

Policy Brief

Who Gets Deepfaked?

Measuring the exposure of European parliamentarians in deepfake pornography

Benjamin Shultz

Executive Summary

- **33x more likely** for women parliamentarians than men to be depicted in or be associated with deepfake pornography, after adjusting for age, migration background, institutional prominence, political lean, and national legal context.
- **138 women parliamentarians** from twenty-two countries were found to be depicted in or be associated with deepfake pornography, compared to nine men, out of 5,872 sitting members audited. This translates to one in fourteen women, versus one in five hundred men.
- **“Gateways”** are newsfeed-like pages of members, which include biographical details and non-nude images, linked directly to AI-nudifier apps, enabling local deepfake production in a few clicks with manipulable images at the ready.
- **Legislation ≠ protection:** countries with deepfake-specific legislation did not show lower exposure in our data, indicating an enforcement gap rather than a policy failure.
- **Search results outlive takedowns:** using the seizure of a prolific deepfake-hosting domain as a controlled test, we identified forty-seven records discoverable using Google from this domain which returned as hits more than a month after the seizure.

This brief summarizes findings from an external audit of AI-generated non-consensual intimate imagery (NCII), colloquially known as deepfake pornography, depicting sitting members of national parliaments across the European Union. We find that exposure to this harm is not randomly distributed: it falls overwhelmingly on women, is amplified by institutional visibility, and is not reliably predicted by whether a country has passed deepfake-specific legislation. We map new forms of exposure that parliamentarians face and argue that takedown-based governance, built around removing content after the fact, cannot reach a harm produced upstream, often using locally downloaded tools. We then propose three concrete points of intervention.

The Problem

Generative AI systems can now produce convincing sexualized imagery of any person whose photograph is public without that person's participation, knowledge, or any expressive act.



Governments around the world are warning parents, schools, and sports teams to exercise caution when posting photos of children and young people, as routine, well-intentioned image-sharing has itself become raw material for AI-generated NCII.¹ For sitting parliamentarians, the problem is similar. The only “act” required is holding office: a single official portrait, or a photograph or recording from an event, published as a routine condition of the role, is sufficient raw material for a generative model to extract and recombine one’s likeness. This is a technological escalation from how vulnerability to online harms has typically been understood and regulated.

In October 2025, [Italy created a criminal offense](#) for deepfake dissemination. Enforcement, however, activates only on the victim’s own complaint, and targets the individual who disseminates content, not the tools or infrastructure that made it possible. [France’s](#) and [Australia’s](#) 2024 deepfake statutes, as well as the [2025 US TAKE IT DOWN Act](#), all share the same underlying architecture: report the content, then remove it and then — sometimes — prosecute. Each of these instruments activates only after a depiction already exists, has already been found, and has usually already circulated. The deepfake-producing ecosystem has, however, [moved into the shadows](#) in the last year, following the takedown and seizure of several prolific publicly available repositories, including [MrDeepFakes](#) and [CFake and SOCFake](#).

The evidence for this is now substantial and convergent, and 2026 has seen myriad research mapping the ecosystem. The Center for Countering Digital Hate [estimates xAI’s Grok chatbot generated](#), in a sudden and uncontrolled rampage, three-million sexualized, non-consensual deepfakes in January, including 23,000 of children — an amount nearly impossible to prosecute in full. In March, Open Measures [documented](#) a functioning commission-based deepfake economy on 4Chan’s requests board, with daily activity nearly tripling from 2025-2026, before the board was shut down in May. In July, research published from AI Forensics shows the most-used image-editing tools on the machine learning community board HuggingFace [comply instantly with undressing prompts](#), with an estimated 95% of sexualization attempts targeting women and 6.7% targeting children. There remains zero safeguards on the platform. And the Institute for Strategic Dialogue found this ecosystem to [live off mainstream infrastructure](#) using app stores, search engines, and payment processors to produce and distribute deepfake NCII.

Despite the documentation of the novel tactics, tools, marketplaces, technical and payment infrastructure involved, little measurement of how this harm is distributed across a defined, comparable population at intergovernmental scale has occurred, nor an assessment of its direct impact on democracy. Elected officials, from [the UK](#), to [the US](#), to [the European Union](#), have been depicted and bravely come forward to testify about how this harm has affected them. Motivated by these accounts and the [already-gendered state of online abuse](#), this brief reports, to our knowledge, the first complete, 27-member-state-audit of AI-generated NCII targeting sitting parliamentarians across the European Union.

Methods

Between 16-20 July we queried the names of 5,872 sitting members of national parliaments from the 27 European Union member states (94.4% coverage) against 160 domains known to host deepfake pornography, “celebrity leaks,” and inappropriate candid images of women in public life (e.g., upskirting). There are approximately 6,218 sitting members in lower and unicameral houses of the European Union’s national parliaments; after excluding temporarily vacant seats

¹ See, e.g., July 2026 warnings from the [Australian](#) and [British](#) governments.



at the time of analysis, as well as gender non-binary, and independent and non-attached members, on account of these groups being too underrepresented to conduct statistical analysis on, we reached a population figure of 5,872 members. We constructed eight Google Custom Search engines, assigning 20 domains to each, and leveraged Google’s application programming interface (API) to pull down results. Our methodology was designed to be non-intrusive: we neither interacted with any producers or distributors of depictive content, nor downloaded or stored depictive content at any stage. Each valid ‘hit’ from the API was coded into one of six categories capturing: 1) whether deepfake NCII was actively hosted; 2) whether a generation tool was linked next to a named individual’s biographical information, with or 3) without raw material present; 4) whether an indexed search record persisted, with or 5) without a named webpage or a production request specific to a parliamentarian; or 6) whether a domain had been formally removed by authorities. The full coding regime can be seen in the appendix.

36 domains yielded hits. Our method was validated against 100 non-existent names, created using a name generator. With the exception of seven name collisions with real adult performers, this test produced no positive results, confirming that the patterns we observe reflect targeted rather than indiscriminate activity. Exposure was then modeled against gender, age, name-perceived migration background,² institutional prominence (current or former cabinet or party leadership), political orientation, and the presence of deepfake-specific legislation in each country.³ *Shortly before the publication of this report, eight domains were seized by the Manhattan (New York) District Attorney’s Office, eliminating the exposure that 19 MPs faced.*

What We Found

1

Women MPs are 33x more likely to face exposure than their male colleagues

Odds ratios show that women are 33-times more likely than their male colleagues to be depicted in or associated with deepfake pornography. Thus, gender is the dominant predictor in the model ($p=0.000$). While women were exposed at scale in all categories, among the nine men, eight were identified in only indexed, non-findable search results. We note that a limited number of active depictions of women (less than ten) were potentially photoshopped rather than AI-generated. However, we counted these nonetheless as non-consensual depictions. In line with [AI Forensics’ research](#) detailing the proliferation of custom-built machine learning models meant to nudify women, we did find one likeness and associated downloadable model on HuggingFace, which we classified as an active depiction. The full count of exposed MPs by category can be seen in the Appendix.

	Not Exposed	Exposed	% Exposed
Women MPs	1,888	138	7.3%
Men MPs	3,837	9	0.2%
Total	5,725	147	2.6%

² Assessing name-perceived migration background across such a wide swath of states with differing definitions of race, ethnicity, and religion, as well as naming conventions, is a challenge; we attempted this to the best of our ability using [Nationalize.io](#).

³ We used a Firth-corrected logistic regression for this analysis specifically because it performs reliably on rare-outcome and group-concentrated data of this kind. Further explanation can be seen in the appendix.

2

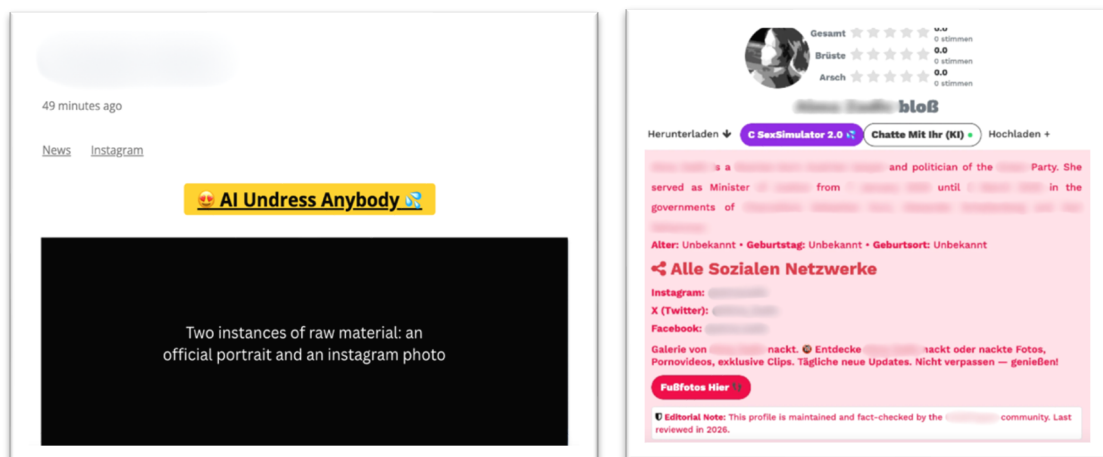
Institutional prominence compounds the gap

Among women parliamentarians specifically, those in senior office (cabinet ministers, party leaders) face substantially higher predicted exposure than their more junior colleagues ($p=0.000$). Because these are positions conferred primarily through political success rather than personal characteristics, this compounding effect is itself a structural pattern: achieving greater public standing carries a cost for women that it does not carry for men in equivalent positions.⁴ We also tested whether political leaning factored into a parliamentarian's probability of being depicted. We found left- ($p=0.003$) and right-leaning ($p=0.038$) parliamentarians to face a higher probability of exposure than their centrist colleagues. This aligns with prior work from the American Sunlight Project [analyzing the US Congress](#), which demonstrated that exposure does not concentrate in any one place along the political spectrum.

3

“Gateways” link raw material and biographic information to “nudifier” apps

We term domains which showcased specific, named pages for parliamentarians and linked directly on these pages to AI-generation tools designed to produce sexualized or nude depictions of subjects as “gateways.” In total, we observed 61 instances of gateways with raw material present and 36 instances of gateways without raw material present. Notably, gateways for three parliamentarians did show active deepfake NCII. Destinations of the gateways we identified appeared to include three notable “nudifier apps,” as well as two in-house generation capabilities. During the course of writing this report, one gateway website removed its out-links to one of the three nudifier apps and two gateways originally observed without raw material gained raw material. Two examples can be seen below of a gateway, respectively, with (obstructed by the black box below), and without, raw material.



On all variations of gateways, there is a specific biography or newsfeed (termed interchangeably), as well as links to parliamentarians' social media accounts. Some also exhibited 'ratings' of women's bodies. To the earlier question of enforcement, it may well be that these sites are not technically in violation of deepfake-specific laws, as they do not show active deepfake NCII. However, even where they do not provide raw material, they provide all of the information and tools necessary for a user to acquire or use raw material to create a non-consensual, sexually explicit deepfake depicting a parliamentarian.

⁴ To validate this finding, we ran our model a second time with an interaction term between gender and institutional prominence, which proved statistically insignificant ($p=0.485$), indicating these variables operate independently.

4

Legislation alone does not predict lower exposure

Counter to expectation, countries with deepfake-specific criminal legislation did not show lower exposure. We do not read this as evidence that legal protection is ineffective. Several countries without a specific deepfake NCII statute, [including the Netherlands](#), actively prosecute this material under existing privacy or image-abuse laws, while some countries with a statute on the books show limited enforcement activity. The relevant policy variable, therefore, is enforcement, not statutory text, a distinction current legal tracking does not consistently capture. To the point of the Netherlands, we retroactively re-coded it as having a deepfake-specific law given its robust enforcement apparatus. We aim to grapple with this dynamic and better measure enforcement in subsequent runs of this study.

5

Indexing is a distinct layer of exposure which outlives takedowns

In the course of planning this study, the prolific deepfake-hosting domain CFake [was seized](#) by US and French authorities. We theorized that parliamentarians' vulnerability to exposure in deepfake NCII was regulated on distinct layers, including — amongst other sociotechnical factors like [gendered attention](#) and [interface accessibility](#) — indexing and hosting infrastructures. As a control, we included CFake in our list of domains and were able to retrieve 47 still-indexed records via the Google Custom Search API, more than a month after the seizure. We were unable to replicate this manually searching via Google; however, the API is [documented](#) as returning a less-localized, more general search result. This finding still requires external replication, but is in line with research showing [de-indexing to constitute its own process](#) in the context of, e.g., the EU GDPR's 'right to be forgotten.' In the case of deepfake-specific laws or proposals to give individuals [copyright over their likeness](#), this finding demonstrates a renewed need to look upstream from just content that may be publicly googleable.

Policy Implications: Governing Upstream

Current platform governance is built primarily around circulation: what gets posted, ranked, amplified, or removed after the fact. This framework is well suited to harms that begin with something a user posted. It has little, however, to offer in solving a harm that affects someone who posted nothing at all, who has had their likeness exploited and co-opted. Our findings point to three upstream points where intervention is possible before a depiction ever exists.

- **Production tools:** require consent-based guardrails at the model and interface level, i.e., restrictions on what a system will generate, rather than relying solely on post-hoc content moderation by hosting platforms. Moreover, rapid adoption of protocols like C2PA, which aim to certify the origin of content, preventing its misuse, may blunt some degree of co-option of parliamentarians' public-facing content into raw material.
- **Hosting and indexing infrastructure:** Search-index records can outlive the content they point to. A domain being taken down does not guarantee a named individual stops appearing in search results for that content. Index persistence, therefore, should be treated as a distinct governance target from the actual hosting of content. Reputational damage, as well as psychological effects from reminders of past abuse, often persist long after the abuse itself has been halted.
- **Enforcement, not just statute:** Legislative trackers that record only whether a statute exists miss the variable that likely better predicts protection. Tracking prosecutions, takedown

response times, and enforcement resourcing would give policymakers a far more accurate picture than statutory presence alone. Furthermore, treating gateways and other ‘toolkits’ as distinct enforcement targets, equivalent to depictive content, would remove by our count at least 97 opportunities for self-assembled deepfake NCII (the total of gateways we found).

Ongoing Monitoring

This brief is the first release of a project funded by Agora Digitale Transformation and the Mercator Foundation, tracking exposure of European public officials in deepfake NCII. Upcoming releases will re-audit the same population using the same methodology, allowing us to observe how exposure moves in response to specific, dateable events (e.g. platform shutdowns, new legislation taking effect, enforcement actions) rather than relying on a single snapshot. The seizure of eight domains shortly before this report’s publication demonstrates the need for continuous monitoring, as well as a limitation of most current evidence in this space, including this baseline release: a one-time measurement cannot distinguish a policy intervention that reduces harm from one that merely displaces it to another platform or jurisdiction. Recent evidence suggests [displacement](#), not elimination, is the more common outcome following platform shutdowns. Subsequent releases in this series will report quarter-over-quarter change in exposure rates; the impact of any identifiable enforcement or platform events occurring within the quarter; and an updated assessment of which EU member states show measurable enforcement activity.

Contact Person at Agora Digitale Transformation



Benjamin Shultz

Data Fellow

benmshultz@gmail.com

agoradigital.de | Agora Digitale Transformation is funded by the Mercator Foundation.



Appendix

The full coding regime we used to classify hits from the Google Custom Search API, along with the count of MPs per category, can be seen here:

Code	State	Interpretation	MP Count
C	Content actively hosted	Depiction exists and is live	38
G	Gateway with active raw material	AI-generation tools linked directly next to raw material, on a page named for an individual	21
G-	Gateway with <i>no</i> raw material	AI-generation tools linked on a page named for an individual, without raw material present	36
S	Name-specific result, indexed and findable	Expressed demand, currently served with an active named page, raw material, or a request in a forum	54
S-	Name-specific result, indexed, <i>not</i> findable	Records persist, remaining indexed, while content may have been removed, deleted, or never existed	135
R	Domain removed	Still-indexed records of known origin; e.g. domain seizure	47

We then ran two Firth-corrected logistic regressions testing the following variables: gender, institutional prominence, name-perceived migration background, the presence of deepfake-specific legislation in a parliamentarian's country, political lean (left and right tested against a centrist reference group), and age. The second of our two models added an interaction term between gender and institutional prominence to test whether the effect of gender on a parliamentarian's probability of being depicted depends on the extent of their institutional prominence. The interaction term was insignificant, indicating that the effects of gender and institutional prominence do not depend on one another. Standard errors were clustered by country.

```

=====
Cluster-robust - Model A
=====

```

	coef	OR	OR_lo_rob	OR_hi_rob	SE_rob	p_orig	p_rob	note
female	3.495	32.955	14.712	73.822	0.411	1.102460e-53	0.000	
high_visibility	2.257	9.558	6.283	14.540	0.214	1.851130e-23	0.000	
migration_background	0.249	1.282	0.530	3.105	0.451	5.095630e-01	0.582	
deepfake_law	1.456	4.291	2.026	9.087	0.383	1.300420e-12	0.000	
ideo_Left	0.773	2.165	1.369	3.424	0.234	3.325720e-03	0.001	
ideo_Right	0.559	1.748	1.111	2.752	0.231	3.780020e-02	0.016	
age	-0.027	0.973	0.953	0.995	0.011	1.941530e-03	0.014	
Intercept	-6.164	0.002	0.001	0.008	0.704	1.150280e-31	0.000	

```

=====
Cluster-robust - Model B
=====

```

	coef	OR	OR_lo_rob	OR_hi_rob	SE_rob	p_orig	p_rob	note
female	3.666	39.090	12.376	123.463	0.587	2.909930e-36	0.000	
high_visibility	2.671	14.454	3.819	54.709	0.679	1.146100e-04	0.000	
migration_background	0.245	1.277	0.530	3.080	0.449	5.151160e-01	0.586	
deepfake_law	1.451	4.269	2.035	8.955	0.378	1.434770e-12	0.000	
female_x_visibility	-0.457	0.633	0.152	2.636	0.728	4.852130e-01	0.530	
ideo_Left	0.769	2.157	1.365	3.409	0.234	3.439990e-03	0.001	
ideo_Right	0.556	1.744	1.110	2.741	0.231	3.841650e-02	0.016	
age	-0.027	0.973	0.953	0.994	0.011	1.863330e-03	0.013	
Intercept	-6.315	0.002	0.000	0.007	0.679	9.989450e-30	0.000	