Reuters Institute Digital News Report 2021. Available at: <https://papers.ssrn.com/abstract=3873260>.

O'Donovan, K., 2024. Supporting the UK General Election in 2024. Google Blog. Available at: <https://blog.google/around-the-globe/google-europe/united-kingdom/google-2024-uk-election-support/> [Accessed December 10, 2024].

Parham, J., 2018. Targeting Black Americans, Russia's IRA Exploited Racial Wounds. Wired.

Paris, B. S., & Donovan, J. M. (2019). Deepfakes and cheap fakes. United States of America: Data & Society. [https://datasociety.net/wp-content/uploads/2019/09/DS\\_Deepfakes\\_Cheap\\_FakesFinal-1-1.pdf](https://datasociety.net/wp-content/uploads/2019/09/DS_Deepfakes_Cheap_FakesFinal-1-1.pdf)

Prior, M., 2014. Post-broadcast democracy: How media choice increases inequality in political involvement and polarizes elections, Cambridge, England: Cambridge University Press.

Quandt, T., 2018. Dark Participation. *Media and communication*, 6(4), pp.36–48. Available at: <https://www.cogitatiopress.com/mediaandcommunication/article/view/1519>.

Rogers, R. & Righetti, N., 2024. Coordinated Inauthentic and Authentic Behaviours Online. A Typology of Attention Hijacking. Available at: <https://aoir.org/aoir2024/>.

Romero-Vicente, A., Miguel, R. & Sessa, M.G., 2024. Coordinated Inauthentic Behaviour (CIB) detection tree, EUDL. Available at: <https://www.disinfo.eu/publications/coordinated-inauthentic-behaviour-detection-tree/>.

Rubin, V.L., 2016. Deception detection and rumor debunking for social media. In The SAGE Handbook of Social Media Research Methods. 1 Oliver's Yard, 55 City Road London EC1Y 1SP: SAGE Publications Ltd, pp. 342–363.

Select Committee on Intelligence, 2019. Russian Active Measures Campaigns and Interference In the 2016 US Election, 116th US Senate Congress.

Silverman, C. & Alexander, L., 2016. How Teens In The Balkans Are Duping Trump Supporters With Fake News. BuzzFeed News. Available at: <https://www.buzzfeednews.com/article/craigsilverman/how-macedonia-became-a-global-hub-for-pro-trump-misinfo>.

Starbird, K., Arif, A. & Wilson, T., 2019. Disinformation as Collaborative Work: Surfacing the Participatory Nature of Strategic Information Operations. *Proc. ACM Hum.-Comput. Interact.*, 3(CSCW), pp.1–26. Available at: <https://doi.org/10.1145/3359229>.

Starbird, K., DiResta, R. & DeButts, M., 2023. Influence and Improvisation: Participatory Disinformation during the 2020 US Election. *Social media + society*, 9(2). Available at: <http://dx.doi.org/10.1177/20563051231177943>.

Vargo, C.J., Guo, L. & Amazeen, M.A., 2018. The agenda-setting power of fake news: A big data analysis of the online media landscape from 2014 to 2016. *New Media & Society*, 20(5), pp.2028–2049. Available at: <https://doi.org/10.1177/1461444817712086>.

Wardle, C. & Derakhshan, H., 2017. Information disorder: Toward an interdisciplinary framework for research and policymaking, Council of Europe. Available at: <http://tverezo.info/wp-content/uploads/2017/11/PREMS-162317-GBR-2018-Report-desinformation-A4-BAT.pdf>.

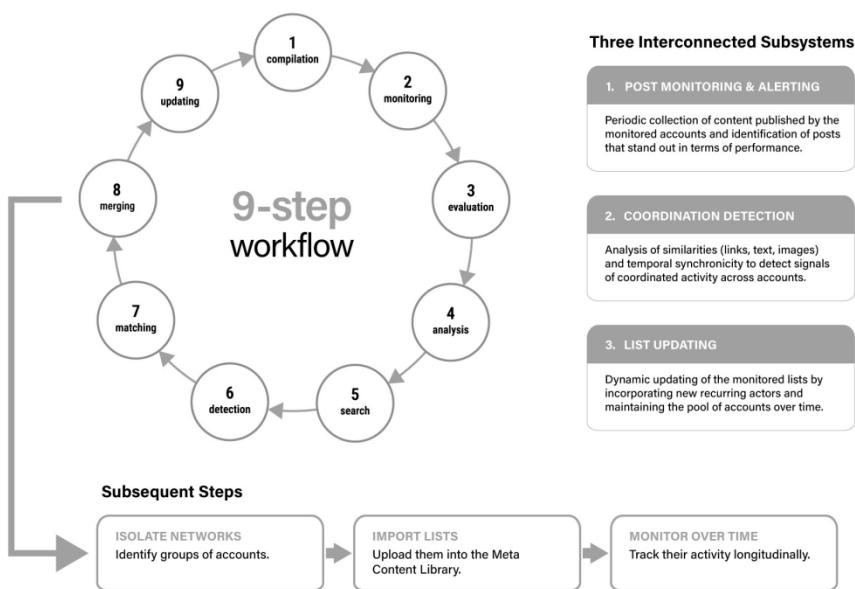
1  
2  
3 Weedon, J., Nuland, W. & Stamos, A., 2017. Information Operations and Facebook. Meta Newsroom.  
4 Available at:  
5 [https://i2.res.24o.it/pdf2010/Editrice/ILSOLE24ORE/ILSOLE24ORE/Online/\\_Oggetti\\_EMBEDDEd/Documenti/2017/04/28/facebook-and-information-operations-v1.pdf](https://i2.res.24o.it/pdf2010/Editrice/ILSOLE24ORE/ILSOLE24ORE/Online/_Oggetti_EMBEDDEd/Documenti/2017/04/28/facebook-and-information-operations-v1.pdf).  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

## Appendix 1

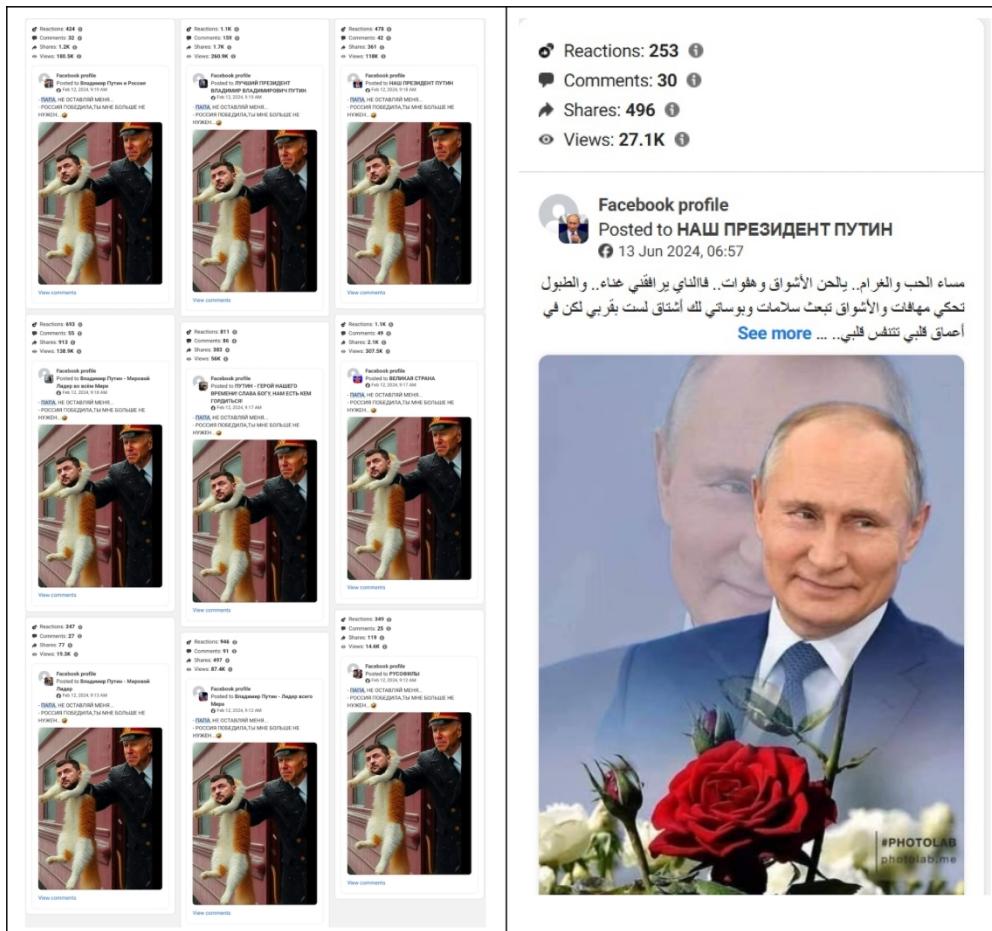
**Table A1. Codebook for Deceptive Information Operations (derived from grounded coding).**

Dimension	Emergent Categories	Description of the Category and/or Indicative Examples
Identity management	<i>Account name changes</i>	Accounts using fluid identities: The group <i>DSTV Premier League Soccerzela</i> changed names 13 times.
	<i>Thematic incoherence</i>	Thematic changes over time or theme declared mismatch with content published. E.g., <i>The Golden Girls Fan Club</i> group was previously linked to the Pakistan Peoples Party (PPP). It later rebranded as a fandom group.
	<i>Indicators of administrator identity</i>	Nationality and profile information of the administrators. For example, ORION STARS, JUWA, GAME VAULT, FIRE KIRIN, MOOLAH - FREEPLAY admin accounts have generally low engagement, few followers, and little content. Many label themselves as “digital creators,” yet their profiles lack meaningful activity.
	<i>Descriptive quantitative data</i>	Number of accounts in the network, Groups members, Pages followers, Number of admins, Creation date, etc.
Behavioral coordination	<i>Bursts of identical posting</i>	WorldRusWorld posted an identical pro-Putin image multiple times.

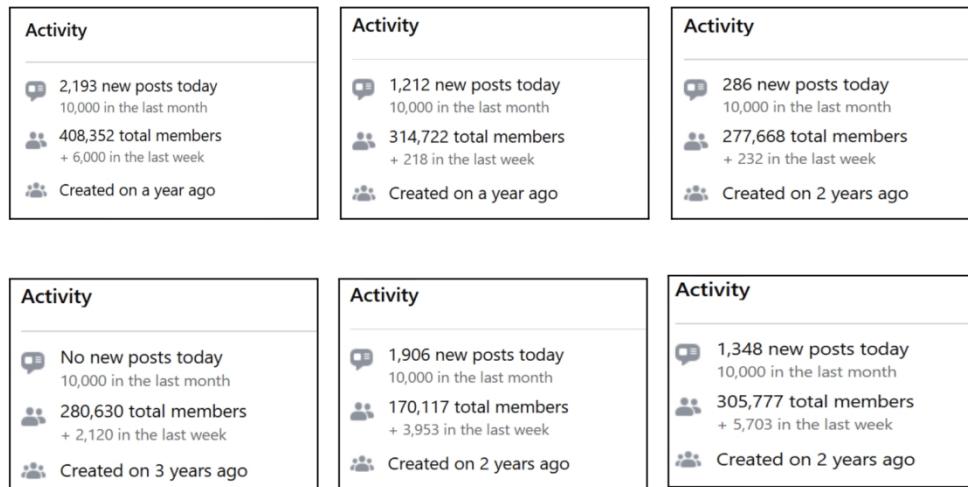
<b>Content manipulation</b>	<i>Posting near-duplicated content in a short time window</i>	The near-duplicated pro-Putin posts were shared by WorldRusWorld within a three-minute window (16:21–16:24 CEST).
	<i>Cross-posting among accounts</i>	WorldRusWorld posts the same content among 10 different groups.
	<i>Use of cloaked or masked links/posts</i>	Use of link shorteners or activity-status posts to link unrelated pages.
	<i>AI-generated visuals</i>	Synthetic or manipulated imagery designed such as “Shark” casino image or heroic Putin portraits.
	<i>Brandjacking</i>	Unauthorized use of known brand names to imply legitimacy. “FREE PLAY ALL GAMES (JUWA CITY ORION STARS / FIRE KIRIN).”
	<i>Narrative mixing</i>	Combination of benign and misleading materials that obscure intent. For example, lifestyle imagery mixed with anti-Ukraine propaganda.
	<i>Unrelated content mixing</i>	Use of innocuous blurbs to mask the content of the post or the external destination such as in the case of sexually explicit material circulated with the BMW history blurb.



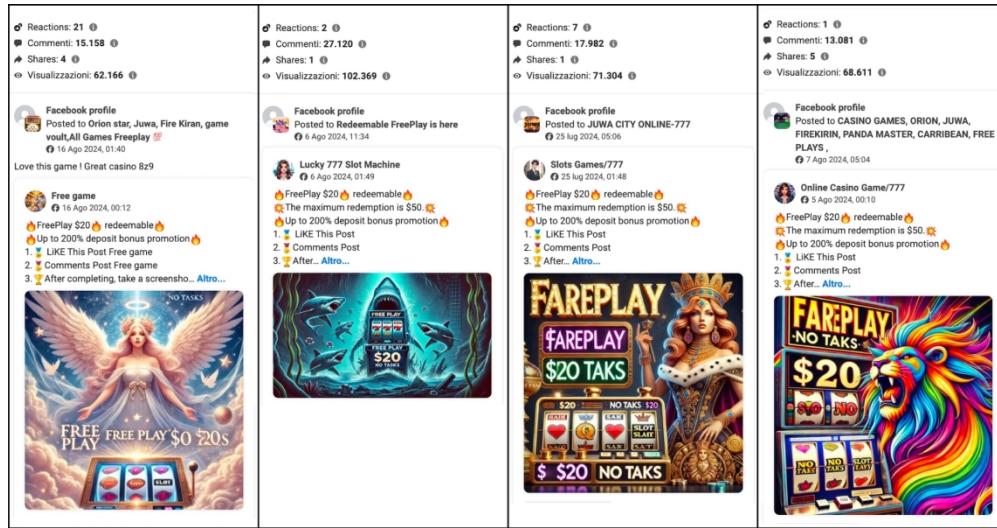
338x235mm (150 x 150 DPI)



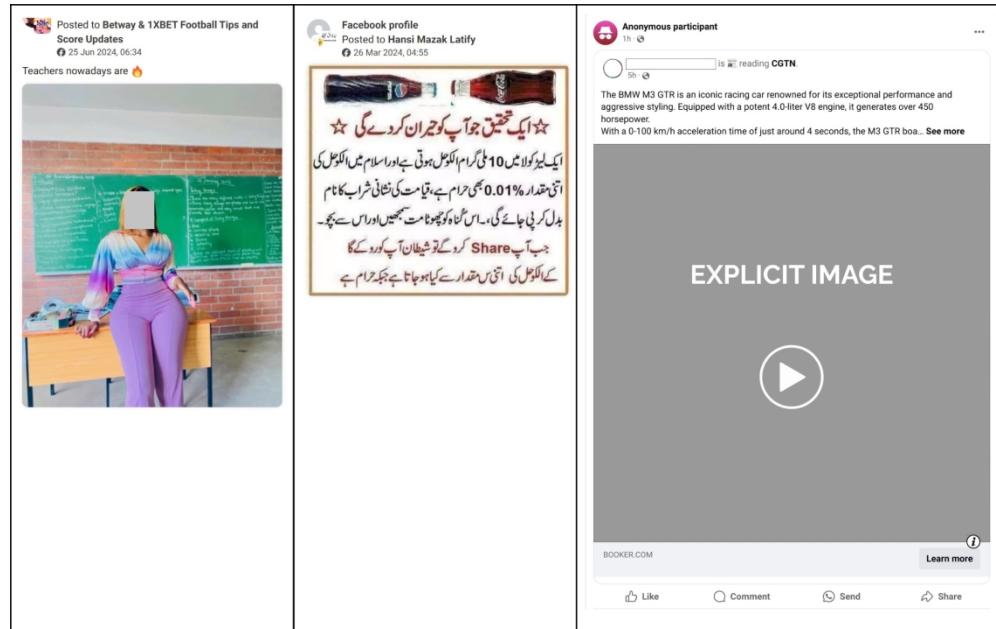
199x186mm (300 x 300 DPI)



321x161mm (144 x 144 DPI)



199x104mm (300 x 300 DPI)



199x126mm (300 x 300 DPI)