

Content-based detection of misinformation expands its scope across politicians and platforms

Sami Nenno^{*1,2}, Cornelius Puschmann³, Kamil Fulawka^{1,4}, Philipp Lorenz-Spreen^{1,4}

¹Center Synergies of Systems, Technical University Dresden, Dresden 01069, Germany

²Alexander von Humboldt Institut for Internet and Society, Berlin 10117, Germany

³ZeMKI, University of Bremen, Bremen 28359, Germany

⁴Center for Adaptive Rationality, Max Planck Institute for Human Development, Berlin 14195, Germany

Abstract

Research on misinformation crucially depends on the methods to detect it. Prior studies have primarily identified content from untrustworthy news sources via their URLs. Yet, since this approach is limited to certain posts, content types, and platforms, new methods beyond URL matching are required. This study applies retrieval-augmented classification (RAC) to detect misinformation in free text, enabling broader platform coverage and more fine-grained content analysis. We analyze a unique dataset of over half a million posts from all German members of parliament and their official party accounts across four platforms – Facebook, X, Instagram, and TikTok. Using RAC, we match these posts with more than 5000 fact-checking articles and 1500 community notes. Our method reveals an absolute amount of misinformation nearly ten times higher than the URL-matching method would capture, while the relative share among all posts with text remains low. However, patterns differ across parties and policy issues, as misinformation is unevenly distributed among political actors. For some parties and specific issues, misinformation shares exceed ten percent. Our findings expand the scope of misinformation research through a text-based matching method, offering a more nuanced distinction between high- and low-prevalence segments, and highlighting the role of populist and conservative parties in spreading misinformation.

Introduction

Misinformation is a pressing concern in contemporary society, particularly when it originates from influential actors such as politicians or parties (1, 2, 3). Although prior research has examined the prevalence of online misinformation, methodological constraints have limited a full understanding of the scope of the issue (4). Most existing studies rely on domain-level approaches (5), which detect links to unreliable news domains but overlook the vast majority of social media posts without URLs. To address these shortcomings, we apply retrieval-augmented classification (RAC), a method that identifies fact-checked claims in posts and can be used on any social media content containing text. Drawing on a dataset of more than half a million posts by German parliamentarians and official party accounts across four platforms (Facebook, X, Instagram, and TikTok), we offer a more detailed account of the prevalence, dynamics, and diffusion of misinformation among political actors.

Our method identifies nearly ten times as many misinformation posts as the domain-level approach. Contrary to earlier research (3), we find that misinformation is disseminated not only by politicians at the political extremes but also by conservative politicians and official party accounts. We also show that the misinformation distribution is heavily skewed: while posts on topics such as culture, education, or transportation contain almost no misinformation across most parties, posts on issues like agriculture, health, energy and defense contain misinformation at rates close to or exceeding ten percent among conservatives, with particularly high values for politicians and parties at the political extremes. This suggests that the danger of misinformation does not stem from its overall presence on social media but is instead concentrated within specific issue domains and among particular groups of actors.

Several studies have examined the spread of online misinformation in the past, both broadly (6, 7, 8, 9) and with a focus on politicians (3, 10, 11). These studies reported figures ranging from 1% to 6% of people’s news diets (5). The low spread of misinformation has fueled a debate about the severity of the problem. Some have argued that the harm of misinformation as depicted in public discourse is detached from empirical research and instead of treating it as a widespread danger, we should rather focus on extreme fringes (12) or promote trust in reliable news (5). Other scholars have opposed this line of reasoning and argued that the low figures might also stem from the quantification method deployed in these studies (4).

Most studies quantifying the prevalence of online misinformation have relied on measuring links to untrustworthy news domains in social media posts (13, 14). This *domain-level approach* has major limitations. It can produce false positives, since not every article from a low-credibility domain constitutes misinformation, and even more false negatives, as it inherently misses content referring to misinformation within the text itself. Because it only captures posts containing URLs, it excludes platforms like Instagram and TikTok, where link sharing is not possible. Even on Facebook and X, news sharing via URLs has always been limited and has declined further in recent years (15). Facebook, for instance, has introduced algorithmic changes reducing the visibility of posts linking to news sources (16). Consistent with this, user self-reports show declining engagement with news content (17). Research on politicians’ sharing of untrustworthy news on X likewise finds that only a small share of their posts includes such links (3). This underscores the growing blind spot of existing misinformation detection methods: the large share of posts without URLs remains unobserved, though many may still contain misinformation.

Because of these constraints, we move beyond the domain-level approach, substantially expanding the scope of analyzable data – from about 4% of posts containing URLs to roughly 92% of posts through a text-based method (see Fig. 1A). We further draw on advances in artificial intelligence to detect misinformation. Recent progress in information retrieval, especially through vector embeddings and large language models (LLMs), has enabled retrieval-augmented generation (18, 19, 20), which enhances LLM outputs with external database information to improve factual accuracy. We adapt this framework as retrieval-augmented classification (RAC) using Llama 3.3 70B (21) to identify fact-checked claims in text (22). This allows us to detect misleading claims in unstructured text by assessing whether a known false or misleading claim is semantically equivalent to the text of a post. If such equivalence exists, we classify the post as false or misleading (see Fig. 1C for an illustration and the methods section for details). As a conservative ground truth for misinformation, we rely on two sources: professional fact-checks (e.g., *Correctiv*, *dpa*, *Faktenfinder*) and tweets flagged as inaccurate by Community Notes (“notes”), a collaborative fact-checking system on Twitter/X (23).

Coverage of Social Media Posts by Method

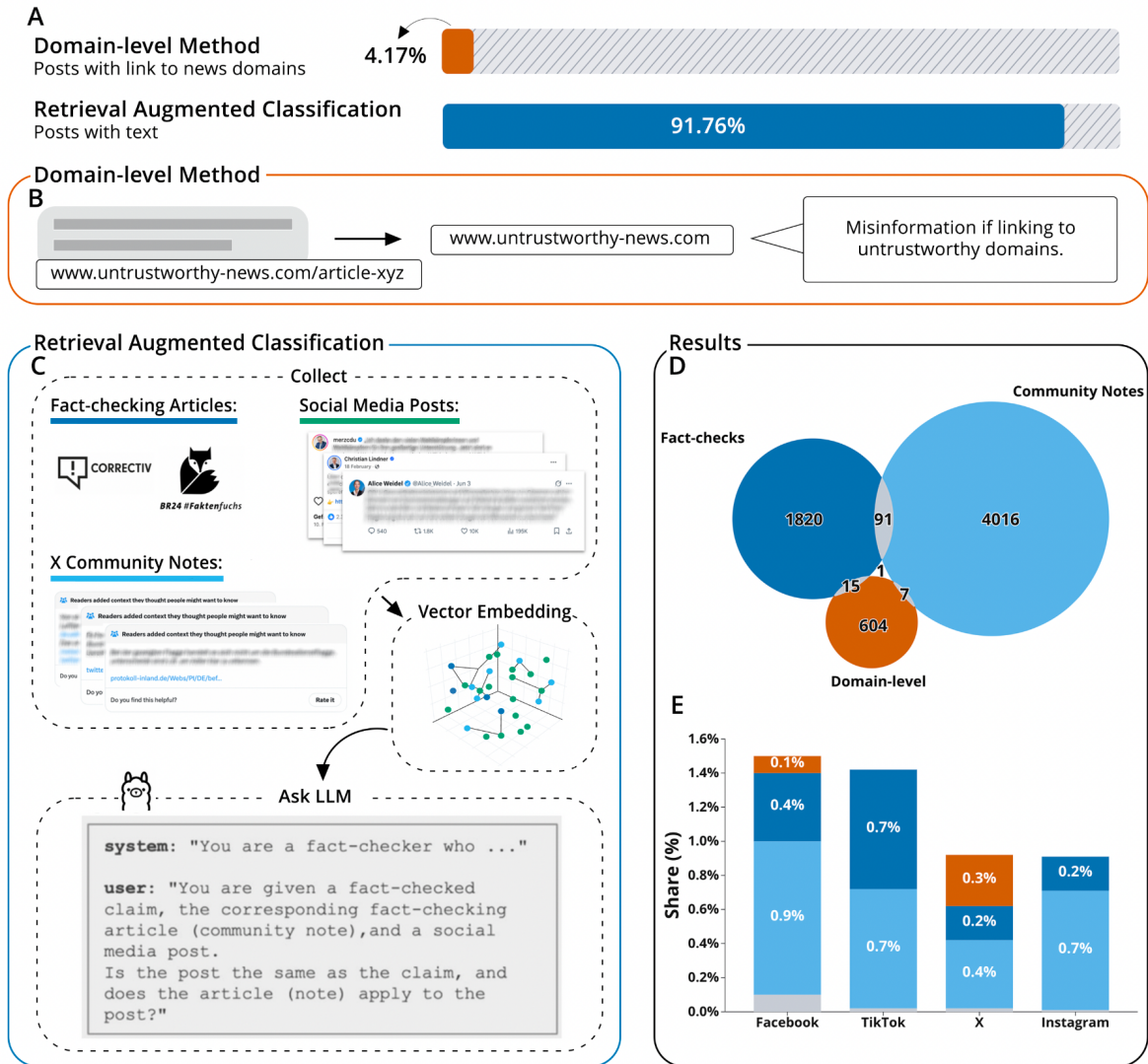


Fig. 1) Scope and workflow of misinformation detection. A) Share of posts analyzable by different methods. B) Workflow of the domain-level approach. C) Workflow of retrieval-augmented classification (RAC). D) Absolute count and overlap of matched misinformation by detection method. E) The share in percentage of posts with misinformation for a platform on all posts on that platform and split by detection method. All numbers refer to posts published since June 2023.

RAC not only applies to a larger volume of posts but also enables more fine-grained analysis of misinformation by identifying it at the instance level rather than through association with an untrustworthy news outlet. It also broadens the range of ground-truth sources usable for misinformation detection – from domain-based labels like NewsGuard (14) to individual fact-checks and other sources such as notes.

We apply RAC to a unique dataset of over half a million posts by all German members of parliament and their official party accounts across four platforms. To address our research objectives, we gradually increase the analytical granularity. We first ask: *What is the overall prevalence of misinformation in posts from politicians on German social media?* Our approach not only provides updated estimates compared to previous methods but also expands the base for these calculations to all text-based posts. We then examine the actors: *How does the prevalence of misinformation vary across parties?* Next, we classify the content of individual posts to ask: *How does misinformation prevalence differ by party and*

topic? Finally, we address the temporal dimension: *In what sequence do individual misinformation stories appear across platforms and parties in relation to different topics?*

Overall, our RAC method identified nearly ten times more posts containing misinformation than the URL-based domain approach – about 6,000 matches compared to roughly 600. When normalized to all posts in our sample, the relative share of misinformation remains low at 1.14%, consistent with previous findings. However, whereas earlier studies measured this share only among posts linking to news domains, our measure applies to all text-based posts. Misinformation is also unevenly distributed across content types and accounts. The overall low prevalence does not imply that misinformation is uniformly scarce; rather, it is concentrated in specific contexts. For some actor–topic combinations, such as posts on agriculture by the CDU, the probability of misinformation approaches 10%, while for many others it is nearly zero. Notably, the far-right AfD alone accounts for about one-third of all misinformation identified.

We advance misinformation research in several ways. Methodologically, we introduce a content-level detection approach that is especially relevant given the decline in news sharing. Substantively, we show that misinformation is not limited to far-right parties and that detailed content analysis is essential to capture its scope across political issues. Finally, we provide new data on the misinformation spread across four platforms and document the dynamics between them.

Results

We compiled a comprehensive dataset of social media posts by members of the German parliament and federal government published between January 2017 and February 2025, covering three national election periods (2017, 2021, 2025) across Facebook, X, Instagram, and TikTok ($n = 2,626,523$). In addition, we collected posts from official party accounts, including subnational branches, over the same period and platforms ($n = 745,642$). After excluding non-political content (e.g., holiday greetings, event announcements), posts predating the introduction of community notes (June 2023), and very short posts (under 100 characters), the final dataset comprised 521,590 posts for analysis.

We aggregated fact-checking articles ($n = 5144$) from the five German-language organizations with the highest monthly output: *Correctiv*, *dpa*, *AFP*, *Faktenfinder*, and *Faktenfuchs*. To provide a complementary perspective to professional fact-checkers, we collected all community notes, formerly known as Birdwatch (23, 24, 25), from X written in German ($n = 1581$). Finally, we also implemented the domain-level strategy using NewsGuard, a service that rates the trustworthiness of online news domains (13, 14).

Figures 1D and 1E show the number of detected posts by method, with examples in Table 1. From June 2023 onward, RAC identified 5,950 matches (1,927 with fact-checks and 4,115 with notes). In contrast, the domain-level method found only 627 posts containing untrustworthy sources. Overlap between methods was minimal (Fig. 1F): only 91 posts matched both fact-checks and notes, while at the domain level, just 15 fact-check and 7 note matches overlapped. Each method therefore captured largely distinct sets of misinformation instances. This demonstrates that the three approaches primarily identified different cases, revealing strong method dependence and underscoring the need for more nuanced and standardized detection to ensure comparability of misinformation prevalence. In the following analyses, we combined fact-check and note matches but excluded domain-level matches.

Match	Party	Post	Rating
-------	-------	------	--------

Fact-check	BSW	“The good news is, peace won’t break out in this region on the very first day.” – ZDF correspondent Theveßen reports on Maybrit Illner, happily noting the upcoming change in Washington, that the war in Ukraine will continue.	The statement by ZDF reporter Elmar Theveßen is presented in a shortened form. He expressed a positive view that the USA did not immediately end its support for Ukraine after Trump’s election.
Fact-check	AfD	The EU’s forced renovation is equivalent to expropriation.	The statements refer to a legislative proposal by the EU Commission. To achieve climate goals, greenhouse gas emissions in the building sector are to be reduced, among other things, through energy-efficient renovations. That expropriation of houses is planned or a usage ban is threatened is not stated in the proposal.
Note	CSU	The agricultural, forestry, and food industries are in crisis. While the global economy and the #economy in the EU are growing, Germany is lagging behind other countries. But instead of supporting our #agriculture, the federal government wants to weaken it.	All members of the audit committee, including members of the CDU/CSU, unanimously approved the abolition of the subsidies in the agricultural sector (e.g. for agricultural diesel).
Note	CDU	“The #citizensbenefit makes up about 10% of the federal budget, that’s transfer payments into unemployment.	In the draft for the 2024 federal budget, 24.3 billion euros are allocated for the citizens’ benefit. That is 5.5% of the total budget of 445.7 billion euros.

Table 1. Examples of matches with fact-checks and notes. The texts of the posts are abbreviated. The ratings are our own paraphrases of the verdicts by fact-checks and notes. For more examples, full texts, and more information on the individual instances, see supplementary material H.

The overall prevalence of misinformation in the postings of politicians

We first examined the overall prevalence of misinformation in politicians’ social media posts. In our sample (June 2023 onward) 5950 posts contained content-level misinformation, which amounted to 1.14% of all posts by German political actors across four platforms – a figure that would drop to only 0.11% using the domain-level approach. Although prevalence varied by platform, it remained below 2% in all cases (Fig. 1E). Facebook showed the highest rate (1.52%), followed by TikTok (1.43%), Instagram (0.97%), and X (0.73%). Of all misinformation posts, 48.18% appeared on Facebook, 30.87%

on Instagram, 14.66% on X, and 6.29% on TikTok. The detection method strongly influenced these results: the domain-level approach could only detect Facebook and X, with most untrustworthy sources found on X (Fig. 1E). In contrast, the RAC method allowed detection on platforms like Instagram and TikTok, where link sharing is not possible.

The prevalence of misinformation for parties

Our second research question addressed the prevalence of misinformation across different political parties. We applied three measures to quantify this prevalence. First, we calculated the share of misinformation among all posts within one party, indicating the extent to which misinformation was part of their social media communication (Fig. 2, top). Second, we measured each party's share on all misinformation posts, showing which parties contributed most (Fig. 2, bottom). Third, we estimated the probability that a post by a specific party contained misinformation, using a multi-level logistic regression model (details of the model in the methods section (Fig. 3).

The relative share of misinformation within each party was highest for the left-conservative populists (BSW, 4.1%), followed by far-right populists (AfD, 2.55%), the Bavarian conservatives (CSU, 2.06%) and the conservatives (CDU, 1.63%). For the other parties, the share was below 1%. For the left-wing Linke it was at 0.53%, for the liberal FDP at 0.51%, for the social democratic SPD at 0.36%, and for the green party Grüne at 0.3%. Taken together, the share of misinformation in German parties' social media communication was highest for populist parties, followed by conservatives.

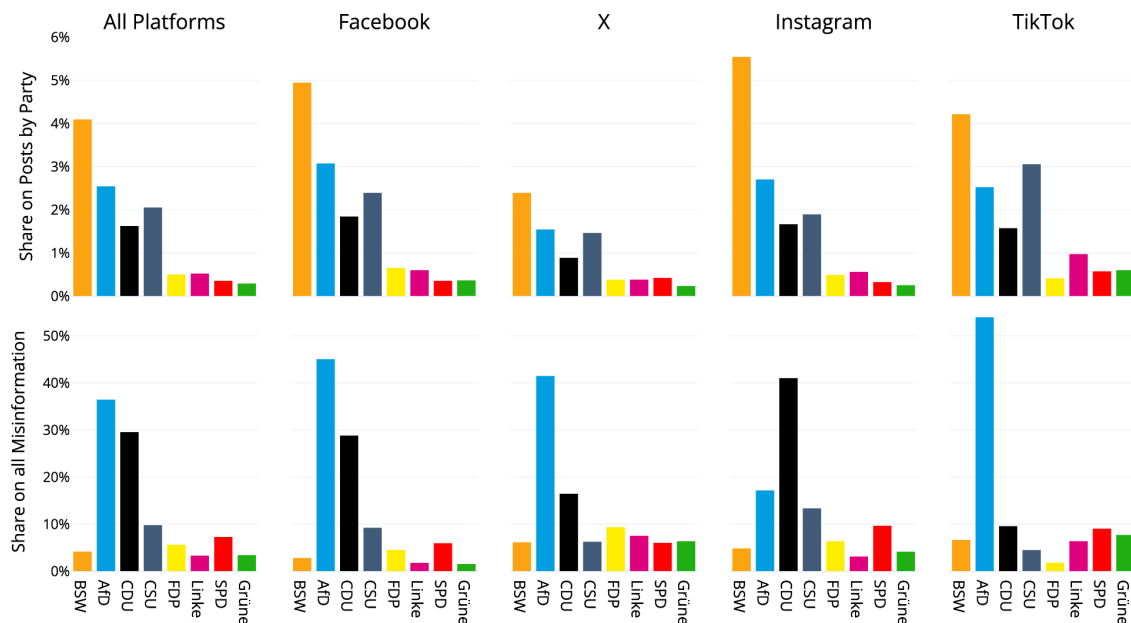


Fig. 2) Top: Percentage of posts that contain misinformation on all posts of the party. Reading example top-left panel: Of all posts by BSW accounts, 4.1% contained misinformation. Bottom: Percentage of posts that contain misinformation on all posts with misinformation. Reading example bottom left panel: Of all posts with misinformation, 4.2% were posted by BSW accounts.

Because parties differed in overall posting activity, the share of misinformation within their communication did not necessarily reflect their contribution to total misinformation (Fig. 2, bottom). This was especially true for BSW: despite the highest misinformation rate, its 17 accounts produced only 1.17% of all posts and just 4.22% of total misinformation. Most misinformation originated from the AfD (36.5%), followed by the CDU (29.6%). The CSU accounted for 9.82% despite only 5.43% of

all posts, while all other parties contributed less than 10%. Notably, the SPD generated the largest share of posts overall (23.07%) but only 7.34% of misinformation.

Across individual platforms, three key patterns emerged. First, on nearly all platforms the AfD contributed the most misinformation, followed by the CDU. Second, this pattern was particularly pronounced on TikTok, where 54.01% of misinformation originated from the AfD. The party's overall activity on TikTok was also higher than elsewhere: while its share of posts reached at most about 20% on Facebook, X, and Instagram, it accounted for nearly one-third (30.57%) on TikTok. Third, on Instagram it was not the AfD but the CDU that contributed the most to the overall level of misinformation (41.05%). This corresponded with a comparatively low posting activity by the AfD on Instagram (6.2%).

To further test the relative contributions of each party to the spread of misinformation, we estimated a Bayesian hierarchical regression model, with a binary misinformation outcome predicted by political party, topic, and the interaction of those two, while controlling for social media platforms. We included a random intercept for each politician to account for individual-level variation. Then, we used the predictive posterior probabilities, which naturally include differences in sample sizes between parties and topics, to test credibility of differences between selected pairwise comparisons (see method section for further details).

Figure 3 shows the posterior probabilities for each party (top left) and each politician that a post contained misinformation, consistent with the descriptive results. Parties formed three distinct groups. The populist AfD and BSW had by far the highest probabilities, credibly exceeding those of the conservative CDU and CSU. In turn, these conservative parties showed credibly higher probabilities than all remaining parties, as indicated by non-overlapping credible intervals of within-party misinformation probabilities (Fig. 3, top left). In relative terms, AfD's probability of posting misinformation was twice that of the CDU (95% CI [1.77, 2.63]); CSU's was twice that of the FDP (95% CI [1.55, 3.06]). These contrasts were even stronger between the first and last groups: AfD's probability was six times higher than the FDP's (95% CI [4.31, 7.09]) and sixteen times higher than the Grüne's (95% CI [10.16, 23.64]).

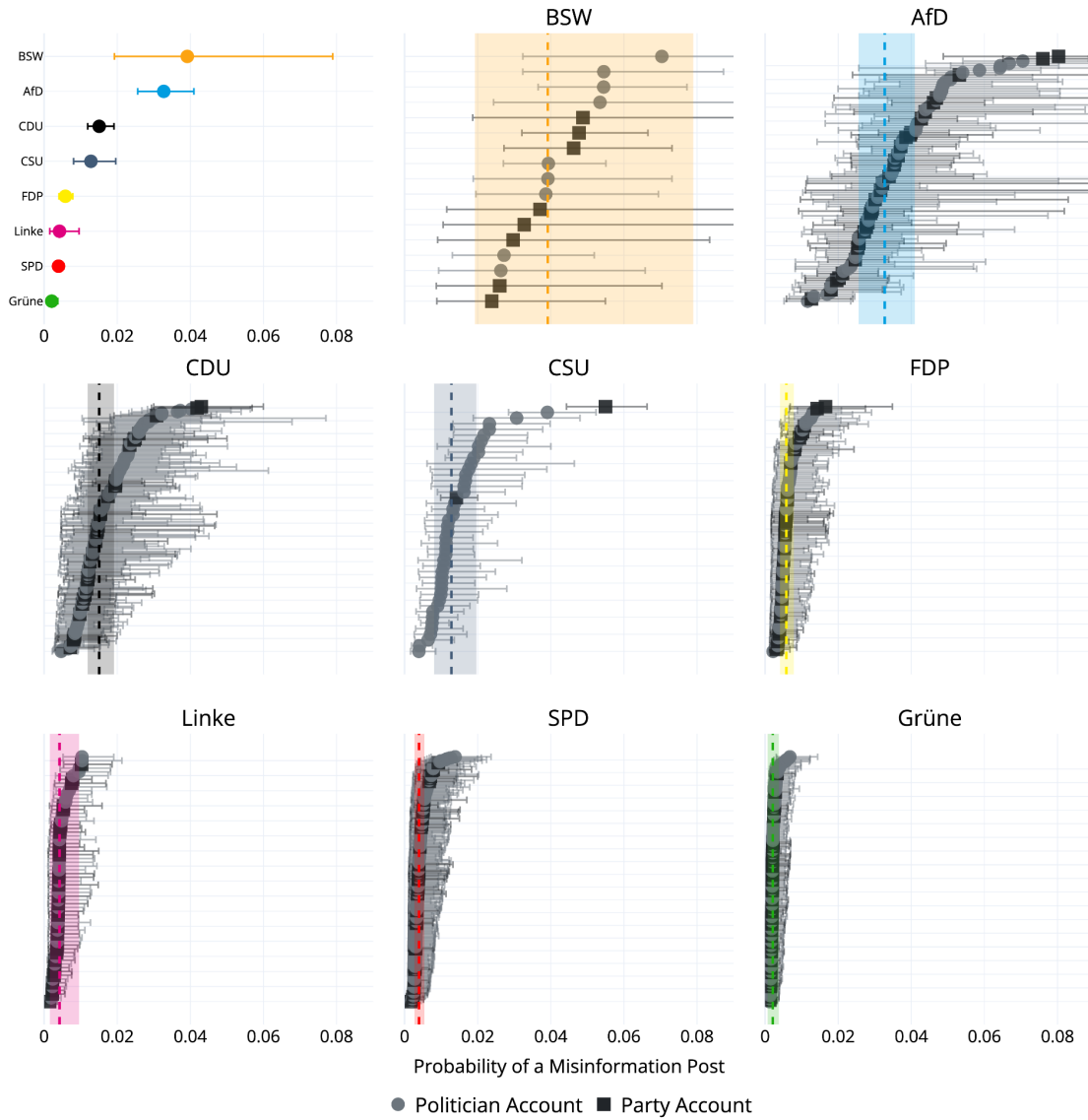


Fig. 3) Predicted probabilities of posts containing misinformation. The top left plot displays the party level posterior median probabilities with 95% credible intervals. These are also displayed as dashed lines and shaded areas in the other panels. The gray and black dots/squares display the median probability of each individual account of a party and their 95% credible interval.

Regarding the probability of misinformation at the account level, we observed three key patterns (Fig. 3). First, across all parties, individual account probabilities followed a fat-tailed distribution, with a few accounts showing much higher rates. Most accounts, however, stayed within their party's overall probability, and at most 4.55% credibly deviated from it. Second, nearly all accounts from the first group (BSW and AfD) had higher probabilities than any from the third group (FDP, Linke, SPD, Grüne); even the lowest-populist accounts exceeded most from the third group. Third, for AfD, CDU, CSU, and FDP, the accounts most likely to post misinformation were not individual politicians but official party accounts. Among the ten accounts with the highest probabilities, three each belonged to AfD and CDU national or state-level organizations; for the CSU it was the party's official account, and for the FDP, two state-level accounts ranked highest.

The prevalence of misinformation by party-topic combinations

Based on the assumption that misinformation is not evenly spread among all types of posts, we classified the posts into 18 policy issues or topics, following the taxonomy of the Comparative Agendas project (26) and using the same Llama model as before (see method section for details). Posts about government operations like budgeting, reforms, or elections were most frequent (30.86%), followed by macroeconomics (8.66%), and international affairs (8.29%).

For misinformation, the topic distribution differed strongly. Of all posts that contained misinformation, 17.19% were about macroeconomics, 12.27% about law and crime, 10.45% about agriculture, and 10.42% about government operations. The remaining topics contributed less than 10% respectively to the overall number of misinformation. When looking at the rate of misinformation per topic, i.e., the percentage of misinformation on a topic on all posts of that topic, the distribution looked different again. Of all posts about agriculture, 6% contained misinformation. This was followed by posts on health (2.98%) and law and crime (2.97%).

Next, we focus on the misinformation rate for different actor-topic combinations, i.e., the percentage of misinformation of a party on a topic on all posts of the same party and topic. The descriptive numbers show a strong variation (see supplementary material D). For example, of all posts by the CDU on agriculture 12.3% were misinformation (BSW 12.1% and CSU 10.5%). And of all posts by the CSU on health the share was 8.9% (CDU 5.9% and AfD 5.5%).

In the following, we report the probability of a post by a party on a policy issue to be misinformation (Fig 4). These figures are based on the same statistical model as before. Note that we combined all topics that had a misinformation rate below 1% and labeled it “other”. These topics were "international affairs", "labor", "environment", "housing", "government operations", "civil rights", "transportation", "education", and "culture". They made up 59.13% of the data and 24.81% of the misinformation.



Fig. 4) Predicted probability of posts containing misinformation by party and policy issue. Each panel displays the median probability and 95% confidence intervals for a party to post misinformation on a specific topic. “Other” includes all topics for which the misinformation rate was below 1%.

The model results reflected the wide variation in prevalence for different topic-actor combinations. For posts on agriculture the probability of misinformation for the CDU was at 0.1 (95%CI: [0.08, 0.11]) and for health the probability for CSU was 0.07 (95%CI: [0.05, 0.09]). On the other hand, the probability was very low for “other” for all of the parties. Even for BSW and AfD, the parties with the highest probability for this category, these probabilities were relatively low. While this was not surprising because the category “other” comprised only topics with low misinformation rates, it underlined that there were many topics for which the probability of misinformation was very low for almost all parties. In the previous section, we showed that three party groups differed credibly in their likelihood of spreading misinformation: AfD and BSW had the highest probability, followed by CDU and CSU, while the remaining parties showed much lower rates. But did this pattern persist across topics? Our results indicate it did not. For many topics, there was no credible difference between BSW and AfD on one side and CDU and CSU on the other (see Supplementary Material D), and in some cases the order even reversed. This was most pronounced in agriculture and health, where CDU and CSU showed the

highest probabilities of misinformation. For most other topics, AfD and BSW ranked highest, though not always credibly above CDU and CSU. Differences were smallest for macroeconomics and law and crime. For energy, AfD differed credibly from the CDU but not from the CSU. In some cases, the divide was not between populists plus conservatives and the rest, but between populists and all other parties combined. For social welfare, defense, and “other,” credible differences appeared only between AfD/BSW and the remaining parties. Notably, for defense, Die Linke was an exception, showing no credible difference from AfD and BSW. In sum, there were meaningful party-specific differences across issue areas. While populist parties still showed the highest misinformation probabilities for most topics, conservative parties exhibited elevated probabilities in certain domains.

The diversity, overlap and order of misinformation across platforms and parties

The previous sections examined the prevalence of misinformation but did not consider whether identical misinformation was reposted or whether each post contained a different story (see Table 1 for four examples). We refer to these stories as unique units of misinformation (UUM) and define a UUM as a claim evaluated by a unique fact-check or note ($n = 688$). Multiple posts of the same UUM are called instances of that UUM. To distinguish this from prevalence, we removed repeated instances by the same party or on the same platform. To control for cross-posting, we also excluded repeated instances by the same actor across platforms. Using RAC, we tracked UUM across platforms and parties, enabling analyses of misinformation diversity (number of UUM per platform or party), overlap (shared UUM between platforms or parties), and order (temporal sequence of UUM across platforms or parties).

More than two thirds of all UUM were posted on Facebook. In other words, only a third of all misinformation stories in our dataset did not appear on this platform. For the other platforms it was significantly less but for X it was still more than half. For parties, the AfD shows the most prominent pattern. For about two thirds of all UUM the AfD posted instances of it. Together with the previous results this emphasizes the AfD’s special role of not only sharing most misinformation posts but also the most misinformation stories among all parties.

We next examine the overlap of UUM. Overall, 46.8% of all UUM appeared on at least two platforms, and 14.4% on all four. Instagram showed the highest overlap with other platforms (88.36%): of 292 UUM, 258 also appeared elsewhere. The corresponding shares were 84.97% for TikTok, 71.46% for X, and 69.83% for Facebook. Across parties, 37.8% of all UUM appeared in posts from at least two parties. The AfD showed the least overlap (40.79%), compared with SPD (69.91%), Grüne (73.63%), BSW (75%), Linke (76.62%), CDU (82.63%), CSU (84.54%), and FDP (87.5%). This separation was not driven by a single topic, though for some issues the AfD dominated. Of 101 UUM on government operations, 71 were posted by the AfD, and 61 of these (85.92%) appeared only in AfD posts. Similarly, AfD-only shares were 92% for civil rights, 86.96% for environment, 73.33% for immigration, 68.75% for health, and 63.51% for law and crime.

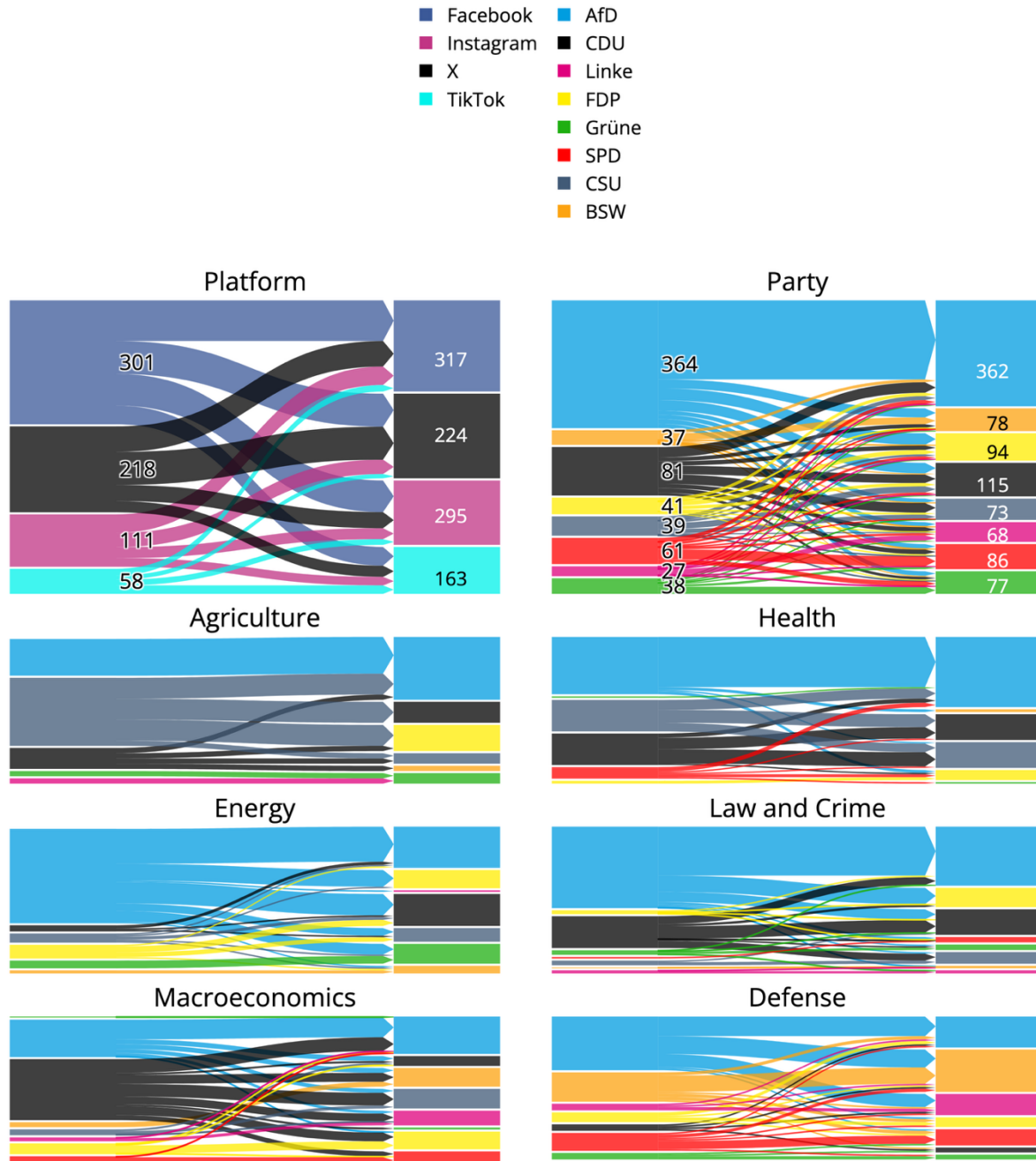


Fig. 5) Temporal order of appearance of unique units of misinformation (UUM, $n = 688$). Top left: flow between platforms. Top right: flow between parties. Bottom panels display flow between parties for different topics. The rectangles on the left side of each panel display the number of UUM that appeared first on the platform or party. The incoming flow displays the number of UUM that appeared first on another platform or appeared only on that platform/party for flows between identical links. Reading example: 301 UUM appeared first on Facebook and 317 UUM that appeared on Facebook appeared either first on another platform or only on Facebook.

Most UUM appeared first on Facebook, followed by X (Fig 5 top left panel). On Instagram more UUM appeared first on Facebook (38.36%) than on Instagram itself (33.55%). Only for X the most UUM on this platform appeared first there (53.79%). Still 31.57% of UUM that appeared on X appeared on Facebook first. For parties, most UUM was posted first by AfD (52.91%), followed by CDU (12.21%). The other parties, respectively, each posted less than 10% first (Fig 5 top right panel). In line with the

low overlap of UUM between AfD and other parties, most of them UUM was also posted first by them (79.84%). For the other parties most of their UUM was first posted by another party. Only for the CDU about half of their UUM was also posted first by them. For all other parties it was less.

With regard to individual topics different patterns emerged. For agriculture most UUM by CSU was also posted first by them and they also posted it before other parties. 80.02% of UUM by CDU on agriculture was first posted by CSU, for FDP it was 60.01% and for AfD 33.33%. A similar pattern could be observed for CSU on health, too. For macroeconomics it was CDU that often posted the UUM first and for energy, law and crime, and defense it was the AfD.

Discussion

The prevalence and dynamics of misinformation on social media platforms are of significant political relevance. This is particularly true for the activities of political actors, whose authority gives them considerable influence in public discourse. While domain-level approaches that match links to untrustworthy news outlets to identify misinformation have limited scope and are increasingly unreliable due to the decline of news sharing, this study advanced beyond link-based methods and examined content-level misinformation. To this end, we applied RAC, an approach that adapts recently developed methods from information retrieval to scale fact-checking, and analyzed more than half a million posts by German politicians and party accounts. Our findings provide new insights into the study of misinformation and underscore the need for a more diverse set of detection methods. In the following, we discuss our results in the context of ongoing debates about the dangers of misinformation, the actors who spread it, and the role of different platforms.

Previous research has found the prevalence of online misinformation to be low, a finding often used to argue that the dangers of misinformation are overstated (5). According to this view, efforts might be better directed toward promoting accurate information. Other scholars – while not rejecting the promotion of accurate information – have suggested that low prevalence estimates may result from the quantification method employed (4). Consistent with this latter argument, we found that the domain-level method excludes the vast majority of social media posts at the outset, whereas our content-level approach identifies roughly ten times more instances of misinformation in absolute terms. Yet, in support of the former argument, we also found that the proportion of misinformation among all text-based posts remains as low as about one percent. This raises the question: which side of the debate do our findings ultimately support?

We argue that the question of prevalence must be addressed differently depending on the broader research interest. When considering all social media posts, the share of misinformation indeed appears low. However, our results suggest not to forget the denominator over the numerator. For certain actors on individual issues, the share of misinformation exceeds ten percent. In other words, the overall prevalence is low not because misinformation is uniformly scarce, but because in some contexts it is nearly absent. We therefore propose distinguishing between high-prevalence and low-prevalence subjects with respect to misinformation. In a small, non-representative sample of misinformation posts, we observed that agriculture and health – issues with substantial shares of misinformation for some parties – often concerned the subsidization of agricultural diesel and the decriminalization of cannabis. These were highly contested topics that dominated public discourse in Germany for some time, which may explain the high prevalence of misinformation.

In this context, it is important to distinguish between prevalence and exposure (and by extension: effect). This study examined the amount of misinformation produced by political actors on social media, which is not equivalent to the amount of misinformation users encounter. Assessing exposure requires different methods. While post view counts offer a useful approximation, RAC is subject to selection bias regarding this metric, as discussed in the limitations section. Ideally, exposure would be measured

by directly tracking user behavior, which was beyond the scope of this study. Nevertheless, prevalence and exposure are not entirely unrelated. At the same time, prevalence provides valuable insights into the supply side of misinformation, specifically the online communication of German political actors.

A previous study found that neither right-wing orientation nor populism alone sufficiently explains higher rates of misinformation spread by politicians; rather, it is their combination that does (3). In contrast, the parties most prone to spreading misinformation in our analysis do not fit neatly into this framework. While the BSW can be described as populist, its economic agenda contains strong left-wing elements. Although previous research suggests that conservative parties such as the CDU and CSU often respond to populist challengers by adopting their strategies (27, 28), they are not full-fledged populist parties like the AfD. Nevertheless, our findings indicate that these parties adopt populist rhetoric, as discussed in the next section. Beyond populism, additional common factors emerge among the parties most likely to post misinformation. A shared characteristic appears to be a culturally conservative orientation. The economically left-wing but culturally conservative BSW showed a high probability of spreading misinformation, whereas the economically right-wing yet culturally progressive FDP showed a low probability. This suggests that a one-dimensional right-left scale obscures important distinctions. Another relevant factor is whether a party is in opposition during the observation period. Although this explanation does not fully align with our results – since Linke was also in opposition but exhibited a relatively low misinformation rate – it remains plausible that holding office versus being in opposition influences a party's inclination toward misinformation.

Our findings align with a broader pattern in which conservatives adopt communication strategies from far-right populist challengers (27, 28), including an increased use of misinformation. This raises the question of how institutionalized misinformation is within conservative party communication. In other words, is it confined to a few radical voices, or does it extend across the party more broadly? We found that the overall misinformation probability for the CDU and CSU is not driven merely by outliers. Most accounts fall within the parties' overall probability. Another indicator of institutionalization is the type of account disseminating misinformation. Surveys among members of parliament indicate that posts are often authored or managed not only by politicians themselves but also by their staff (29). Since there is no central party office reviewing posts before publication, both politicians and their staff retain considerable autonomy over content. By contrast, official party accounts are likely subject to greater scrutiny. This makes it especially notable that several official CDU and CSU accounts rank among those with the highest probability of spreading misinformation. This suggests that misinformation is increasingly embedded in the communication strategy of these parties.

Previous research has shown that while German populists and conservatives do not differ greatly in the amount of misinformation they spread, they differ in its content (30). Conservatives tend to use misinformation to attack political opponents, whereas populists target democratic institutions. Although our study did not examine this dimension directly, we found that individual units of misinformation overlap far less between the AfD and other parties than among the other parties themselves. This suggests that the AfD's misinformation is more detached, while conservative misinformation aligns more closely with that of other parties.

Another key contribution of our study is the examination of misinformation on Instagram and TikTok. Since these platforms do not support link sharing, the domain-level approach cannot be applied – yet our findings indicate this is a significant limitation, as misinformation is prevalent on both. Somewhat unexpectedly, the share of misinformation on Instagram and TikTok is even higher than on X. Furthermore, these platforms play a distinctive and, in some respects, contrasting role for individual parties. While misinformation on TikTok is more strongly driven by the AfD than on other platforms, Instagram is the only platform where the CDU, rather than the AfD, contributes the most to misinformation.

We also found that the same misinformation is frequently disseminated not only on a single platform but across several, and in some cases on all of them. Because we excluded identical misinformation posted by the same actors on different platforms, this pattern cannot be explained by individual actors' cross-posting content. While our study does not establish causality, the results suggest that misinformation often originates on one platform and is subsequently propagated to others. This, however, warrants further investigation. In summary, the RAC method enabled us to use a broader range of fact-checked claims as sources and apply them to a much larger set of posts. This approach revealed several nuances in how German politicians spread misinformation. In the future, the method can be applied to diverse datasets and to other countries or platforms, provided that verified claims and textual posts are available.

Methods

Social Media Data Collection

We compiled a comprehensive dataset of social media posts by members of the German parliament and federal government published between January 2017 and February 2025, spanning three national election periods (2017, 2021, 2025), across Facebook, X, Instagram, and TikTok ($n = 2,626,523$). See Table 2. In parallel, we collected posts from official party accounts, including sub-national party branches, over the same period and platforms ($n = 745,642$). We compiled the list of relevant accounts using the public speakers database (31). Since this database does not include politicians elected in 2017, we manually added them based on a list of names from Wikipedia (32). Politicians elected for the first time in 2025 were not included.

After exclusion of non-political content, such as holiday greetings and event announcements, posts from before the adoption of community notes (June 2023) and very short text (below 100 characters), to ensure relevance, we filtered out non-political content – such as holiday greetings and event announcements – retaining only posts classified as political. For a long-term analysis, see supplementary material E. For the analysis in the subsequent sections, we focused on posts from June 2023 onward, when notes became widely adopted in Germany. We further excluded all posts with text below 100 characters as these could not reliably be matched with fact-checks and notes. The dataset for analysis consisted of a total of 521,590 posts.

	<i>n</i>	<i>SPD</i>	<i>CDU</i>	<i>Grüne</i>	<i>AfD</i>	<i>FDP</i>	<i>Linke</i>	<i>CSU</i>	<i>BSW</i>
<i>X</i>	1,125,232	17.0	14.0	23.2	15.8	16.0	10.6	2.1	1.3
<i>Facebook</i>	931,530	23.4	21.8	8.2	21.1	13.7	6.2	4.9	0.8
<i>Instagram</i>	605,616	27.5	22.4	15.4	7.8	13.8	6.4	6.0	0.7
<i>TikTok</i>	32,905	20.1	8.8	15.9	35.9	6.0	9.0	2.3	2.0
<i>Overall</i>	2,695,283	21.6	18.5	16.2	16.1	14.6	8.1	4.0	1.0

Table 2. Number of political posts per platform and absolute between January 1st 2017 and February 28th 2025. Shares per party are percentages.

Posts from Facebook and Instagram were collected via Meta's content library (33), which grants access to content from verified personal accounts or pages, personal Facebook accounts with at least 25,000 followers, Facebook pages with at least 15,000 followers, or Instagram accounts with at least 25,000 followers.

Posts on X were partially retrieved via the API. The API-based data, used in Lasser et al. (2022), includes posts up to 11 February 2023. Missing accounts were collected using the Zeeschuimer browser plugin (34). TikTok posts, including associated audio, were collected using yt-dlp (35). Audio streams were transcribed into text using whisper-large-v3 (36).

Retrieving Fact-checks and Community Notes

Germany has a well-established fact-checking ecosystem. Journalists from independent outlets, news agencies, and public broadcasters publish several hundred fact-checks each month, covering a wide range of claims made on social media and in the news (37, 38, 39). We aggregated fact-checking articles ($n = 5144$) from the five German-language organizations with the highest monthly output: *Correctiv*, *dpa*, *AFP*, *Faktenfinder*, and *Faktenfuchs*. These sources represent the diversity of the German fact-checking landscape, including independent organizations (Correctiv), fact-checking units within news agencies (dpa, AFP), and public broadcasters (Faktenfinder, Faktenfuchs) (47). Correctiv, dpa, AFP, and Faktenfuchs are signatories of the international fact-checking network’s (IFCN) Code of Principles, and the first three participate in Meta’s third-party fact-checking program. For the collection we used custom scrapers for each of the organizations.

Notes, formerly known as Birdwatch, were introduced on X (formerly Twitter) in 2022. In Germany, they began to gain broader use starting in June 2023. Although notes are not authored by professional journalists, existing research suggests that the wisdom of the crowd can scale up the identification of misinformation (23, 24, 25). We include notes to provide a complementary perspective to that of professional fact-checkers and to broaden the scope of misinformation sources. Notes are publicly available and regularly updated (40). We filtered for German-language entries using the *fast-langdetect* Python library (41) and collected all notes written in German ($n = 1581$). While German is also spoken in Austria and Switzerland, the majority of notes are likely intended for the German context due to population size.

Community notes Retrieval-Augmented Classification

Content-level detection typically relies on machine learning classifiers trained on examples of misinformation. These classifiers use lexical and other surface-level features that tend to be associated with misinformation (42, 43, 44). However, such features are neither necessary nor sufficient to determine whether a post qualifies as misinformation. To qualify as misinformation, a claim must be demonstrably false or deceptive – something that cannot be reliably detected using text features alone. In contrast, by matching posts with fact-checks or notes, retrieval-augmented classification (RAC) could also assess whether the post is false or deceptive.

To match posts with fact-checks and community notes, we sampled a subset of social media posts containing at least 100 characters – a threshold chosen to approximate a minimum length for meaningful content in the context of misinformation. We generated vector embeddings for each post using *jina-embeddings-v3*, applying the same process to claims from fact-checking articles and community notes. Cosine similarity was computed between all post embeddings and all fact-check or note embeddings. We only considered pairs for which the post was published six months before or after the fact-check or note. Pairs with a similarity score of at least 0.65 were retained. We chose this threshold because it guaranteed some degree of similarity but was liberal enough to not miss out on possible matches. The process resulted in 101,416 pairs of posts and fact-checks and 59,233 for notes. These candidate pairs were then evaluated using Llama-3.3, which received the full text of the post and the corresponding fact-check or note. The model was instructed not only to assess semantic similarity but also to determine whether the fact-check or note was applicable to the post.

A match was defined as a case in which both the post and the claim could be evaluated by the same fact-checking article or community note. This definition was explicitly encoded in the prompt provided to the model. However, we employed a three-level hierarchical matching scale. An implicit match was assigned when the post and the claim followed a similar narrative or shared a general thematic overlap, making a match likely even if the connection was not directly stated. An explicit match was identified when the post and the claim conveyed the same core meaning using different phrasing; although they differed in wording and structure, their semantic content was identical. The criteria for a literal match

were met when the post and the claim were nearly identical in wording, with only minimal variations such as synonyms or slight rewordings. The scale was hierarchical in nature, meaning that each higher level subsumed the criteria of the levels below, but only the highest applicable level was assigned.

To test the reliability of RAC, we recruited participants to annotate pairs of posts and fact-checks or community notes, resulting in a sample of 86 annotated items. For these items, the majority of annotators (three to five per item) agreed on 82.56% of cases. In other words, RAC correctly identified misinformation in the vast majority of instances (see Supplementary Material C for details).

Domain-Level Matching

To identify posts linking to news domains and assess their trustworthiness, we used NewsGuard's domain ratings (version dated 28 February 2025). For each Facebook and X post, we detected the presence of links and resolved shortened URLs using Python's requests library. Posts were classified as linking to untrustworthy domains if at least one URL directed to a domain with a NewsGuard score below 60.

Political and Topic Classification

To classify posts as political, we used Llama-3.3. We prompted it to identify posts as political if they dealt with political institutions and actors, public policy and governance, international and geopolitical affairs. Posts about, for example, entertainment, lifestyle, or sports were classified as not being political. We used Llama-3.3 to classify each post according to the policy issues defined by the Comparative Agendas Project (26). To reduce complexity, we decreased the number of possible topics from 21 to 18 and merged, for example, foreign trade with macroeconomics. We validated the classifications against a human baseline, which resulted in an F1 score of 0.75. For the task to classify any topic in the post and not necessarily the most important one, the model was correct in 86% of the cases. These results made us confident that our classifier was reliable. An additional factor for this conclusion was that the least reliable predictions were for the topics that were less frequently misinformation and merged to "other" for the statistical analysis (see supplementary material C for further details).

Statistical Modeling

We modeled the prevalence of misinformation using a Bayesian multilevel logistic regression. The model predicted a binary misinformation outcome based on political party, topic, and their interaction, while controlling for social media platforms. A random intercept for each political actor was included to account for individual-level variation. The model was fit using the *brms* R package with weakly informative priors to regularize the estimates: a Normal(-4.6, 1.5) prior on the intercept, reflecting the low (~1%) base rate of misinformation, and Student-t(7, 0, 2.5) priors for all other coefficients. Estimation used two MCMC chains of 5,000 iterations (1,000 warmup), thinned by four. Convergence was confirmed with R-hat values of 1.00. The model's discriminative ability was confirmed by applying an optimal classification threshold of 0.012, which yielded a sensitivity of 0.78 and a specificity of 0.77. A full description of the statistical parameters can be found in the data repository.

For inference and visualization, we generated predictions from the full model's posterior distribution, accounting for the uncertainty in our estimates. Specifically, we calculated the expected probability of misinformation at several levels of aggregation: for each political party (marginalizing over topics), for each specific party-topic interaction, and for each individual political actor. In all figures and inferential analyses, these posterior distributions are summarized by their posterior median, which serves as the point estimate, and their 95% credible intervals (CIs), defined by the 2.5th and 97.5th posterior percentiles. To infer credible differences between groups (e.g., between two political parties), we computed the sample-by-sample posterior distribution of the difference in their probabilities. A credible difference was inferred if the 95% credible interval of this difference distribution excluded zero. Note

that in Bayesian estimation the imbalances in the number of observations across subgroups are naturally reflected in the width of the corresponding posterior distributions.

Limitations

The first limitation concerns the completeness of the dataset on social media posts by politicians and parties. For Facebook and Instagram, data were collected using Meta's content library. We are confident that all posts meeting the platform's criteria were included. For X (formerly Twitter), a large portion of the data was obtained via the API; however, after its closure in 2023, additional data had to be collected. The same applies to TikTok. Although we employed measures to estimate data completeness, full coverage cannot be guaranteed.

Additional limitations arise from both the methodology and the nature of the underlying datasets. Our approach relied on matching social media posts to external sources, which means that the results may inherit the biases of those sources. In particular, it remains unclear whether the observed topic distribution, temporal trends (supplementary material E), and engagement metrics reflect the actual dissemination of misinformation by political actors or merely the distribution and timing of fact-checking articles and community notes. For example, the high incidence of misinformation related to macroeconomics may be due to the salience of the topic for political misinformation – or simply because it attracted a disproportionately high number of fact-checks and notes. To assess this limitation, we compared the topic distribution of the fact-checks (and notes) with the distribution of matched posts (supplementary material F). We found only weak correlations, with the most prominent misinformation topics diverging notably from those most frequently addressed in the fact-checks and notes. This suggests that our findings reflect the actual importance of specific topics in political misinformation. However, the analysis of misinformation over time is more complex. The weekly number of fact-checks was a strong predictor of the weekly number of matched posts, whereas this effect was less pronounced for notes. We therefore advise caution when interpreting temporal trends and recommend considering the matching rate, which is less sensitive to fluctuations in the publication volume of fact-checks and notes. Furthermore, we did not perform any analysis of the engagement metrics (views, reactions, comments, etc.) of the misinformation. Fact-checks and notes focus on the most viral content on social media. It is therefore likely that the posts that are matched have a high level of engagement. Since there were no comparable checks as for topic and temporal trends, we decided not to include engagement in the analysis.

Finally, the credibility of notes may be questioned, as they are authored by users rather than professional journalists. While some research supports the potential of crowd-based verification (23, 24), other studies suggest that the selection of posts for annotation may be influenced by ideological motives (24). For fact-checks, recent findings indicate a decline in the focus on politicians' claims, with increased focus on viral misinformation from anonymous sources (38, 45, 46). Given the often outlandish content, politicians may avoid engaging with it, which could explain the relatively low number of matches with fact-checking articles.

Data availability

The datasets generated and analysed during the current study are available in the OSF repository: <https://doi.org/10.17605/OSF.IO/GBDY6>. Due to copyright and data protection restrictions, we cannot make the full text or URLs of social media posts publicly available. Instead, we provide post-level metadata that enables reproduction of the figures and analyses presented in the study.

Code availability

All Python and R scripts that were used for this study are available at <https://doi.org/10.17605/OSF.IO/GBDY6>. They can be used to reproduce the figures and statistical analysis but not the LLM analysis (see data availability).

References

1. Slothuus, R., & Bisgaard, M. (2021a). How Political Parties Shape Public Opinion in the Real World. *American Journal of Political Science*, 65(4), 896–911. <https://doi.org/10.1111/ajps.12550>
2. Slothuus, R., & Bisgaard, M. (2021b). Party over Pocketbook? How Party Cues Influence Opinion When Citizens Have a Stake in Policy. *American Political Science Review*, 115(3), 1090–1096. <https://doi.org/10.1017/S0003055421000332>
3. Törnberg, P., & Chueri, J. (2025). When Do Parties Lie? Misinformation and Radical-Right Populism Across 26 Countries. *The International Journal of Press/Politics*, 19401612241311886. <https://doi.org/10.1177/19401612241311886>
4. Ecker, U. K. H., Tay, L. Q., Roozenbeek, J., van der Linden, S., Cook, J., Oreskes, N., & Lewandowsky, S. (2025). Why misinformation must not be ignored. *American Psychologist*, 80(6), 867–878. <https://doi.org/10.1037/amp0001448>
5. Acerbi, A., Altay, S., & Mercier, H. (2022). Research note: Fighting misinformation or fighting for information? *Harvard Kennedy School Misinformation Review*. <https://doi.org/10.37016/mr-2020-87>
6. Allen, J., Howland, B., Mobius, M., Rothschild, D., & Watts, D. J. (2020). Evaluating the fake news problem at the scale of the information ecosystem. *Science Advances*, 6(14), eaay3539. <https://doi.org/10.1126/sciadv.aay3539>
7. Altay, S., Nielsen, R. K., & Fletcher, R. (2022). Quantifying the “infodemic”: People turned to trustworthy news outlets during the 2020 coronavirus pandemic. *Journal of Quantitative Description: Digital Media*, 2. <https://doi.org/10.51685/jqd.2022.020>
8. Grinberg, N., Joseph, K., Friedland, L., Swire-Thompson, B., & Lazer, D. (2019). Fake news on Twitter during the 2016 U.S. presidential election. *Science*, 363(6425), 374–378. <https://doi.org/10.1126/science.aau2706>
9. Guess, A., Nagler, J., & Tucker, J. (2019). Less than you think: Prevalence and predictors of fake news dissemination on Facebook. *Science Advances*, 5(1). <https://doi.org/10.1126/sciadv.aau4586>
10. Lasser, J., Aroyehun, S. T., Simchon, A., Carrella, F., Garcia, D., & Lewandowsky, S. (2022). Social media sharing of low-quality news sources by political elites. *PNAS Nexus*, 1(4), pgac186. <https://doi.org/10.1093/pnasnexus/pgac186>
11. Mosleh, M., & Rand, D. G. (2022). Measuring exposure to misinformation from political elites on Twitter. *Nature Communications*, 13(1), 7144. <https://doi.org/10.1038/s41467-022-34769-6>
12. Budak, C., Nyhan, B., Rothschild, D. M., Thorson, E., & Watts, D. J. (2024). Misunderstanding the harms of online misinformation. *Nature*, 630(8015), 45–53. <https://doi.org/10.1038/s41586-024-07417-w>
13. Lin, H., Lasser, J., Lewandowsky, S., Cole, R., Gully, A., Rand, D. G., & Pennycook, G. (2023). High level of correspondence across different news domain quality rating sets. *PNAS Nexus*, 2(9), pgad286. <https://doi.org/10.1093/pnasnexus/pgad286>

14. Lühring, J., Metzler, H., Lazzaroni, R., Shetty, A., & Lasser, J. (2025). Best practices for source-based research on misinformation and news trustworthiness using NewsGuard. *Journal of Quantitative Description: Digital Media*, 5. <https://doi.org/10.51685/jqd.2025.003>
15. Trilling, D., Kulshrestha, J., Vreese, C. de, Halagiera, D., Jakubowski, J., Möller, J., Puschmann, C., Stępińska, A., Stier, S., & Vaccari, C. (2022). Is sharing just a function of viewing? The sharing of political and non-political news on Facebook. *Journal of Quantitative Description: Digital Media*, 2. doi: 10.51685/jqd.2022.016.
16. Talaga, S., Wertz, E., Batorski, D., & Wojcieszak, M. (2025). Changes to the Facebook Algorithm Decreased News Visibility Between 2021-2024 (No. arXiv:2507.19373). arXiv. <https://doi.org/10.48550/arXiv.2507.19373>
17. Altay, S., Fletcher, R., & Nielsen, R. K. (2024). News participation is declining: Evidence from 46 countries between 2015 and 2022. *New Media & Society*, 14614448241247822. <https://doi.org/10.1177/14614448241247822>
18. Choi, E. C., & Ferrara, E. (2024). Automated Claim Matching with Large Language Models: Empowering Fact-Checkers in the Fight Against Misinformation. *Companion Proceedings of the ACM Web Conference 2024*, 1441–1449. <https://doi.org/10.1145/3589335.3651910>
19. Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W., Rocktäschel, T., Riedel, S., & Kiela, D. (2020). Retrieval-augmented generation for knowledge-intensive NLP tasks. *Proceedings of the 34th International Conference on Neural Information Processing Systems*, 9459–9474.
20. Reimers, N., & Gurevych, I. (2019). Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, 3980–3990. <https://doi.org/10.18653/v1/D19-1410>
21. Grattafiori, A., Dubey, A., Jauhri, A., Pandey, A., Kadian, A., Al-Dahle, A., Letman, A., Mathur, A., Schelten, A., Vaughan, A., Yang, A., Fan, A., Goyal, A., Hartshorn, A., Yang, A., Mitra, A., Sravankumar, A., Korenev, A., Hinsvark, A., ... Ma, Z. (2024). The Llama 3 Herd of Models (No. arXiv:2407.21783). arXiv. <https://doi.org/10.48550/arXiv.2407.21783>
22. Yu, G., Liu, L., Jiang, H., Shi, S., & Ao, X. (2023). Retrieval-Augmented Few-shot Text Classification. In H. Bouamor, J. Pino, & K. Bali (Eds.), *Findings of the Association for Computational Linguistics: EMNLP 2023* (pp. 6721–6735). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2023.findings-emnlp.447>
23. Allen, J., Arechar, A. A., Pennycook, G., & Rand, D. G. (2021). Scaling up fact-checking using the wisdom of crowds. *Science Advances*, 7(36), eabf4393. <https://doi.org/10.1126/sciadv.abf4393>
24. Martel, C., Allen, J., Pennycook, G., & Rand, D. G. (2024). Crowds Can Effectively Identify Misinformation at Scale. *Perspectives on Psychological Science*, 19(2), 477–488. <https://doi.org/10.1177/17456916231190388>
25. Wojcik, S., Hilgard, S., Judd, N., Mocanu, D., Ragain, S., Hunzaker, M. B. F., Coleman, K., & Baxter, J. (2022). Birdwatch: Crowd Wisdom and Bridging Algorithms can Inform Understanding and Reduce the Spread of Misinformation (No. arXiv:2210.15723). arXiv. <https://doi.org/10.48550/arXiv.2210.15723>
26. Comparative Agendas. (n.d.). Retrieved May 14, 2025, from <https://www.comparativeagendas.net/pages/master-codebook>
27. Kollberg, M. (2025). Does mainstream populism work? Populist rhetoric and the electoral fortunes of mainstream parties. *Political Science Research and Methods*, 1–13. <https://doi.org/10.1017/psrm.2025.14>

28. Licht, H., Abou-Chadi, T., Barberá, P., & Hua, W. (2025). Measuring and Understanding Parties' Anti-elite Strategies. *The Journal of Politics*, 87(1), 275–290. <https://doi.org/10.1086/730711>
29. Bauer, P. C., Ecker, A., Imre, M., Landesvatter, C., & Malich, S. (2023). Who tweets, and how freely? Evidence from an elite survey among German politicians. *Research & Politics*, 10(1), 20531680221144237. <https://doi.org/10.1177/20531680221144237>
30. Nenno, S. (2025). Separate worlds of misinformation. An explorative study of checked claims in German public broadcast news and talk shows. *Information, Communication & Society*, 0(0), 1–17. <https://doi.org/10.1080/1369118X.2025.2561030>
31. Schmidt, J.-H., Merten, L., & Münch, F. V. (2024). DBoS-data (Version 2) [Dataset]. Open Science Framework. <https://doi.org/10.17605/OSF.IO/SK6T5>
32. Liste der Mitglieder des Deutschen Bundestages (19. Wahlperiode). (2025). In Wikipedia. [https://de.wikipedia.org/w/index.php?title=Liste_der_Mitglieder_des_Deutschen_Bundestages_\(19._Wahlperiode\)&oldid=254538521](https://de.wikipedia.org/w/index.php?title=Liste_der_Mitglieder_des_Deutschen_Bundestages_(19._Wahlperiode)&oldid=254538521)
33. Meta Platforms, Inc., (February 2025). Meta Content Library version v5.0 <https://doi.org/10.48680/meta.metacontentlibrary.5.0>.
34. Digitalmethodsinitiative/zeeschuimer. (2025). [JavaScript]. Digital Methods Initiative. <https://github.com/digitalmethodsinitiative/zeeschuimer> (Original work published 2021)
35. Yt-dlp/yt-dlp. (2025). [Python]. yt-dlp. <https://github.com/yt-dlp/yt-dlp> (Original work published 2020)
36. Radford, A., Kim, J. W., Xu, T., Brockman, G., McLeavey, C., & Sutskever, I. (2022). Robust Speech Recognition via Large-Scale Weak Supervision (No. arXiv:2212.04356). arXiv. <https://doi.org/10.48550/arXiv.2212.04356>
37. Humprecht, E. (2019). Where 'fake news' flourishes: A comparison across four Western democracies. *Information, Communication & Society*, 22(13), 1973–1988. <https://doi.org/10.1080/1369118X.2018.1474241>
38. Nenno, S. (2025). Fact-checks as Data Source? Content Analysis of Fact-checking Articles in Germany between 2019 and 2023. *Computational Communication Research*, 7(1), Article 1. <https://doi.org/10.5117/CCR2025.1.6.NENN>
39. Rieger, J., Hornig, N., Flossdorf, J., Müller, H., Mündges, S., Jentsch, C., Rahmenführer, J., & Elmer, C. (2023). Debunking Disinformation with GADMO: A Topic Modeling Analysis of a Comprehensive Corpus of German-language Fact-Checks. In S. Carvalho, A. F. Khan, A. O. Anić, B. Spahiu, J. Gracia, J. P. McCrae, D. Gromann, B. Heinisch, & A. Salgado (Eds.), *Proceedings of the 4th Conference on Language, Data and Knowledge* (pp. 520–531). NOVA CLUNL, Portugal. <https://aclanthology.org/2023.ldk-1.56/>
40. Downloading data. (n.d.). Retrieved May 14, 2025, from <https://communitynotes.x.com/guide/en/under-the-hood/download-data>
41. fast-langdetect: Quickly detect text language and segment language (Version 0.3.2). (n.d.). [Computer software].
42. Damstra, A., Boomgaarden, Hajo G., Broda, Elena, Lindgren, Elina, Strömbäck, Jesper, Tsifti, Yariv, & and Vliegenthart, R. (2021). What Does Fake Look Like? A Review of the Literature on Intentional Deception in the News and on Social Media. *Journalism Studies*, 22(14), 1947–1963. <https://doi.org/10.1080/1461670X.2021.1979423>
43. Lebernegg, N., Eberl, Jakob-Moritz, Tolochko, Petro, & and Boomgaarden, H. (2024.). Do You Speak Disinformation? Computational Detection of Deceptive News-Like Content Using Linguistic and Stylistic Features. *Digital Journalism*, 0(0), 1–24. <https://doi.org/10.1080/21670811.2024.2305792>

44. Wirtschafter, V., Pereira, F. B., Bueno, N., Pavão, N., Santos, J. P. O. dos, & Nunes, F. (2024). Detecting Misinformation: Identifying False News Spread by Political Leaders in the Global South. *Journal of Quantitative Description: Digital Media*, 4. <https://doi.org/10.51685/jqd.2024.007>
45. Cazzamatta, R. (2024). The Content Homogenization of Fact-Checking Through Platform Partnerships: A Comparison Between Eight Countries. *Journalism & Mass Communication Quarterly*, 10776990241261725. <https://doi.org/10.1177/10776990241261725>
46. Graves, L., Bélair-Gagnon, V., & Larsen, R. (2023). From Public Reason to Public Health: Professional Implications of the “Debunking Turn” in the Global Fact-Checking Field. *Digital Journalism*, 0(0), 1–20. <https://doi.org/10.1080/21670811.2023.2218454>
47. Graves, L., & Cherubini, F. (2016). *The Rise of Fact-Checking Sites in Europe*. Oxford: Reuters Institute for the Study of Journalism Publications

Acknowledgments

The authors thank Jana Lasser for sharing data with us, Isin Kosemen for the design of the figures, Jeff Speaks for advice on identical meaning, and Michael Günther for advice on vector embeddings. The authors gratefully acknowledge the computing time made available to them on the high-performance computer at the NHR Center of TU Dresden. This center is jointly supported by the Federal Ministry of Education and Research and the state governments participating in the NHR (www.nhr-verein.de/unsere-partner).

Author contributions

Conceptualization: S.N., C.P, P.L.-S. Methodology: S.N. Software, formal analysis, investigation, and visualization: S.N., K.F. All authors wrote the manuscript.

Competing interests

The authors declare no competing interests.

Materials & Correspondence

Correspondence and requests for materials should be addressed to Sami Nenno.