



INVESTOR ALLIANCE
FOR HUMAN RIGHTS
AN INITIATIVE OF ICCR

Webinar: Disinformation, Deepfakes, and Distrust: Tech companies' responsibility in restoring election integrity

13 February 2024

11 am – 12 pm ET

Panelists



- **Alexandra Pardal** is the Interim Co-Executive Director of Digital Action, leading campaigns, programmes, partnerships, and fundraising. She has over 24 years' experience in international non-profits and politics, leading campaigns, investigations, and policy reforms. Her work has generated positive impacts including reforms to a billion-euro EU trade and development programme and financial restitution to communities in Africa and been reported on by international media outlets like The Times, The Guardian, Le Monde, El Pais and Süddeutsche Zeitung.
- **Tim Harper**, Senior Policy Analyst, Democracy and Elections at Center for Democracy and Technology leads the organization's efforts to fight election disinformation, support technology that bolsters a fair and secure vote, and build trust in elections. He most recently served as Content Policy Associate Manager, Political Advertising at Meta, where he developed and implemented electoral, political, and social issues advertising policies for Meta's products and services globally. He focused on content policies to ensure advertising transparency, protect elections from foreign interference, and prevent election delegitimization around the world.
- **Jeff Allen** is the Co-founder and Chief Research Officer for the Integrity Institute. He has worked on multiple sides of the internet information ecosystem: on publishers who are trying to maximize the traffic they get from platforms, on platforms themselves, and on political organizations and campaigns just trying to navigate the online spaces. While at Facebook, he worked on tackling systemic issues in the public content ecosystems of Facebook and Instagram, developing strategies to ensure that the incentive structure that the platforms created for publishers was in alignment with Facebook's company mission statement.
- **Lauren Compere** is the Managing Director and Head of Stewardship & Engagement at Boston Common Asset Management. Lauren leads Boston Common's global shareholder engagement and stewardship program. Lauren sits on the board of the Global Network Initiative (GNI). She serves on the Advisory Council for the Investor Alliance for Human Rights, an ICCR project, and is the Co-Chair for the International Corporate Governance Network (ICGN) Natural Capital Committee.



Global Coalition for Tech Justice and the Year of Democracy 2024

#YearOfDemocracy



2024 elections

When it comes to the widespread use and impact of social media, 2024 will be the make-or-break year for democracy globally

- In 2024, the 'Year of Democracy', +65 elections take place as US Presidential and European Parliamentary elections converge with polls in India, South Africa, Mexico and some 50 other countries.
- There is a **global equity crisis at the centre of tech accountability. Companies are barely investing in harm assessment and mitigation** outside the US and English language content. Yet as the companies' influence on information ecosystems in Africa, Asia and Latin America grows, their **impacts are becoming more serious, more extensive, and more frequent.**
- 2024 offers a once-in-a-generation opportunity in addressing tech harms across the global majority, as well as to enhance the connection between ad tech design, democratic rights, and the prevention of violence.
- The [Global Coalition for Tech Justice](#) has over 200 member organisations and experts in over 55 countries. Our initiative – 'Protect People and Elections, Not Big Tech' – aims to catalyse a global movement - from civil society to advertisers, shareholders, and policymakers - leveraging their respective power, influence and voice in favour of equitable and effective platform safeguards to protect democratic and human rights during the 2024 global election cycle.



Social media impacts in Latin America

Brazil congress attack: what we know so far

President Lula denounces 'fanatical fascists' and announces federal security intervention after Bolsonaro supporters storm political buildings in capital

● **Full report: Bolsonaro supporters storm congress, presidential palace and supreme court**



📹 Supporters of Brazil's former president Jair Bolsonaro on the roof of the national congress building in Brasilia after storming it on Sunday. Photograph: Eraldo Peres/AP

Misinformation spreads on social media in Chile ahead of vote on new constitution

Support is dropping amid a proliferation of "half truths" on social media about what's in the document, and voters are expected to reject it, say experts.



ARGENTINA > OPINION 1

Power of social media on display in Argentina's elections

The unexpected success and popularity of Javier Milei's political party sparked immediate comparisons with Donald Trump and Jair Bolsonaro, prompting questions about the digital landscape



Presidential candidate Javier Milei holds a press conference in Buenos Aires; October 2023. AGUSTIN MARCIAN (REUTERS)



Social media impacts in the Middle East and North Africa



Tunisia

"The publishing of my personal information on social media platforms in a dangerous and improper manner by the government and others led to real-world threats in my own home"

"Social media companies must safeguard all minorities and individuals who critique the government and other regimes. If they fail to take resolute and tangible actions to stop this situation, and I firmly believe this is not a mere farce, they will become complicit in this violence."

Rania Amdouni, Tunisian human rights activist



Social media impacts in Africa

Unmoderated hate speech leading to real-world violence



PRESS RELEASE | JULY 28, 2022

Facebook approves ads calling for ethnic violence in the lead up to a tense Kenyan election



ARTICLE | JUNE 20, 2023

“We need to kill them”: Xenophobic hate speech approved by Facebook, TikTok and YouTube

October 31, 2023, Index Number: AFR 25/7292/2023

Ethiopia: ‘A death sentence for my father’: Meta’s contribution to human rights abuses in northern Ethiopia

In November 2020, a brutal conflict broke out in northern Ethiopia. In 2021, whistle-blower Frances Haugen said that the Facebook platform was “literally fanning ethnic violence” in places like Ethiopia. This report is based on an investigation into Meta’s role in the serious human rights abuses perpetrated against the Tigrayan community between 2020 and 2022. It reveals the devastating impacts that the Facebook platform’s surveillance-based business model had in the context of Ethiopia’s armed conflict. It shows that, despite the company’s claim to have learned the lessons of its contribution to the atrocities against the Rohingya in 2017, many of the same systemic failures occurred again in Ethiopia.



Facebook failing to check hate speech, fake news in India: Report

Leaked papers show company is selective in curbing hate speech, misinformation and inflammatory posts, particularly anti-Muslim content.



Zuckerberg, right, hugs Modi at Facebook headquarters in Menlo Park, California [File: Jeff Chiu/AP]

Social media impacts in Asia

Pakistan:

'We're seeing a rise in harmful content online particularly with respect to gendered disinformation with the use of false information manipulating the masses and the use of sexist and misogynistic narratives that are perpetuating harms. These attacks are well coordinated and planned and seem to be adding to offline harms of individuals, particularly women human rights defenders and journalists.'

Nighat Dad, Executive Director, Digital Rights Foundation

Indonesia

"I can see an increase in hate speech and misinformation related to the election, especially in local dialects on social media platforms like Facebook groups, TikTok, X, and YouTube" Nuurrianti Jalli of Oklahoma State University.

Rohingya

The online hate campaign turning Indonesians against Rohingya refugees

As Indonesia's presidential election approaches, a campaign of disinformation against Rohingya is spreading across social media



Ethnic Rohingya disembark from their boat upon landing in north Aceh, Indonesia. Photograph: Rahmat Mirza/AP



Systemic platform failures with cuts to trust + safety, inadequate safeguards, disclosures, oversight and accountability across global majority regions

- Inadequate investments in trust + safety, in localised policies, content moderation and enforcement of violations of platform policies
- Algorithms favour engagement over harm mitigation
- Lack of investment in human rights risk assessment and mitigation
- Large discrepancies in tools and safeguards outside the United States
- Poor transparency and disclosures
- Politicians largely get a free pass with disinformation and hate speech
- Platform monitoring by journalists, researchers and civil society watchdogs increasingly restricted by platforms
- Poor to no accountability for harms



What investors can do

Request full disclosure of risk assessment and mitigation relating to 2024 elections in markets where they have users, including responses to [10-point asks](#):

- Safeguards for compliance with international human rights and electoral standards, human rights impact assessments
- Disclosures of trust + safety spending globally, regionally and nationally, including number of content moderators per language/dialect/region in Africa, the Middle East, Asia and Latin America.
- Full spectrum of harm mitigation tools and measures for all elections, not just the US, operational at each stage of electoral process including post-ballot where risks of post-election violence loom
- Increase local expertise and stakeholder engagement.
- Resourcing of partnerships with fact-checkers, independent media, civil society and other bodies that protect electoral integrity.
- Publishing all contacts with governments, and all government requests to suppress speech and surveillance demands, where permitted by law. Eliminating policy exemptions for politicians and ensuring fact-checking of political ads. Transparency of political advertising for oversight and legal compliance..
- Evidence they enable public oversight and transparency, including data access and training for researchers, civil society, independent media and election monitors.



www.yearofdemocracy.org

[Digital Action](#), *convenor and organizer of*
Global Coalition for Tech Justice

info@digitalaction.co

Generative AI and 2024 Elections: Risks and Policy Recommendations

February 13, 2024

Deepfakes in the News



politics

SCOTUS

Congress

Facts First

2024 Elections

DeSantis campaign posts fake images of Trump hugging Fauci in social media video

The Washington Post
Democracy Dies in Darkness

Fake Biden robocalls urge Democrats not to vote in New Hampshire primary

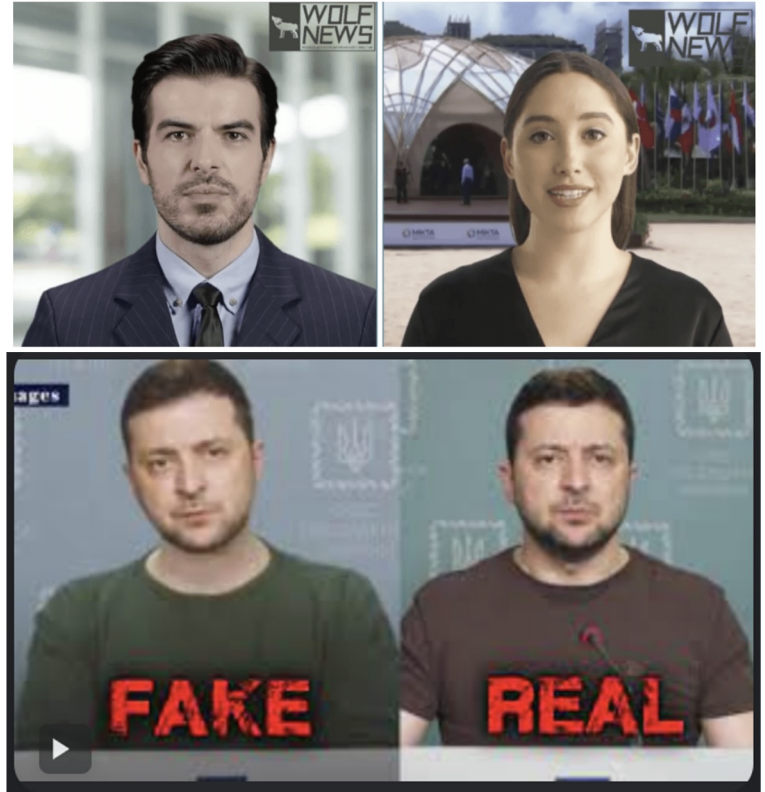
State's attorney general announced it is investigating several complaints of an AI-generated voice impersonating President Biden

AI Risks: Deepfakes

What is a deepfake?

An image, video or audio recording that has been convincingly altered and manipulated to misrepresent someone as doing or saying something that was not actually done or said

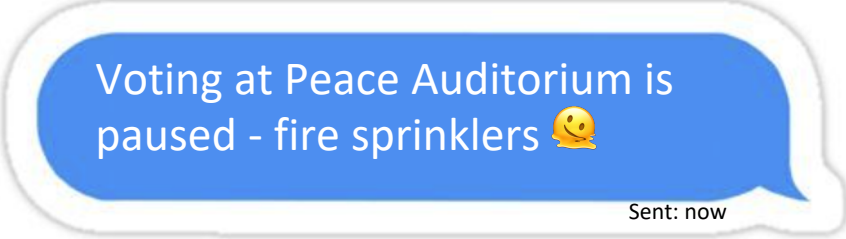
Can you Spot the deepfake?



AI Risks: Hyperlocal Misinfo

Hyper-targeted, personalized messaging

Generative AI facilitates the generation and spread of hyperlocal misinfo at scale. Take the example of polling conditions at individual precincts on Election Day. Bad actors could use publicly available precinct location data and phone numbers to distribute AI-generated content with highly specific claims about named polling places.



Voting at Peace Auditorium is paused - fire sprinklers 🤔

Sent: now

AI Risks: Language-Minority Influence



Study Suggests We Should Worry About Political Microtargeting Powered by Generative AI

TIM BERNARD / FEB 7, 2024

Ang daming krimen at sabi bila, bumoto. Bakit?



There is a lot of crime and they say vote. Why?

AI Risks: Cybersecurity



**Election
Processes**



**Election
Offices**



**Election
Officials**



**Election
Vendors**

Call to Action

We can't admire the problem (or even wait for legislation), we need concrete action now. And we'll focus on key stakeholders - the AI developers, distribution platforms, political campaigns, and election officials.

HOW

Action With Accountability
=
Civil Society Engagement
+
Oversight and Transparency

WHO

Four Key Stakeholders
=
Model Developers
+
Social Media Platforms
+
Politicians and Campaigns
+
Election Officials

CONTEXT

Platform Readiness
+
Partisan Court Cases and Investigations
+
Widening Distribution Channels

Policy Recommendations - AI Developers

- Watermarking and Content Provenance
- Content rules against generating harmful text/audio/video/images
- Restrictions on use for personalized persuasion in campaigns
- Product interventions on election-related queries
- Promote authoritative information in answers and interface
- Label potential misinformation and provide context
- Disclosures and researcher access
- Know Your Customer requirements



Policy Recommendations - Social Media

- New social media platforms (Twitch, Discord) need to rapidly scale up Trust and Safety commitments
- AI content - Usage policies, labeling, content credentialing
- AI-specific Policies aren't enough - Elections policies are needed
- Renew investment in Trust and Safety
- Support authoritative information and election officials safety and communication



Policy Recommendations - Candidates/Campaigns

- Pledges against use of deepfakes and manipulated media
- Utilize cybersecurity tools to prevent leaks and hacks
- Adopt authentication techniques like watermarks on official statements
- There's a key role for professional associations like the American Association of Political Consultants (AAPC)



Policy Recommendations - Election Officials

- Cybersecurity measures - transition to .gov domains; train up on phishing; adopt content authentication in ballot samples and press releases; authenticate authoritative info and official docs; implement technical controls
- Build relationships with local media and platforms to quickly disseminate accurate information
- Utilize crisis communications playbooks to debunk mis/disinfo
- Utilize social media account tools



THANK YOU | [CDT.ORG](https://cdt.org)

Integrity Institute

Elections Program 2024



About the Integrity Institute

Part professional community. Part think tank.

Our community is made up of “integrity professionals”, people who have fought the harms that platforms are causing to people, societies, and democracies.

We have 350+ members with experience at 60+ online platform companies

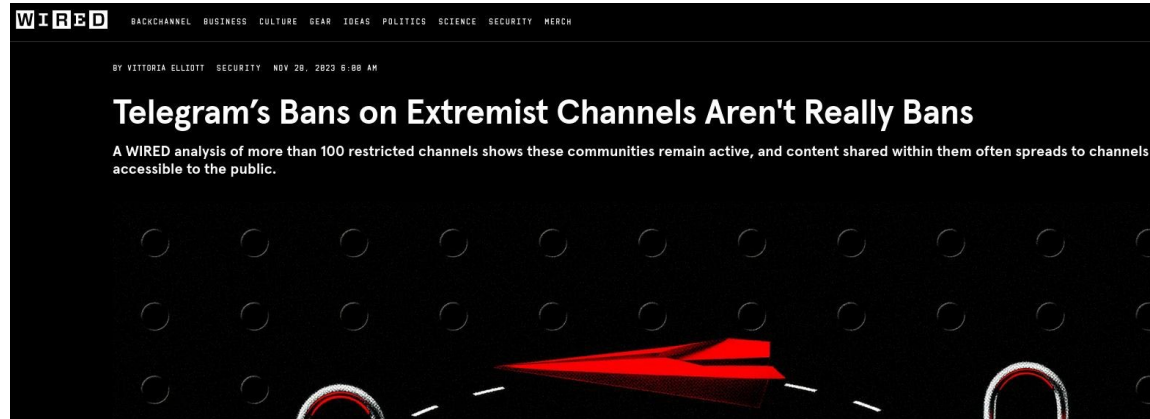
We bring that expertise and experience to the stakeholders in the social internet

- ▷ Social internet companies
- ▷ Civil society and advocacy orgs
- ▷ Policymakers



General Industry Trends We See

- ▶ **The tech layoffs will have an impact**
 - We've seen layoffs from the major companies
 - This has impacted policy teams, user researchers, and partnership managers
- ▶ **More diversity in platforms**
 - TikTok much bigger than in 2020
 - Mid size platforms, Discord, Twitch, are playing a bigger role
 - They don't have quite as much experience as Meta, Google
- ▶ **Telegram is a major platform in many countries around the world and doesn't do integrity**
 - It doesn't care about the impact it has on democracies



Our Elections Work

- ▷ We want to see platforms
 - ▷ Follow best practices in protecting elections
 - ▷ Partnering with civil society
 - ▷ Partnering with researchers
- ▷ We work towards this by
 - ▷ Working with our community to establish best practices
 - ▷ Connecting our community to external organizations and partners
 - ▷ Conducting original research to track platforms performance



Best Practices in Elections

- ▶ Last year, we prepared two briefings broadly covering election work



- ▶ This year: narrower topics, deeper scope



DSA Guidance Acting as Best Practices

- ▷ Regulatory pressure to follow best practices is coming!
- ▷ The Digital Services Act (EU law) requires that very large platforms adhere to [their guidance](#) on ensuring their platforms are not harming the democratic process
- ▷ Asks platforms to follow industry standards, including those established by the Integrity

Institute

PRESS RELEASE | Publication 08 February 2024

Commission is gathering views on draft DSA guidelines for election integrity

Today, the Commission has launched a public consultation to seek views on draft Digital Services Act (DSA) guidelines on the integrity of electoral processes.

These are the first ever guidelines under Article 35 of the DSA and aim to present Very Large Online Platforms and Search Engines with best practices and possible measures to mitigate systemic risks on their platforms that may threaten the integrity of democratic electoral processes.

The draft guidelines provide examples of potential mitigation measures relating to election-related risks, specific mitigation measures linked to Generative AI content, the planning of risk mitigation measures before or after an electoral event and specific guidance for the European Parliament elections.

Margrethe Vestager, Executive Vice-President for a Europe Fit for the Digital Age, said:

In 2024, we have elections at local, regional, national, and European levels. For this to go well we need have to public discussions – and a lot of those are happening online. With the Digital Services Act, we have concrete tools to work together with online platforms. We can address the emerging online risks to electoral processes, like deep fakes. So we can enable people, in a safe way, to engage, discuss and make up their minds without illegal interference.



iStock photo Getty images plus

Related topics

[Media](#) [Online platforms and e-commerce](#)

[Disinformation](#) [Digital Services Act Package](#)



DSA Guidance Acting as Best Practices

- ▷ Regulatory pressure to follow best practices is coming!
- ▷ The Digital Services Act (EU law) requires that very large platforms adhere to their guidance on ensuring their platforms are not harming the democratic process
- ▷ Asks platforms to follow industry standards, including those established by the Integrity

Institute

PRESS RELEASE | Publication 08 February 2024

Commission is gathering views on draft DSA guidelines for election integrity

Today, the Commission has launched a public consultation on draft Digital Services Act (DSA) guidelines on

These are the first ever guidelines under Article 3 of the Digital Services Act (DSA) on Platforms and Search Engines with best practices on their platforms that may threaten the integrity of

The draft guidelines provide examples of potential risks, specific mitigation measures linked to General Provisions and measures before or after an electoral event and special elections.

Margrethe Vestager, Executive Vice-President for

(15) Mitigation measures could draw, in particular, on industry standards established through the Code of Practice on Disinformation, other relevant EU industry codes, such as the Code of conduct on countering hate speech online, and from existing best practices such as those documented in the Content-Agnostic Election Integrity Framework for Online Platforms ⁽¹³⁾ and the Election Integrity Programme of the Integrity Institute ⁽¹⁴⁾ as well as recommendations from civil society, such as those from the Civil Liberties Union for Europe and European Partnership for Democracy ⁽¹⁵⁾.

In 2024, we have elections at local, regional, national, and European levels. For this to go well we need have to public discussions – and a lot of those are happening online. With the Digital Services Act, we have concrete tools to work together with online platforms. We can address the emerging online risks to electoral processes, like deep fakes. So we can enable people, in a safe way, to engage, discuss and make up their minds without illegal interference.

Related topics

[Media](#) [Online platforms and e-commerce](#)

[Disinformation](#) [Digital Services Act Package](#)



DSA Guidance Acting as Best Practices

- ▷ Regulatory pressure to follow best practices is coming!
- ▷ Few key things to highlight from draft guidance
 - ▷ “Analysing and appropriately moderating virality”: EU interested in tackling the problem of virality. Platform design heavily impacts virality.
 - ▷ “Demonetisation of disinformation content”: EU will be paying attention to the content that is monetized on the platform (Creator funds, partnership programs, etc). This is a key area to pay attention to, platforms funding disinformation is high risk situation.
 - ▷ “Third party scrutiny and research”: EU expecting companies to work with researchers and external experts
 - ▷ “Put in place an incident response mechanism”: EU expecting robust operations inside companies to both prepare (including red teaming) and respond to rapidly changing events.

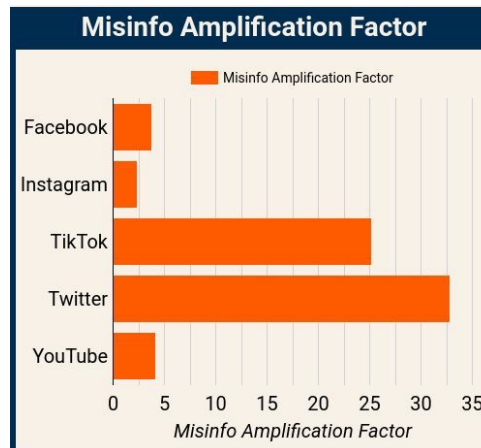
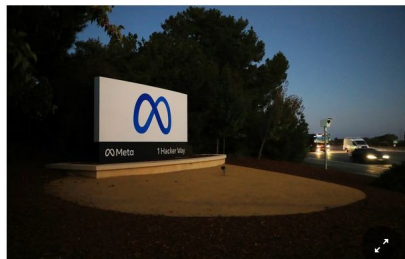


Original Research

- ▷ Misinformation Amplification Analysis and Monitoring
- ▷ In 2022, told the story “The design choices that platforms make strongly impact how much they amplify misinformation”
- ▷ This year, want to explore, “As bad as things are in the US, how bad do they get abroad?”
 - ▷ Study how much the platforms amplify misinformation across multiple countries
 - ▷ Will we see that misinfo is amplified more in countries that don’t get as much attention from the platforms?
 - ▷ Or in countries with more fragile news ecosystems?
- ▷ With the right partner, we might be able to go beyond misinformation. Hate speech amplification analysis?

How Social Media Amplifies Misinformation More Than Information

A new analysis found that algorithms and some features of social media sites help false posts go viral.



Summary

- ▷ 2024 will be different from 2020
 - ▷ Wave of elections globally
 - ▷ Industry trends: New platforms, layoffs
 - ▷ Regulation is coming
- ▷ Integrity Institute has many community members who have deep experience protecting elections on social media platforms
- ▷ We want to help platforms play a more positive role around the world
- ▷ Do this by
 - ▷ Giving advice, guidance, and best practices from our experienced members
 - ▷ Connecting our community to other orgs protecting elections
 - ▷ Tracking how platforms are performing and connecting that to company decisions



Some resources for reference

- [Seismic Shifts: How Economic, Technological, and Political Trends are Challenging Independent Counter-Election-Disinformation Initiatives in the United States](#), Center for Democracy and Technology
- [2024: Year of Democracy Inception Phase Report](#), Digital Action
- [Campaign Asks: A Call for “Big Tech Action Plans”](#), Digital Action
- [Elections Integrity Best Practices: Responsibly Supporting Elections on Online Platforms](#), Integrity Institute
- [Defining and Achieving Success in Elections Integrity](#), Integrity Institute
- [Study Suggests We Should Worry About Political Microtargeting Powered by Generative AI](#), Tech Policy Press
- [Commission is gathering views on draft DSA guidelines for election integrity](#), European Commission
- [Investors Say Tech Companies are Failing to Address Systemic Human Rights Risks Inherent in Business Models and Exacerbated by AI](#), Investor Alliance for Human Rights



QUESTIONS

Investor Alliance for Human Rights

A collective action platform connecting institutional investors with tools and strategies to promote corporate respect for human rights.

Join Us

Sign up to become a member: <https://bit.ly/2IOYDA9>

For more information, contact: [Anita Dorett \(adorett@iccr.org\)](mailto:adorett@iccr.org)