INACH

Bringing the Online In Line with Human Rights

DARE TO BE GREY

beyond black-and-white thinking

Mapping Online Hate in the Dutch 2025 Election Campaign

By Adinde Schoorl & Jordy Nijenhuis

2026

TABLE OF CONTENTS

International Network Against Cyber Hate – INACH

INACH was founded in 2002 to use intervention and other preventive strategies against cyber hate. The member organisations are united in a systematic fight against cyber hate, for example as complaints offices, monitoring offices or online help desks. In their respective countries, they provide important contacts for politicians, internet providers, educational institutions, and users.

Funding for INACH is provided by its members, the European Commission and other donors. The International Network Against Cyber Hate (INACH) unites multiple organizations from the EU, Africa, Israel, Russia, United Kingdom, North Macedonia, Albania and South America. While serving as a network of online complaints offices, INACH today pursues a multi-dimensional approach of educational and preventive strategies.

This publication has been produced with the financial support of the Citizens, Equality, Rights and Values (CERV) Programme of the European Union. The contents of this publication are the sole responsibility of the International Network Against Cyber Hate and can in no way be taken to reflect the views of the European Commission.

Supported by the Citizens, Equality, Rights and Values (CERV) Programme of the European Union

Abstract

This study examines the role, intensity, and development of online hate narratives during the 2025 Dutch parliamentary elections, with a particular focus on how digital ecosystems shape political discourse and public perception. The research shows that xenophobic, anti-immigrant, anti-Muslim, and anti-LGBT+ narratives were consistently present and often strategically amplified throughout the campaign. This also included the increased use of synthetic propaganda and AI-generated images and messages. The analysis demonstrates that political actors themselves play a major role in driving, amplifying, or legitimising these narratives. Posts, speeches, and campaign messaging from certain parties circulated widely across social media platforms, triggering algorithmic amplification and reinforcing echo chambers. As a result, harmful ideas shaped the broader public debate and contributed to the normalisation of concepts such as Great Replacement and remigration, which have shifted from fringe rhetoric into mainstream political communication.

The report concludes with a set of recommendations to strengthen platform accountability, support civil society monitoring efforts, encourage responsible political communication, and enhance user awareness—ultimately contributing to a safer and more democratic online environment in the Netherlands.

Introduction

'It's the internet, stupid.' In a recent article in the Dutch newspaper NRC, Francis Fukuyama argued that the global surge of populism cannot be understood through traditional explanations alone (such as economic grievances, cultural backlash, or migration debates). Instead, he places the digital ecosystem at the centre of this rise, pointing to the decisive role of social media platforms, algorithmic amplification, and online echo chambers. His conclusion is clear: to understand contemporary populist narratives, one must analyse the online environments in which they are produced and circulated.

Our own long-term monitoring of online hate speech confirms this. A significant share of online hate is politically driven: it is generated by political actors, weaponised during public debates, and concentrated around political themes. Building on this insight, the present study examines the 2025 Dutch parliamentary elections, mapping how online hate and disinformation evolved throughout the electoral campaign and identifying the narratives that shaped the digital political landscape.

Our research methodology combines manual analysis with AI-driven monitoring using the EOOH dashboard. Based on keyword detection, the AI scanned social media platforms and scored the toxicity levels of the related content. On the other hand, we have followed content created by politicians on their social media pages.

Previous research findings from both INACH and EOOH indicate that the majority of online hate is disseminated by populist right-wing actors targeting LGBTQIA+ individuals and immigrants. That is why our analysis focuses on these themes partly. On the other hand, we have also used neutral words as 'verkiezingen' (elections) en 'nederlanders' (the Dutch) to see what kind of narratives are online in general. The purpose of this analysis is to show how hate shapes political campaigns, using data-driven insights to illustrate key patterns of toxicity and its amplification, and includes some recommendations on how this should be countered. Because online hate is constantly evolving and takes different shapes, tailored and effective strategies are needed.

Without reliable data on it and clear insights into how this content emerges and spreads, one cannot develop any policies or implement measures to address it.

Context

The Netherlands has a multi-party parliamentary democracy, where coalition governments are the norm since no single party holds a majority in the 150-seat House of Representatives. This complex political system, designed to encourage pluralism and compromise, has also led to frequent cabinet collapses, lengthy coalition talks, and rising voter frustration. In recent years, Dutch politics has become highly fragmented, with over 20 parties in parliament ranging from progressive greens to far-right populists. Key issues dominating the political landscape include migration, housing, climate policy, healthcare, and public trust in institutions.

The elections in October 2025 took place amid ongoing political fragmentation. The last two governments collapsed over issues related to migration policy. These recurring crises have contributed to historically low public trust in politics and institutions, with many citizens feeling disconnected from policymakers (Metro Nieuws 2025). Against this backdrop, it is particularly relevant to monitor how these sensitive issues are discussed and politicized in the run-up to the 2025 elections.

Social media and digital media play a central role in shaping public opinion by framing and amplifying narratives that dominate online discourse. Algorithms prioritise engagement (clicks, shares, and reactions) over accuracy or balance, which means that content provoking strong emotions, especially anger or fear, spreads fastest. This revenue-driven model rewards polarisation, as negative or divisive posts keep users active and advertising profits high. As a result, social media platforms don't just reflect societal debates but actively construct and reinforce online narratives, influencing how people perceive political issues, identity, and truth itself. That is why it is so important to track and analyse online behavior in the context of elections.

The Dutch media landscape is generally diverse, professional, and guided by strong journalistic standards, yet it is not immune to the spread and amplification of hateful narratives. Traditional outlets such as NOS, NRC, and Volkskrant maintain fact-based reporting, but the fragmentation of the media ecosystem through online platforms, partisan outlets, and influencer-driven commentary, has blurred the lines between

journalism and opinion. Certain populist or alternative media platforms (e.g., GeenStijl, Ongehoord Nederland, De Dagelijkse Standaard) have played a role in normalising xenophobic or anti-immigration frames, often presenting them as “taboo-breaking” or “anti-elite” viewpoints.

Mainstream media are sometimes drawn into these debates by the logic of controversy and click-driven visibility, inadvertently reinforcing polarising narratives. Combined with the dynamics of social media, where emotionally charged and divisive content spreads fastest, the Dutch media environment increasingly mirrors and magnifies online hate trends, particularly around migration, Islam, gender, and political distrust.

With over twenty years of experience monitoring online hate, we have observed a clear shift from explicitly illegal hate speech toward broader hateful narratives that encompass conspiracy theories, nationalistic symbols, and other forms of coded hostility. These trends are amplified by technological developments such as generative AI and increasingly exploited by politicians to mobilise voters. Additionally, it is far more difficult to counter them than traditional hate speech, as they often fall outside the scope of legal removal requests to platforms.

Developments in online hate

Online hate constantly evolves, shifts in form and changes along with technological progress. A few new developments linked to political campaigns stand out and are important to review;

Synthetic propaganda and generative AI

Generative Artificial Intelligence (AI) and synthetic propaganda are relatively new developments in the digital information sphere. While propaganda itself is as old as politics, its combination with generative cheap, fast, and widely accessible AI tools, has transformed its scale and impact, making it a powerful instrument for spreading hateful and manipulative narratives, including those circulated by political actors. Synthetic propaganda can be defined as "the deliberate attempt to shape perceptions and influence the understanding of issues in line with the intentions of the sender through generative AI content," encompassing deepfakes, cheapfakes, and other AI-generated media, but with a stronger focus on emotional manipulation and political use (Marcus Bösch, 2025).

In essence, synthetic propaganda refers to content created or altered by AI that mimic authentic human communication with the aim of misleading the public or distorting reality. The traditional press machinery of myth-making has evolved into automated ecosystems of interlinked bots that amplify and imitate human behaviour, blurring the boundary between synthetic and organic communication (Zhou et al., 2020; Vosoughi et al., 2018). Such techniques have already appeared in various geopolitical contexts, from election interference in Brazil and India to COVID-19 misinformation in the United States and the EU (Albanese et al., 2023).

Synthetic propaganda appears in many forms and is often saturated with nationalistic symbols, colours, and emotionally charged imagery. These materials may not even attempt to appear real. Instead, their purpose is to provoke emotion, pride, fear, or outrage, shaping public perception and fuelling polarization regardless of factual accuracy.

In the Netherlands, several clear examples of synthetic propaganda have already emerged. One case involved a Facebook page posing as a fan account of Geert Wilders, titled “Wij doen geen aangifte tegen Geert Wilders” (“We do not report Geert Wilders”). The page shared AI-generated images depicting blond children dressed in Dutch national colours, seemingly threatened by men of colour, accompanied by captions such as “Send back all criminal asylum seekers.” Journalistic investigations later revealed that the page was in fact managed by two members of the PVV party of Geert Wilders, who also distributed AI-generated images targeting Frans Timmermans, leader of the GroenLinks-PvdA party (Volkskrant 2025). Further research showed that half of all AI-generated campaign content circulating during that period originated from the PVV itself (Volkskrant 2025).

Generated AI is not only used for images or videos but also in music. A far-right AI-generated song titled ‘Wij zeggen nee, nee, nee, tegen een AZC’ (“We say no, no, no, against an asylum-seeker centre”), which targets asylum-seekers, has climbed into the Dutch charts, placing number 28 on the Single Top 100 and gaining millions of streams. Despite clear xenophobic messaging, the streaming platform Spotify has declined to remove the track, stating that it does not breach its hate-speech policy._The development raises urgent questions about how synthetic propaganda, AI music in this case, can be used to steer public sentiment and bypass regulatory mechanisms (AD 2025). Even when the music eventually was deleted from Spotify, the music was widely shared across social media platforms.

Remigration and the Great Replacement

The term “remigration” is increasingly used by far-right and populist parties as a euphemism for mass deportation, disguising its racist and exclusionary origins. While

these terms were once limited to fringe online spaces like 4Chan or Gab, both remigration and the Great Replacement have now become mainstream political buzzwords used to legitimize ideas of ethnic cleansing and forced repatriation. The language deliberately masks violent intent. Words like “programme”, “plan”, or “commissioner for remigration” sound bureaucratic and policy-oriented, yet they reduce moral outrage and enable politicians to present deportation as routine governance rather than persecution.

A 2024 Guardian investigation traced how “remigration” spread across Europe: in Germany, tens of thousands protested after it was revealed that AfD members attended a Potsdam meeting discussing deportations; in Austria, the FPÖ proposed appointing an EU “remigration commissioner”; and in eastern Germany, AfD posters read “Summer, Sun, Remigration” next to a “deportation airline.” The term has also crossed continents; Donald Trump referred to “remigration” when vowing to return migrants if re-elected (Guardian 2024).

Researcher Julia Ebner (Institute for Strategic Dialogue) notes that the term’s apparent neutrality conceals its true meaning, allowing extremists to mainstream what amounts to ethnic expulsion. Reflecting this manipulation of language, the German jury of the annual Unwort des Jahres (“Un-word of the Year”) campaign selected Remigration as the 2023 winner, condemning it as a distortion that hides the intent to strip people of rights and forcibly remove them (Bild 2024). Flyers in Germany have even depicted fake plane tickets with the slogan “Only remigration can still save Germany”, a stark reminder of how dangerous narratives can be normalised through sanitised language (AP News 2025).

The “Great Replacement” theory, coined by French writer Renaud Camus, falsely claims that native European populations are being deliberately replaced by non-European, mainly Muslim immigrants. This idea has evolved from a fringe conspiracy into a mainstream political narrative. Research shows that this discourse has been strategically adopted by far-right and populist leaders, such as Viktor Orbán and Robert Fico, who frame migration as an existential demographic threat to Europe. This process of mainstreaming such narratives makes racist and xenophobic ideas appear legitimate,

often wrapped in the language of “cultural preservation” or “population policy.” The Dutch intelligence service (AIVD) confirms that extremist groups interpret immigration as a planned “demographic attack” orchestrated by elites, sometimes described as a “Jewish conspiracy”, with politicians, media, and academics allegedly complicit (AIVD 2024).

Digital media have accelerated this normalisation: online platforms amplify conspiratorial content through emotionally charged visuals, memes, and algorithms that reward outrage. As a result, the Great Replacement functions as a flexible narrative; explicitly racist in extremist circles, coded and “cultural” in populist rhetoric, and implicitly framed as “rational concern” in mainstream discourse. Scholars like Ekman warn that such normalisation threatens democratic institutions and social cohesion by embedding symbolic violence against minorities into everyday political communication (Ekman 2022).

In parallel, right-wing populism across Europe, including the Netherlands, thrives on this narrative. Leaders like Geert Wilders exploit fears of Islamisation and “migration crises” to sustain a sense of perpetual emergency; a political strategy that keeps public anxiety alive and ensures continued voter mobilisation. As Ayhan Kaya notes, populists not only exploit migration-related fears but actively manufacture and sustain them, turning the crisis itself into a source of political power (Kaya 2025).

Insights on different narratives during the elections

This analysis combined a mixture of automated scraping and manual research to ensure both breadth and contextual depth. Using the European Observatory on Online Hate (EOOH) monitoring dashboard, data was analysed from multiple social media platforms through customised search channels. Each channel was configured with specific keywords and search terms related to key themes such as migration, identity, and politics, allowing for targeted retrieval of relevant posts and comments. This setup enabled systematic tracking of online discourse, while manual verification and annotation provided qualitative insights into tone, framing, and narrative patterns that automated tools alone could not fully capture. The AI built for the dashboard analyses and determines every scraped comment and determines its level of toxicity. Here is an overview of the four created channels for this part of the research:

Elections Netherlands 2025 (verkiezingen)

To capture the broader online conversation about the upcoming elections, this channel was set up using search terms related to “elections,” “parliament,” and other political keywords. It aimed to identify how users discuss electoral processes, parties, and key political figures online.

Foreigners / Refugees

Immigration has been one of the most divisive issues in Dutch politics, leading to the fall of the Rutte IV government in 2023 after coalition parties failed to agree on migration policy. It remained the dominant topic in the subsequent elections, which resulted in a major victory for Geert Wilders’ radical right-wing PVV. In 2025, Wilders again dissolved the government, claiming insufficient progress in restricting immigration. Given this political context, this channel focused on online debates and narratives surrounding migration—an issue expected to remain central in the current election cycle.

LGBT+

Radical right-wing parties often direct online hostility toward two main target groups—immigrants/refugees and the LGBT+ community. Accordingly, this channel focused on content related to anti-LGBT+ rhetoric and hate, mapping how such narratives circulate within Dutch online spaces.

Dutch People (Nederlanders)

During the initial baseline monitoring on anti-refugee hate, the term “Nederlanders” (“Dutch people”) appeared frequently in the word cloud. Although it is a general term that can yield a high volume of irrelevant results, a dedicated channel was nonetheless created to examine how national identity is invoked in online discourse, particularly in contrast to immigrant groups.

Elections Netherlands 2025

Date	25 August - 30 October
Amount	239,451
% that was toxic of the total dataset	26,948 (11.3%)
Average toxicity score	0.18 (0-1)
Platforms where the data was mostly found	Twitter (74%), Youtube (9%) and Facebook (5%)
Messages with high toxicity are more frequent on	Twitter
Often involved	Politics
% of the posts have a toxicity score of 0.8 or higher (on a scale of 0-1)	0.3%

% contained violent language	16%
Keywords	Politiek, verkiezingen, Nederland

The dataset contains 239,451 messages, of which 11.3% (26,948) were classified as toxic. While this means that the majority of posts are not overtly hateful, the absolute number of toxic posts is still significant, especially during a concentrated electoral period. The average toxicity score of 0.18 suggests that most content sits in a "grey zone", neither fully benign nor explicitly hateful, consistent with patterns of borderline speech often found in political debates. Only 0.3% of posts reached a toxicity score of 0.8 or higher, indicating extreme hate speech is limited in frequency but still measurable. This small fraction is important because such highly toxic content often circulates widely and shapes narratives disproportionately. About 16% of posts contained violent expressions, showing that while direct hate may be rare, aggressive framing (calls for harm, metaphors of violence) is a substantial feature of political discourse online, potentially normalizing hostility and dehumanization.

Below is the word cloud generated with the data from this channel. Logically, the word 'verkiezingen' (elections) has appeared the most, as 'politiek' (politics), 'democractie' (democracy) and 'politici' (politicians). But it is interesting to see other words appear often in this channel that are not necessarily naturally related to the topic of elections. 'Islamisering' (islamisation) was used 378 times and refers to anti-muslim hatred, or hoofddoek (head scarf) was mentioned 248 times, israel (397 times mentioned) and genocide (254 times mentioned) show that migration-related hate seeps into general political discourse. Interestingly enough also vrouwen (women) (532

times mentioned) appeared here

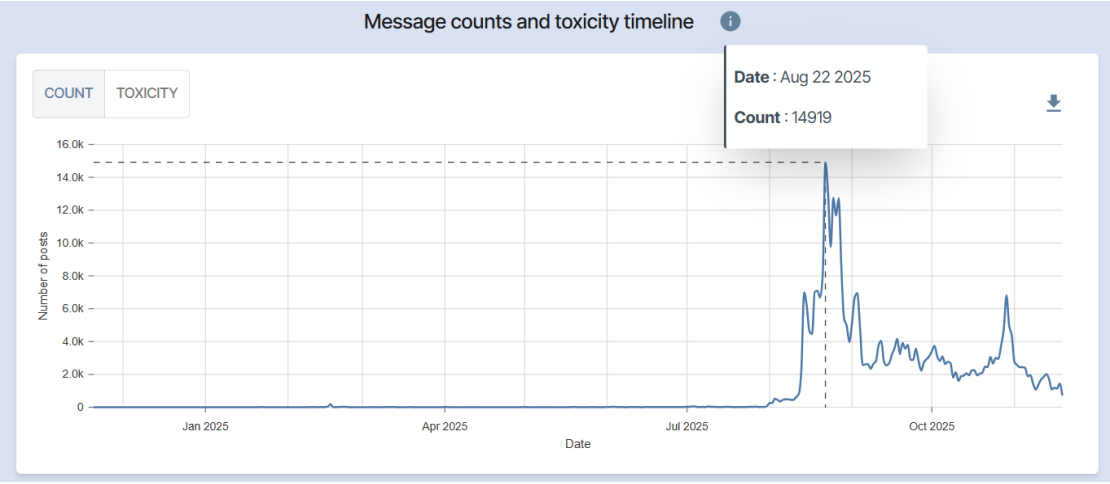
frequently.

By far, most of the toxic content appeared on X, 74%, underlining the platform's role as the epicentre of election-related hostility. The dataset on X contains 188,820 posts, of which 21,211 (11.2%) are toxic. About 13% of the toxic messages contain violent language.

Across platforms, toxicity was consistently tied to political discussions, though its intensity varied by medium. On YouTube, 8.4% of 22,854 posts were toxic, with an average toxicity of 0.16 and frequent use of political terms like politiek (politics), verkiezingen (elections), and Nederland (the Netherlands). Facebook showed the highest toxicity rate (14.8%) and more violent language (23%), often combining political debate with references to Israël, suggesting overlap with global conflict narratives. Instagram had the lowest toxicity (4.9%, average 0.12), but even here most hateful comments (87%) were political. Telegram, while small in volume, displayed the most extreme tone (13.8% toxic; average 0.24), dominated by conspiracy-related keywords such as coronadictatuur (corona dictatorship) and antifa. Overall, political polarization was the main driver of online toxicity across all platforms, with Telegram amplifying the most radical narratives.

While Facebook contributes fewer posts overall, it has the highest share of toxic content (14.8%) and a relatively high average toxicity score (0.22). This suggests that users on

Facebook post fewer but more aggressive comments, often in polarized or conspiratorial political discussions. Only 4.9% toxic posts and a low average toxicity (0.12) indicate that visual platforms like Instagram host less explicit hate speech. However, the fact that nearly 9 out of 10 toxic posts relate to politics (87%) suggests political cynicism and hostility do spill over even in lifestyle-oriented spaces, though in a less aggressive tone. Despite the small dataset on Telegram (80 posts), 13.8% were toxic with the highest average toxicity score (0.24) and 1.2% scoring above 0.8, much higher than any other platform. The keywords (coronadictatuur, antifa, controlled) reveal a conspiratorial ecosystem where election-related discussions overlap with anti-establishment and anti-democratic rhetoric typical of far-right or conspiracy groups. The chart below shows there was already a peak in the amount of messages in this channel around 22 August 2025. The amount of messages was even higher than around the date of the elections itself in October; 14,919 messages. That is interesting considering this channel has keywords associated with the elections and late August did not mark the official start of the campaign period yet. However, in the night of 19 to 20 August 2025 a girl was raped and murdered in Amsterdam, an asylum seeker was arrested as suspect. Judging the content around this peak date, a lot of messages talk about the murder and call upon voting for the PVV who promises to close the border. In other words, political narratives and online hate narratives started way before the campaigning period.



Foreigners / refugees

Date	25 August - 30 October
Amount	116,702 posts
% that was toxic of the total dataset	23,283 were toxic, 20%
Average toxicity score	0.28 (0-1)
Platform	Twitter (77%), Youtube (8%) and Facebook (4%)
Messages with high toxicity are more frequent on	Twitter
Often involved	Politics
% of the posts have a toxicity score of 0.8 or higher (on a scale of 0-1)	0.6%
% contained violent language	14%
Keywords	islamisering, buitenlanders, omvolking

Out of 116,702 posts, 20% (23,283) were classified as toxic — a relatively high share, suggesting that hostile or aggressive rhetoric was a structural feature of discussions in this channel. The average toxicity score of 0.28 (on a 0–1 scale) confirms that this dataset was not dominated by mild irony or debate but by consistently negative, confrontational language. 0.6% of posts reached a toxicity score above 0.8, meaning several hundred posts contained explicit hate or incitement. While numerically small, this level of extreme content is significant because such posts often set the tone and are widely shared or quoted. The fact that 14% of posts contained violent language — such as calls for harm or militarized metaphors — suggests that aggression and

dehumanization are normalized within this thematic area, particularly when directed at out-groups like migrants or Muslims. This indicates that the discourse in this channel was not just politically toxic but ideologically charged, fusing nationalism, Islamophobia, and conspiracy thinking.

By far, most of the toxic content appeared on X. This dataset contains 95,198 posts, of which 19,881 (20.9%) are toxic. About 12% of the toxic messages contain violent language.

However, across all platforms, online hate was strongly linked to political discussions and anti-migrant narratives. On YouTube, 16% of 9,919 posts were toxic, often focused on Nederland (the Netherlands), buitenlanders (foreigners), and islamisering (islamisation), with a moderate average toxicity score (0.26) and 14% containing violent language. TikTok showed lower levels of toxicity (6.3%, score 0.16), but recurring stereotypes such as gelukszoekers (opportunist) and islamisering (islamisation) suggest subtle normalization of xenophobic ideas. Instagram had a higher toxicity rate (13.5%, score 0.22) and more violent language (20%), with frequent mentions of criminelen (criminals) linking migration to crime. Though the Telegram dataset was small, it showed the most extreme content (16.4% toxic, 44% violent language), featuring openly conspiratorial and eliminationist terms like omvolking (Great Replacement) and deportatie (deportation). Overall, the same xenophobic frames appeared across all platforms, but their tone and intensity differed — from more coded hostility on mainstream networks to overt extremism on fringe spaces.

Across all platforms, the most common keywords — buitenlanders (foreigners), islamisering (Islamisation), gelukszoekers (fortune seekers / opportunists), criminelen (criminals), and omvolking (Great Replacement) — point to a shared narrative portraying migrants and Muslims as a demographic and cultural threat to the Netherlands. The data illustrates a continuum of hate; from subtle hostility and “coded” Islamophobia on mainstream platforms (YouTube, Instagram) to openly extremist speech on Telegram. The ideological core remains the same, a fear of cultural invasion and loss of national identity. But the tone shifts from normalized resentment to explicit calls for exclusion or violence.

Below is the word cloud of this channel. One can see the words most associated with foreigners / immigrants online and it is easily noticeable that conspiracy theories like islamisation (word in dutch: islamisering, mentioned 7012 times) and Great Replacement (in dutch: omvolking, mentioned 4853 times) are often linked with immigration. The strongest frames are security/criminality (with the dutch word criminelen (criminals) mentioned 5663 times), cultural incompatibility (islamisering (islamisation), sharia, christenen (christians)), and existential threat (omvolking (Great Replacement), overbevolking (overpopulation)). Mentions of links (leftwing), rechts (rightwing), politiek (politics) show that this hate is embedded in a broader political narrative of "us versus them." Compared to the other clouds, this one has the densest cluster of dehumanizing and fear-inducing terms.

LBGTQ+

Date	August 25 - 30 October 2025
Amount	19,212

% that was toxic of the total dataset	7,257 37.8%
Average toxicity score	0.43 (0-1)
Platforms	Twitter (85%), Youtube (4%) and Facebook (2%)
Messages with high toxicity are more frequent on	Twitter
Often involved	Sexism
% of the posts have a toxicity score of 0.8 or higher (on a scale of 0-1)	1.6%
% contained violent language	8%
Keywords	homo, vrouwen, mietjes

With 37.8% of 19,212 posts classified as toxic and an average toxicity score of 0.43, this dataset stands out as one of the most hostile compared to other monitored themes (e.g. migration or general politics). This indicates that discussions around gender, sexuality, or perceived “masculinity” sparked more overt hate and aggression than other political topics. Although overall toxicity is high, only 8% of posts contained violent language, suggesting that most hostility took the form of mockery, slurs, and humiliation rather than explicit calls for harm. This fits with patterns of performative misogyny and casual homophobia, where insults are used to enforce gender norms and reject inclusivity. The share of 1.6% of posts with toxicity scores above 0.8, though small, shows that a subset of users engaged in explicit hate or dehumanization. It contributes to a climate where violent misogyny and homophobia are normalized through repetition.

On X, the dataset contains 19,480 posts, of which 7,917 (40.6%) are toxic, which is an extremely high percentage. By far, most of the toxic content appeared on X. However,

all three other platforms, YouTube, Facebook and TikTok, show very high levels of toxicity compared to the migration-related datasets.

On YouTube, 27.6% of 764 posts were toxic (average score 0.36), mostly involving sexist and homophobic language such as homo (gay) and mietjes (sissies), often linked to right-wing identity. Facebook showed the highest toxicity (38.5%, score 0.43) and the most extreme hate (2.1% above 0.8), frequently targeting gay and transgender people through overtly sexist and degrading comments. TikTok had lower toxicity (18.9%, score 0.33), but still showed strong gender hostility (95% categorized as sexism), occasionally mixed with conspiratorial or religious tones (e.g. Illuminati, Allah). Overall, these platforms reveal a pattern where homophobia and sexism were used to mock or delegitimize gender diversity, reinforcing culture-war narratives about masculinity, morality, and national identity.

The consistently recurring keywords (homo, mietjes (sissies), transgender) make it clear that this dataset reflects direct homophobic and transphobic hostility, not just political disagreement. The frequent tagging of these messages under “Sexism” (79–95%) suggests that hate against LGBTQ+ groups is deeply tied to gendered, patriarchal, and hyper-masculine discourses.

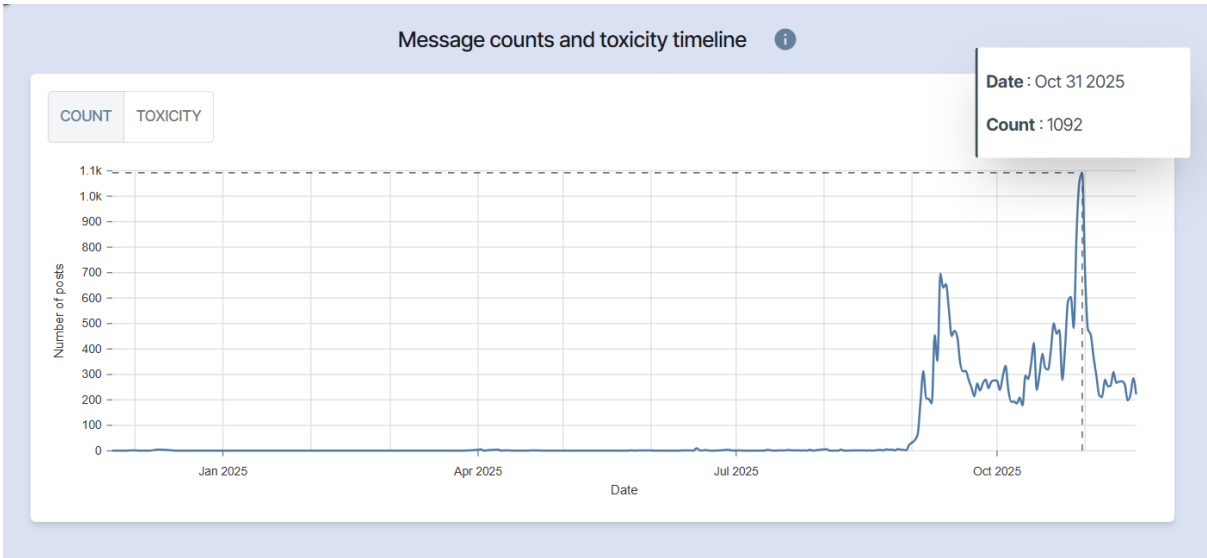
With 38.5% toxic posts and the highest average toxicity score (0.43), Facebook stands out as a hotspot for explicit hate speech. The presence of 2.1% of posts scoring above 0.8 indicates a high concentration of extreme hate, often more personal and demeaning. The overlap between homo, transgender, and mietjes points to mockery of gender non-conformity and rejection of progressive values, possibly reflecting generational or ideological divides among Facebook users.

Although its average toxicity (0.36) is slightly lower than Facebook’s, 27.6% toxic posts on YouTube is still substantial. The association with the keyword rechts (right-wing) is notable, it hints that homophobic language is being politicized, possibly used to assert traditional or nationalist masculinity in opposition to perceived “leftist” or “woke” values. The combination of homo and mietjes reflects how online ridicule often frames LGBTQ+ identities as symbols of weakness or moral decay, not just difference.

Although only 18.9% of posts are toxic on TikTok (lower than Facebook and YouTube), the appearance of Illuminati and Allah among keywords signals a conspiratorial and religious undercurrent in online homophobia. This suggests cross-contamination between far-right and religiously conservative discourses, where LGBTQ+ rights are framed as “unnatural,” “elite-driven,” or “anti-religious.”

Below, one can see the word cloud of this channel. Most used words in this channel are logically ´homo´ (gay), transgender and vrouwen (women). However, other words are less likely to appear here but interestingly enough do appear here; ‘moslims’ (muslims, mentioned 513 times), antifa (mentioned 552 times), israel (mentioned 214 times), ‘hoofddoek’ (headscarf, mentioned 196 times) and ‘joden’ (jews, mentioned 356 times). Hate in this channel focuses on moral decay and gender ideology, frequently connecting LGBT+ identities to indoctrinatie (indoctrination), pedofiel (pedophile), ziek (sick), doodstraf (death penalty). There’s overlap with anti-Islamic and anti-left rhetoric (moslims (muslims), links (left), antifa), suggesting convergence of culture-war frames where sexual diversity is framed as both a leftist and societal threat.

It is interesting to see that the chart below shows a peak in content in the days after the elections. In the days after the elections it became clear that D66 won the elections and that there is a high chance that the leader of D66, Rob Jetten, might become prime minister of the Netherlands. The messages in this channel on the days after the elections focus on the possibility of having a gay prime minister.



Dutch (Nederlanders)

Date	15 September - 30 October
Amount	23,092
% that was toxic of the total dataset	3,097 (13.4%)
Average toxicity score	0.20 (0-1)
Platform	Twitter (74%), Youtube (10%) and Tiktok (7%)
Messages with high toxicity are more frequent on	Twitter
Often involved	Politics

% of the posts have a toxicity score of 0.8 or higher (on a scale of 0-1)	0.4%
% contained violent language	7%
Keywords	nederland, linkse, dom

With 13.4% of 23,092 posts classified as toxic and an average toxicity score of 0.20, this dataset reflects a lower intensity than migration- or gender-related hate but still indicates that one in seven political posts carried some form of insult or aggression. Only 0.4% of posts reached a toxicity score above 0.8, meaning that severe or dehumanizing hate speech was rare.This reinforces the idea that polarization in Dutch political conversations operates more through mockery, dismissal, and ideological labeling than through explicit incitement or threats.

7% of posts contained violent wording, showing that while physical threats are infrequent, aggressive metaphors and anger remain part of the political vocabulary online.

On X, the dataset contains 20,701 posts, of which 3,009 (14.5%) are toxic. About 5% of the toxic messages contain violent language.

Across platforms, political discussions during the Dutch elections showed moderate but consistent levels of toxicity. On YouTube, 11.7% of 2,780 posts were toxic (average score 0.20), with keywords like Nederland (Netherlands), agenda 2030, and onzin (nonsense) pointing to conspiracy-tinged criticism of politics and global institutions. Facebook displayed a similar toxicity rate (11.5%, score 0.19) but more violent language (9%), often linking politics to xenophobic and anti-left rhetoric through terms such as buitenlanders (foreigners) and linkse (left). TikTok had lower toxicity (4.1%, score 0.10) but repeated racialized references to blanke Nederlanders (white Dutch), suggesting the normalization of nationalist identity talk in short-form political content. On Instagram, 12.8% of 516 posts were toxic (score 0.20), mixing political hostility with references to blanke (white) and Israël, reflecting the spread of nationalist and international grievance narratives.

Overall, toxicity across these platforms reflected a climate of political distrust, mild xenophobia, and conspiracy-influenced rhetoric, rather than overt hate speech.

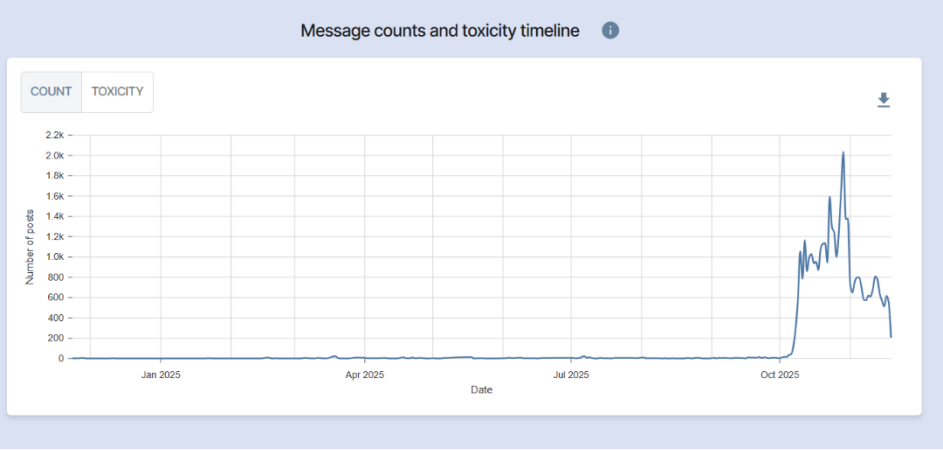
By far, most of the toxic content appeared on X. On the other three platforms, YouTube, Facebook and Instagram, toxicity levels are relatively stable: around 11–12% on YouTube, Facebook, and Instagram, and lower on TikTok (4.1%). This indicates that while overt hate was not dominant, a consistent layer of antagonism and distrust ran through online political discussions. The average toxicity scores (0.19–0.20) further suggest that this hostility took the form of mockery, sarcasm, and contempt, rather than explicit hate speech.

On YouTube, the keyword agenda 2030 stands out. This is a reference to the UN Sustainable Development Goals, often reinterpreted by conspiracy theorists as part of a "globalist" or "elite" agenda. Violent expressions remain limited (3–9% across platforms), yet the steady recurrence of political contempt and nationalist keywords shows how subtle antagonism has become normalized. The conversation tone is less about open aggression and more about resentment and moral outrage.

Below, one can find the word cloud from this channel. Interesting here is that as well words appear as 'moslim' (muslim), islamisering (islamisation, mentioned 2022 times), omvolking (Great Replacement, mentioned 178 times), vluchtelingen (refugees, mentioned 220 times) and allochtonen (foreigners, mentioned 110 times). This is also confirmed by the other data in this channel.

The discourse centers on national identity and purity, contrasting "echte Nederlanders" (real Dutch) with foreigners and Muslims. Terms like 'ons land' (our country), vaderland (homeland), landverraders (traitors), and 'de echte Nederlanders' (the real Dutch) highlight strong in-group/out-group framing. It overlaps substantially with the migration channel but adds a layer of national victimhood and racialized belonging.

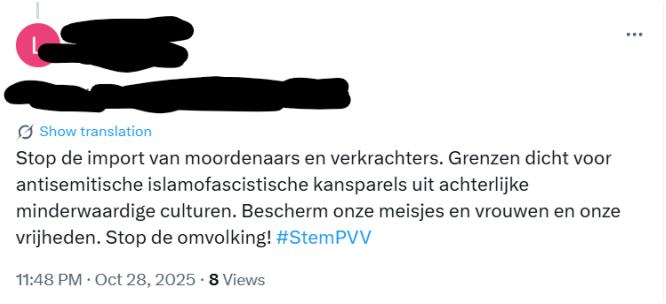
The chart below clearly shows the peak in content around the election date.



Findings

The analysis of Dutch-language online channels reveals that xenophobic narratives dominate nearly all monitored spaces, regardless of topic. Even channels focused on anti-LGBT+ hate prominently feature terms like “omvolking” (“replacement”) and “islamisering” (“Islamization”) in their word clouds, showing how far-right, anti-immigrant rhetoric infiltrates different discussions.

Across all monitored channels, the discourse repeatedly returns to asielzoekers (asylum seekers), buitenlanders (foreigners), and islamisering (islamisation), even when the initial topic is unrelated. Searches for neutral keywords such as Nederlanders (Dutch) still produce anti-foreigner hate and references to the Great Replacement, underscoring how deeply these ideas are embedded.



In Dutch society, “Islamization” has become a normalized conspiracy theory, often treated as fact and frequently used by politicians online and in the media. For example, Chris Stoffer, leader of the Christian party SGP, said in the newspaper NRC that “the Netherlands seems like an Islamic society” (NRC 2025). Lidewij de Vos, leader of Forum voor Democratie (FvD), told de Volkskrant that omvolking (Great Replacement) is “just a word in the dictionary,” downplaying its extremist origins (Volkskrant 2025). Geert Wilders shared a misleading video on X showing a Muslim crowd allegedly “awaiting orders” to take over Europe. The content was later debunked

as footage from a religious procession, yet his retweet was viewed over 224,000 times (X 2025).

Judging the content of the discussions online, anti-muslim hate in the Netherlands is multilayered, often blending anti-refugee hae, xenophobia, anti-Arab sentiment, and racism. Political figures routinely exploit such sentiments for electoral gain. When a 17-year-old Dutch girl was murdered by an asylum seeker in Amsterdam in August 2025, several politicians used the tragedy to argue that mass immigration endangers Dutch citizens. Geert Wilders cited “data” from researcher Jan van de Beek proving that muslims are overrepresented in criminality in the Netherlands. Beek´s work has been widely criticized for selective interpretation and ideological bias (BNNVARA 2024).
During the election campaign, several anti-immigration protests escalated into violence. Some demonstrators carried a variation of the Dutch flag once used by the Dutch Nazi Party, while others were seen giving the Hitler salute. The protests involved clashes with police, significant damage to public property, and even an attack on a political party’s office. These events were widely circulated online and largely coordinated through digital networks. Many of the groups behind them—operating under names like Defend Holland or Defend the Netherlands—use mainstream platforms to spread synthetic propaganda, conspiracy theories, and calls to stop the so-called Islamization of Dutch society, blending online disinformation with real-world mobilization.
These examples demonstrate how fear-driven narratives of Islamization and migration are politically weaponized, shaping public discourse far beyond fringe extremist spaces.
While the LGBT+ channel contains around 30% toxic content, much of it consists of hostile or incoherent posts full of profanity. The channel shows high levels of toxicity, often directed at public figures such as Rob Jetten. The amount of content increased in the days after the elections and was mostly directed at Jetten and his sexual identity.
Another interesting phenomenon is the amount of repeated messages. Due to the availability of tools to send out a sequence of copied messages, the dataset shows the same content several times on different dates. This message for example appeared about 15 times in the scraped data:

Across all datasets, politics was the most frequent context of toxicity, regardless of the topic, from migration (islamisation, Great Replacement) to gender (gay, sissies) and general debate (left, stupid).This shows that hate and hostility were rarely isolated prejudices; instead, they were woven into political identity conflicts. Gender and sexuality topics triggered the highest proportion of toxicity, showing that sexism and homophobia were even more aggressively expressed than xenophobia.

Across datasets, 5–20% of toxic messages contained violent expressions, though explicit calls for violence were rare. However, phrases like “deporteren”(deport them), “kapot slaan” (beat them), “weg ermee” (get rid of it) normalize aggression and dehumanization, turning symbolic violence into part of everyday online talk.

Despite differences, nearly every dataset included references to: the Netherlands, real Dutch people, white Dutch people, foreigners, islamisation and Great Replacement. These recurring frames link xenophobia, nationalism, and culture-war language into a shared meta-narrative: “The Netherlands (and its people) are under threat, from outsiders, progressive elites, or moral decay.” This framing unites anti-migrant, anti-LGBT+, and anti-left sentiments under one ideological umbrella, reflecting a mainstreamed far-right worldview.

By far, most of the content appears on X. This reinforces the finding that real-time, high-engagement platforms with low moderation are especially prone to toxic political discourse.

Comparing online discourse around Dutch party leaders

In the lead-up to, and during, the Dutch general elections, we also looked at social media posts to assess the tone of public online discourse around the country's political party leaders. The keywords we used in these channels were the last names of the leaders. This analysis focuses on social media messages published between August 25 and November 10, 2025. Data was collected using the multilingual AI-driven monitoring system and filtered for Dutch-language posts referencing individual party leaders. Platforms included Reddit, X/Twitter, 4chan, Gab, Telegram, YouTube, Facebook, Instagram, TikTok, 9gag, Threads, Bluesky, with Twitter accounting for the majority of toxic content across all datasets.

Only leaders with a dataset larger than 2,000 posts were included in the comparative analysis to ensure meaningful results. Datasets smaller than this threshold (Mirjam Bikker, Laurens Dassen, Chris Stoffer, and Eddy van Hijum) were excluded from the insights and visualisations below.

Data

Keyword analysis: Yeşilgöz

Dilan Yeşilgöz is the leader of the VVD party. This dataset covers messages posted between 1 September and 10 November 2025, containing 9,345 posts, of which 894 (9.6%) were identified as toxic. The majority of content originated from Twitter (80%), followed by YouTube (9%) and TikTok (3%). The average toxicity score is 0.17, and 0.1% of posts reached a toxicity score of 0.8 or higher. Highly toxic messages appeared predominantly on Twitter (71.4%). Toxicity was most frequently associated with political themes (64%), and 6% of toxic posts contained violent language. Common keywords occurring in toxic messages include antifa, kaag, and nederland.

Keyword analysis: Jetten

Rob Jetten is the leader of the D66 party. This dataset comprises 68,182 posts, of which 9,401 were identified as toxic, representing 13.8% of all collected messages. The majority of content originated from Twitter (82%), followed by YouTube (8%) and GeenStijl (3%).

The average toxicity score across the dataset is 0.20, and 0.4% of all posts reached a toxicity level of 0.8 or higher, indicating extreme hostility. Highly toxic content appears predominantly on Twitter (81.9%).

Toxic messages are most frequently associated with political discussions (56%), and approximately 4% contain violent language. Common keywords in toxic posts include “kaag,” “rob jetten,” and “homo,” reflecting the political and identity-based targeting patterns observed during the monitoring period.

Keyword: Timmermans

An analysis of online messages posted between 25 August 2025 and 10 November 2025 shows a total dataset of 94,677 posts, of which 12,524 (13.2%) contain toxic language. The majority of these posts originated from Twitter (76%), followed by YouTube (12%) and TikTok (4%).

The overall average toxicity score is 0.20, and 0.4% of all posts reach a toxicity level of 0.8 or higher, indicating extreme hostility. Highly toxic content appears predominantly on Twitter (83.7%). Much of the toxicity relates to political discourse (64%), and around 7% of the toxic messages include violent language. Frequently occurring keywords in these toxic posts include “nederland”, “linkse”, and “Frans Timmermans”, highlighting how political identity and polarization shape the online conversations surrounding him during the campaign period.

Keyword: Wilders

Geert Wilders is the leader of the PVV party. Between 25 August and 10 November 2025, a total of 130,540 online messages were collected for analysis. Of these, 15,986 posts (12.2%) were identified as toxic. The majority of the content originated from Twitter (72%), followed by YouTube (12%) and TikTok (7%). The dataset shows an average

toxicity score of 0.19, with 0.3% of posts reaching a very high toxicity level of 0.8 or above. Highly toxic messages appeared predominantly on Twitter (73.6%). Toxicity in this dataset is most frequently linked to political themes (62%), and 13% of toxic posts contain violent language. Common keywords occurring in toxic messages include "nederland," "linkse," and "links."

Keyword: Bontenbal

Henry Bontenbal is the leader of the CDA party. An analysis of messages posted between 25 August 2025 and 10 November 2025 shows a dataset of 30,756 posts, of which 3,705 (12%) were classified as toxic. The majority of content originated from Twitter (86%), followed by YouTube (6%) and Facebook (2%). The dataset has an average toxicity score of 0.19, and 0.2% of posts reached a high-toxicity threshold of 0.8 or above. Highly toxic messages appeared disproportionately on Twitter (81.1%). Toxicity was most commonly linked to political discussions (58%), and 2% of toxic posts contained violent language. Frequently recurring keywords in toxic messages include "nederland," "homo," and "henri bontenbal."

Keyword: Jimmy Dijk

Jimmy Dijk is the leader of the SP party. An analysis of messages posted between 1 September 2025 and 10 November 2025 shows a total of 4,499 posts, of which 521 (11.6%) were classified as toxic. The majority of the data originated from Twitter (77%), followed by YouTube (9%) and Facebook (4%). The average toxicity score across all messages was 0.18, with 0.2% of posts reaching a toxicity score of 0.8 or higher. Highly toxic messages appeared most frequently on Twitter (81.8%). Toxic content in this dataset predominantly fell under the category of Ridicule (53%), and approximately 4% of toxic messages contained violent language. Common keywords associated with toxic messages included: "jimmy dijk," "clown," and "belachelijk."

Keyword: Lidewij de Vos

Lidewij de Vos is the leader of the FVD party. An analysis of messages posted between 1 September 2025 and 10 November 2025 shows a total dataset of 13,137 posts, of which 973 posts (7.4%) were classified as toxic. The majority of content originated from Twitter (74%), followed by YouTube (13%) and TikTok (8%). The average toxicity score across the dataset is 0.13, and 0.2% of all posts reached a toxicity score of 0.8 or higher. Highly toxic messages appear most frequently on Twitter (80.0%). Toxic content is predominantly associated with political themes (62%), and approximately 4% of toxic posts contain violent language. Common keywords recurring in toxic messages include “antifascisten,” “antifa,” and “onzin.”

Keyword: Ouwehand

This analysis covers messages posted between 25 August 2025 and 10 November 2025. The dataset contains 12,367 posts, of which 1,749 (14%) were identified as toxic. Most messages originated from Twitter (83%), followed by YouTube (8%) and GeenStijl (4%). The average toxicity score across the dataset is 0.20, with 0.5% of posts reaching a toxicity level of 0.8 or higher. Highly toxic messages appear predominantly on Twitter (77.4%). Toxicity most frequently relates to political topics (54%), and approximately 3% of toxic messages contain violent language. Commonly occurring keywords in toxic messages include “Esther Ouwehand,” “Hamas,” and “Kaag.”

Keyword: Eerdmans

An analysis of messages posted between 1 September 2025 and 10 November 2025 shows a total of 11,179 posts, of which 1,252 (11.2%) contain toxic language. The majority of content was collected from Twitter (81%), followed by YouTube (10%) and TikTok (2%). The average toxicity score across the dataset is 0.18, with 0.2% of posts reaching a toxicity score of 0.8 or higher. Highly toxic messages appear most frequently on Twitter (71.4%). Toxic content is most commonly associated with political themes (54%), and approximately 3% of toxic posts contain violent language. Frequently occurring keywords in toxic messages include: joost eerdmans, plas, and rechts.

Keyword: van Baarle

Between 25 August 2025 and 10 November 2025, a total of 10,456 messages were collected, of which 1,589 posts (15.2%) were identified as toxic. The majority of the content originated from Twitter (70%), followed by YouTube (17%) and TikTok (7%). The average toxicity score was 0.20, and 0.6% of messages reached a toxicity level of 0.8 or higher. Highly toxic content appeared predominantly on Twitter (88.5%). Toxicity was most frequently associated with political discussions (49%), and approximately 9% of toxic messages contained violent language. Common keywords found in toxic posts included “hamas,” “nederland,” and “genocide.”

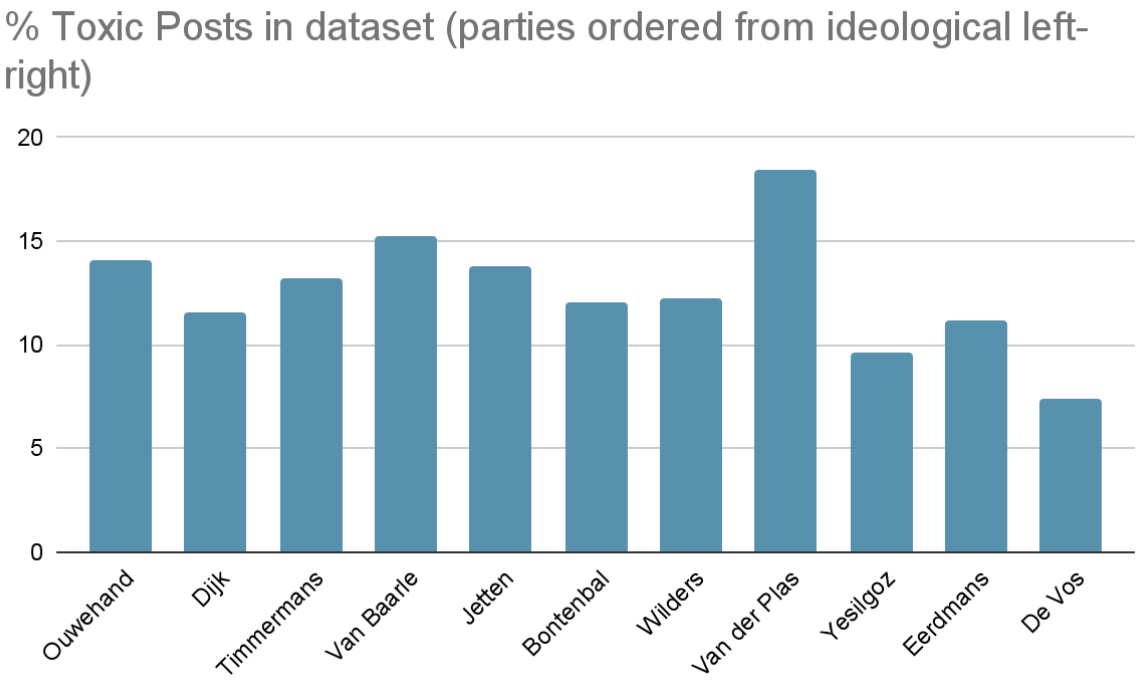
Keyword: van der Plas

Analyzing messages from 1 September 2025 to 10 November 2025, this dataset contains 7,246 posts, of which 1,334 (18.4%) were classified as toxic. The majority of content was collected from Twitter (83%), followed by YouTube (9%) and TikTok (2%). The average toxicity score across the dataset is 0.37, and 0.3% of posts reached a high toxicity level of 0.8 or above. Highly toxic messages appeared predominantly on Twitter (95.0%). Toxic content most frequently related to political themes (50%), and approximately 4% of toxic posts contained violent language. Commonly occurring keywords in toxic messages included “plas,” “Caroline van der Plas,” and “Israël.”
Of the leaders with sufficient data, Rob Jetten, Frans Timmermans, and Geert Wilders were among the most frequently mentioned. Their datasets also showed elevated levels of toxicity, though not always in equal measure. The table below summarizes the main indicators.

Leader	% Toxic Posts	Avg. Toxicity	% Highly Toxic (≥0.8)	% Violence
Jetten	13.8 %	0.20	0.4 %	4 %
Timmermans	13.2 %	0.20	0.4 %	7 %
Wilders	12.2 %	0.19	0.3 %	13 %

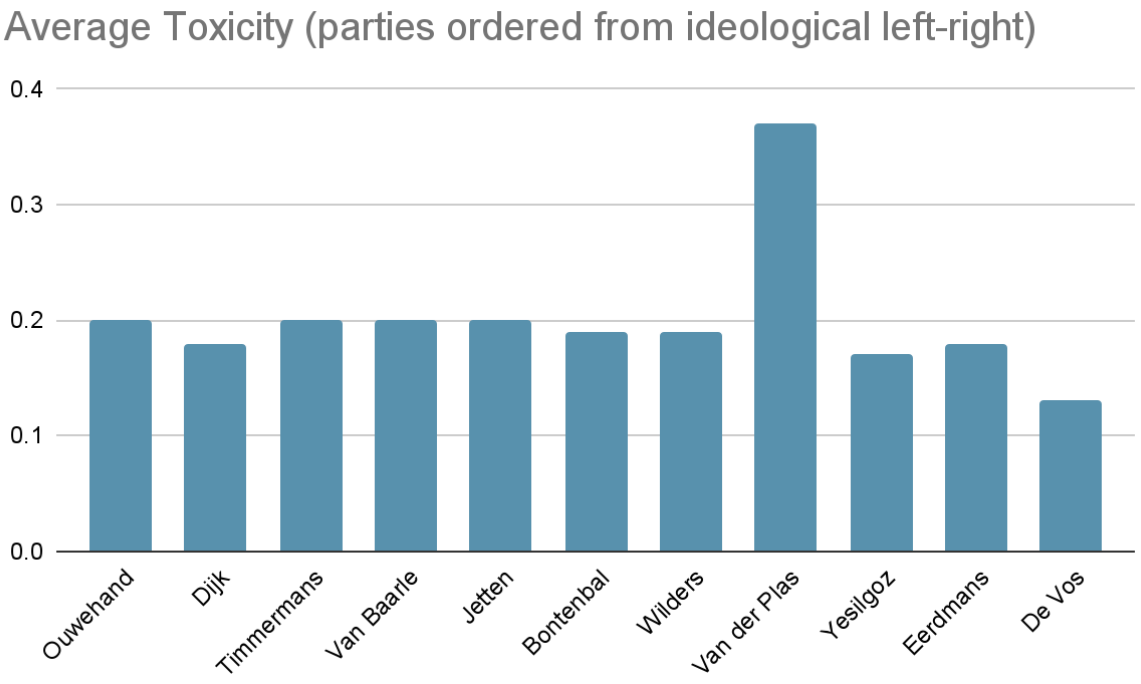
Ouwehand	14.1 %	0.20	0.5 %	3 %
Bontenbal	12 %	0.19	0.2 %	2 %
Van der Plas	18.4 %	0.37	0.3 %	4 %
Van Baarle	15.2 %	0.20	0.6 %	9 %
Eerdmans	11.2 %	0.18	0.2 %	3 %
Yesilgöz	9.6 %	0.17 %	0.1 %	6 %
De Vos	7.4 %	0.13	0.2 %	4 %

In the graph below one can see the percentage of toxic posts in the dataset per politician. Van der Plas received the highest percentage of toxic posts.

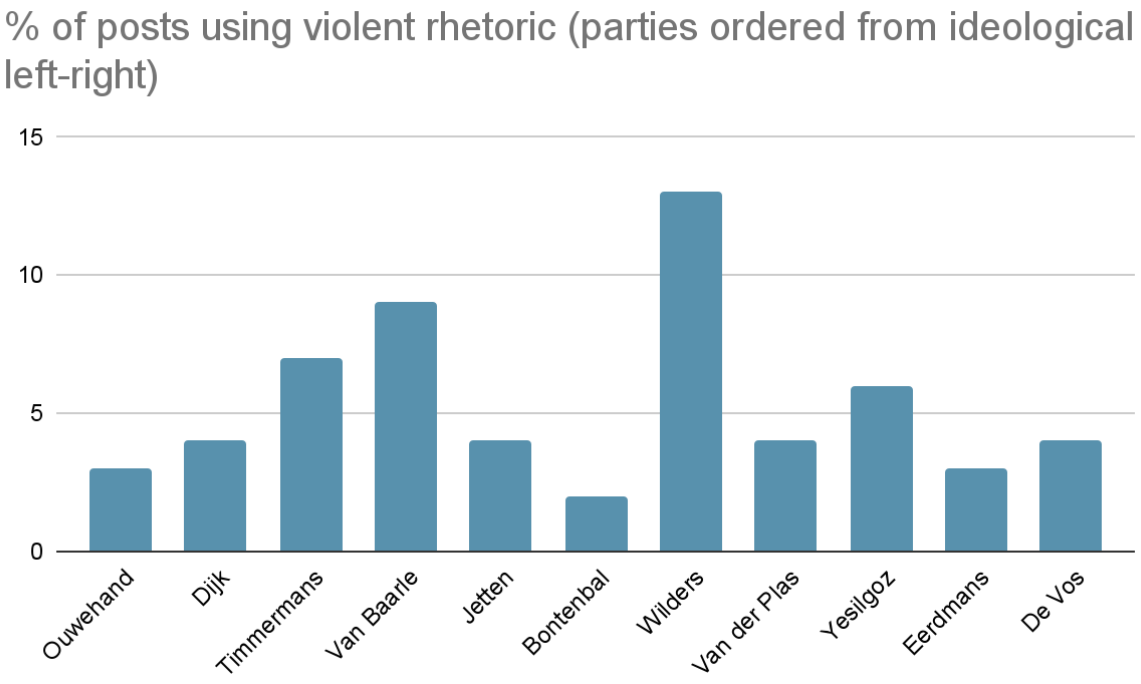


In the graph below, the average toxicity per politician is shown. Again, Van der Plas received the highest average toxicity in the content about

her.



In the graph below the percentage of posts using violent rethoric per politician is visualized. Here, Wilders by far received the highest amount of violent rhetoric.



Findings

Higher Toxicity on the Left, But with Important Nuances

Contrary to the assumption that far-right figures attract the most hostility online, some of the highest levels of toxicity were aimed at left-leaning leaders. Esther Ouwehand (PvdD), Frans Timmermans (GroenLinks-PvdA), and Rob Jetten (D66) all registered toxicity levels above 13%. Much of this toxicity focused on perceived ideological stances and personal identity. For example, Jetten's dataset contained frequent homophobic slurs, with keywords such as "kaag" and "homo" appearing disproportionately in toxic posts.

Right-Wing Leaders: High Volume, Lower Toxicity

Geert Wilders and Dilan Yeşilgöz were among the most mentioned leaders but had relatively lower toxicity rates (12.2% and 9.6% respectively). However, this does not mean the tone of discussion was more civil. Posts about Wilders, for instance, contained the highest proportion of violent rhetoric (13%) of any included dataset.

Ouwehand, Van der Plas, and Van Baarle Stand Out

Esther Ouwehand (PvdD) and Caroline van der Plas (BBB) stood out with some of the highest toxicity and average toxicity scores, especially in relation to their overall visibility. Stephan van Baarle's dataset included the highest rate of highly toxic posts (0.6%), and a relatively high share of violent language (9%).

Platform Dynamics Drive Polarisation

Across all datasets, Twitter/X dominated as the platform with the highest share of toxic messages, accounting for over 70% of toxicity. This aligns with earlier EOOH findings and reflects the high-velocity, low-moderation nature of the platform.

Narrative Triggers and Toxic Keywords

Frequently recurring terms in toxic posts—such as “linkse,” “antifa,” “nederland,” and “kaag”—signal how public figures can become symbols within broader ideological battles. These keywords function as triggers in both right- and left-wing commentary.

Conclusion

Overall, the data demonstrates that xenophobic and conspiratorial narratives—particularly those centered on Islamization, asylum seekers, and the Great Replacement—have become deeply embedded in Dutch online discourse, permeating discussions far beyond extremist spaces and appearing even in conversations unrelated to migration. These narratives are amplified by political actors and exploited during moments of social tension, contributing to a climate in which anti-Muslim sentiment, xenophobia, and hostility toward minorities are normalized. Toxicity across datasets is consistently tied to political identity conflicts, while gender- and sexuality-related topics generate the highest levels of abuse, indicating that sexism and homophobia remain potent drivers of online hate. Although explicit calls for violence are rare, repeated dehumanizing and aggressive language signals a troubling normalization of symbolic violence. The dominance of X in the dataset highlights how low-moderation, high-velocity platforms act as accelerators of toxic political speech. Together, the findings illustrate a mainstreamed far-right narrative ecosystem in which fears of cultural collapse, demographic threat, and moral decline bind various forms of hate into a cohesive worldview that increasingly shapes public debate in the Netherlands.

The comparative analysis shows that online toxicity during the 2025 Dutch elections was widespread, and not confined to one side of the political spectrum. While progressive leaders like Jetten and Timmermans received higher shares of toxic commentary, right-wing figures like Wilders and Van der Plas encountered more violent and emotionally charged attacks. The findings also reveal how identity, ideology, and platform dynamics intersect to shape public discourse. With Twitter/X continuing to dominate as the most toxic platform, and specific keywords driving ideological aggression, the digital landscape remains highly volatile.

One of the insights from this analysis is how narratives, once they are embedded in the public discourse, continue to shape online toxicity, regardless of current political leadership. The recurring presence of terms like “Kaag” in toxic posts about D66’s Rob Jetten, despite Sigrid Kaag no longer leading the party, highlights how certain names

become lasting symbols within polarised political storytelling. Symbolic figures like Kaag and Timmermans serve as shorthand for broader grievances, whether about migration, gender, or elite distrust, and are easily reactivated in each new campaign cycle.

Generative AI tools risk accelerating this trend: with the ability to mass-produce images, memes, and synthetic quotes, even outdated or debunked narratives can be repackaged for virality. When combined with emotionally charged rhetoric, GenAI can help political actors (or their supporters) revive and remix old polarising frames, further distorting public debate. The findings around keywords like “Kaag,” “homo,” “linkse,” and “nederland” show how toxicity is not just reactive, it is part of a narrative infrastructure that GenAI could help entrench.

Online hate does not exist in a vacuum. Broader (global) narratives, especially those tied to conflict, identity, and injustice, shape and amplify discourse. During the Dutch elections, we observed that terms like Israel, Gaza, and genocide appeared across multiple datasets, especially in relation to leaders such as Van Baarle and Ouwehand. This illustrates how international events bleed into national conversations, fueling polarisation and intensifying digital hostility. As political discourse becomes more globalised and emotionally charged, systematic monitoring of these online patterns becomes essential, not only to understand electoral dynamics, but to safeguard democratic debate.

Recommendations

As this report has shown, online hate narratives played a significant and recurring role throughout the 2025 Dutch parliamentary elections. They shape public debate, influence perceptions of minority groups, and amplify polarisation. Monitoring this alone is not enough; these dynamics require a coordinated response from policymakers, platforms, civil society, and the wider public. To address the spread and impact of online hate in political contexts like the 2025 Dutch elections, we propose a multi-stakeholder approach. Each actor in the digital ecosystem (social media platforms, regulators, civil society, media, and political institutions) has a specific role to play. Below are targeted recommendations for each group:

1. Social Media Platforms: Build Safer Digital Spaces

- **Prioritise context-aware moderation**: Develop moderation tools that better detect coded language and polarising narratives during election periods. This includes hate that falls outside traditional illegal speech but fuels discrimination or incites hostility.
- **Enable meaningful counterspeech**: Amplify diverse and constructive voices by adjusting algorithms to highlight trusted messengers, verified information, and context-rich content.
- **Improve transparency**: Provide real-time dashboards and public access to data on content removals, strikes, and political advertising during elections.
- **Restrict coordinated inauthentic behaviour**: Proactively detect and deplatform networks engaged in repetitive hate messaging, AI-generated content abuse, and harassment campaigns.

2. Digital Services Act (DSA) Implementation

- **Enforce systemic risk mitigation obligations**: Ensure Very Large Online Platforms (VLOPs) conduct and publish meaningful risk assessments before and after elections, including their impact on public discourse and hate speech.

- **Strengthen DSC engagement**: The Dutch Digital Services Coordinator (DSC) should actively monitor platform compliance and facilitate coordination between NGOs, researchers, and tech companies.
- **Mandate algorithmic transparency**: Platforms should be required to disclose how recommendation systems influence polarising content, especially on high-volume platforms like X/Twitter.
- **Clarify thresholds for enforcement**: The EU Commission and DSCs should define what constitutes sufficient mitigation, including removing harmful trends, suppressing virality of synthetic propaganda, and correcting platform inaction.

3. Public and Civil Society: Build Digital Resilience

- **Scale digital literacy programs**: Teach users to identify synthetic propaganda, emotionally manipulative narratives, and disinformation techniques, particularly targeting younger and vulnerable groups.

- **Empower users to take control**: Raise awareness of DSA rights such as viewing content chronologically, turning off algorithmic feeds, and flagging harmful content effectively.
- **Support grassroots counterspeech**: Invest in training and infrastructure for community-driven responses to hate speech that use persuasive storytelling, humour, and credible messengers.
- **Fund watchdogs and researchers**: Ensure long-term support for independent organisations that monitor hate trends, platform accountability, and electoral manipulation.

4. Media and Journalism: Strengthen Ethical Framing

- **Avoid amplification**: Media should resist framing polarising narratives as legitimate political debate when rooted in disinformation or dehumanisation.
- **Contextualise hate incidents**: Reports on online hate should trace its origins, intent, and links to coordinated campaigns, not just surface-level quotes.
- **Promote narrative nuance**: Elevate stories that challenge binary thinking, expose manipulation, and feature marginalised voices.
- **Hold platforms and politicians accountable**: Investigative journalism should continue to uncover the political actors and networks behind hate campaigns.

5. Political Institutions: Set Democratic Standards

- **Adopt a Code of Conduct**: Political parties and candidates should publicly commit to a code of ethics that rejects hate speech, disinformation, and intimidation.
- **Disavow synthetic manipulation**: Ban the use of generative AI for deceptive political messaging, including synthetic videos, images, or soundbites.
- **Train political communicators**: Provide capacity-building on responsible digital campaigning, narrative framing, and social media ethics.
- **Create oversight mechanisms**: Develop independent electoral monitoring bodies to track online campaign behaviour and intervene when thresholds are crossed.

Sources

Synthetic Propaganda, AI & Disinformation

- EU DisinfoLab (2024). Synthetic Propaganda: Generative AI and the Future of Political Communication (4 September).
  EU DisinfoLab - 4 September: Synthetic Propaganda – Generative AI and the Future of Political Communication

- Pakina, A. K. (2024). AI-Driven Disinformation Campaigns: Detecting Synthetic Propaganda in Encrypted Messaging via Graph Neural Networks. ResearchGate.
  https://www.researchgate.net/.../AI-Driven-Disinformation-Campaigns-Detecting-Synthetic-Propaganda-in-Encrypted-Messaging-via-Graph-NNeural-Networks.pdf

- Volkskrant (2025). Twee PVV-Kamerleden vallen met nepbeelden anoniem Timmermans aan; GroenLinks-PvdA doet aangifte.
  https://www.volkskrant.nl/politiek/twee-pvv-kamerleden-vallen-met-nepbeelden-anoniem-timmermans-aan-groenlinks-pvda-doet-aangifte~b6958ea1/?utm_campaign=shared_earned&utm_medium=social&utm_source=copylink

- Volkskrant (2025). AI-campagne in opkomst: De helft van AI-berichten komt van PVV.
  https://www.volkskrant.nl/verkiezingen-tweede-kamer/ai-campagne-in-opkomst-de-helft-van-ai-berichten-komt-van-pvv~b460de2b/

National Security & Threat Assessments

- AIVD (2024). AIVD Annual Report 2024.
  https://www.aivd.nl/documenten/jaarverslagen

- Algemeen Dagblad (AD) (2024). Rechts-radicaal AI-liedje tegen AZC's bestormt hitlijsten; Spotify weert nummer niet.
  Tech | AD.nl

Migration, Remigration Narratives & Far-Right Messaging

- AP News (2024). Germany: Migration, Far-Right Election Dynamics & Deportation Debates.
  A far-right party uses a mock plane ticket to campaign for deportations in Germany | AP News

- Bild (2024). Remigration ist „Unwort des Jahres".
  Jury-Entscheidung: Remigration ist „Unwort des Jahres" | Leben & Wissen | BILD.de

- The Guardian (2024). How "Remigration" Became a Buzzword for the Global Far Right.
  How remigration became a buzzword for global far right | The far right | The Guardian
  https://www.theguardian.com/

- Ekman, M. (2022). The Instrumentalisation of Migration in the Populist Era. Mixed Migration Centre.
  https://mixedmigration.org/the-instrumentalisation-of-migration-in-the-populist-era/

- Kaya, A. (2022). The Instrumentalisation of Migration in the Populist Era. Mixed Migration Centre.
  https://mixedmigration.org/the-instrumentalisation-of-migration-in-the-populist-era/

- Joop / BNNVARA. Migratiemagneet Nederland? Jan van de Beek winkelt selectief in de cijfers.
  Joop - BNNVARA

Dutch Political Context, Public Trust & Election Climate

- Metro Nieuws (2025). Nederlanders hebben een historisch laag vertrouwen in de politiek.
  https://www.metronieuws.nl/in-het-nieuws/binnenland/2025/10/vertrouwen-in-politiek-bereikt-historisch-dieptepunt-xyz/

- *NRC (2025). Fukuyama, F. "It's the internet, stupid."*
  https://www.nrc.nl/nieuws/2025/10/09/its-the-internet-stupid-a4908843

Statements by Political Actors

- NRC. Verkiezingsspeech Chris Stoffer (SGP): "Het lijkt wel een soort islamitische samenleving te worden."
  Verkiezingsspeech Chris Stoffer - SGP

- Volkskrant. Lidewij de Vos (interview). "Waar bent u op uit? Dat u kunt opschrijven dat Forum van azc's gevangenissen wil maken?"
  Lidewij de Vos: 'Waar bent u op uit? Dat u kunt opschrijven dat Forum van azc's gevangenissen wil maken?' | de Volkskrant

- X / Twitter. Geert Wilders (2025). “Melbourne, Australië: ‘Wij wachten op bevelen’.”
  (Viewed 224.3K times).
  https://x.com/geertwilderspvv/status/1955512782738296976

- Facebook. Geert Wilders (2025). “Waarom durft niemand te benoemen dat we geen…”
  https://www.facebook.com/geertwilders/posts/pfbid02nXzbftz9BnWBbbpfVZYitibafMyzUEbyx3fz4GgifKfArdTSdUScSmd6yUzd2Coml