DGAP MEMO
German Council on Foreign Relations e. V.
No. 1 | February 2024

The AI Election Year: How to Counter the Impact of Artificial Intelligence

Dr. Katja Muñoz

In 2024, democracy takes center stage with over 70 elections worldwide. Already, the strategic manipulation of voters is intensifying, and is supercharged by generative AI. This manipulation of information space poses a significant threat. In this "AI election year," Germany must confront the disparity head-on, cracking down on perpetrators to mitigate their influence on geopolitical conflicts, public opinion, and electoral outcomes. This urgency of this issue demands an assertive defense strategy.

THE CONTEXT MATTERS: SOCIAL MEDIA ECOSYSTEMS ARE TRANSFORMING

In 2024, the current geopolitical disruptions and scheduled elections will converge in an information environment that is ever more decentralized, fractured, and politicized. Moreover, many actors are competing for dominance within it. Also, social problems have increased polarization and a lack of trust in public institutions. Populism is on the rise, independent institutions are weaker, and there is broad and growing disillusionment. And then, on top of all these dynamics, there is the factor of AI. This confluence of crises make democracies in Germany and the EU vulnerable.

With nearly five billion users, the social media landscape is undergoing significant transformation. The rise of numerous newcomers who quickly gain popularity segues with Twitter's sale or TikTok's success. This dynamic, known as the great decentralization, involves users across the political spectrum migrating to smaller alternative platforms or the fediverse, a decentralized network of interconnected social media platforms, such as Rumble, Truth Social, Mastodon, Threads, and Blue Sky. The outcome: a growing number of people increasingly occupying distinct online spaces.

The growing gap is exacerbated by algorithmically created echo chambers, namely algorithmically constructed walled bubbles that isolate individuals within their own ideological bubbles. This deepens divisions even within generations.

Parallel to this transformation, there has been a gradual paradigm shift in how we perceive information. The shift in mindset – moving from understanding information as a tool or weapon, such as disinformation, to recognizing it as a theater of war – carries significant implications. First, it entails understanding that the battle for influence extends into the digital realm, much like conflicts in air, space, and sea. It thus broadens the perspective on modern cognitive warfare. It highlights the interconnectedness of information operations with traditional military domains. Second, it allows analysts to study influence operations with a wider scope, on different levels, and devise hybrid strategies to counteract adversarial information. This bolsters our defense against information-based threats.

The surge in synthetic content using genAI – strategically disseminated to sway public opinion – marks a pivotal shift in the battle for the integrity of information space. Since reliable information is the basis for informed opinions and thus essential for democratic discourse, collective efforts must develop strategies that safeguard the integrity of information space.

GENAI BOOSTS INFLUENCE OPERATIONS

Germany, the EU, and the US pride themselves on the importance of information space integrity. But they face significant threats from domestic and/or foreign actors with diverse interests, including state bureaucracies, non-state actors driven by economic or political incentives, digital mercenaries, political parties, and populists. Most attacks demonstrate a high expertise in cognitive psychology and behavioral science. Threat actors possess a sophisticated understanding of how influence works, and the effects of inundating individuals with diverse information to exploit their biases.

In fact, past influence operations concluded that there are playbooks guiding violent mobilization, both online and offline. These operations adhere to hybrid cross-platform strategies that require a high level of understanding of platform dynamics and their algorithms, as well as how influence works.

Access to genAI supercharges tactics and strategies in influence operations, enhancing the persuasiveness of posts using synthetic content. A few user-provided prompts greatly reduce the cost of creating sophisticated synthetic material. According to a recent European External Action Service on Foreign Information Manipulation Intervention report, it represents an evolution in the arsenal of threat actors because it boosts existing strategies. However, this is not the defining factor in current influence campaigns.

The problem at hand is hidden, covert, professionalized, and sustained influence operations. The focus should be on concerted attempts to manipulate the health of our information environment. This is confirmed by recent successful detection efforts, such as the 40,000 inauthentic accounts on X that disseminated disinformation and propaganda seen by hundreds of millions of people in the first two days after the Hamas attack on Israel. Other examples include the presence of 600,000 inauthentic accounts on Facebook waiting to be deployed, or the recent detection of 50,000 inauthentic accounts on X conducting disinformation campaigns in Germany.

These examples underscore the necessity to move away from framings that define the primary problem as disinformation. Such framings limit the scope of any defense strategy to counter these sophisticated attempts at manipulating information space and rendering populations vulnerable.

One notable potential threat is genAI-bots, which can establish direct one-on-one relationships with target audiences. It is well established that influence flows through social connections and is linked to human emotions. This is precisely why influencers wield so much impact that translates into mobilization potential. Such a bot would be extremely difficult to detect as an influencing factor because it can communicate exclusively through direct message. In the upcoming elections such a bot could leverage relationships to subtly introduce ideas, influence the salience of issues, and manipulate biases. Thus far, analysts have not identified these types of bots engaging in influence operations, but they will be part of influence portfolios in the future.

GERMANY AND THE EU ARE NOT DEFENSELESS

In 2024, the EU parliament, the US presidential elections, and the state elections in Thuringia, Saxony and Saxony-Anhalt could reshape Germany's political landscape. We know that Germany will also be the target of election interference in the form of information manipulation, similar to the Doppelganger campaign that has been ongoing for almost two years. To uphold robust information integrity in Germany and the EU, it is imperative to counter attacks on our hearts and minds – and in alignment with our democratic values. It is commendable that the German Federal Foreign Office publicly communicated the latest detection of 50,000 inauthentic accounts. Transparency is key in fostering trust.

But current efforts are not enough. A serious imbalance exists when it comes to punishing perpetrators. There is still time to put an assertive defense strategy in place that counters influence campaigns and protects German and EU elections. It is a long-term commitment starting months before voting; it requires strong political will as well as collective action.

RECOMMENDATIONS FOR AN ASSERTIVE DEFENSE STRATEGY

Tech companies and social media platforms

• Tech companies are implementing measures to contribute to election integrity. Initiatives such as water marking, content provenance, and enhanced usage policies for campaigning demonstrate promise in tracking and authenticating digital content. However, their effectiveness is limited, as they can be circumvented by simple techniques like screenshots, etc. Robust solutions to identify synthetic content are imperative and demand continuous innovation.

• Social media platforms are referred to as the dissemination vector in respect to synthetic content. To address the rampant abuse of genAI content on platforms, a multifaceted approach is essential. Implementing a mandate for a more robust identity infrastructure is crucial, finding the delicate balance between protecting user privacy and strengthening trust. An increase of sign-up barriers on platforms is

necessary to deter the use of inauthentic accounts in influence campaigns.

• Integration of more digital nudging mechanisms should be required, guiding users toward responsible and ethical engagement while navigating AI-generated content.

• Imposing limitations on the reshare button is another critical step that Frances Haugen suggested, which prevents the rapid dissemination or virality of potentially harmful content.

• It is essential to rebuild more rigorous moderation practices, put pressure on closing loopholes, and restore of trust and safety measures specific to platforms.

Germany, the EU, and NATO:

• Effective legislation plays a crucial role in implementing measures that target possible genAI-boosted influence operations. The successful implementation of the Digital Services Act across Europe is paramount to ensure data access to enhance detection efforts on platforms.

• Germany also needs to adopt more asymmetric tactics that are flexible and innovative to impose greater costs and risks on threat actors involved in influence operations. This includes implementing no-fly lists and sanctions, applying existing criminal laws, restricting access to Western finance, and deplatforming.

• Collaboration between tech companies and platforms needs to be enforced, given the increased fragmentation and decentralization of the social media ecosystem. Establishing repositories to share information would be an essential initial step, enabling more effective and collaborative efforts to combat the misuse of genAI.

• Innovative media literacy programs should include educating the public on recognizing synthetic content and rhetorical signifiers that are used in linguistic polarization attempts to enhance societal resilience.

• In response to geopolitical disruptions, supranational organizations like the UN and the EU, but also NATO and individual states like Germany, can play a more proactive role. They cannot allow information space to be flooded by narratives pushed by threat actors, as happened in the wake of the attack on Israel. Contingency plans should be in place to actively engage on platforms to not allow a vacuum that can be used to shape perceptions. A recent successful example of countering specific narratives to outwit algorithms is the actions taken by "Swifites."

• Today, influencers are regarded as trusted sources, often surpassing the reach and trust of traditional media and official state accounts. During times of geopolitical uncertainty, collaborating with them provides targeted access to specific demographic segments and enhances the dissemination of accurate information.

DGAP

Advancing foreign policy. Since 1955.

Rauchstraße 17/18
10787 Berlin
Tel. +49 30 254231-0
info@dgap.org
www.dgap.org
@dgapev

The German Council on Foreign Relations (DGAP) is committed to fostering impactful foreign and security policy on a German and European level that promotes democracy, peace, and the rule of law. It is nonpartisan and nonprofit. The opinions expressed in this publication are those of the author(s) and do not necessarily reflect the views of the German Council on Foreign Relations (DGAP).

DGAP receives funding from the German Federal Foreign Office based on a resolution of the German Bundestag.

Publisher
Deutsche Gesellschaft für
Auswärtige Politik e.V.

ISSN 749-5542

Editing Paul Hockenos

Layout Luise Rombach