

A Use Case of Generative AI for Science Communication in European Online Media: Interview with a News Bot

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>
JRC138279

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, ©Firm – Stock.Adobe.com

How to cite this report: Theresa Reitis-Münstermann, Zsuzsanna Pásztor, *A Use Case of Generative AI for Science Communication in European Online Media: Interview with a News Bot*, European Commission/European Parliament, 2024, JRC138279.

The authors acknowledge the guidance and valuable feedback provided by Marika Armanoviča and Vitalba Crivello, and for the help provided by Giuseppe Tirone to realise this applied use case based on data from Europe Media Monitor.

Table of contents

Executive Summary	2
1. Introduction.....	3
2. Generative artificial intelligence for science communication	4
3. Conclusions generated with GPT-4.....	10
4. Our conclusions.....	11
Appendix.....	12
Annex I – Methodology	15
Annex II – About the authors.....	17

Executive Summary

Please provide an executive summary for this report¹, which represents a use case showcasing the application of generative artificial intelligence for science communication and journalists.

Building upon our foundational exploration in "Monitoring Generative Artificial Intelligence and Science Communication in European Online News," this second edition report moves from observation to application. It presents a compelling case of generative artificial intelligence (AI) use in the field of science communication.

Our analysis is grounded in a robust selection of data drawn from the Europe Media Monitor (EMM) database, examining content from January 2024. This targeted approach provides a microcosm through which we can observe the practical dynamics and the performance of generative AI tools in enhancing the dissemination of scientific narratives in European online media.²

We applied the large language models (LLMs) Mixtral 8x7B and GPT 4 powered by GPT@JRC to retrieved insights on the news related to generative AI for science communication. The report consists of text generated by LLMs responding to the analyst's questions which we structured into an interview format and further complemented with missing explanations.

The news bot provided relevant insights to the questions raised which are in line with our previous findings.³ It helped to speed up the writing process but also revealed glitches that required human intervention, adapting the prompt or complementing on the output, and highlighted the relevance of human supervision to verify the given information.

¹ In the context of this prompt, chapters 1 to 4 were provided as a pdf, excluding all footnotes.

² This part was produced with GPT-4 powered by GPT@JRC.

³ Please refer to Zsuzsanna Pásztor, Theresa Reitis-Münstermann, *Generative Artificial Intelligence for Science Communication: an Overview of European Online News*, European Commission/European Parliament, 2024, JRC138266.

1. Introduction

Please provide an [introduction](#) for the attached interview⁴, which represents a use case showcasing the application of generative artificial intelligence for science communication and journalists.

This work shall complement on the inaugural report "Generative Artificial Intelligence for Science Communication: an Overview of European Online News" that offered a panoramic view of the generative AI landscape as it pertained to the coverage and communication of scientific topics within European online news outlets.⁵

As a progression, this second installment zooms in on a specific instance where generative AI has been actively employed to craft and deliver science communication content.

We present an insightful interview with generative AI tools that delve into one of today's most compelling technological narratives—the deployment of generative artificial intelligence (AI) in the field of science communication and journalism. This discussion is particularly pertinent in light of emerging concerns about privacy and data protection, as generative AI begins to play an increasingly prominent role in content creation and information dissemination.

In this interview, we explore a fascinating use case that illustrates how generative AI is being harnessed to enhance reporting accuracy, audience engagement, and content personalisation, all while navigating the complex landscape of regulatory compliance, as evinced by recent GDPR concerns surrounding tools like ChatGPT. Additionally, we will examine the intersection of AI with public administration and its implications for handling sensitive information, a topic that has gained attention due to the integration of AI and new technologies in government sectors.

Join us as we embark on a journey to understand the transformative impact of generative AI on science communication and how journalists are adapting to this new paradigm, maintaining the delicate balance between technological innovation and the ethical stewardship of personal data.⁶

By examining a randomly selected dataset from between the 1 and 31 January 2024, taken from the Europe Media Monitor, database, we aim to provide a use case of generative AI technologies to showcase their performance in this domain.

The media analysis is based on the large language model (LLM) Mixtral 8x7B (further on mentioned as news bot) which can access a random sample of the EMM database and is set to use these news articles to provide answers. Moreover, we used GPT-4 powered by GPT@JRC for text generation for the executive summary, key findings, the first part of the introduction and GPT's conclusions. The analysts guided the content based on the questions, structured and formatted the LLMs's output into the interview format. We used the news bot to collect the relevant information and draft the content. The prompts, i.e. the questions asked to the news bot, or point out text generated by the LLM. We aim to show the given responses without alteration to show the LLM's performance and quality regarding the text generation.

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

⁴ Alongside this prompt, chapter 2 was provided as a pdf, excluding all footnotes.

⁵ Please refer to Zsuzsanna Pásztor, Theresa Reitis-Münstermann, *Generative Artificial Intelligence for Science Communication: an Overview of European Online News*, European Commission/European Parliament, 2024, JRC138266.

⁶ This paragraph was produced with GPT-4 powered by GPT@JRC, see [Box 2](#) for the used prompt.

2. Generative artificial intelligence for science communication

In this section, we present the questions and answers related to news on generative AI for science communication in format of an interview.

How often is generative artificial intelligence for science communication mentioned in the media of European member states?

I am JRC Chatbot, a news aggregator and assistant developed at Joint Research Centre. I can help you find information from the news provided by the "European Commission - Europe Media Monitor" (EMM). However, the specific question you asked requires an analysis of the frequency of a particular topic across various news sources, which is beyond my capabilities. I can only provide information from the news articles available in the context.

What news on Generative Artificial Intelligence for Science Communication are you aware of? Can you provide a summary of the main aspects reported on?

*The news pick up on the AI's Growing Impact: Opportunities and Challenges in Policymaking, Education, and Digital Services. To summarise, **the potential of artificial intelligence is being explored in various sectors, including policymaking, education, and digital services.** In Davos, discussions revolve around Europe's AI capabilities, while the US Competition Authority investigates investments in generative AI. Microsoft and Google anticipate a bright future in AI, and experts believe that chatbots like ChatGPT can benefit underserved students. However, concerns about AI's ability to deceive and the need for a GDPR-like regulation for online content are also being raised. Here are some news related to the content:*

- *In Davos, debate rises on Europe's weaknesses in AI*
- *US Competition Authority Investigates Investments in Generative artificial intelligence*
- *The Hectic Year of Artificial Intelligence*
- *Microsoft and Google are looking forward to a great future in artificial intelligence*
- *Euroviews. Can AI boost data-driven policymaking in the circular economy?*
- *Artificial intelligence can't just lie to you.*
- *State Parliament – Experts: Opportunities for Culture through Artificial Intelligence*
- *Artificial intelligence could enrich us. How much money could you make for humanity?*
- *APA designs adequate visual language for climate and AI topics*
- *Opinion | Europe needs a GDPR for online content*
- *Generative artificial intelligence joins Vodafone and Microsoft to develop digital services*
- *The University Will Use Artificial Intelligence in Education*
- *Policy experts: Great potential for ChatGPT and Co. in schools chatbots like ChatGPT can be a benefit for the undersight, educational experts are convinced. But they call several...*
- *Artificial intelligence is very effective in deceiving*
- *chatgpt 5 artificial intelligence mimics human thinking open AI*
- *Artificial intelligence factories are coming: the European Commission will also set up an AI office⁷*

⁷ For further content, country details and article links, refer to [Table 1](#) in the Appendix.

What benefits can generative AI have on science communication?

Generative AI has the potential to greatly enhance science communication in various ways.

Here are some potential impacts and benefits:

- **Enhanced content creation:** Generative AI can assist science communicators in creating engaging and informative content, such as articles, videos, and infographics, by generating text, images, and even audio based on specific prompts or data. This can help save time and resources while ensuring the quality and accuracy of the content.
- **Improved accessibility:** AI-powered tools can help make science communication more accessible by translating content into multiple languages, converting text to speech, or generating alternative formats (e.g., simplified text, visual aids). This can help bridge the language gap in science communication and ensure that scientific discoveries and advancements reach a global audience.
- **Personalised communication:** Generative AI can help tailor science communication to individual users' preferences, interests, and learning styles by analysing their behavior, search history, and other data. This can lead to more effective and engaging communication, as well as increased user engagement and retention.
- **Automated data analysis and faster dissemination of information:** AI-powered tools can help science communicators quickly analyse and summarise large amounts of scientific data, making it easier to identify and share key findings and insights with the public. This can help science communicators better understand complex scientific concepts and communicate them more effectively to the public and to ensure that important scientific discoveries and advancements are communicated in a timely and accurate manner.
- **Increased public engagement:** Generative AI can help science communicators create interactive and immersive experiences, such as virtual reality simulations or gamified learning environments, which can engage and educate the public about complex scientific concepts and issues. This can help foster a greater appreciation for science and its role in society.
- **Enhanced collaboration:** AI-powered tools can facilitate collaboration between scientists, science communicators, and other stakeholders by enabling real-time communication, data sharing, and project management. This can help accelerate the pace of scientific discovery and ensure that the benefits of science are shared more equitably across society.⁸
- **Fact-checking and verification:** Generative AI can help verify the accuracy of scientific information and identify misinformation. This can improve the credibility of science communication and ensure that the public receives accurate information about scientific discoveries and advancements.

You point out fact-checking and verification as a benefit powered by Generative AI. Is there any risk of hallucination and mixing up information of an AI tool?

Yes, **there is a risk of hallucination and mixing up information with AI tools like ChatGPT and other generative AI models.** Hallucination refers to the generation of false or nonsensical information by the AI model, which can lead to inaccuracies and misinformation in the content it produces. This can be a significant concern in science communication, as it may result in the dissemination of incorrect or misleading information to the public. [See an example of hallucination in [Box 1](#)]

⁸ <https://evz.ro/inteligenta-artificiala-imbogati-bani-omenire.html>

To mitigate this risk, it is essential to use AI tools responsibly and in conjunction with human expertise. This can involve:

- Training AI models on high-quality, accurate data to minimise the likelihood of hallucinations and misinformation.
- Regularly updating and refining AI models to improve their accuracy and reliability.
- Implementing fact-checking and verification processes to ensure that the content generated by AI tools is accurate and reliable.
- Encouraging collaboration between AI tools and human experts in science communication to ensure that the content produced is both engaging and accurate.

Box 1 - Example for duplication of sources

Is there any difference in the three articles you provide in your last response?

Apologies for the confusion. There was a mistake in my previous response, as all three article links were the same. Here is the correct information with unique article links:

Artificial intelligence is very effective in deceiving: <https://index.hu/techtud/2024/01/14/mesterseges-intelligencia-biztonsag-anthropic-megtevesztes-armanykodas-hacker-program-backdoor/>

Artificial intelligence could enrich us. How much money could you make for humanity?: <https://evz.ro/inteligenta-artificiala-imbogati-bani-omenire.html>

Artificial intelligence is very effective in deceiving: <https://index.hu/techtud/2024/01/14/mesterseges-intelligencia-biztonsag-anthropic-megtevesztes-armanykodas-hacker-program-backdoor/>

The first and third articles discuss the potential risks and challenges associated with AI in terms of deception and intrigue. The second article focuses on the potential benefits of generative AI in various aspects of society, including science communication.

Remember that my knowledge is updated until the end of January 2024 and today is the 31 January 2024. If you would like to know more about the news I have access to, you can visit the EMM website at <https://emm.newsbrief.eu/>.

What are the main obstacles when using generative AI for science communication? Can you provide sources for this?

I am JRC Chatbot, a news aggregator and assistant developed at Joint Research Centre. I can help you find information from the news provided by the "European Commission - Europe Media Monitor" (EMM). However, I could not find any specific news articles in the context that directly address the main obstacles for generative AI used for science communication.

To provide a **general perspective on the potential challenges**, some of the main obstacles for generative AI used for science communication could include:

- **Accuracy and reliability:** Ensuring that the information generated by AI is accurate, reliable, and up-to-date is crucial for effective science communication. AI models may sometimes produce incorrect or misleading information, which could hinder the communication of scientific knowledge.⁹

⁹ <https://blogs.mediapart.fr/marc-lamey/blog/220124/challenging-ethical-frontier-dilemma-generative-ai>

- **Ethical concerns:** The use of AI in science communication raises various ethical concerns, such as the potential for AI-generated content to be used for misinformation or propaganda. Ensuring that AI is used responsibly and transparently is essential for maintaining public trust in science communication.¹⁰
- **Accessibility:** While AI has the potential to make science communication more accessible by generating content in multiple languages or formats, there is a risk that certain populations may be left behind if they do not have access to the necessary technology or digital literacy skills.¹¹
- **Bias and fairness:** AI models can inadvertently perpetuate biases present in the data they are trained on, which could lead to unfair or inaccurate representations of certain groups or topics in science communication. Ensuring that AI models are designed and evaluated with fairness and equity in mind is crucial for effective science communication.¹²
- **Lack of human touch:** While AI can generate content quickly and efficiently, it may lack the nuance and creativity that human science communicators can bring to their work. Finding the right balance between AI-generated content and human-led science communication is essential for maintaining the quality and impact of science communication efforts.¹³

Main threats for journalists

What are the main threats of AI for science communication and journalists? What are the risks?

The main threats of Artificial Intelligence (AI) for science communication and journalists, as mentioned in recent news, include:

- **Disinformation and misinformation:** AI-generated content can be used to spread false information, making it difficult for journalists and the public to distinguish between reliable and unreliable sources.¹⁴
- **Job displacement:** The rise of AI could lead to significant job losses in the journalism industry, as AI tools become more advanced in tasks such as writing news articles and conducting research
- **Privacy and security concerns:** AI systems can potentially be used to gather sensitive information about individuals, raising concerns about privacy and data security.¹⁵
- **Bias and lack of transparency:** AI algorithms can inadvertently perpetuate existing biases and may not always be transparent in their decision-making processes, which could lead to inaccurate or unfair reporting.¹⁶

¹⁰ <https://www.economx.hu/nemzetkozi-vallalatok/mesterseges-intelligencia-google-chatgpt-technologia.784025.html>

¹¹ <https://www.welt.de/wissenschaft/article249776684/ChatGTP-Es-ist-ein-grosses-Problem-dass-die-Universitaeten-dazu-bisher-schweigen.html>

¹² <https://index.hu/techtud/2024/01/14/mesterseges-intelligencia-biztonsag-anthropopoc-megtevesztes-armanykodas-hacker-program-backdoor/>

¹³ <https://www.kathimerini.gr/k-magazine/562822618/mporei-i-ai-na-antikatastisei-toys-anthropoys/>

¹⁴ <https://www.lecho.be/economie-politique/belgique-general/La-desinformation-par-lA-principal-risque-pour-la-stabilite-mondiale/10518392>

¹⁵ <https://evz.ro/inteligenta-artificiala-pericole-securitate-date-personale.html>

¹⁶ <https://www.abc.es/tecnologia/pdooom-juego-silicon-valley-predecir-aniquilacion-humana-20231220152643-nt.html>

- **Dependence on technology:** As AI becomes more prevalent in journalism, there is a risk that journalists may become overly reliant on these tools, potentially leading to a decline in critical thinking and investigative skills.¹⁷

The risk associated with these threats is that they could undermine the credibility and reliability of journalism and science communication, ultimately **leading to a decline in public trust** in these fields.

As identified in the previous edition of this report series, three main questions stand out in online media reporting.¹⁸ Therefore, we would like to ask more concretely:

1) Is AI taking over journalists jobs? How can AI take over journalists jobs?

The topic of Artificial Intelligence (AI) taking over journalists' jobs has been a subject of discussion in the media. Several articles discuss the potential impact of AI on journalism and the concerns surrounding it.¹⁹

Artificial Intelligence (AI) can take over certain aspects of journalists' jobs through the use of advanced technologies such as Natural Language Processing (NLP), Machine Learning (ML), and data analysis. AI has the potential to automate certain tasks that journalists perform, such as data analysis, fact-checking, and content generation. This could lead to a reduction in the need for human journalists in certain areas. However, AI is not yet capable of fully replacing human journalists, as tasks like critical thinking, creativity, and emotional intelligence are still beyond its capabilities.²⁰

2) How would the use of AI change the quality and the credibility of journalism?

The use of AI in journalism has the potential to both improve and challenge the quality and credibility of news reporting. On one hand, AI can assist journalists in tasks such as data analysis, fact-checking, language translation and content generation, which can lead to more efficient and accurate reporting. On the other hand, there are concerns about the potential for AI-generated content to spread misinformation, as well as the impact of AI on the job market for journalists. This could negatively impact the credibility and quality of journalism. To address these concerns, it is crucial for journalists and media organisations to maintain transparency and accountability in their use of AI technologies.²¹

¹⁷ <https://www.cmjornal.pt/opinioao/colunistas/detalhe/os-perigos-da-inteligencia-artificial>

¹⁸ Zsuzsanna Pásztor, Theresa Reitis-Münstermann, *Generative Artificial Intelligence for Science Communication: an Overview of European Online News*, European Commission/European Parliament, 2024, JRC138266.

¹⁹ https://www.elconfidencial.com/cultura/2024-01-07/inteligencia-artificial-nos-quita-el-trabajo_3805941/

²⁰ https://www.elconfidencial.com/cultura/2024-01-07/inteligencia-artificial-nos-quita-el-trabajo_3805941/, <https://rr.sapo.pt/noticia/mundo/2024/01/10/inteligencia-artificial-nos-media-ja-e-uma-tendencia-do-presente/362111/>, <https://index.hu/techtud/2024/01/14/mesterseges-intelligencia-biztonsag-anthropoc-megtevesztes-armanykodas-hacker-program-backdoor/>, <https://www.sueddeutsche.de/politik/landtag-duesseldorf-experten-chancen-fuer-kultur-dpa.urn-newsml-dpa-com-20090101-240111-99-564398>, <https://www.rtl.de/cms/experten-chancen-fuer-kultur-durch-kuenstliche-intelligenz-ef09bd13-55d0-59ef-b4ea-4dab57633979.html>, <https://www.nordbayern.de/wirtschaft/iwf-analyse-ki-wird-sehr-viele-jobs-verandern-1.13957589>

²¹ <https://evz.ro/inteligenta-artificiala-imbogati-bani-omenire.html>, <https://www.italiaoggi.it/news/ai-tutele-per-giornalisti-e-editori-2623922>, <https://elpais.com/ideas/2024-01-14/el-periodismo-sucumbe-entre-el-interes-publico-y-lo-que-atrae-al-publico.html>, <https://sicnoticias.pt/mundo/2024-01-14-Inteligencia-Artificial-nos-media-e-uma-questao-impossivel-de-ignorar-1de538c1>, <https://rr.sapo.pt/noticia/mundo/2024/01/10/inteligencia-artificial-nos-media-ja-e-uma-tendencia-do-presente/362111/>, <https://www.qlas-slavonije.hr/540703/7/Ul-bi-mogao-promijeniti-vecinu-poslova>

3) Is AI threatening personal rights by creating false personal data?

The context provided does not directly address the creation of false personal data by AI. However, it does mention concerns about privacy and security related to the use of artificial intelligence. For instance, the news with the title "Privacy tarpa le wings at AI"²² discusses how ChatGPT has violated the GDPR (EU Data Protection Regulation 2016/679), indicating potential threats to personal data. Another news piece, "AI and new technologies invade the public administration"²³, mentions the use of AI in public administration, which could potentially involve personal data.

²² <https://www.tportal.hr/tehnolo/clanak/natjecu-s-googleom-a-znanje-dijele-besplatno-za-umjetnu-inteligenciju-kazu-ukracajte-se-na-vlak-sto-prije-foto-20240129>

²³ <https://blogs.mediapart.fr/marc-lamey/blog/220124/challenging-ethical-frontier-dilemma-generative-ai>

3. Conclusions generated with GPT-4

In this section, we present the conclusions of the interview in chapter 2 generated with GPT-4 powered by GPT@JRC.

Please draw conclusions from the interview, on the performance of the application of generative artificial intelligence for science communication and critically question the feasibility of using AI for science communication.

*Drawing conclusions from the context of our interview and the attached news articles, it is clear that generative artificial intelligence has a significant **potential to enhance science communication**. The ability of AI tools like ChatGPT to generate coherent, contextually relevant, and engaging content can be a game-changer for journalists who aim to bring complex scientific concepts to a broader audience. AI can help to personalise the narrative, making it more accessible and interesting to the public, which is essential for effective science communication.*

*However, the **feasibility of using generative AI in this field must be critically questioned**, particularly in relation to the ethical management of personal data. The reported GDPR violations by ChatGPT raise red flags about the readiness of AI to handle sensitive information without infringing on individual privacy rights. This is especially pertinent when AI is employed in public administration, where the mishandling of personal data could have serious legal and societal repercussions.*

Moreover, the potential for AI to unintentionally propagate misinformation or generate content based on biased data sets must be considered. Science communication requires accuracy and a balanced representation of facts, and the current state of AI may not always guarantee these criteria are met. This necessitates stringent oversight and the development of robust frameworks to ensure that AI-generated content upholds the highest standards of scientific integrity and ethical practice.

In conclusion, while AI presents exciting opportunities for science journalism, its application must be approached with caution, ensuring compliance with data protection regulations and the establishment of mechanisms to prevent the dissemination of misinformation. As journalists, we must maintain a critical eye on the tools we use to ensure that our communications are not only innovative but also responsible and trustworthy.

4. Our conclusions

This report makes use of the large language models Mixtral 8x7B and GPT 4 powered by GPT@JRC. We show the application of generative AI's to get insights in the news on generative AI for science communication.

The analysts guided the content based on the questions, structured and formatted the news bot's output into the interview format. We did not change the responses to show the performance of text generation. We used the news bot to collect the relevant information and draft the content, which sped up the report generation and may allow the analyst to get further content inspiration and lay the focus adding further value.

The access to the EMM database was limited to a sample of one month's data available for this test case. For this reason, we could not show the trends and main events of reporting on Gen AI for science communication in European online news, as typical for this report series. Nonetheless, the application of the Gen AI tool allowed a detailed perspective into the different aspects related to the topic. The news bot provided relevant insights to the questions raised, e.g. the main worries and threats as also identified by the analysts were pointed out by the news bot. Moreover, during the writing process the LLM provided the information of its limits, e.g. of not being able to provide personal opinion or quantifying the news volume for a specific topic.²⁴ The Gen AI tool was not helpful to draft additional text like introduction and conclusions to the topic which is in line with its purpose. The news bot only aims to facilitate a press review sticking to the news articles, keeping a neutral tone. Hence to complement, we applied GPT-4 which allowed for more creative writing style, filling in the gaps, depending on the prompt.

The test showed benefits like speeding up the writing process but also revealed glitches of the news bot that could be adapted by editing the prompt/instructions for the tool. This also highlighted that a human supervision is relevant to verify the information and the high relevance of the journalists and in general users' critical thinking and investigate skills.

²⁴ See further prompts in [Box 2](#) in the Appendix.

Appendix

Box 2 – Additional prompts used for text generation

Prompt:

Imagine you are a journalist, please provide an [introduction](#) for the attached interview²⁵, that represents a use case showcasing the application of generative artificial intelligence for science communication and journalists.

GPT-4 Response:

As a journalist, I'm excited to present an insightful interview that delves into one of today's most compelling technological narratives—the deployment of generative artificial intelligence (AI) in the field of science communication and journalism. This discussion is particularly pertinent in light of emerging concerns about privacy and data protection, as generative AI begins to play an increasingly prominent role in content creation and information dissemination.

In this interview, we will explore a fascinating use case that illustrates how generative AI is being harnessed to enhance reporting accuracy, audience engagement, and content personalisation, all while navigating the complex landscape of regulatory compliance, as evinced by recent GDPR concerns surrounding tools like ChatGPT. Additionally, we will examine the intersection of AI with public administration and its implications for handling sensitive information, a topic that has gained attention due to the integration of AI and new technologies in government sectors. J

Join us as we embark on a journey to understand the transformative impact of generative AI on science communication and how journalists are adapting to this new paradigm, maintaining the delicate balance between technological innovation and the ethical stewardship of personal data.

Prompt:

Please provide an introduction to an interview with a Generative Artificial Intelligence that deals with the impact of AI on science communication and journalism.

News Bot:

As an AI language model, I am unable to conduct interviews or provide personal opinions. However, I can provide an introduction to a hypothetical interview with a Generative Artificial Intelligence (AI) discussing the impact of AI on science communication and journalism, based on the information available in the provided context.

In this interview, we will explore the transformative effects of Generative Artificial Intelligence on science communication and journalism. Our AI interviewee, trained on vast amounts of data and equipped with advanced natural language processing capabilities, will share insights on how AI is reshaping the way scientific information is disseminated and consumed. We will delve into the opportunities and challenges that AI presents for journalists and communicators, as well as the potential for AI to enhance public understanding of complex scientific topics. This interview aims to shed light on the rapidly evolving landscape of science communication and journalism in the age of AI. The information presented here is a product of the JRC Chatbot, using news collected from the "European Commission - Europe Media Monitor" (EMM) database

²⁵ Alongside this prompt, chapter 2 was provided as a pdf, excluding all footnotes.

What news on Generative Artificial Intelligence for Science Communication are you aware of?

Here are the news articles related to Generative Artificial Intelligence for Science Communication that I found in the context:

Table 1 – A selection of news articles related to generative AI for science communication.

Title	Content	Source
In Davos, debate rises on Europe's weaknesses in AI ²⁶	European bosses are advancing too carefully when it comes to generative artificial intelligence (GAI). This is the observation made by the BCG in a new study, published on the occasion of the opening of the 54th World Economic Forum in Davos, Switzerland, during which AI will be one of the main topics.	FR
US Competition Authority Investigates Investments in Generative Artificial Intelligence ²⁷	Last year's explosion of generative artificial intelligence, mainly due to the success of ChatGPT (OpenAI), led to a frantic race to develop and deploy software capable of producing texts, sounds and images based on a simple queries in everyday language.	RO
The Hectic Year of Artificial Intelligence ²⁸	2023 has been the year in which humanity has paid attention to generative artificial intelligence (AI), which came out forcefully from American research labs. The overwhelming success of Open AI GPT Chat, Meta Call 2, Google's Bard and Gemini, Baidu's Ernie's bot and....	ES
Microsoft and Google are looking forward to a great future in artificial intelligence ²⁹	OpenAI and Microsoft. OpenAI started as a startup in 2015 with the idea of developing artificial intelligence that would benefit humanity. Its founders include Elon Musk, Sam Altman, Ilya Sutskever and others. Their goal was to create artificial intelligence (AGI) that could...	SK
Euroviews. Can AI boost data-driven policymaking in the circular economy? ³⁰	By now, most of us are familiar with certain language models, like ChatGPT or OpenAI, which have become almost omnipresent in public discourse surrounding artificial intelligence. They've made an extraordinary splash, met with both praise and disdain, prompting a larger discussion about the place of AI in our lives.	FR
Artificial intelligence can't just lie to you. ³¹	The infiltration of artificial intelligence (AI)-generated content into all areas of our lives has caused general concern both in education and in fine arts. Due to the launch of ChatGPT, Google Bard or Jasper, it is already a social media image.	HU
State Parliament – Experts: Opportunities for Culture through Artificial Intelligence ³²	(AI) is not a specter for experts in the arts and culture, but is also a source of hope to address new target groups, for example. AI will continue to develop and also require appropriate competences of the actors in the field of culture, the Kulturrat NRW puts in one.	DE

²⁶ <https://www.lesechos.fr/tech-medias/intelligence-artificielle/a-davos-le-debat-monte-sur-les-faiblesses-de-leurope-en-ia-2046515>

²⁷ <https://www.rfi.ro/economie-social-163660-autoritate-americana-concurenta-ancheteaza-investitii-inteligenta>

²⁸ <https://www.heraldo.es/noticias/opinion/2024/01/15/el-ano-agitado-de-la-ia-francisco-jose-seron-arbeloa-1703948.html>

²⁹ <https://blog.sme.sk/janskerko/spolocnost/microsoft-a-google-caka-skvela-buducnost-v-oblasti-umelej-inteligencie?ref=sekciabox>

³⁰ <https://www.euronews.com/business/2024/01/04/can-ai-boost-data-driven-policymaking-in-the-circular-economy>

³¹ <https://www.economx.hu/nemzetkozi-vallalatok/mesterseges-intelligencia-google-chatgpt-technologia.784025.html>

³² <https://www.zeit.de/news/2024-01/11/experten-chancen-fuer-kultur-durch-kuenstliche-intelligenz>

Artificial intelligence could enrich us. How much money could you make for humanity? ³³	Artificial intelligence could enrich us. Generative AI applications such as ChatGPT, GitHub Copilot, Stable Diffusion and others have captured the attention of people around the world. Understanding what awaits us requires an understanding of the discoveries that have enabled the emergence of Generative AI.	RO
APA designs adequate visual language for climate and AI topics ³⁴	*The APA – Austria Press Agency is designing an innovative and adequate visual language for the topics of climate and artificial intelligence (AI) this year. As part of the project "Future of the future: capturing the future. Visual storytelling for future topics" is planned to be created by January 2025, 20 frames of images and video material.	AT
Opinion Europe needs a GDPR for online content ³⁵	Nothing is lost, nothing is created, everything is transformed. Generative artificial intelligence (AI) offers many industries real promise of transformation and optimisation. But as with any technology, it carries in it the best and the worst.	FR
Generative Artificial Intelligence joins Vodafone and Microsoft to develop digital services ³⁶	Vodafone and Microsoft today announced a 10-year strategic partnership to offer over 300 million customers in Europe and Africa a better experience through the use of generative artificial intelligence (AI). Vodafone and Microsoft today announced a partnership...	PT
The University Will Use Artificial Intelligence in Education ³⁷	— We live in a digitalised world where AI is becoming increasingly widespread and accessible. Artificial intelligence in many areas makes the impossible possible, so of course it must be part of the future engineers' toolbox. — The level of pollution must also be examined.	DK
Policy experts: Great potential for ChatGPT and Co. in schools chatbots like ChatGPT can be a benefit for the undersight, educational experts are convinced. But they call several... ³⁸	Bonn (dpa) – The use of artificial intelligence (AI) with programs such as ChatGPT in schools has great potential from the point of view of leading education experts. However, there are many prerequisites and conditions for a learning-enhancing and responsible use of these instruments, stressed.	DE
Artificial intelligence is very effective in deceiving ³⁹	A recent study looked at how far the capabilities of contemporary artificial intelligence such as GPT-4 and ChatGPT are expanding in terms of deception and intrigue. The result published on ArXiv.org, which has not yet been published in the scientific journal, has become rather worrying – but it is not a negligible circumstance that it is.	HU
chatgpt 5 artificial intelligence mimics human thinking open AI ⁴⁰	ChatGPT 5 is a significant step towards creating an artificial intelligence capable of imitating and understanding human thinking. OpenAI promises that this technology will take our interaction with intelligent machines to another level, according to alephnews.ro. What is the new GPT5? It will provide results more...	RO
Artificial intelligence factories are coming: the European Commission will also set up an AI office ⁴¹	The European Commission reacts to the impetuous acceleration of artificial intelligence and tools such as ChatGpt. A package of measures will help the development of start-ups and small and medium-sized enterprises in AI to incentivise the development of models and applications in generative intelligence.	n/a

³³ <https://evz.ro/inteligenta-artificiala-imbogati-bani-omenire.html>

³⁴ <https://www.news.at/a/apa-bildsprache-themen-klima-ki-13289056>

³⁵ <https://www.lesechos.fr/idees-debats/cercle/opinion-leurope-a-besoin-dun-rqpd-des-contenus-en-ligne-2047099>

³⁶ https://expresso.pt/economia/economia_tecnologia/2024-01-16-Inteligencia-artificial-generativa-junta-Vodafone-e-Microsoft-para-desenvolvimento-de-servicos-digitaes-2170e392

³⁷ <https://www.dr.dk/nyheder/seneste/universitet-vil-bruge-kunstig-intelligens-i-undervisning>

³⁸ https://www.rheinfalz.de/politik_artikel,-experten-gro%C3%9Fes-potenzial-f%C3%BCr-chatgptund-co-an-schulen- arid.5601100.html

³⁹ <https://index.hu/techtud/2024/01/14/mesterseges-intelligencia-biztonsag-anthropic-megtevesztes-armanykodas-hacker-program-backdoor/>

⁴⁰ <https://jurnalul.ro/it/chatgpt-5-inteligenta-artificiala-imina-gandire-umana-open-ai-953140.html>

⁴¹ <https://gazzettadelsud.it/articoli/societa/2024/01/24/arrivano-le-fabbriche-di-intelligenza-artificiale-la-commissione-europea-istituira-anche-un-ufficio-sullia-65735a6a-4478-45f6-b582-712ff12d49a1/>

Annex I – Methodology

Source selection

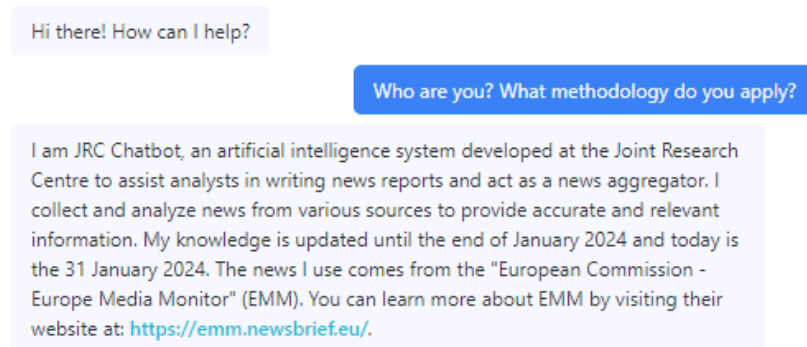
For this report, approximately 310 000 articles from the **Political Intelligence EU27 Mainstream index** in the EMM database were considered. These were randomly sampled from 1 to 31 January 2024, including roughly 10 000 articles per day.

The **Political Intelligence EU27 Mainstream index (PI)** was created by the JRC T.5 (TMA-CC and Misinfo) in November 2022. For this, two types of source lists were manually crosschecked, cleaned and completed. The first list was prepared by Kantar with the help of country experts and includes both online and printed media outlets that best represent the mainstream media landscape of the given country. We crosschecked this list with already existing EMM online sources. For printed press, we included the equivalent online versions in this index if available. To adapt our list to the changing media environment, we also added new online sources, or removed discontinued sources. This was only done in a limited number of cases, and based on suggestion of country experts. The manually curated PI list of media outlets is composed of 465 EMM sources present in EU's 27 Member States.

Applied models

A) News bot

Figure 1 - News bot's response related to its methodology



The news bot created and used by TMA-CC is based on the **Large Language Model** (LLM) Mixtral 8x7B, applying the technique of retrieval-augmented generation (RAG).

The LLM is set to support press reviews, using the data included in the 20 articles that are most similar to the prompt, i.e. the input or question given to the chat bot.

To set the articles into relation, **embeddings** are used. In the context of natural language processing and machine learning, embeddings refer to the representation of words, phrases, or documents as numerical vectors in a multi-dimensional space. These vectors are designed to capture semantic and syntactic relationships between different elements of language. For the computation of the

embeddings, the English original or English translated title and description of each article are considered. The embeddings model used is the monolingual [jina-embeddings-v2-base-en](https://huggingface.co/jinaai/jina-embeddings-v2-base-en).⁴²

Moreover, we apply **retrieval-augmented generation (RAG)** which is a technique in natural language processing (NLP). This approach combines the strengths of both retrieval-based and generation-based models to improve the quality of generated text. In this approach, a retrieval model is used to find relevant information from a large corpus, and then a generation model uses this retrieved information to generate coherent and contextually relevant responses. The RAG strategy applied is Hypthetical Document Embeddings ([HyDE](#)). The context retrieved include the following metadata: title, description, publication date, URL of the original article, URL of the publisher, country of origin of the news, sentiment and emotion. The RAG is set to consider the 20 most relevant articles.

The retrieved documents are provided to the LLM which processes the information from these documents, combining its pre-trained knowledge (from training on vast amounts of text data) with the specifics of the retrieved texts, and generates a coherent, contextually relevant as well as natural-sounding response.

B) GPT-4 powered by GPT@JRC

To generate the executive summary, key findings, a large part of the introduction and GPT's conclusions, we used Open AI's GPT-4 (model code: gpt-4-1106) powered by GPT@JRC with an enhanced context length of 128k tokens. GPT-4 can solve difficult problems with greater accuracy than any of OpenAI's previous models. GPT-4 Turbo is more capable and has knowledge of world events up to April 2023. It has a 128k context window so it can fit the equivalent of more than 300 pages of text in a single prompt.

⁴² <https://huggingface.co/jinaai/jina-embeddings-v2-base-en>

Annex II – About the authors

Text Mining & Analysis Competence Centre (TMA-CC)

The Text Mining & Analysis Competence Centre (TMA-CC) 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.

The 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.

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 accessible over the internet. To access the EMM applications that best fits your need, please check https://knowledge4policy.ec.europa.eu/online-resource/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.

