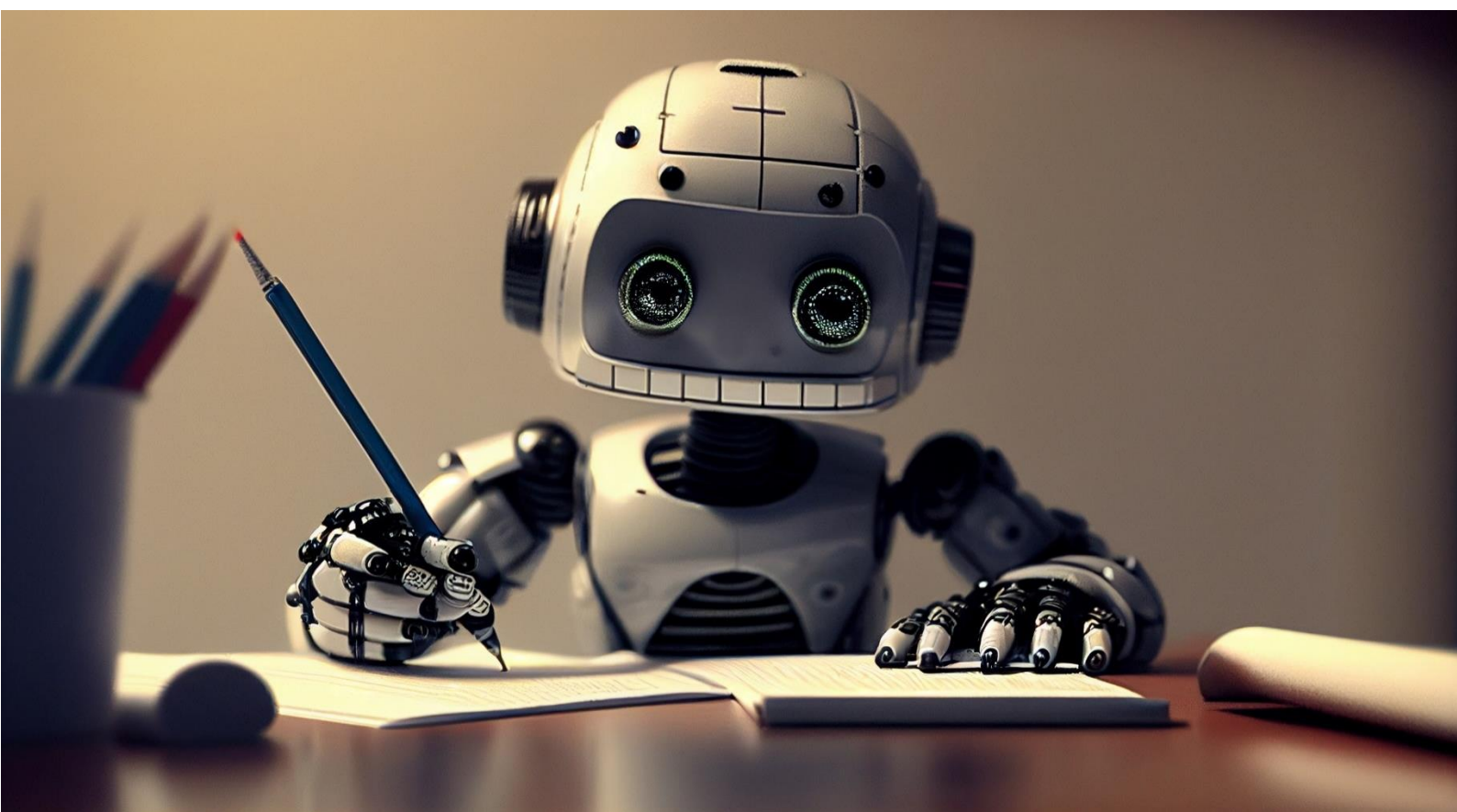


Generative Artificial Intelligence for Science Communication: an Overview of European Online News

2024



This is a joint publication of the Joint Research Centre (JRC), the European Commission's science and knowledge service, and the European Science-Media Hub, in the Scientific Foresight Unit at the European Parliamentary Research Services (EPRS) of the European Parliament. This report aims to provide evidence-based scientific support to the European policymaking process. It is prepared for, and addressed to, the Members and staff of the European Parliament as background material to assist them in their parliamentary work. The scientific output expressed does not imply a policy position of the European Commission nor the European Parliament. Neither the European Commission nor the European Parliament nor any person acting on behalf of these are responsible for the use that might be made of this publication. For information on the methodology and quality underlying the data used in this publication for which the source is neither Eurostat nor other Commission/Parliament services, users should contact the referenced source. The designations employed and the presentation of material on the maps do not imply the expression of any opinion whatsoever on the part of the European Union concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries.

Contact information:

European Commission Joint Research Centre
Directorate for Digital Transformation and Data
Unit text and data mining

Email: Bertrand.DE-LONGUEVILLE@ec.europa.eu

European Parliament
European Parliamentary Research Services
Panel for the Future of Science and Technology
Unit European Science-Media Hub

Email: Vitalba.CRIVELLO@europarl.europa.eu
Marika.ARMANOVICA@europarl.europa.eu

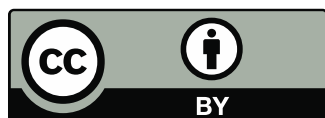
EU Science Hub

<https://joint-research-centre.ec.europa.eu>

JRC138266

Ispra: European Commission, 2024
Brussels: European Parliament, 2024

© European Union, 2024



The reuse policy of the European Commission documents is implemented by the Commission Decision 2011/833/EU of 12 December 2011 on the reuse of Commission documents (OJ L 330, 14.12.2011, p. 39). Unless otherwise noted, the reuse of this document is authorised under the Creative Commons Attribution 4.0 International (CC BY 4.0) licence (<https://creativecommons.org/licenses/by/4.0/>). This means that reuse is allowed provided appropriate credit is given and any changes are indicated.

Reproduction and translation for non-commercial purposes are authorised, provided the source is acknowledged and the European Parliament is given prior notice and sent a copy.

For any use or reproduction of photos or other material that is not owned by the European Union permission must be sought directly from the copyright holders. The European Union does not own the copyright in relation to the following elements:

- Cover/back page illustration, © Emmy Ljs / stock.adobe.com

How to cite this report: Zsuzsanna Pásztor, Theresa Reitis-Münstermann, *Generative Artificial Intelligence for Science Communication: an Overview of European Online News*, European Commission/European Parliament, 2024, JRC138266.

The authors acknowledge the guidance and valuable feedback provided by Marika Armanoviča, Vitalba Crivello and Svetla Tanova-Encke as well as for the help provided by Francesco Cabras to process the relevant data.

Table of contents

- Key Findings..... 2
- Introduction..... 3
 - 1. Overview of reporting on Artificial Intelligence related to Science 4
 - 2. Generative AI and journalism 9
- Conclusions..... 11
- Appendix..... 12
- Annex I – Methodology 16
- Annex II - About the authors 17

Key Findings

The purpose of this document is to illustrate, using Text Mining techniques, what topics related to generative artificial intelligence (GenAI) in the context of science communication are discussed in the European online media landscape. It provides an overview of the most recurrent viewpoints and reporting trends related to this topic. The sources of our analysis include the most important online mainstream media outlets in the 27 Member States.

This report summarises the findings of the analysis on data collected from 1 January 2023 till 10 February 2024.

Key findings:

- The topic of “GenAI for science communication” in quantitative terms is a niche issue in European mainstream online media.
- On a broader scope, “AI and science” are more frequently mentioned in online news (12 000 articles). In this context, “ChatGPT” is the most widely mentioned topic. But also the art and creative sector is widely covered and the media analysis confirms a strong agreement of the various stakeholders including tech experts and policymakers on the need for AI regulation.
- Related to “**GenAI**/GPT related to journalism”, three main issues stand out in online media reporting on the risks of using AI in journalism: it can replace journalists’ jobs, it can spread credible fake news and disinformation on large scale harming credibility of online sources, and it can threaten personal rights by creating false information about real persons.

Introduction

Generative Artificial Intelligence (GenAI) represents a revolutionary advancement in the field of technology. Distinguished from traditional Artificial Intelligence (AI), GenAI is specifically designed to produce new, original content, such as images, text, and sounds, with the potential of becoming indistinguishable from those created by humans. Large language models, such as ChatGPT-3, have further advanced the capabilities of GenAI, raising additional considerations about their impact on science communication and the broader societal landscape.

The recent launch of widely accessible GenAI tools has sparked a surge in media coverage, igniting discussions about its profound impact on everyday life and scientific research. As GenAI continues to influence various aspects of society, including scientific communication and research, it presents both opportunities and challenges in upholding accuracy, transparency, and comprehensibility.

This analysis of European mainstream media seeks to delve into the evolving role of GenAI in the context of science communication. The data collection of news related to GenAI for science communication (mentioning both concepts in title and/or text of an article) yielded only four articles on the selected EMM sources.¹ This could be due to Mainstream media using the term AI when they referring to GenAI, although GenAI is a specific concept as explained above.

Therefore, we investigated AI in the context of science more in general in the first step. In the first section of this report, we aim to provide a brief overview of the focal points in media coverage on AI and science in general.

As we noticed the frequent use of GPT in the context of GenAI, zoomed in on the media attention to GenAI/GPT and journalism. In the second section, we focus on the key negative and positive aspects, delving into the portrayal of these topics in news articles and their implications.

By examining how the media portrays AI in the context of science and GenAI for journalism, we can better understand the public perception of the issue and the role of media in shaping that perception. News articles play a crucial role in informing the public about the various angles of using GenAI for science communication, providing insights, and facilitating discussions about the opportunities and challenges associated with this technology.

The analysis was performed on European mainstream online articles published between 1 January 2023 and 10 February 2024. To identify the most relevant topics, we apply an automated topic-clustering algorithm that unveils the main events and narratives appearing in media. Data is collected by Europe Media Monitor (EMM) on a selection the most important EU27 online media sources, including 581 media outlets from all Member States².

For a more specific description of the methodology and applied technology, please refer to the Annexes.

¹ These covered the “Tendencias 2023” event held in Madrid or the “Athens Science Festival2024” held in Greece, or initiatives, such as “TeresIA” that aims to facilitate translations and generation of scientific terminology, and “Turacoz” providing GenAI in medical communication.

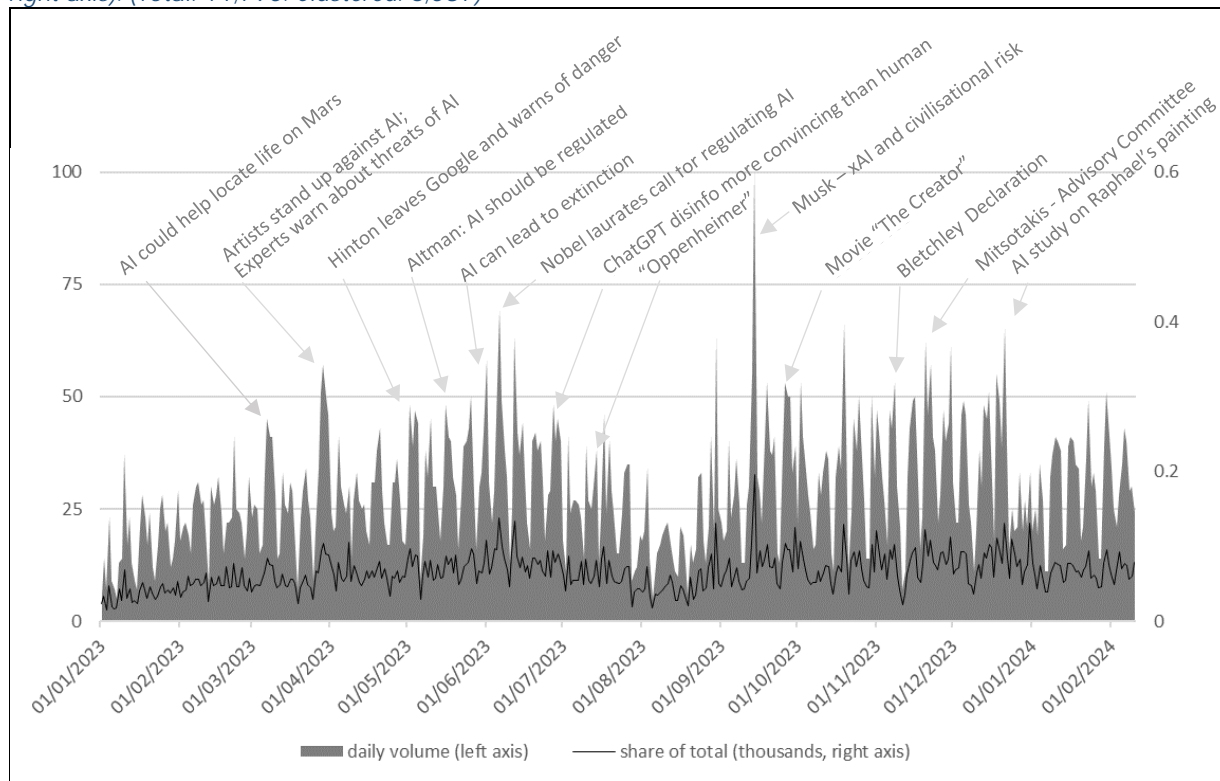
² The number of sources has been weighted based on the countries’ population.

1. Overview of reporting on Artificial Intelligence related to Science

In this chapter, we provide an overview of media reporting about AI in the context of Science.

Figure 1 displays the full dataset containing more than 11 000 articles that discuss AI and science, from different perspectives.³ It presents the total number of articles published per day (left axis) and the daily reporting peaks as share of the total mainstream media news of each day (right axis). The figure shows some effects of normal seasonality, with lower reporting volumes during holiday periods, in August and around the winter break in 2023. (The leading news will be explained later in this chapter).

Figure 1 – Reporting volume on AI and science over time (total volume – left axis, proportion per thousand – right axis). (Total: 11,790. Clustered: 6,087)

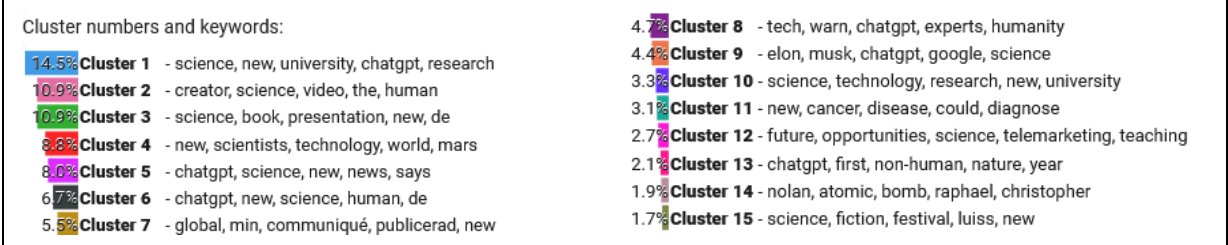


Source: EMM "EE24campaign" index. Period: 2023-01-01 to 2024-02-10. Daily aggregated data.

³ For a description of the keyword-based article selection article selection and full query (Query 1), please consult the Annex I and Appendix.

Second, to identify the main topics and events covered by the media, we applied an automated clustering algorithm on the titles of the articles.⁴ The most frequent keywords of the main topics and the news groups' share are displayed in Figure 2. The keyword list unveils a large share of articles including GenAI with the prevalence of ChatGPT, which appears to be a cross-cutting issue in several news groups (clusters 1, 5, 6, 8, 9 and 13).⁵

Figure 2 - Main clusters of "AI and science" articles.



Source: EMM "EE24 campaign" index. Period: 2023-01-01 to 2024-02-10.

Upon analysing the main reporting peaks (see Figure 1) and considering the insights gained from the clustering process, it has been observed that, albeit not directly linked to scientific work, one of the most recurrent topics in the media pertains to AI in the creative media sector, including art and cinema as a means of communication. For instance, artists opposing the use of AI that poses a threat to the industry, and the release of the science fiction movie "The Creator" (that tells the story of a battle between humans and AI robots) were widely covered (cluster 2). Filmmaker Christopher Nolan's take on AI in the film industry is featured in news related to cinema. He contends that the story of the atomic bomb presented in his film "Oppenheimer" is useful for understanding the dilemmas associated with AI. Furthermore, the use of AI in history of art to reveal the authorship information on Raphael's painting also received great publicity (cluster 14).

In addition, topic of AI assisting scientists in complex tasks is covered as well. AI helping to look for signs of extra-terrestrial life received great publicity (cluster 4), but the value of this technology in healthcare is also documented in media. Particular attention was paid to its use in improving mammography screening process and other high-end medical imaging equipment (cluster 7) or in predicting diseases such as cancer (cluster 11). Furthermore, the value of AI in predicting meteorological phenomena is recognised in media.

News about ChatGPT

We further analysed those clusters for which "ChatGPT" is among the most frequent keywords. To better understand the different angles of the selected clusters, first, we applied an automated analysis showing the "framing" of the title and the first sentences of the articles. Information on main entities and source country of the articles was added as well. (Results are shown in the Appendix' additional figures).

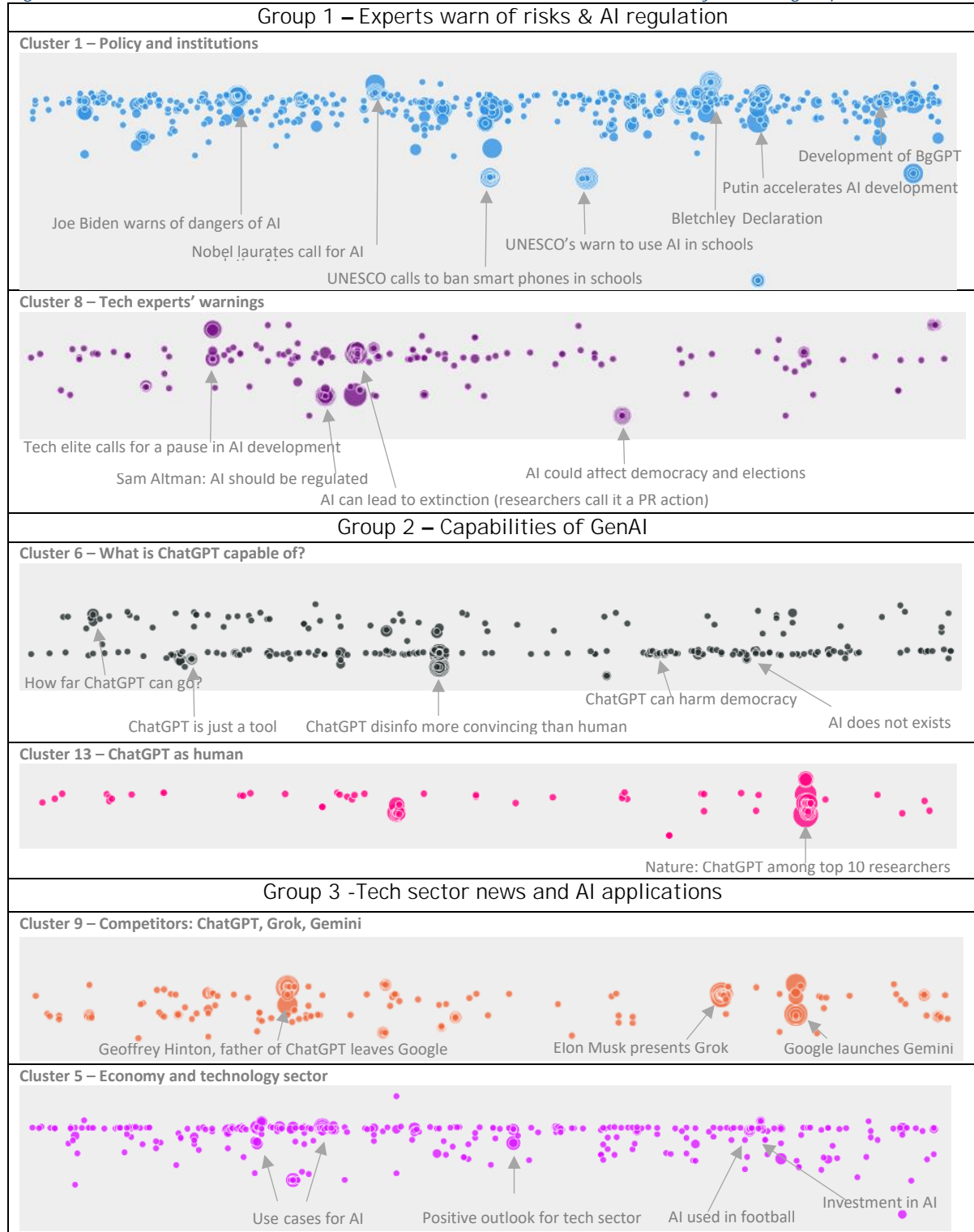
Subsequently to complete the analysis, we conducted a more in-depth content analysis of the articles, and we grouped the clusters in three separate categories:

⁴ Note: Clustering was performed on headlines. Headlines not similar to any of the groups are not considered in the distribution across topics. "Similarity, includes also the similar length of titles and mentions of names.
⁵ Other models of GenAI seem to be less represented in media, such as Gemini, DialoGPT, BlenderBot or Grok.

- 1) Experts warning of risks & AI regulation (clusters 1 and 8),
- 2) Capabilities of GenAI (clusters 6 and 13),
- 3) Tech sector news (cluster 5 and 9).

Figure 3 displays the bubble charts for these news groups. The bubbles represent single titles sorted on a horizontal timeline. The vertical distance portrays the semantic distance, i.e. the closer they are to each other, the more similar they are; e.g. concentric circles represent almost identical headlines.

Figure 3 – Volume and most mentioned entities for ChatGPT-related clusters by defined group



Source: EMM "EE24campaign" index. Period: 2023-01-01 to 2024-02-10. Selection of articles displayed for each cluster.

Experts warning of risks & AI regulation (Group 1)

We found main headlines mostly revolving around primarily the tech elite expressing serious concerns and calling for AI regulation (cluster 8). The most striking news included:

- The open letter signed by the “tech elite” calling for a pause of AI development. (The original letter, signed by thousands of experts and researchers along with Elon Musk, asked to pause “the training of AI systems more powerful than GPT-4” for at least 6 months, in order to slow down the “out-of-control” race for developing these systems, that undermines responsible planning. Especially the central claim of the letter “AI systems with human-competitive intelligence can pose profound risks to society and humanity” is highlighted in media reporting⁶ (29 March 2023).
- OpenAI CEO Sam Altman urges AI regulation in the US Senate hearing (16 May 2023).
- OpenAI (Sam Altman), Google DeepMind (Demis Hassabis), Anthropic (Dario Amodei) signed a statement issued by the Center for AI Safety’s. They warn about the risks of AI, as future systems could be as deadly as pandemics and nuclear weapons. The statement was also supported by prominent AI researchers Geoffrey Hinton and Yoshua Bengio, as well as executives at Microsoft and Google (30 May 2023).

On the contrary, reactions from politicians and institutions seem to be more intense in volume in a more recent period that suggest top AI experts’ role in triggering state interventions to regulate AI (cluster 1):

- Joe Biden warns of potential social risks posed by AI (4 April 2023).
- Nobel laureates advocate for AI regulation in order to harness its useful potential but prevent its harmful misuse, e.g. in war scenarios (5 June 2023).
- UNESCO warning about the potential risks of the use of phones and AI in schools (26 July and 7 September).
- The Bletchley Declaration “establishes a shared understanding of the opportunities and risks posed by frontier artificial intelligence”. (The declaration was signed by 28 countries and the European Union at the AI Safety Summit (1 November 2023).
- Advisory committee in Greece was set up by Greek PM Mitsotakis, to better understand how to exploit AI’s potential in several sectors and fields, from schools to tax evasion or to locate sources of fire by civil protection service.
- Putin accelerates a Russian AI development
- BgGPT developed in Bulgaria.

Capabilities of GenAI (Group 2)

These events described under Group 1, triggered a wider public debate about what GenAI is, what is it capable of and how could affect society more in general (cluster 6 and 13). We found a whole spectrum of answers to these questions with very distant points of view.

Some articles argue that ChatGPT is “just a tool”, while others highlight its similarity and impact on human reasoning. For examples, in this group of news we found that (1) reporting about ChatGPT generated disinformation more convincing to the human brain, and (2) ChatGPT appeared as the first non-human on Nature’s list of most important researchers. Media also sheds light on the threats to

⁶ <https://futureoflife.org/open-letter/pause-giant-ai-experiments/>

democracy, including the worst case scenario, i.e. AI taking control of humans. On the other extreme of this dimension of news, we found sporadic articles referring to an idea that “AI does not exist”.

Tech sector news (Group 3)

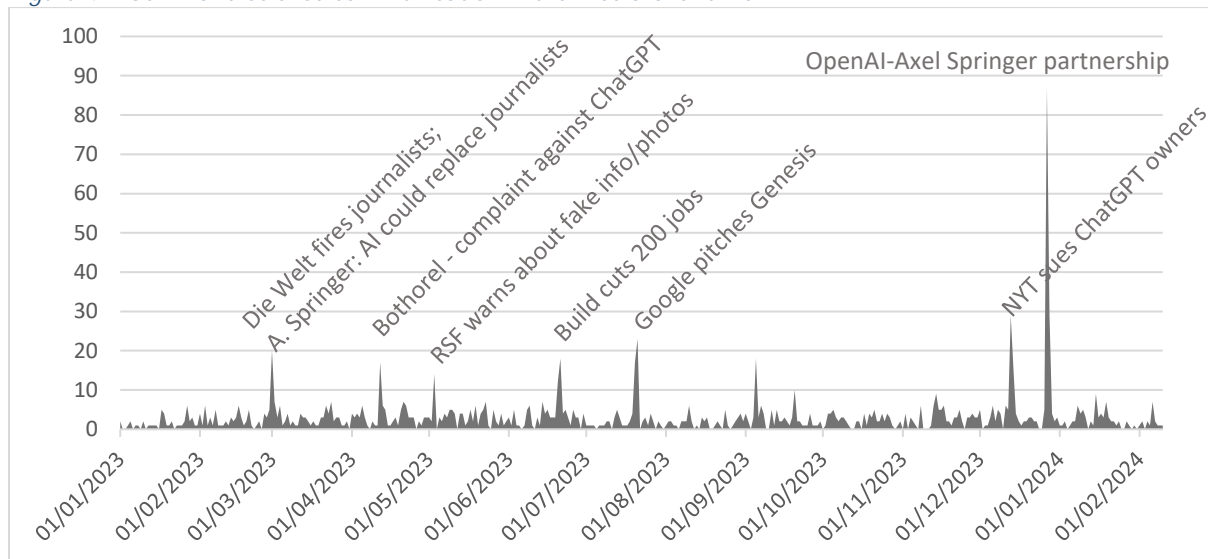
This group of news helps to keep track of main news from the technology sector. It includes the presentation of Grok by Elon Musk, and Gemini by Google, or the resignation of Geoffrey Hinton from Google, the godfather of AI, regretting his work as the intelligence being developed is very far from human intelligence and calling for AI regulation (cluster 9).

Moreover, visions and initiatives are subject of media reporting showing the huge potential to develop AI further in different sectors. Overall, the media reporting on a promising and positive outlook which results in increased investment for the technology sector (cluster 5).

2. Generative AI and journalism

GenAI's text generating capabilities present a potential that is increasingly capturing the attention of various sectors. Experts agree that GenAI will have major implications and will fundamentally transform not only the news ecosystem⁷, but also academia and science communication.⁸ Nevertheless, according to our analysis, the issues of AI for science communication is still a niche in European online media. Science communication is rarely covered, therefore for the analysis we chose a wider scope, and we focus on GenAI represented by ChatGPT and journalism. Results are shown in Figure 4.⁹

Figure 4 – GenAI and science communication in the media over time



Timeline with relevant news volumes. EMM "EE24campaign" index. Period: 2023-01-01 to 2024-02-10. Daily aggregated data.

The main news items published in mainstream online media are touching upon around three main questions that we demonstrate in the subsequent paragraphs:

Q1: Is AI taking over journalists jobs?

The global partnership signed between Axel Springer and OpenAI received most attention in the context of AI & journalism. The agreement foresees that ChatGPT users receive summaries of selected global news content from Axel Springer's media brands (e.g. Politico, Business Insider, Bild, Die Welt)¹⁰ (27-28 December 2023). This first news items related to Die Welt/Axel Springer were published earlier that year, with titles suggesting AI could replace journalists¹¹ (1 March 2023), while later, German Bild cutting 200 jobs because some roles will be taken over by AI made to the

⁷ E.g. <https://reutersinstitute.politics.ox.ac.uk/news/ai-and-journalism-whats-next>

⁸ E.g. https://jcom.sissa.it/article/pubid/JCOM_2202_2023_Y02/

⁹ For a description of the keyword-based article selection article selection and full query (Query 2), please consult the Annex I and Appendix.

¹⁰ See <https://www.axelsspringer.com/en/axel-springer-and-artificial-intelligence> and <https://openai.com/blog/axel-springer-partnership>

¹¹ <https://www.theguardian.com/technology/2023/mar/01/german-publisher-axel-springer-says-journalists-could-be-replaced-by-ai>

headlines¹² (20 June). Related to job cuts, Federal Minister of Labour and Social Affairs of Germany - Hubertus Heil – said that by 2035, there will no longer be a workplace without AI applications. Given the changes of the labour market, he took a rather stance on the issue and anticipated both the promotion of the safe use of AI, and strong regulation in this field where major risks may arise as well (5 September).

Q2: How would it change the quality and the credibility of journalism?

On the opposite, Reporters Without Borders (RSF) warn in the 2023 edition of World Press Freedom Index that journalism threatened by AI & fake news. They highlight that *“the disinformation industry disseminates manipulative content on a huge scale, as shown by an investigation by the Forbidden Stories consortium (...) and now AI is digesting content and regurgitating it in the form of syntheses that flout the principles of rigour and reliability”* and draw attention also to the spread of AI generated fake photos produced by the 5th version of Midjourney (an AI programme that generates very high-definition images in response to natural language requests¹³. (3 May 2023)

The New York Times first reported on Google pitching “Genesis” to major American media organisation (21 July 2023). News highlighted how this technology may soon transform the journalism industry, and main benefits and dangers were mentioned¹⁴. Google introduced the tool as a “helpmate” for journalists, that can help to find titles, writing styles, or could create summaries. At the same time, affidability and reliability of such tools were questioned: if misused, it could easily spread misinformation or damage reputation of the news sources that use it. Later that year, New York Times also took concrete action and sued Microsoft and OpenAI for copyright infringements, blaming them for using millions of their articles and posts to train AI technologies, without compensating the publishers, in addition to creating a competitor and thus causing actual damages (13 December 2023).

Q3: Is AI threatening personal rights by creating false personal data?

In addition to journalism, generative AI raised issues also under the protection of personal data and rights. Following other complaints by private persons in different countries, Éric Bothorel French politician filed a complaint too against ChatGPT with the French Data Protection Agency CNIL under article 5.1.d of the general data protection regulation, as it is not compatible with personal rights, due to numerous inaccuracies and “hallucinations” of the information generated by ChatGPT. This news was surfacing in French media (12 April 2023).

¹² <https://www.theguardian.com/world/2023/jun/20/german-tabloid-bild-to-replace-range-of-editorial-jobs-with-ai>

¹³ <https://rsf.org/en/map-2023-world-press-freedom-index>

¹⁴ <https://www.nytimes.com/2023/07/19/business/google-artificial-intelligence-news-articles.html>

Conclusions

After the launch of ChatGPT at the end of November 2022, we have seen a surge of public debate about AI applications. In this report, we combined manual content analysis with automated analysis to provide a summary of the main news published in European online media around artificial intelligence in the context of science more broadly.

One of the most recurrent topics pertains to AI posing a threat to the creative media sector. Moreover, its use in the realm of history of art receives great publicity. Besides, the topic of AI assisting scientist is covered frequently. Overall, we identified that “ChatGPT” is among the most frequent keywords.

As compared to the richness of European online media, we found that in quantitative terms “GenAI for science communication” is still a niche issue which could be due to the media using the term AI and GPT for the concept of GenAI.

Therefore, we provided a deep dive into a somewhat broader topic “GenAI/GPT and journalism”. In this field, three main threats emerged from our analysis: 1) AI is already taking over the jobs of journalists, 2) AI generated hallucinations and disinformation can threaten credibility and reliability of the sector if not used carefully, 3) AI is threatening personal rights by creating false personal data.

On a positive note, the results of our analysis show the vast societal agreement for a strong need of the regulation of GenAI.

Appendix

Query 1: AI and science

+(title:(“artificial intelligence” OR “изкуствен интелект” OR “umělá inteligence” OR “kunstig intelligens” OR “künstliche Intelligenz” OR “τεχνητή νοημοσύνη” OR “inteligencia artificial” OR “tehisintellekt” OR “tekoäly” OR “intelligence artificielle” OR “intleacht shaorga” OR “umjetna inteligencija” OR “mesterséges intelligencia” OR “intelligenza artificiale” OR “dirbtinis intelektas” OR “mākslīgais intelekts” OR “intelligenza artificjali” OR “kunstmatige intelligentie” OR “sztuczna inteligencja” OR “inteligência artificial” OR “intelligența artificială” OR “umelá inteligencia” OR “umetna inteligenca” OR “artificiell intelligens”))

text:(“artificial intelligence” OR “изкуствен интелект” OR “umělá inteligence” OR “kunstig intelligens” OR “künstliche Intelligenz” OR “τεχνητή νοημοσύνη” OR “inteligencia artificial” OR “tehisintellekt” OR “tekoäly” OR “intelligence artificielle” OR “intleacht shaorga” OR “umjetna inteligencija” OR “mesterséges intelligencia” OR “intelligenza artificiale” OR “dirbtinis intelektas” OR “mākslīgais intelekts” OR “intelligenza artificjali” OR “kunstmatige intelligentie” OR “sztuczna inteligencja” OR “inteligência artificial” OR “intelligența artificială” OR “umelá inteligencia” OR “umetna inteligenca” OR “artificiell intelligens”))

+(title:(“science” OR “наука” OR “věda” OR “videnskab” OR “Wissenschaft” OR “επιστήμη” OR “ciencia” OR “teadus” OR “tiedettä” OR “eolaíocht” OR “nauka” OR “tudomány” OR “scienza” OR “mokslas” OR “zinātnes” OR “xjenza” OR “wetenschap” OR “ciência” OR “știință” OR “veda” OR “znanost” OR “vetenskap”))

text:(“science” OR “наука” OR “věda” OR “videnskab” OR “Wissenschaft” OR “επιστήμη” OR “ciencia” OR “teadus” OR “tiedettä” OR “eolaíocht” OR “nauka” OR “tudomány” OR “scienza” OR “mokslas” OR “zinātnes” OR “xjenza” OR “wetenschap” OR “ciência” OR “știință” OR “veda” OR “znanost” OR “vetenskap”))

Query 2: GenAI/GPT and journalism/science communication

+(title:(“journalism” OR “science communication” OR “журналистика” OR “научна комуникация” OR “žurnalistiky” OR “vědecká komunikace” OR “journalistik” OR “videnskabelig kommunikation” OR “Journalismus” OR “wissenschaftliche Kommunikation” OR “δημοσιογραφία” OR “επιστημη ΕΠΙΚΟΙΝΩΝΙΑ” OR “periodismo” OR “comunicación científica” OR “ajakirjandus” OR “teaduskommunikatsioon” OR “journalisme” OR “communication scientifique” OR “iriseoireacht” OR “cumarsáid eolaíochta” OR “novinarstvo” OR “znanstvena komunikacija” OR “újságírás” OR “tudományos kommunikáció” OR “giornalismo” OR “comunicazione scientifica” OR “žurnalistikos” OR “mokslinė komunikacija” OR “žurnālistikas” OR “zinātniskā komunikācija” OR “ğurnalizmu” OR “komunikazzjoni tax-xjenza” OR “journalistiek” OR “wetenschappelijke communicatie” OR “dziennikarstwo” OR “komunikacja naukowa” OR “jornalismo” OR “comunicação científica” OR “jurnalism” OR “comunicarea științifică” OR “žurnalistika” OR “vedecká komunikácia” OR “znanstveno komuniciranje” OR “vetenskaplig kommunikation” OR “journalismi” OR “tiedeviestintä”))

text:(“journalism” OR “science communication” OR “журналистика” OR “научна комуникация” OR “žurnalistiky” OR “vědecká komunikace” OR “journalistik” OR “videnskabelig kommunikation” OR “Journalismus” OR “wissenschaftliche Kommunikation” OR “δημοσιογραφία” OR “επιστημη ΕΠΙΚΟΙΝΩΝΙΑ” OR “periodismo” OR “comunicación científica” OR “ajakirjandus” OR “teaduskommunikatsioon” OR “journalisme” OR “communication scientifique” OR “iriseoireacht” OR “cumarsáid eolaíochta” OR “novinarstvo” OR “znanstvena komunikacija” OR “újságírás” OR “tudományos kommunikáció” OR “giornalismo” OR “comunicazione scientifica” OR “žurnalistikos” OR “mokslinė komunikacija” OR “žurnālistikas” OR “zinātniskā komunikācija” OR “ğurnalizmu” OR “komunikazzjoni tax-xjenza” OR “journalistiek” OR “wetenschappelijke communicatie” OR “dziennikarstwo” OR “komunikacja naukowa” OR “jornalismo” OR “comunicação científica” OR “jurnalism” OR “comunicarea științifică” OR “žurnalistika” OR “vedecká komunikácia” OR “znanstveno komuniciranje” OR “vetenskaplig kommunikation” OR “journalismi” OR “tiedeviestintä”))

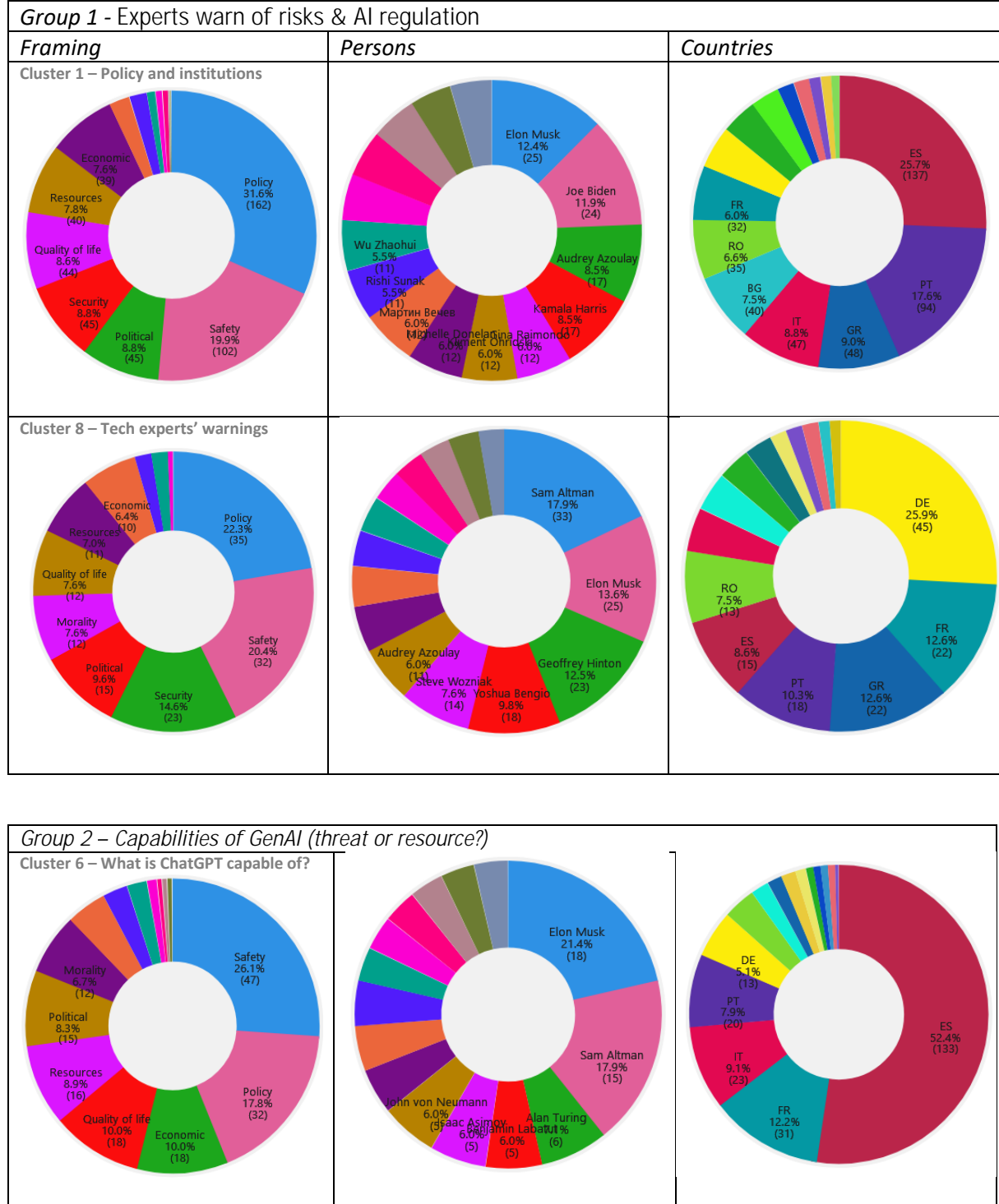
+(title:(“GPT” OR “chat gpt” OR “chatgpt” OR “gpt AI” OR “ai-generated content” OR “GenAI” OR “GenAI4EU” OR “generative ai” OR “generative artificial intelligence” OR “съдържание, генерирано от ИИ” OR “генеративна

ai" OR "генеративен изкуствен интелект" OR "obsah generovaný umělou inteligencí" OR "generativní ai" OR "generativní umělá inteligence" OR "AI-genereret indhold" OR "generativ ai" OR "generativ kunstig intelligens" OR "KI-generierte Inhalte" OR "generative künstliche Intelligenz" OR "περιεχόμενο παραγόμενο από ΤΝ" OR "γενετικό ai" OR "Τεχνητή Νοημοσύνη" OR "contenido generado por IA" OR "ai generativo" OR "inteligencia artificial generativa" OR "tehisintellekti abil loodud sisu" OR "generatiivne ai" OR "generatiivne tehisintellekt" OR "tekoälyn tuottama sisältö" OR "generatiivinen ai" OR "generatiivinen tekoäly" OR "contenu généré par l'IA" OR "ai générative" OR "intelligence artificielle générative" OR "inneachar arna ghiniúint ag an intleacht shaorga" OR "ai giniúnach" OR "intleacht shaorga ghiniúnach" OR "sadržaj generiran umjetnom inteligencijom" OR "generativní ai" OR "generativna umjetna inteligencija" OR "mesterséges intelligenciával előállított tartalom" OR "generatív ai" OR "generatív mesterséges intelligencia" OR "contenuti generati dall'IA" OR "generativo ai" OR "intelligenza artificiale generativa" OR "dirbtiniu intelektu sukurtas turinys" OR "generacinis ai" OR "generacinis dirbtinis intelektas" OR "MI radīts saturs" OR "ģeneratīvais ai" OR "mākslīgais intelekts" OR "kontenut iggenerat mill-IA" OR "ai generattivi" OR "intelligenza artificjali generattiva" OR "door AI gegeneerde inhoud" OR "generatieve ai" OR "generatieve kunstmatige intelligentie" OR "treści generowane przez sztuczną inteligencję" OR "generatywny ai" OR "sztuczna inteligencja generatywna" OR "conteúdos gerados por IA" OR "inteligência artificial geradora" OR "conținutul generat de IA" OR "inteligența artificială generativă" OR "obsah vytvorený umelou inteligenciou" OR "generatívne ai" OR "generačná umelá inteligencia" OR "vsebine, ki jih ustvarja umetna inteligenca" OR "generativna umetna inteligenca" OR "AI-genererat innehåll" OR "generativ artificiell intelligens"))

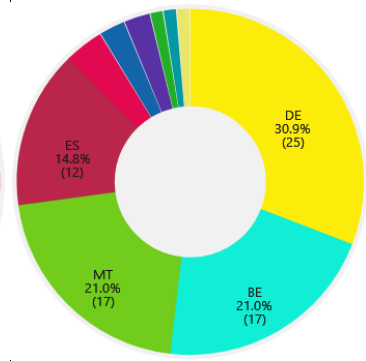
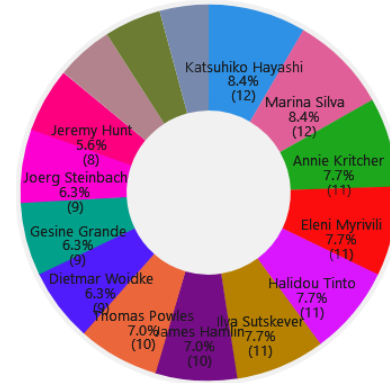
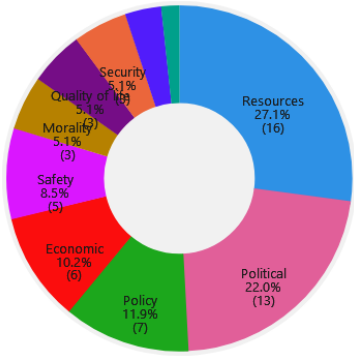
text: ("GPT" OR "chat gpt" OR "chatgpt" OR "gpt AI" OR "ai-generated content" OR "GenAI" OR "GenAI4EU" OR "generative ai" OR "generative artificial intelligence" OR "съдържание, генерирано от ИИ" OR "генеративна ai" OR "генеративен изкуствен интелект" OR "obsah generovaný umělou inteligencí" OR "generativní ai" OR "generativní umělá inteligence" OR "AI-genereret indhold" OR "generativ ai" OR "generativ kunstig intelligens" OR "KI-generierte Inhalte" OR "generative künstliche Intelligenz" OR "περιεχόμενο παραγόμενο από ΤΝ" OR "γενετικό ai" OR "Τεχνητή Νοημοσύνη" OR "contenido generado por IA" OR "ai generativo" OR "inteligencia artificial generativa" OR "tehisintellekti abil loodud sisu" OR "generatiivne ai" OR "generatiivne tehisintellekt" OR "tekoälyn tuottama sisältö" OR "generatiivinen ai" OR "generatiivinen tekoäly" OR "contenu généré par l'IA" OR "ai générative" OR "intelligence artificielle générative" OR "inneachar arna ghiniúint ag an intleacht shaorga" OR "ai giniúnach" OR "intleacht shaorga ghiniúnach" OR "sadržaj generiran umjetnom inteligencijom" OR "generativní ai" OR "generativna umjetna inteligencija" OR "mesterséges intelligenciával előállított tartalom" OR "generatív ai" OR "generatív mesterséges intelligencia" OR "contenuti generati dall'IA" OR "generativo ai" OR "intelligenza artificiale generativa" OR "dirbtiniu intelektu sukurtas turinys" OR "generacinis ai" OR "generacinis dirbtinis intelektas" OR "MI radīts saturs" OR "ģeneratīvais ai" OR "mākslīgais intelekts" OR "kontenut iggenerat mill-IA" OR "ai generattivi" OR "intelligenza artificjali generattiva" OR "door AI gegeneerde inhoud" OR "generatieve ai" OR "generatieve kunstmatige intelligentie" OR "treści generowane przez sztuczną inteligencję" OR "generatywny ai" OR "sztuczna inteligencja generatywna" OR "conteúdos gerados por IA" OR "inteligência artificial geradora" OR "conținutul generat de IA" OR "inteligența artificială generativă" OR "obsah vytvorený umelou inteligenciou" OR "generatívne ai" OR "generačná umelá inteligencia" OR "vsebine, ki jih ustvarja umetna inteligenca" OR "generativna umetna inteligenca" OR "AI-genererat innehåll" OR "generativ artificiell intelligens"))

Additional figures

Figure 5 – Framing, main persons and source countries for clusters containing ChatGPT in top keywords.

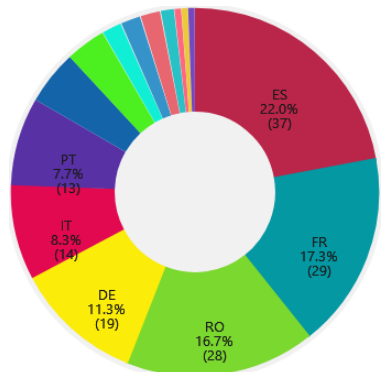
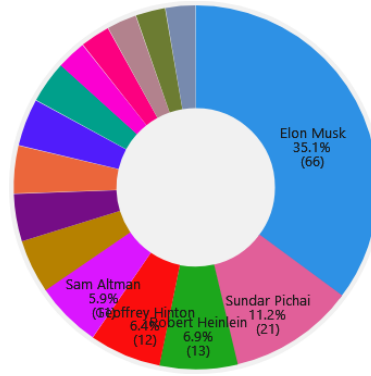
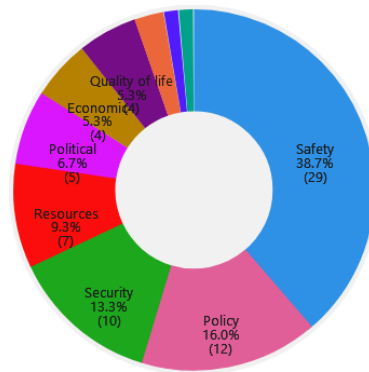


Cluster 13 – ChatGPT as human

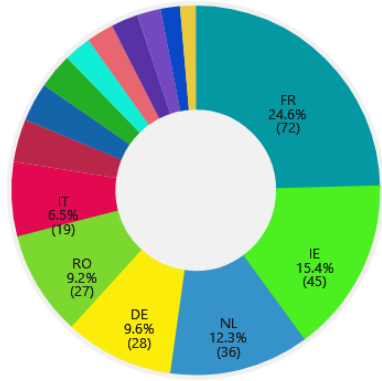
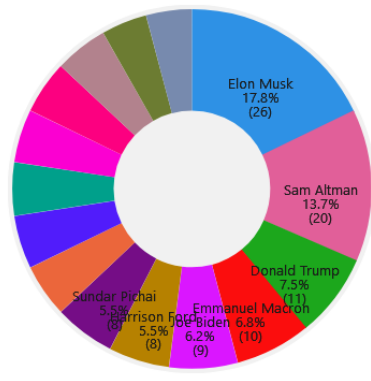
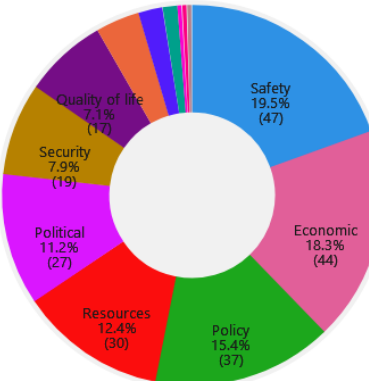


Group 3 – Tech sector news and AI applications

Cluster 9 – Competitors: ChatGPT, Grok, Gemini



Cluster 5 – Economy and technology sector



Annex I – Methodology

Source selection

The manually curated list of Mainstream media outlets used for this report is composed of 581 EU27 online media sources monitored by Europe Media Monitor (EMM). This sample of sources represents the most popular open news websites in each EU country in terms of visitors. The number of outlets per country was weighted based on the member states' population. The name of the index being used is [pi_ee24campaign].

Keyword-based article selection

Articles relevant to the topic are selected by searching for a pre-established set of relevant combined keywords that appear in the title or text of the article. To obtain the selected keywords in all official EU languages for this report, we translated the keywords automatically from English to all other EU24 languages by using the European Commission's E-Translation services.

Automated Clustering

Clusters are computed by a community-based algorithm on the original language titles of the articles and are based on semantic similarity. Stopwords were not removed from the titles but are not shown in the keyword section. The model transforms sentences into language-independent vectors, suiting our need to compare articles coming from the various EU countries. It is also easy to use and easy to deploy thanks to its python implementation.

In order to facilitate the analysis, automatically translated English titles are generated. The minimum cluster size is set to 5 articles. The maximum cluster size is not fixed. The colours of the bubbles represent the cluster, the size of the cluster has no meaning, it is purely a graphical solution since many of them are concentric or very close and otherwise wouldn't be visible. Distance from other items is representing semantic similarity. The clustering algorithm can process maximum 10.000 articles per input combination. Depending on the distribution of its neighbours, a document may belong to a cluster or not. In other words, not all articles can be clustered.

Automated framing detection

Automated detection of framing dimensions constituted an extension of our analysis conducted in this report. This work was based on in-house machine learning classifiers that detect framing dimensions in the articles. The machine-learning algorithm trained on a multilingual corpus.

Framing refers to the perspective under which an issue or a piece of news is presented. We consider 14 frames: *Economic, Capacity and resources, Morality, Fairness and equality, Legality, constitutionality and jurisprudence, Policy prescription and evaluation, Crime and punishment, Security and defence, Health and safety, Quality of life, Cultural identity, Public opinion, Political, External regulation and reputation.*

For more information see the JRC Technical Report: News Categorization, Framing and Persuasion Techniques: Annotation Guidelines: https://knowledge4policy.ec.europa.eu/text-mining/news-categorization-framing-persuasion-techniques-annotation-guidelines_en

Named entities recognition

Names Entities Recognition has been done using the NEROne module which has been developed in the context of the Europe Media Monitoring project. For detailed information, see: *Guillaume Jacquet, Jakub Piskorski, and Sophie Chesney. 2019. Out-of-context fine-grained multi-word entity classification. In Proceedings of the 34th ACM/SIGAPP Symposium On Applied Computing (SAC 2019).*

Annex II - About the authors

Text Mining & Analysis Competence Centre (TMA-CC)

DG JRC's Text Mining & Analysis Competence Centre (TMA-CC) of the European Commission is specialised in making sense of vast amounts of text through computing and analytics. It is an in-house consultancy and innovation service supporting EU Institution's policymakers, investigators and analysts in their knowledge-intensive tasks by providing consultancy and advanced analytical tools in the field of text mining.

TMA-CC aims to be an incubator for text mining ideas and solutions within the European Commission – a single focal point where expertise is concentrated, then shared and applied to the benefit of its clients.

Please check https://knowledge4policy.ec.europa.eu/text-mining_en for more information on TMA-CC.

European Science and Media Hub

The European Science-Media Hub (ESMH), operating under the political responsibility of the European Parliament Panel for the Future of Science and Technology (STOA), is a platform to promote networking, training and knowledge sharing between the European Parliament, the scientific community and the media.

The ESMH creates a network among policy-makers, scientists and media involving science, academia, educational and research entities, and professional associations of journalists and scientists. For journalists and media representatives, the ESMH organises training sessions and workshops on current technological developments, both as subjects of their reporting and as means of facilitating their work. Via media monitoring and media intelligence tools, the ESMH follows the most popular topics in the field of science and technology on different platforms including journals, newspapers and social media.

The ESMH makes information available to journalists, other media and citizens about new scientific developments, as well as about scientific topics that attract media attention, and promotes information based on evidence.

Check <https://sciencemediahub.eu/> for more information, methodology and technology.

For an overview of all science media intelligence reports produced in collaboration between TMA-CC and ESMH, please check <https://sciencemediahub.eu/science-media-intelligence-reports/>.

Europe Media Monitor

Europe Media Monitor (EMM) is a tool developed and maintained by the Text and Data Mining Unit of the Joint Research Centre (JRC) of the European Commission. The main purpose of EMM is to provide monitoring of a large set of online media, reducing the information flow to manageable proportions by clustering related news, categorising articles and applying Language Technology tools to derive further metadata, such as recognising and disambiguating entities in the text, extracting quotes by and about people, applying sentiment/tonality analysis and more. Since 2018, EMM can also grab content from live Twitter streams.

A lot of EMM's functionalities are freely available. To access the tool that best fits your need, please check https://knowledge4policy.ec.europa.eu/text-mining/topic/europe-media-monitor-emm_en

GETTING IN TOUCH WITH THE EU

In person

All over the European Union there are hundreds of Europe Direct centres. You can find the address of the centre nearest you online (european-union.europa.eu/contact-eu/meet-us_en).

On the phone or in writing

Europe Direct is a service that answers your questions about the European Union. You can contact this service:

- by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),
- at the following standard number: +32 22999696,
- via the following form: european-union.europa.eu/contact-eu/write-us_en.

FINDING INFORMATION ABOUT THE EU

Online

Information about the European Union in all the official languages of the EU is available on the Europa website (european-union.europa.eu).

EU publications

You can view or order EU publications at op.europa.eu/en/publications. Multiple copies of free publications can be obtained by contacting Europe Direct or your local documentation centre (european-union.europa.eu/contact-eu/meet-us_en).

EU law and related documents

For access to legal information from the EU, including all EU law since 1951 in all the official language versions, go to EUR-Lex (eur-lex.europa.eu).

Open data from the EU

The portal data.europa.eu provides access to open datasets from the EU institutions, bodies and agencies. These can be downloaded and reused for free, for both commercial and non-commercial purposes. The portal also provides access to a wealth of datasets from European countries.

