Committee: Youth Assembly on Topic 1 (YA1)

Topic: Discussing the use of AI-generated content on social media

Student Officer: Eva Andreopoulos, Nikitas Stamatopoulos

Position: Co-Heads

Honourable delegates,

My name is Nikitas Stamatopoulos, and I will be serving as a Co-Head in the 13th CSMUN Youth Assembly on Topic 1. I am a tenth-grade student attending the German School of Athens, and it is my honour to have this position.

Firstly, I want to say that joining MUN may seem unfamiliar at first, but that's natural when trying something new. Secondly, I would like to remind you to enjoy this experience, as MUN, other than being a school club, is also a place for you to meet people, engage with them and learn how to confidently speak in front of an audience. Lastly, I want to welcome you to this committee, encourage you to engage with this study guide, but also conduct your own research on this topic, which is of great importance. Through this committee, you will practice problem-solving skills, learn how to speak in front of a crowd, and respond to situations using proper etiquette.

I truly hope that you will all have an amazing experience in the CSMUN conference and come to love MUN as a whole. I also urge you to ask me questions regarding the topic and the guide via my email, which is: stamatopoulosnikitas@gmail.com. If you have any other questions, please contact the Head of the Youth Assembly on Topic 1, Rea Karvouni, at rea.karvouni@gmail.com.

Sincerely,

Nikitas

Dear Delegates,

My name is Eva Andreopoulos, I am a Year 12 student at Campion School, and I am honoured to serve as a Co-Head for this committee! I'm beyond excited to welcome you to what I know will be an educational and highly rewarding conference. Being a part of this committee means a lot to me - not only because of the incredible debates and teamwork that will be fostered in our Youth Assembly, but because of the opportunities given to each delegate to develop confidence, critical thinking, and global awareness. Our topic of AI in social media is a highly relevant issue in our current age, affecting everyone, so I expect to see collaboration between all of you to produce thorough resolutions and have fruitful discussions on our topic. Our study guide is here to guide and aid you, but it is extremely important that you conduct your own research, too.

Whether you're new to MUN or a seasoned delegate, in this committee, my fellow Student Officers and I will be here to support you in any way possible. You can always approach me with any questions, concerns, or even suggestions you have. My goal is to make sure you feel welcomed, prepared, and ready to participate fullheartedly in our committee! I can't wait to see the energy you'll bring, the solutions you'll propose, and the diplomacy you'll demonstrate. Let's make this a memorable experience for everyone involved!

If you have any questions regarding the topic or study guide, my email is eandreopoulos@campion.edu.gr. If you have any other questions, please contact the Head of the Youth Assembly on Topic 1, Rea Karvouni, at rea.karvouni@gmail.com.

Sincerely,

Eva

Topic Introduction

Artificial intelligence (AI) is increasingly shaping the creation and distribution of content on social media platforms.[1] With the rapid growth of generative AI technologies[2], content can now be produced by individuals, corporations, or automated systems, and often without clear disclosure of its origins, which means that the content is often indistinguishable from human-created content.[3] This AI-generated content includes realistic deepfake videos, AI-generated images, synthetic voices, and text-based content such as captions or articles.[4]

A common example includes videos falsely depicting political figures making statements that influence public perception or affect political events.[5] On a personal level, AI can be used to create fabricated content intended to damage someone's reputation, relationships, or career[6], such as through AI-generated content of the person committing actions they otherwise wouldn't do. These developments raise serious concerns around misinformation, manipulation, and consent in digital spaces, for all of society worldwide, both on and off the internet.

This issue directly relates to this year's Conference Theme, *Free Will in the Age of Artificial Intelligence*, as the growing presence of AI-generated media on social platforms risks undermining our ability to make informed decisions and express original views. Manipulated content can distort reality and erode trust in public discourse, reducing the space for authentic free speech.

---

[1] "The Impact of Artificial Intelligence on Social Media Content", "ResearchGate 2024"
https://www.researchgate.net/publication/380000781_The_Impact_of_Artificial_Intelligence_on_Social_Media_Content
[2] "The State of AI in 2023", "McKinsey, 2023"
https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-in-2023-generative-ais-breakout-year
[3] "Generative AI and Copyright", "European Parliament, 2025"
https://www.europarl.europa.eu/RegData/etudes/STUD/2025/774095/IUST_STU(2025)774095_EN.pdf
[4] "Children and Deepfakes", "European Parliament, 2025"
https://www.europarl.europa.eu/RegData/etudes/BRIE/2025/775855/EPRS_BRI(2025)775855_EN.pdf
[5] "Tiktok users being fed misleading election news, BBC finds", "BBC, 2024"
https://www.bbc.com/news/articles/c1ww6vz1l81o
[6] "Media Manipulation and Disinformation: The role of AI in fake news creation", "ResearchGate, 2025"
https://www.researchgate.net/publication/390040541_Media_Manipulation_and_Disinformation_The_Role_of_AI_in_Fake_News_Creation_and_Its_Impact_on_Trust_in_Media

This is made more concerning with synthetic media expanding at a rapid pace, driven by major technology firms such as Meta, TikTok, OpenAI, and Google. These companies are developing increasingly advanced tools with limited regulatory oversight.[7]

Currently, there are no unified global standards to ensure transparency or ethical use of AI-generated content[8] on social media. Without such measures, misinformation, copyright infringement, identity theft, and job displacement become risks that can catalyse further issues, such as economic disruption, reduced literacy rates, lawsuits, and more.

This issue is especially relevant to youth, who are typically early adopters of new technologies and highly active on social platforms. AI is already impacting areas central to young people's lives, including education, communication, and creative industries.[9]

According to Pew Research (2024), over 60% of global social media users have encountered AI-generated content, often unknowingly.[10]

AI's seemingly quiet integration into our lives raises the pressing concern of how we balance innovation with safety and authenticity, whilst maintaining the level of ethicality required.

Definition of key concepts

AI-generated content

"Content created entirely or partially by artificial intelligence systems without direct human authorship, including text, images, video, or audio."[11]

---

[7] "Why and how is the power of Big Tech Increasing in the Policy Process?", "Oxford Academic, 2024"
https://academic.oup.com/policyandsociety/article/44/1/52/7636223
[8] "Developing global standards for AI accountability", "ResearchGate, 2023"
https://www.researchgate.net/publication/387090664_DEVELOPING_GLOBAL_STANDARDS_FOR_AI_ACCOUNTABILITY_IN_CONTENT_GENERATION
[9] "Artificial Intelligence and the Future of Teaching and Learning", "US Education, 2023"
https://www.ed.gov/sites/ed/files/documents/ai-report/ai-report.pdf
[10] "What web browsing data tells us about how AI appears online", "Pew Research Centre, 2025"
https://www.pewresearch.org/data-labs/2025/05/23/what-web-browsing-data-tells-us-about-how-ai-appears-online/
[11] "What is AI Generated Content?", "IBM, 2024"
https://www.ibm.com/think/insights/ai-generated-content

Deepfake

"A video of a person in which their face or body has been digitally altered so that they appear to be someone else, typically used maliciously or to spread false information."[12]

Algorithmic amplification

"The process by which social media algorithms prioritise and promote certain content, leading to greater visibility and influence."[13]

Generative AI tools

"Generative AI tools are applications powered by machine learning models that can produce new content- text, images, music, or video based on user prompts."[14]

Misinformation

"False or inaccurate information that is spread without the intent to deceive- often due to misunderstanding or error."[15]

Disinformation

"Deliberately false or manipulated information created and shared with the intention of deceiving or harming individuals, groups, or institutions."[16]

---

[12] on, Insights. "The Cyber Helpline." The Cyber Helpline, 19 Apr. 2024, www.thecyberhelpline.com/helpline-blog/2024/4/19/understanding-deepfakes-insights-on-detection-and-prevention.

[13] "Algorithmic Amplification", "Sustainable Directory, 2025" https://climate.sustainability-directory.com/term/algorithmic-amplification/#:~:text=Algorithmic%20Amplification%20can%20be%20defined,belief%20systems%2C%20and%20potentially%20impacting

[14] "What is Generative Ai?", "Coursera, 2024" https://www.coursera.org/articles/what-is-generative-ai?utm_medium=sem&utm_source=gg&utm_campaign=b2c_emea_x_multi_ftcof_career-academy_

[15] American Psychological Association. "Misinformation and Disinformation." American Psychological Association, 2025, www.apa.org/topics/journalism-facts/misinformation-disinformation.

[16] Cambridge Dictionary. "DISINFORMATION | Meaning in the Cambridge English Dictionary." Dictionary.cambridge.org, 2019, dictionary.cambridge.org/dictionary/english/disinformation.

Synthetic Media

"Media content that has been algorithmically generated, modified, or synthesised by AI technologies, such as deepfakes or virtual humans."[17]

Digital Literacy

"The ability to find, evaluate, and communicate information using digital platforms and technologies responsibly."[18]

Ethical AI Use

"The development and application of artificial intelligence in a manner that aligns with principles of transparency, fairness, accountability, and respect for human rights."[19]

Background Information

Historical Background

While generative AI has only recently gained mainstream attention, its origins date back seventy years ago.[20] Early examples include computer-generated speech systems developed in the 1960s[21], such as the first rudimentary vocal synthesisers - and self-sustaining algorithms created in early machine learning research.[22] AI was meant to mimic human characteristics such as problem-solving and reasoning, but still be controlled by humans.[23] However, the current wave of generative AI emerged more prominently after 2022, with the public release of tools like ChatGPT, DALL-E,

---

[17] "What is Synthetic Media", "Synesthesia Glossary, 2025"
https://www.synthesia.io/glossary/synthetic-media
[18] "Digital Literacy Definition", "UNESCO, 2022"
https://uis.unesco.org/en/glossary-term/digital-literacy
[19] "What is Ethical AI Use?", "SAP, 2024"
https://www.sap.com/resources/what-is-ai-ethics
[20] "History of AI", "Tableau, 2024"
https://www.tableau.com/data-insights/ai/history
[21] "A brief history of generative AI", "Dataversity, 2024"
https://www.dataversity.net/a-brief-history-of-generative-ai/
[22] "A Systemic Review of AI based music generation", "ScienceDirect, 2022"
https://www.sciencedirect.com/science/article/pii/S0957417422013537
[23] A Systemic Review of AI based music generation", "ScienceDirect, 2022"
https://www.sciencedirect.com/science/article/pii/S0957417422013537

Midjourney, and other advanced models capable of producing text, visuals, audio, and video with minimal human input.[24]

Since then, the use of generative AI in social media has grown rapidly. From AI-generated TikTok videos to automated captions, fake news articles, and deepfake images, the presence of synthetic media on popular platforms has expanded dramatically.[25] According to Grand View Research, the generative AI market is projected to exceed $1811 billion USD by 2030[26], signalling its growing influence across industries, especially in digital media and communication.

Youth Engagement and Accessibility

Young people play a central role in both consuming and producing AI-generated content. According to Statista, 35% of Gen Z social media users report using AI tools, such as ChatGPT or image generators, to create posts, videos, or other online materials.[27] This generation is highly digitally literate and quick to adopt new technologies. As AI tools become more user-friendly, free, or built directly into social media platforms, their accessibility continues to increase, particularly for youth.[28]

Platforms like TikTok, Instagram, YouTube, and Snapchat have integrated AI-based features such as automatic captioning, video editing, AI avatars, and recommendation algorithms.[29] Users are also exposed to synthetic content in the form of AI-written comments, artificially generated photographs, and videos of public figures saying things they never actually said. Furthermore, the Future of Privacy Forum reported that only 14% of AI-generated videos online are clearly labelled as synthetic - even on platforms that claim to require disclosure.[30]

---

[24] "What is generative AI?", "McKinsey and Company, 2024"
https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-generative-ai
[25] "The Spread of Synthetic Media on X", "Misinformation Review, 2024"

https://misinforeview.hks.harvard.edu/article/the-spread-of-synthetic-media-on-x/
[26] "Artificial Intelligence Market Summary", "Grand View Research, 2024"
https://www.grandviewresearch.com/industry-analysis/artificial-intelligence-ai-market
[27] "US teens using AI", "Statista, 2024"
https://www.statista.com/statistics/1500204/us-teens-generative-ai-tools-usage-frequency/
[28] "The AI Generation, Youth in the Artificial Intelligence Era", "United Nations, 2025"
https://unric.org/en/the-ai-generation-youth-in-the-artificial-intelligence-era/
[29] "How AI is changing short video apps like tiktok", "OyeLabs, 2025"
https://oyelabs.com/how-ai-content-is-changing-short-video-app-like-tiktok/
[30] "Synthetic Content: Exploring the risks", "Future of Privacy Forum, 2024"
https://fpf.org/wp-content/uploads/2024/10/Synthetic_Content_Report_October_2024.pdf

Logically, young people are also most susceptible to the risks and negative impacts of AI in social media as well. AI can rapidly generate convincing yet entirely false narratives, not only around elections, health, or global conflicts, but also on a personal level. Synthetic media can be weaponised to harass, shame, or impersonate individuals, often anonymously and at scale.[31] AI systems trained on biased data may also reinforce harmful stereotypes, especially in content recommendation algorithms. For example, studies have shown that AI image tools often overrepresent certain gender and racial stereotypes in generated images.[32] AI is also prone to making mistakes and handing out false information[33], which, if we are not aware of, can be harmful. The OECD Global Youth Survey (2023) found that 47% of youth aged 15-24 struggle to distinguish between real and AI-generated content online.[34]

Case Studies

Deepfake Political Video

The implications of AI-generated content are already being shown in real-world scenarios. Two days before Slovakia's parliamentary election (held 30 September 2023), an AI-manipulated audio clip circulated on social media purporting to record Progressive Slovakia leader Michal Šimečka and a journalist discussing election-rigging.[35] The clip spread during Slovakia's legally mandated pre-election "silence" period and was difficult to debunk quickly.[36] Fact-checkers and platform teams concluded the audio showed signs of AI manipulation, but the timing, format, and existing political

---

[31] "Increasing Threat of deepfake identities", "Department of Homeland security, 2023"
https://www.dhs.gov/sites/default/files/publications/increasing_threats_of_deepfake_identities_0.pdf
[32] "The Complex World of AI Failures", "Univio, 2024"
https://www.univio.com/blog/the-complex-world-of-ai-failures-when-artificial-intelligence-goes-terribly-wrong/
[33] "The Complex World of AI Failures", "Univio, 2024"
https://www.univio.com/blog/the-complex-world-of-ai-failures-when-artificial-intelligence-goes-terribly-wrong/
[34] "How's Life for children in the digital age?", "OECD, 2025"
https://www.oecd.org/en/publications/how-s-life-for-children-in-the-digital-age_0854b900-en/full-report/the-impact-of-digital-activities-on-children-s-lives_4df70664.html
[35] "Election Disinformation in Slovakia", "EU, 2023"
https://ec.europa.eu/newsroom/edmo/newsletter-archives/52231
[36] "Election Disinformation in Slovakia", "EU, 2023"
https://ec.europa.eu/newsroom/edmo/newsletter-archives/52231

context amplified its impact and raised alarms about how generative AI can be weaponised in electoral contests.[37]

AI-Generated Celebrity Endorsements

AI-generated celebrity endorsements have emerged as a major misuse of generative technology, with deepfake images, videos, and cloned voices of public figures being used to promote products, scams, or political agendas without consent.[38]

In early 2024, fake ads on Facebook and Instagram used AI to mimic Taylor Swift, Selena Gomez, Jennifer Lopez, and others in a false Le Creuset Cookware giveaway.[39]

This shows AI's ability to convincingly replicate voices and appearances, which can undermine consumer trust, damage reputations, and threaten the $21 billion influencer economy[40], while exploiting gaps in platform detection, verification, and legal protections.

Benefits of AI on Social Media

However, despite the risks associated with AI, it also offers many opportunities. Whether for creativity, learning, activism, education, brainstorming ideas, or artistically, AI's greatest advantage is that it is a blank canvas. Furthermore, AI can help to analyse large datasets of work that would otherwise be too time-consuming for people to achieve, help personalise experiences on social media, better the efficiency of online customer service, as well as help enhance user experience through personalised content, and improve content moderation.[41]

---

[37] "Election Disinformation in Slovakia", "EU, 2023"
https://ec.europa.eu/newsroom/edmo/newsletter-archives/52231

[38] "The Emergence of Deepfake Technology", "Technology Innovation Management Review, 2019"
 https://www.timreview.ca/article/1282

[39] "Incident 626", "AI Incident Database, 2024"
https://incidentdatabase.ai/cite/626/

[40] "Influencer Industry has Ad Fraud Problem", "Business Insider, 2024"
https://www.businessinsider.com/influencer-industry-marketing-fraud-discrimination-unethical-deals-content-creators-brands-2024-4

[41] "How AI is enhancing internet awareness and connectivity", "jeebrnet, 2024"
https://jeebr.net/internet-awareness/ai-in-internet-awareness/

2016	Through the use of AI-powered bots in the US elections, it is demonstrated how AI can sway public opinion and, thus, influence election results.[42] This marks the first global warning of AI and disinformation.
2019	Deepfake videos begin being posted online.[43]
18 October 2022	The "AI Bill of Rights" is introduced by the US as a non-binding framework which emphasises accountability and transparency across all AI systems.[44]
28 April 2023	The EU passes the AI Act, regulating transparency in generative AI tools.[45]
July 2023	China enacts laws mandating the labelling of AI content.[46]
30 October 2023	Japan leads the G7 Hiroshima AI process and promotes international collaboration on AI safety. Japan links Western and Asian

---

[42] *"The Role of AI Twitter Bots Used During U.S. Elections: A study of how malicious Twitter bots play a role in increasing digital conflict and potentially influence voter perceptions leading up to a U.S election." "NSU Florida 2021"* https://nsuworks.nova.edu/shss_dcar_etd/185

[43] History of deepfakes https://www.realitydefender.com/insights/history-of-deepfakes

[44] *"Blueprint for an AI Bill"* "The White House 2022" https://bidenwhitehouse.archives.gov/ostp/ai-bill-of-rights

[45] EU AI Act "European Commission 2023" https://artificialintelligenceact.eu/the-act

[46] China's mandates "*South China morning Post*"https://www.scmp.com/news/china/politics/article/3302477/china-mandates-labels-all-ai-generated-content-fresh-push-against-fraud-fake-news

	democracies on the issue.[47]
Early 2024	Meta and OpenAI introduce watermarking for AI-generated media.[48]

Major countries, Organisations and Alliances

United States of America (USA)

The USA, which shelters the largest AI, technology and social media companies globally, such as Meta, OpenAI, Google and X, is aware of the risks of deepfakes, algorithmic bias and disinformation.[49] All of the aforementioned can be a direct result of the use of AI-generated content; for that reason, federal regulation is still developing. Initiatives like the AI Bill of Rights and voluntary agreements between tech firms and the White House are indicators of the growing awareness of the risks of deepfakes, algorithmic bias and disinformation. These agreements include investing in the development of mechanisms to label AI content, such as watermarking, which also benefits American AI companies, which can copyright the generated output of the AI bots.[50] This measure aims at protecting internet users, since watermarks are a clear indicator of AI generation and thus benefit both sides. What's important to note is that the USA is a dominant force in shaping global standards, due to the influence it has through its global companies and large corporations.

Further reading: https://www.whitehouse.gov/ostp/ai-bill-of-rights/

---

[47] Japan in the G7 https://www.mofa.go.jp/policy/economy/summit/hiroshima23/en
[48] "Labeling AI-Generated Images on Facebook, Instagram and Threads" *"Meta 2024"* https://about.fb.com/news/2024/02/labeling-ai-generated-images-on-facebook-instagram-and-threads
[49] How the US Public and experts view artificial intelligence "Pew Research Centre 2025" https://www.pewresearch.org/internet/2025/04/03/how-the-us-public-and-ai-experts-view-artificial-intelligence
[50]*"Voluntary Commitments from Leading Artificial Intelligence Companies" "Harvard Law Review july 21, 2023" https://harvardlawreview.org/print/vol-137/voluntary-commitments-from-leading-artificial-intelligence-companies-on-july-21-2023*

China

Due to China's political status, it maintains strict information control policies while balancing this with innovation. China has passed specific rules requiring the labelling of AI-generated content and completely banning synthetic content[51], due to its risk of subverting state power. These have been applied on their most popular platforms like Baidu, WeChat and TikTok.[52] Compared to the EU and the USA, which have followed more liberal paths, China's actions illustrate an emphasis placed on state security.

Further   reading:
https://www.reuters.com/technology/china-publishes-rules-generative-ai-2023-07-13/

https://ucigcc.org/blog/chinas-uncharacteristic-approach-to-artificial-intelligence-ai-development/#:~:text=China%20is%20working%20to%20develop,national%20firms%20to%20fulfill%20production.

European Union (EU)

The EU, a leading force when it comes to regulating AI and digital content, has introduced the 2024 AI Act, which demands transparency and categorises AI systems based on their displayed risk.[53] The issue of generative AI is of pressing importance due to its ability to mass-produce and spread disinformation in many different forms at alarming rates. The Digital Services Act (DSA), which is another EU initiative on the topic, holds social media platforms accountable for harmful content, misinformation and algorithmic amplification of false narratives displayed on their platforms.[54] As a result, social media platforms are legally liable for the content on their platforms and are susceptible to legal action to be taken against them if they do not adequately monitor what is being posted.[55]

---

[51] China's AI regulations *"Naaia 2023"*
https://naaia.ai/china-ai-regulation/#:~:text=Key%20principles%20of%20AI%20regulation,in%20the%20Chinese%20regulatory%20plan.
[52] China's mandates *"South China morning Post"https://www.scmp.com/news/china/politics/article/3302477/china-mandates-labels-all-ai-generated-content-fresh-push-against-fraud-fake-news*
[53] EU Ai Act "European Commission 2023"
https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai#:~:text=The%20AI%20Act%20is%20the,in%20AI%20across%20the%20EU.
[54] AI Act "Digital Services Act 2024"
https://www.eu-digital-services-act.com/#:~:text=The%20Digital%20Services%20Act%20is,4.
[55] AI Act "Digital Services Act 2024"
https://www.eu-digital-services-act.com/#:~:text=The%20Digital%20Services%20Act%20is,4.

This urges them to apply countermeasures such as the obligatory labelling of AI content and the limitations placed on it. The DSA is effective as it obliges social media platforms to take action.

Further reading: https://artificialintelligenceact.eu

Previous attempts to solve the issue

G7 Hiroshima AI Process (2023)

Led by Japan, the G7 Hiroshima AI process was established in October 2023, aiming to develop common guidelines for generative AI.[56] In order to combat false information, the G7 framework consisted of an analysis of generative AI's risks and challenges and AI guiding principles for all actors in the AI development field.[57] Lastly, they introduced "the Hiroshima Process International Code of Conduct for Organisations Developing Advanced AI Systems".[58] Even though this Act is not legally binding, it demonstrates how a nationwide consensus is being reached on the issue, allowing for future cooperation on the topic. Additionally, it takes action in the developmental stage of AI, ensuring that it is created ethically and meets the worldwide safety standards.

AI Content Regulation in China

In 2023, China introduced the Interim Measures for Generative AI Services, with the goal of having all AI content be labelled and prohibiting any AI content that might undermine national security.[59] Some of these measures include using open-source AI software, which means that anyone can review the code behind an AI software and check it for malicious intent.[60] This way, it is ensured that AI

---

[56] "EuropeanCommisison"
https://digital-strategy.ec.europa.eu/en/library/g7-leaders-statement-hiroshima-ai-process
[57] "EuropeanCommisison"
https://digital-strategy.ec.europa.eu/en/library/g7-leaders-statement-hiroshima-ai-process
[58]"EuropeanCommisison"
https://digital-strategy.ec.europa.eu/en/library/g7-leaders-statement-hiroshima-ai-process
[59] China's Interim Measures "*China law translate*"https://www.chinalawtranslate.com/en/generative-ai-interim/#:~:text=(1)%20Uphold%20the%20Core%20Socialist,to%20carry%20out%20unfair%20competition;
[60] China's Interim Measures "*China law translate*"https://www.chinalawtranslate.com/en/generative-ai-interim/#:~:text=(1)%20Uphold%20the%20Core%20Socialist,to%20carry%20out%20unfair%20competition;

development is proper and controlled. Contrary to the process of the G7, these regulations act as the first legally binding national laws, which have been specifically placed to combat AI content[61].

Possible solutions

Promoting Transparency

To successfully address an issue like AI generated content on social media, it is important to promote transparency, as such it should be made mandatory for AI generated content to be labeled, in order to prevent any kind of misinformation of the users or influence their opinion by something completely baseless This can be achieved by various methods, such as digital watermarking for deepfakes, which is easily noticeable by users. This is especially crucial during democratic occasions like elections and public emergencies. This way, users know what to believe and what not, thus being able to critically assess the information they come across, making this an effective measure when tackling this issue.[62]

Supporting Ethical AI Development and Open Research

Tackling one of the root causes of the issues associated with AI-generated content on social media, it is essential to facilitate the ethical development of AI through the creation of guidelines that researchers must follow on an international scale. Supporting open research promotes traceability and understanding behind AI tech. This approach prevents misuse before it reaches the general public and additionally fosters collaboration between developers and the public, leading to a broader feeling of trust. These processes can be monitored by ethics boards that can be established for this purpose. Nevertheless, it is important to note that various problems can come up due to companies not wanting external cooperation and interference, making this solution challenging to implement on a large scale.

---

[61] China's Interim Measures "China law translate"https://www.chinalawtranslate.com/en/generative-ai-interim/#:~:text=(1)%20Uphold%20the%20Core%20Socialist,to%20carry%20out%20unfair%20competition;
[62]*"Proposal for a Regulation Laying Down Harmonised Rules on Artificial Intelligence"* *"European Commission,* April"2021*"https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-laying-down-harmonised-rules-artificial-intelligence*

Establishing Global Governance Mechanisms

Mechanisms for global governance are needed, in which standards can be established concerning transparency and labelling of AI content. These standards should be created by multiple third-party bodies in order to promote objectivity and ensure that all countries can comply with said standards. Such organisations could include the European AI Office, which is the centre of AI expertise in the EU.[63]

---

[63] AI alliance members "AI alliance" https://thealliance.ai

Bibliography

Anderson, Ketty. DEVELOPING GLOBAL STANDARDS for AI ACCOUNTABILITY in CONTENT GENERATION. 16 Dec. 2023, www.researchgate.net/publication/387090664_DEVELOPING_GLOBAL_STANDARDS_FOR_AI_ACCOUNTABILITY_IN_CONTENT_GENERATION.

Brundage, Miles, et al. "Toward Trustworthy AI Development: Mechanisms for Supporting Verifiable Claims." arXiv preprint, arXiv:2004.07213, 2020, https://arxiv.org/abs/2004.07213.

Cardona, Miguel, et al. Artificial Intelligence and the Future of Teaching and Learning Insights and Recommendations. May 2023, www.ed.gov/sites/ed/files/documents/ai-report/ai-report.pdf.

Chavda, Janakee. "What Web Browsing Data Tells Us about How AI Appears Online." Pew Research Center, 23 May 2025, www.pewresearch.org/data-labs/2025/05/23/what-web-browsing-data-tells-us-about-how-ai-appears-online/.

Civit, Miguel, et al. "A Systematic Review of Artificial Intelligence-Based Music Generation: Scope, Applications, and Future Trends." Expert Systems with Applications, vol. 209, Dec. 2022, p. 118190, www.sciencedirect.com/science/article/pii/S0957417422013537.

Cohen, Marshall. "Fake Celebrity Endorsements Become Latest Weapon in Misinformation Wars, Sowing Confusion ahead of 2024 Election." CNN, 22 Aug. 2024, edition.cnn.com/2024/08/22/media/fake-celebrity-endorsements-social-media-2024-election-misinformation.

Corsi, Giulio, et al. "The Spread of Synthetic Media on X." Harvard Kennedy School Misinformation Review, vol. 5, no. 3, 3 June 2024, misinforeview.hks.harvard.edu/article/the-spread-of-synthetic-media-on-x/, https://doi.org/10.37016/mr-2020-140.

Content Exploring the Risks, Technical Approaches, and Regulatory Responses US POLICY. Oct. 2024, fpf.org/wp-content/uploads/2024/10/Synthetic_Content_Report_October_2024.pdf.

Coursera, Staff. "What Is Generative AI? Definition, Applications, and Impact." Coursera, 2024, www.coursera.org/articles/what-is-generative-ai?utm_medium=sem&.

DATAVERSITY, Foote, Keith D. "A Brief History of Generative AI." DATAVERSITY, 5 Mar. 2024, www.dataversity.net/a-brief-history-of-generative-ai/.

Department of Homeland Security. "Increasing Threat of Deepfake Identities." 2023, www.dhs.gov/sites/default/files/publications/increasing_threats_of_deepfake_identities_0.pdf.

Devine, Curt, et al. "A Fake Recording of a Candidate Saying He'd Rigged the Election Went Viral. Experts Say It's Only the Beginning." CNN, Feb. 2024, edition.cnn.com/2024/02/01/politics/election-deepfake-threats-invs.

European Commission. Digital Services Act Package. 2023, https://digital-strategy.ec.europa.eu/en/policies/digital-services-act-package.

European Commission. Proposal for a Regulation Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act). 21 Apr. 2021, https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-laying-down-harmonised-rules-artificial-intelligence.

European Parliament, Mar, Negreiro. "EN BRIEFING Children and Deepfakes." European Parliament, July 2025, www.europarl.europa.eu/RegData/etudes/BRIE/2025/775855/EPRS_BRI(2025)775855_EN.pdf.

Europa.eu. "Election Disinformation in Slovakia – Disinfo Bulletin 9 April." Europa.eu, 2024, ec.europa.eu/newsroom/edmo/newsletter-archives/52231.

Floridi, Luciano, et al. "AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations." Minds and Machines, vol. 28, no. 4, 2018, pp. 689–707. Springer, https://doi.org/10.1007/s11023-018-9482-5.

Grand View Research. "Artificial Intelligence Market Size, Share | AI Industry Report, 2025." Grand View Research, 2024, www.grandviewresearch.com/industry-analysis/artificial-intelligence-ai-market.

Incidentdatabase.ai. "Incident 626: Social Media Scammers Used Deepfakes of Taylor Swift and Several Other Celebrities in Fraudulent Le Creuset Cookware Giveaways." Incidentdatabase.ai, 2 Oct. 2023, incidentdatabase.ai/cite/626/.

Jeebr Internet. "How AI Is Enhancing Internet Awareness and Connectivity." Internet Awareness, 31 May 2024, jeebr.net/internet-awareness/ai-in-internet-awareness/.

JURI COMMITTEE. Generative AI and Copyright Training, Creation, Regulation Policy Department for Justice, Civil Liberties and Institutional Affairs Directorate-General for Citizens' Rights. July 2025, www.europarl.europa.eu/RegData/etudes/STUD/2025/774095/IUST_STU(2025)774095_EN.pdf.

Khanal, Shaleen, et al. "Why and How Is the Power of Big Tech Increasing in the Policy Process? The Case of Generative AI." Policy & Society (Print), vol. 44, no. 1, 27 Mar. 2024, https://doi.org/10.1093/polsoc/puae012.

McKinsey & Company. "The State of AI in 2023: Generative AI's Breakout Year." McKinsey, 1 Aug. 2023, www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-in-2023-generative-ais-breakout-year.

McKinsey & Company. "What Is Generative AI?" McKinsey & Company, 2 Apr. 2024, www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-generative-ai.

Ministry of Foreign Affairs of Japan. G7 Hiroshima Leaders' Statement on the Hiroshima AI Process. Oct. 2023, https://www.mofa.go.jp/files/100489998.pdf.

Mucci, Tim. "AI Generated Content." Ibm.com, 27 Nov. 2024, www.ibm.com/think/insights/ai-generated-content.

OECD. "How's Life for Children in the Digital Age?: The Impact of Digital Activities on Children's Lives." OECD, 2025, www.oecd.org/en/publications/how-s-life-for-children-in-the-digital-age_0854b900-en/full-report/the-impact-of-digital-activities-on-children-s-lives_4df70664.html.

Oyelabs (https://www.facebook.com/kinganuragjain). "How AI Content Is Changing Short Video App like Tiktok." Oyelabs.com, 4 Apr. 2025, oyelabs.com/how-ai-content-is-changing-short-video-app-like-tiktok/.

Pal, Mayuree, and C P Rashmi. "Media Manipulation and Disinformation: The Role of AI in Fake News Creation and Its Impact on Trust in Media." Igi-Global.com, 2025, pp. 333–358, www.igi-global.com/chapter/media-manipulation-and-disinformation/373021, https://doi.org/10.4018/979-8-3693-6675-2.ch013.

Reuters. "China Publishes Rules for Generative AI Services." Reuters, 13 July 2023, https://www.reuters.com/technology/china-publishes-rules-generative-ai-2023-07-13/.

Saad, Ali, et al. "The Impact of Artificial Intelligence on Social Media Content." ResearchGate, Science Publications, 22 Apr. 2024, www.researchgate.net/publication/380000781_The_Impact_of_Artificial_Intelligence_on_Social_Media_Content.

SAP. "What Is AI Ethics? The Role of Ethics in AI." Sap.com, 9 Aug. 2024, www.sap.com/resources/what-is-ai-ethics.

Spring, Marianna. "TikTok: AI Fakes, Abuse and Misinformation Pushed to Young Voters." BBC.com, 2 June 2024, www.bbc.com/news/articles/c1ww6vz1l81o.

Statista. "U.S. Teens Generative AI Usage Frequency 2024 | Statista." Statista, 2024, www.statista.com/statistics/1500204/us-teens-generative-ai-tools-usage-frequency/.

Sustainability Directory. "Algorithmic Amplification → Term." Climate → Sustainability Directory, 22 Mar. 2025, climate.sustainability-directory.com/term/algorithmic-amplification/.

Tableau. "What Is the History of Artificial Intelligence (AI)?" Tableau, Salesforce, 2024, www.tableau.com/data-insights/ai/history.

Tangalakis-Lippert, Katherine. "The $21 Billion Influencer Industry Has an Ad Fraud Problem." Business Insider, 2024, www.businessinsider.com/influencer-industry-marketing-fraud-discrimination-unethical-deals-content-creators-brands-2024-4.

Taylor, Derrick Bryson. "Tom Hanks Warns of Dental Ad Using A.I. Version of Him." The New York Times, 2 Oct. 2023, www.nytimes.com/2023/10/02/technology/tom-hanks-ai-dental-video.html.

UNESCO. Recommendation on the Ethics of Artificial Intelligence. 2021, https://unesdoc.unesco.org/ark:/48223/pf0000380455.

UNESCO. "Digital Literacy." Uis.unesco.org, 22 June 2022, uis.unesco.org/en/glossary-term/digital-literacy.

United Nations. "The AI Generation: Youth in the Artificial Intelligence Era." United Nations Western Europe, 14 July 2025, unric.org/en/the-ai-generation-youth-in-the-artificial-intelligence-era/.

United States, Executive Office of the President. Blueprint for an AI Bill of Rights: Making Automated Systems Work for the American People. The White House, Oct. 2022, https://www.whitehouse.gov/ostp/ai-bill-of-rights/.

19

UNIVIO. "AI Failures: Learning from Common Mistakes and Ethical Risks." Univio.com, 2024, www.univio.com/blog/the-complex-world-of-ai-failures-when-artificial-intelligence-goes-terribly-wrong/.

Vinuesa, Ricardo, et al. "The Role of Artificial Intelligence in Achieving the Sustainable Development Goals." Nature Communications, vol. 11, no. 1, 2020, https://doi.org/10.1038/s41467-019-14108-y.

Westerlund, Mika. "The Emergence of Deepfake Technology: A Review." Technology Innovation Management Review, vol. 9, no. 11, 2019, www.timreview.ca/article/1282.

Synthesia. "What Is Synthetic Media? | Synthesia Glossary." Synthesia.io, www.synthesia.io/glossary/synthetic-media.