BOOK OF ABSTRACTS

SCIENCE COMMUNICATION IN THE AGE OF ARTIFICIAL INTELLIGENCE

Annual Conference of the
"Science Communication" Division of the
German Communication Association

June 6-7, 2024

ETH Zurich Main Building
Rämistrasse 101, 8092 Zurich

Credit: shutterstock

Organized by:

Mike S. Schäfer, Sabrina H. Kessler, Daniela Mahl & Sophia C. Volk

Supported by:

Roche

Bilendi & respondi

DGPuK

SGKM SSCM SSCM SACM

Swiss National Science Foundation

ETH zürich

Table of contents

Parallel Panel I: AI DISCOURSES IN THE NEWS MEDIA

Thursday, June 6, 10:00 – 11:00
Chair: Friederike Hendriks
Room: HG E 1.2

Visualizations of invisible technologies. How German print media illustrate articles on artificial intelligence

Carolin Moser, Tabea Lüders, Melanie Leidecker-Sandmann

Introduction
Next year's conference of the science communication division of DGPuK deals with the topicof generative AI. One of the proposed perspectives submissions may focus on is analyses of "public communication about generative AI in legacy media" (CfP). This perspective is what our submission follows. We ask, how six national print media in Germany illustrate articles on artificial intelligence (AI) and whether in their figurative representation of AI recurring visual frames can be recognized.

Relevance
We consider this question to be relevant as pictures are an integral part of journalistic reporting (Renner 2013) and as the visual is a central part of social construction of reality (Lucht et al. 2013). Although AI is expected to be more and more integrated into our everyday lives, public's knowledge of AI is at best "patchy" (Nader et al. 2022), which is whyimages of AI may affect public perception of AI as well as expectations, fears or hopes about AI (Cave et al. 2018; Kong 2019). Although visual elements in media coverage have increased significantly since the 19th century and although research has shown that images are "powerful framing tools"[1] (Rodriguez & Dimitrova 2011), researchers complain the marginality of the image as a centralresearch object (Grittmann & Lobinger 2011; Schnettler & Bauernschmidt 2018).

Theoretical Background and State of Research
Theoretically, our analysis follows the visual framing theory. Despite an increasing number of framing studies, the "question of how issues are framed through images that stand alone or accompany text has remained relatively under-researched" (Rodriguez & Dimitrova 2011).
Visual framing is understood as a process in which specific images or aspects of images on atopic are selected, while others are neglected, which can imply a certain meaning or interpretation to the recipient (Geise & Lobinger 2015).
While textual communication about generative AI in news media has already been analyzed(e.g. Kieslich et al. 2022; Ouchchy et al. 2020; Sun et al. 2020; Vergeer 2020), images of AIin news coverage were – if at all – only recorded as additions. We are only aware of a singleconference proceeding that focuses on AI in news photographs in the U.S. and China (Kong2019). Beyond that, to our knowledge there are only few studies that analyze (mostly fictional) visual AI narratives in literature and film qualitatively (e.g. Cave et al. 2018; Herrmann 2018; Xanke & Bärenz 2012).

Method
We have conducted a quantitative visual content analysis (according to Grittmann & Lobinger, 2011) of illustrated German national print media coverage on AI between January 1 and December 31, 2019. We chose 2019 as period of analysis because it was named the 'Science Year of Artificial

[1] When textual and visual framing are in conflict, visual frames often win (Rodriguez & Dimitrova 2011).

Intelligence' by the Federal Ministry of Education and Research (BMBF 2021). It can therefore be assumed that the topic of AI was considerably covered during this year. As news media titles, we chose six (leading) national quality newspapers and news magazines in Germany, namely: Süddeutsche Zeitung (SZ), Frankfurter Allgemeine Zeitung (FAZ), Die Welt (DW), taz, Der Spiegel, and Die Zeit. All illustrated articles during the analysis period that contained the keywords 'artificial intelligence' or 'AI'[2] in their headline or subtitle (focus on articles that cover AI as a main topic) were selected for analysis. Our search resulted in 225 news articles, of which 127 were illustrated (56%) with n = 150 imagesin sum (some articles contained more than one image). For each image we captured categories (e.g. medium, publishing date, department, image genre etc.) as well as content categories. Per image, several image subjects (human, robot, computer etc.) with multiple specifying subcategories could be coded, for example if an image shows a human and a robotat the same time, as well as one image topic (again with multiple subcategories; e.g. man as role model, international competition, etc.)[3]. The intercoder reliability values (two coders) forthese variables ranged between 0.703-1 (Krippendorff's alpha).
Finally, a quantitative image type analysis according to Grittmann and Ammann (2011), which bundles all image subjects and topics with the same content, was used to identify visualframes in news media coverage.

Results

By far the most images of AI were published by the Süddeutsche Zeitung (n = 74), least by Der Spiegel (n = 7), most often in the economics department (29%), followed by the science department (18%) and feature sections (17%). The vast majority of AI images were photographs (67%), with illustrations the second most common (19%). Interestingly, AI is most often illustrated by pictures of humans (44%; see also Kong 2019), followed by robots (16%; however, only once or twice in Die Zeit, Der Spiegel and taz), computers (9%; relatively often in Der Spiegel), and cultural objects from arts, literature, or music (8%; relatively often in SZ and Die Zeit). Thematically, the images and associated articles most often dealt with specific applications of AI (29%) or portrayed human experts in AI discourse (25%). Novel AI inventions were also visualized relatively frequently (15%; e.g. smart home technologies). In sum, six visual frames were identified in news media coverage: potential uses of AI (43%; most often: taz, Die Welt, Die Zeit), hazard potential of AI (22%; most often: Der Spiegel), (international) competition (17%; most often: SZ, DW), cultural and artistic debate (6%; only DW, SZ), further development of AI (5%; most often: FAZ, taz), human role model (3%; most often: Die Zeit, FAZ, DW), and other (4%). Thus, news media coverage in 2019 seemed to paint a relatively differentiated picture of AI, with potential benefits of AI for humans appearing to be more in focus than potential risks (see also Kong 2019).

References

BMBF (2021). *Wissenschaftsjahr 2019*. https://www.wissenschaftsjahr.de/2019/indexb6b3.html?id=657 (December 27, 2021).

Cave, S., Craig, C., Dihal, K., Dillon, S., Montgomery, J., Singler, B., & Taylor, L. (2018).Portrayals and perceptions of AI and why they matter. *Apollo-University of Cambridge Repository*. https://royalsociety.org/-/media/policy/projects/ai-narratives/AI-narratives-workshop-findings.pdf (December 5, 2023).

Geise, S., & Lobinger, K. (2015). Zur Einleitung: Visual Framing als zentrales Forschungsfeld der Visuellen Kommunikationsforschung. In Geise, S., & Lobinger, K.(Eds.), *Visual Framing. Perspektiven und Herausforderungen der Visuellen Kommunikationsforschung* (pp. 9-18). Köln: Halem.

Grittmann, E., & Ammann, I. (2011). Quantitative Bildtypenanalyse. In Petersen, T., & Schwender, C. (Eds.), *Die Entschlüsselung der Bilder: Methoden zur Erforschung visuellerKommunikation: ein Handbuch* (pp. 163-178). Köln: Halem.

Grittmann, E., & Lobinger, K. (2011). Quantitative Bildinhaltsanalyse. In Petersen, T., & Schwender, C. (Eds.), *Die Entschlüsselung der Bilder: Methoden zur Erforschung visuellerKommunikation: ein Handbuch* (pp. 147-162). Köln: Halem.

Hermann, I. (2023). Artificial intelligence in fiction: between narratives and metaphors. *AI &Society*, 38, 319-329.

Kieslich, K., Došenović, P., & Marcinkowski, F. (2022). *Everything, but hardly any science fiction. A topic analysis of German*

[2] German search string: "Künstliche* Intelligenz" OR *KI*

[3] In sum, n = 218 image subjects have been coded on the top level (with n = 258 subcategories).

media coverage of AI. https://www.researchgate.net/publication/365033703_Everything_but_hardly_any_science_fiction (December 1, 2023).

Nader, K., Toprac, P., Scott, S., & Baker, S. (2022). Public understanding of artificial intelligence through entertainment media. *AI & Society*. https://doi.org/10.1007/s00146-022-01427-w

Ouchchy, L., Coin, A., & Dubljević, V. (2020). AI in the headlines: the portrayal of theethical issues of artificial intelligence in the media. *AI & Society*, 35, 927–936.

Sun, S., Zhai, Y., Shen, B., & Chen, Y. (2020). Newspaper coverage of artificial intelligence:A perspective of emerging technologies. *Telematics and Informatics*, 53, 101433.

Kong, Y. (2019). Artificial Intelligence in News Photographs: A Cross-Cultural Visual Content Analysis. *Proceedings of the 37th ACM International Conference on the Design ofCommunication*, Article Number 48, 1-3. https://dl.acm.org/doi/abs/10.1145/3328020.3353905

Lucht, P., Schmidt, L.-M., & Tuma, R. (Eds.) (2013). *Visuelles Wissen und Bilder des Sozialen: aktuelle Entwicklungen in der Soziologie des Visuellen*. Wiesbaden: VS Verlag für Sozialwissenschaften.

Renner, K. N. (2013). Journalistische Wirklichkeitserzählungen und fotografische Bilder. *Diegesis: Interdisziplinäres E-Journal für Erzählforschung*, 2, 33–50.

Rodriguez, L., & Dimitrova, D. (2011). The levels of visual framing. *Journal of VisualLiteracy*, 30(1), 48-65.

Schnettler, B., & Bauernschmidt, S. (2018). Bilder in Bewegung: Visualisierungen in der Wissenschaftskommunikation. In Müller, M. R., & Soeffner, H. G. (Eds.), *Das Bild als soziologisches Problem. Herausforderungen einer Theorie visueller Sozialkommunikation* (pp. 197-208), Weinheim: Beltz/Juventa.

Vergeer, M. (2020). Artificial Intelligence in the Dutch Press: An Analysis of Topics andTrends. *Communication Studies*, 71(3), 373-392.

Xanke, L., & Bärenz, E. (2012). Künstliche Intelligenz in Literatur und Film – Fiktion oderRealität? *Journal of New Frontiers in Spatial Concepts*, 4, 36-43.

AI coverage of legacy and alternative news media. Comparing framing and choice of sources

Markus Schug, Helena Bilandzic, Susanne Kinnebrock, Lena Sedlmeier

Artificial Intelligence (AI) is increasingly part of controversial societal and media discourses. As media coverage of AI technologies is "an important indicator of the central issues, actors, frames, and evaluations attached to a technology, and a critical arena where stakeholders negotiate future pathways for AI and its role in societies" (Brause et al., 2023, p. 2), it also became an important research object for communication scholars. However, existingstudies of AI media representations often focus only on established legacy media while neglecting other influential communicators as well as comparative approaches (Brause et al., 2023). In this paper, we argue that current phenomena such as counter-knowledge, alternative facts, the spread of pseudoscience (see Eslen-Ziya, 2022; Marwick & Partin, 2022; Merkley, 2020) and a growing "hyperaccessibility of expertise" (Brubaker, 2021, p. 75) require research on AI media representations to include 'alternative' news media and compare them to legacy media. According to Holt et al. (2019), alternative news media "represent a proclaimed and/or (self-)perceived corrective," pretending to report or re-narrate news disregarded by"mainstream" institutions (Boberg et al., 2020; Holt et al., 2019, p. 862). By comparing, we seek to better understand the public discourse on AI and acknowledge the remarkable role that alternative news media can play in some people's media repertoires on societally relevant topics(Viehmann et al., 2020).

We expect alternative news media to contradict the AI coverage of legacy media regarding framing and choice of sources (Bingaman et al., 2021; Chuan et al., 2019; Fast & Horvitz, 2017; Nisbet, 2009; Strekalova, 2015; Vicsek, 2011). We first ask how AI coverage inlegacy and alternative news media differs regarding risk and benefit framing (RQ1a) and the respective framing of time horizons (RQ1b), expecting alternative news media to focus more on risks than on benefits of AI and to frame societal consequences of AI rather in immediate than in distant future time horizons. Second, we ask in how far AI coverage in legacy and alternative news media differs regarding the choice of sources (RQ2), assuming that alternative news media use other sources than legacy media to substantiate their assessments.

We conducted a qualitative thematic content analysis (Kuckartz, 2018) of news articles derived from exemplary German legacy (*Der Spiegel, Die Welt, Die ZEIT*) and alternative newsmedia (*Junge Freiheit, NachDenkSeiten, Epoche Times, Tichys Einblick, Rublikon, Compact, Telepolis*). Both legacy and alternative media contain publications of divergent political positions (left- and right-leaning). The sampling was carried out in the spring of 2022, before ChatGPT went public. The sample covers a period of one year (2021), leaving us enough margin to build artificial weeks balancing seasonal fluctuations in media coverage (Frost & Carter, 2020). We used broad keywords (German translations for "AI", and "Artificial Intelligence") to scan the results for relevant articles focusing on societal influences of AI. We designed four artificial weeks (legacy media: half year I, half year II; alternative news media: half year I, half year II), with each week consisting of 24 articles representing four articles per journalistic day unit (Monday–Friday, weekend). Due to smaller fluctuations regarding periodic publication in alternative news media, $N = 93$ articles were included in the final sample. According to the technique suggested by Kuckartz (2018), we first determined main categories from the literature (deductive step) and used them to perform initial coding of the material. The text parts collected within those main categories were then split into subcategories, leading to a refinement of the category system. In a final coding step, the differentiated category system was used for a renewed coding of the entire material.

Results suggest that for RQ1a, risk framing is much more present in alternative news media as they address a broader spectrum of possible areas of societal AI risks and misuse. Likewise, in their AI reporting, alternative news media emphasized concerns associated with concepts such as 'posthumanism' or 'transhumanism' as well as conspiracy theories such as an AI-induced 'Great

Reset'. Benefit frames, on the other hand, were more present in the reportingof legacy media, which often – albeit cautiously and under certain conditions – referred to the potential of AI, e.g., in the context of space research or general natural sciences. Both forms ofmedia framed AI in their coverage predominantly in the "distant" time horizon, with alternativemedia tending to emphasize future threats and possible AI dystopias, and legacy media, again, being cautiously optimistic emphasizing the social and economic potential of AI (RQ1b). Regarding RQ2, we found that legacy media predominantly used scientific studies and industry reports as content references. Conversely, alternative news media frequently relied on articles of legacy media, re-framing and criticizing their contents. Additionally, they used a great variety of other sources including dystopian novels or YouTube videos to enrich their critique.

In summary, alternative news media, in a way, produced a counter-'mainstream' AI coverage, augmenting the discourse on uncertainty, doubts, and confusion regarding AI, and additionally undermining professional accounts in legacy media. Moreover, audiences exposedto both sources may perceive a deep societal discord regarding AI. In the discussion, we will elaborate on such possible implications of our results for audience views regarding public perceptions of AI more in detail.

References

Bingaman, J., Brewer, P. R., Paintsil, A. & Wilson, D. C. (2021). "Siri, show me scary imagesof AI": Effects of text-based frames and visuals on support for artificial intelligence. *Science Communication*, *43*(3), 388–401. https://doi.org/10.1177/1075547021998069

Boberg, S., Quandt, T., Schatto-Eckrodt, T., & Frischlich, L. (2020). *Pandemic populism: Facebook pages of alternative news media and the corona crisis—A computational content analysis*. SSRN. https://arxiv.org/abs/2004.02566

Brause, S. R., Zeng, J., Schäfer, M. S., & Katzenbach, C. (2023). Media representations of artificial intelligence. In S. Lindgren (Ed.), *Handbook of critical Studies of artificial intelligence*. Edward Elgar Publishing.

Brubaker, R. (2021). Paradoxes of populism during the pandemic. *Thesis Eleven*, *164*(1), 73– 87. https://doi.org/10.1177/0725513620970804

Chuan, C.-H., Tsai, W.-H. S., & Cho, S. Y. (2019). Framing artificial intelligence in Americannewspapers. In V. Conitzer, G. Hadfield, & S. Vallor (Eds.), *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society* (pp. 339–344). https://doi.org/10.1145/3306618.3314285

Eslen-Ziya, H. (2022). Knowledge, counter-knowledge, pseudo-science in populism. In H. Eslen-Ziya & A. Giorgi (Eds.), *Populism and science in Europe* (pp. 25–41). Palgrave Macmillan. https://doi.org/10.1007/978-3-030-97535-7_2

Fast, E. & Horvitz, E. (2017). Long-term trends in the public perception of artificial intelligence. *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*,*31*(1), 963–969. https://arxiv.org/pdf/1609.04904.pdf

Frost, E. K. & Carter, S. M. (2020). Reporting of screening and diagnostic AI rarely acknowledges ethical, legal, and social implications: a mass media frame analysis. *BMCMedical Informatics and Decision Making*, *20*(1), Article 325. https://doi.org/10.1186/s12911-020-01353-1

Holt, K., Figenschou, T. U., & Frischlich, L. (2019). Key dimensions of alternative news media. *Digital Journalism*, *7*(7), 860–869. https://doi.org/10.1080/21670811.2019.1625715

Marwick, A. E., & Partin, W. C. (2022). Constructing alternative facts: Populist expertise and the QAnon conspiracy. *New Media & Society*. Advance online publication. https://doi.org/10.1177/14614448221090201

Merkley, E. (2020). Anti-intellectualism, populism, and motivated resistance to expertconsensus. *Public Opinion Quarterly*, *84*(1), 24–48. https://doi.org/10.1093/poq/nfz053

Nisbet, M. C. (2009). Framing science: A new paradigm in public engagement. In L. Kahlor & P. A. Stout (Eds.), *New agendas in communication. Communicating science* (pp. 40– 67). Routledge.

Strekalova, Y. A. (2015). Informing dissemination research: A content analysis of U.S. newspaper coverage of medical nanotechnology news. *Science Communication*, *37*(2),151–172. https://doi.org/10.1177/1075547014555025

Vicsek, L. (2011). Costs and benefits of stem cell research and treatment: Media presentation and audience understanding in Hungary. *Science Communication*, *33*(3), 309–340. https://doi.org/10.1177/1075547010389820

Viehmann, C., Ziegele, M., & Quiring, O. (2020). Gut informiert durch die Pandemie? Nutzungunterschiedlicher Informationsquellen in der Corona-Krise [Well informed through thepandemic? Usage of different information sources in the Corona crisis]. *Media Perspektiven*, *2020*(10/11), 556–577.

Politicisation and polarisation on an industry-shaped ground: Public discourse and actors around AI in the French press and social media (2012-2022)

Panos Tsimpoukis, Nikos Smyrnaios, Pierre Ratinaud

Ever since the launch of ChatGPT, there has been a tremendous buzz in the public sphere regarding the applications of artificial intelligence (AI), the potential risks it poses to society, and the advantages it offers. Nevertheless, the public debate surrounding artificial intelligence has begun many years prior (Crépel & Cardon, 2022). A very large part of the discourse around artificial intelligence seems to revolve around the threats and promises that this technology brings to society (Shaping AI, 2023). In order to delve further into this question in the particular context of France, we conducted research through a robust method (Smyrnaios & Ratinaud, 2017) including lexicometric and network analysis of 24.055 articles from national and regional press and 3.599.000 Twitter posts mentioning artificial intelligence during the period 2012-2022. Our results reveal political polarisation around facial recognition technology, dominance of AI companies in social media discussions, but also a major shift in the discourse and the key actors debating around this technology during the period of 2017-2018.

In the French press, this shift coincided with the announcement by Emmanuel Macron of the “France AI” strategic plan, which aimed to support the development of artificial intelligence in France, foster a startup-friendly environment, and introduce regulations for the technology's development. Since the announcement of this plan in 2018, we have observed a striking increase in the number of articles in the press. Notably, new topics have gained prominence, such as the geopolitics of artificial intelligence, discussions about BATX (Aidu, Alibaba, Tencent, Xiaomi) - which are the Chinese counterparts to GAFAM-, artificial intelligence as an economic driver for the country, data protection, and funding allocated to create an environment conducive to the development of artificial intelligence. This shift in discourse followed statements by leaders of major countries in international economic forums, highlighting the pivotal role that artificial intelligence is expected to play in the future (Bareis & Katzenbach, 2022). In order to explore which AI actors were mentioned in the articles along time, we applied the Named Entities Recognition (NER) process with a view to extract all the person and organization entities. Our results show that, although during the period 2012-2015 the Big Tech companies dominated in the discourse, from 2016 onward, two other main poles appear: French government actors and French AI ecosystem actors - the defense industry included. The presence of these actors is strengthened after 2020, along with a distinct pole of European stakeholders, linked to the EU efforts to draw up regulatory frameworks on AI.

On X (formerly known as Twitter), concerns such as “transhumanism” and potential dangers like autonomous weapons or algorithmic bias began to receive significant attention from 2016 onwards, shortly after scientists began warning of the potential dangers of artificial intelligence development. However, the substantial change in 2018 was the widespread appearance of lobbyists, marketing specialists, consultants, influencers and think tanks who shifted the focus towards marketing AI solutions, applications of AI in healthcare, and the promotion of conferences and discussions about artificial intelligence. After 2020, we observe likewise the emergence of actors from the defense sector and from the AI ecosystem of Africa. The aforementioned actors are not limited to sharing news about innovation in AI, but also disseminated news articles related to the ethics and stakes of artificial intelligence development. These findings imply that the AI industry remained dominantnot only in the press discourse (Brennen et al., 2018), but also in digital arenas.
Nevertheless, from 2021 onward, we observed a polarisation both in the press and on Twitter, with

the entry of new actors criticizing facial recognition. This criticism was linked to the Global Security Law, voted in 2021, which allowed police live-feed access to body cameras and drone footage. The concern centered on the possibility that the collected images could be processed using facial recognition software. On Twitter, this criticism involved left-wing actors and political parties,far-right actors and political parties, and conspiracy theorists, while in the press, for the first time, the names of opposition politicians were frequently mentioned in articles about artificial intelligence.

Our results indicate that criticism of artificial intelligence became more grounded in actual applications of this technology, much later than when AI-related policies were initiated by the French state and considerably after the lobbyists had established their presence in the digital public sphere. Given that different stakeholders have an interest in shaping the AI-related debates to their advantage and steer public attention in directions that benefit them (Richter et al., 2023), and that lobbying may have played an important role in shaping the regulatory frameworks AIDA (ArtificialIntelligence in a Digital Age) and AI Act (Schyns, 2023), we argue that our findings could give further insight on how different actors may have influenced the discourse around artificial intelligence over a time period of ten years.

References

Bareis, J., & Katzenbach, C. (2022). Talking AI into Being: The Narratives and Imaginaries of National AI Strategies and Their Performative Politics. *Science Technology and Human Values*, *47*(5), 855–881. https://doi.org/10.1177/01622439211030007

Brennen, A. J. S., Howard, P. N., & Nielsen, R. K. (2018). An Industry-Led Debate: How UK Media Cover Artificial Intelligence. *Reuters Institute for the Study of Journalism Fact Sheet*, *December*, 1–10.

Crépel, M., & Cardon, D. (2022). Robots vs algorithmes. *Réseaux*, *N° 232-233*(2), 129–167. https://doi.org/10.3917/res.232.0129

Dandurand, G., Blottière, M., Jorandon, G., Gertler, N., Wester, M., Chartier-Edwards, N., Roberge,J., & McKlve, F. (2022). Training the News : Coverage of Canada's AI Hype Cycle (2012 – 2021). *Shaping 21st-Century AI*.

Richter, V., Katzenbach, C., & Schäfer, M. (2023). Imaginaries of Artificial Intelligence. In *Handbook of Critical Studies of Artificial Intelligence*.https://doi.org/https://doi.org/10.26092/elib/2190

Shaping AI. (2023). *Approcher l'IA par l'analyse des récits médiatiques*. Sciences Po Médialab. https://medialab.github.io/ShapingAI/#medias

Smyrnaios, N., & Ratinaud, P. (2017). The Charlie Hebdo Attacks on Twitter: A Comparative Analysis of a Political Controversy in English and French. Social Media + Society, 3(1). https://doi.org/10.1177/2056305117693647

Parallel Panel I: TOWARD STANDARDS FOR COMMUNICATION WITH AI

Thursday, June 6, 10:00 – 11:00
Chair: Jana Egelhofer
Room: HG E 1.1

Developing quality criteria for AI reporting: a modular design for journalism practice and science communication

Tobias Kreutzer, Marcus Anhäuser, Holger Wormer

Sparked by the disruptive potential of large language models such as ChatGPT, DALL-E and Midjourney, generative AI as a public topic has seen a "meteoric rise" (Schäfer, 2023) from the end of 2022 on. The current boom in reporting on AI topics has been accompanied by discussions about how AI applications can and should be used in journalism and communication (Graßl et al., 2022; Stray, 2021). Media practitioners have reacted by publishing guidelines on the editorial use of AI (Becker, 2023), including Reporters Without Borders, the German Press Agency (dpa) and Der Spiegel.

Although first field surveys of AI media representations raised awareness of the dominance of "affirmative positions" towards AI in media coverage as well as a potentially distorting media focus on "established, institutional and often economic stakeholders" (Brause et al., 2023, p. 272), an equally fundamental debate about how journalists should critically report on topics and applications based on artificial intelligence and what criteria determine the quality of good AI reporting is still missing. The UNESCO handbook on "Reporting on Artificial Intelligence" (Jaakkola, 2023) mostly addresses specialized communicators who intensively research AI topics. So far, there has been given little attention to non-specialist journalists outside of technology departments who soon will bedealing with AI applications more often. As a rare example, a white paper by the AI + Automation Lab of German broadcasting company Bayerischer Rundfunk identifies "Blackbox Reporting" as a challenge for journalists covering AI.

Developing topic-specific quality criteria for AI reporting

AI applications affect various aspects of life and society, posing a particular challenge: Medicine, business, social life, politics, sport – journalists confronted with AI applications touching on "their" areas hardly possess the knowledge of a specialized technology journalist. In our project we therefore want to develop quality criteria for AI reporting, which provide orientation to various professional profiles in the field. Building on widely accepted general quality criteria for journalism (e.g. Meier, 2019), international science communication researchers have contributed to a growing body of work on the operationalization of topic-specific quality criteria for science journalism, which are necessary for an in-depth-understanding of particular topics. Such specific criteria – for example adapted to environmental and health journalism – provide journalists with more exact research questions especially relevant to certain topic areas. Starting with a focus on health and medical journalism (Schwitzer, 2008; Wilson et al., 2009) the research field hasproduced follow-up work taking into account both the heterogenous field of science communication actors and the field-specific differences in science journalism quality (e.g., Lilienthal et al., 2014, Rögener et al., 2015, Anhäuser et al., 2020). The resulting criteria catalogues work as practice tools for research and critical reporting in journalistic day-to- day work. They are characterized by a modularization into general journalistic, general science journalistic and discipline-specific (e.g. medical, environmental) journalistic aspects. The criteria include evaluation categories such as the "classification of the level of evidence", the "mentioning of possible side effects" or the "mentioning of conflicts of interest" in a given article.

Based on these approaches, we are developing a set of criteria that will help non- specialist journalists and communicators ask topic-specific critical questions and research fundamental aspects of applications based on artificial intelligence and machine learning. This seems necessary as AI-based applications have started to affect all areas of society, which means that journalists from all departments increasingly encounter such AI-based applications in their reporting.

Our criteria should also help to determine the journalistic quality of AI coverage and engage in constructive media monitoring. Our methodology follows the development process of the catalogues for environmental and nutrition reporting mentioned above. We first surveyed the field through a structured inventory of research as well as practical literature to identify existing recommendations for AI reporting. A first literature review could not identify any systematic approaches to developing quality criteria for reporting on AI topics: The existing literature is dominated by (few) general works on the content and evaluative tone of the reporting (e.g. Vergeer, 2020; Brantner & Saurwein, 2021). At least the discourse about "algorithmic accountability reporting" (Diakopoulos, 2014) as a complex and often investigative form of reporting has produced recommendations for journalism practice that can be used as a starting point.

We are currently testing the transferability of single modules from the criteria lists mentioned above. Two block seminars with a total of 36 participants already set the frame for the co-productive development of AI reporting criteria between journalism students and AI experts. In parallel, we reach out to AI scientists from various fields of computer science and statistics for qualified and constructive assessments of frequent deficits in journalistic reporting on AI applications. We are, however, aware that the experts' expectations do not necessarily have to be realistic with regards to journalism practice. To test the practicability of proposed criteria and foster further acceptance, a broader journalistic focus group is included in the process. A survey among journalists specializing in AI coverage will ask about aspects particularly important for AI reporting. Finally, the results of the seminars and expert surveys are compiled into an applicable catalogue of criteria.

References

Anhäuser, M., Wormer, H., Viciano, A., & Rögener, W. (2021). Ein modulares Modell zur Qualitätssicherung im Medizin- und Ernährungsjournalismus. Bundesgesundheitsblatt, Gesundheitsforschung, Gesundheitsschutz, 64(1), 12-20.

Bayerischer Rundfunk, AI + Automation Lab. (2023, August). Blackbox Reporting: Wie Journalist*Innen über KI und Algorithmen berichten können. https://interaktiv.br.de/paper/AI-Automation-Lab_Blackbox-Reporting.pdf

Becker, K. B. (2022). Neues Spiel, neue Regeln. Eine Untersuchung von redaktionellen Richtlinien für den Umgang mit Künstlicher Intelligenz im Newsroom. Journalistik.Zeitschrift für Journalismusforschung, 2, 6. Jg., 142-163.

Brantner, C., & Saurwein, F. (2021). Covering Technology Risks and Responsibility: Automation, Artificial Intelligence, Robotics, and Algorithms in the Media. International Journal of Communication, 15, 5074-5098.

Brause, S. R.; Zeng, J.; Schäfer, M.; Katzenbach, C. (2023). Media Representations of Artificial Intelligence. Surveying the Field. In: Lindgren, S. Handbook of Critical Studies of Artificial Intelligence. Cheltenham: Edward Elgar Publishing Ltd, Epub ahead of print.

Diakopoulos, N. (2014). Algorithmic accountability reporting: On the investigation of black boxes. Columbia University Libraries.https://academiccommons.columbia.edu/doi/10.7916/D8ZK5TW2.

Graßl, M., Schützeneder, J., & Meier, K. (2022). Künstliche Intelligenz als Assistenz: Bestandsaufnahme zu KI im Journalismus aus Sicht von Wissenschaft und Praxis. Journalistik: Zeitschrift für Journalismusforschung, 5(1), 3-27.

Lilienthal, V., Reineck, D., & Schnedler, T. (Eds.). (2014). Qualität im Gesundheitsjournalismus: Perspektiven aus Wissenschaft und Praxis. Springer-Verlag.

Meier, K. (2019). Quality in Journalism. In T. P. Vos, F. Hanusch, D. Dimitrakopoulou, et al. (Eds.), The International Encyclopedia of Journalism Studies (pp. 1-8). Wiley.

Rögener, W., & Wormer, H. (2017). Defining criteria for good environmental journalism and testing their applicability: An environmental news review as a first step to more evidence based environmental science reporting. Public Understanding of Science, 26(4),418-433.

Schäfer, M. S. (2023). The Notorious GPT: science communication in the age of artificial intelligence. Journal of Science Communication, 22(2), Y02.

Schwitzer, G. (2008). How do US journalists cover treatments, tests, products, and procedures? An evaluation of 500 stories. PLoS medicine, 5(5), e95.

Stray, J. (2021). Algorithms, Automation, and News. Routledge.

UNESCO, Jaakkola, M. (Ed.) (2023). A Handbook for Journalism Educators Reporting onArtificial Intelligence. https://unesdoc.unesco.org/ark:/48223/pf0000384551.

Vergeer, M. (2020). Artificial intelligence in the Dutch press: An analysis of topics andtrends. Communication Studies, 71(3), 373-392.

Wilson, A., Bonevski, B., Jones, A., & Henry, D. (2009). Media reporting of healthinterventions: signs of improvement, but major problems persist. PloS one, 4(3), e4831.

Promoting quality communication of AI: insights from an AI research project

Alessandra Fornetti, Ilda Mannino, Folco Soffietti

Discussion of Artificial Intelligence (AI) has increased sharply since 2009 with both optimistic and pessimistic views. In recent years many specific concerns related to AI advancements (e.g. ethics, negative impact on job availability, misinformation, inequalities) have been growing (Fast & Horvitz, 2017).

We outline a science communication research agenda built around three key questions emerging from the discussion: What is the role of AI researchers in promoting a reliable, trustworthy, and effective communication of AI? What are the elements in AI communication strategies that can support a constructive engagement with AI? How can quality communication of AI support the current debate on science communication, especially in relation to scientists' engagement?

This paper uses the MUHAI project – a highly-technical research project on Human-centric AI, involving renowned AI scientists[4] – as case study to investigate the three research questions above. Results are drawn from direct interactions with AI scientists involved in the project, data related to the project communication activities, and existing literature on the topic.

Interviews with key actors, and specifically MUHAI researchers provide relevant insights to define how to promote quality communication of AI. With specific reference to the role of AI researchers, the objective is to understand:

a) AI researchers' understanding of the role of public communication to support the impactof their research and influence how societies respond to technological advances;

b) if and how the experience in communicating MUHAI's research has changed AI researchers' awareness of the role of communication for end-users and society;

c) how scicommers can support AI researchers in communicating AI.

Based on the above, three main aspects on quality AI communication emerge and are described in the paper:

1. **AI researchers** can play a crucial role in enhancing the understanding of AI and its implications through public communication, thus making evident the value and the impact that AI can have on society (The Royal Society, 2020). AI researchers are however not always aware of the importance and role that a quality communication of AI can have on society. Moreover, they do not often feel confident also due to their lack of skills (Mannino et al, 2021). As a consequence, they do not engage in public communication activities that could, instead, bring a benefit to their own work, to the AI field as a whole, and to society.
2. **Art** is a positive and highly-engaging lever for involving non-experts in AI discussion, both on digital media and on the traditional press. As such it can be an effective way to reach directly the wider audience and to attract both journalists and scientists from non-AI specific disciplines. The MUHAI project opened up a practice-based reflection on the display and communication of AI inits ethical, poetic, and historic aspects, despite being a relatively young technological innovation.
3. **The visual communication of AI** constructed on the depiction of science fiction characters, robot-like, in cold, blue colors shape the public perception of AI, which, most of the time, prompt a pessimistic narrative (Brauner et al, 2023). Similarly, the **language used in AI** research domain is often taken from the everyday vocabulary (e.g. understanding, meaning, memory…): if, on the one side, this brings AI closer to reality, on the other it possibly augments lay people's perception of AI as a substitute for humans. It thus risks to increase the distance between researchers and citizens, pushing them to take a further step back from understanding AI.

[4] Luc Steels, scientific coordinator of MUHAI project, is EurAI Distinguished Service Award 2022

The results of the analysis of this specific project are seen as a starting point to help shaping a wider research agenda on AI communication for the current science communication landscape. They also contribute to understand how the debate on AI communication support the ongoing discussions on the role of researchers and research institutions (e.g. training, reward) as it also emerged from the Concluding Statement of therecent PCST Venice Symposium.

References

Brauner P, Hick A, Philipsen R and Ziefle M (2023.) What does the public think about artificial intelligence? — A criticality map to understand bias in the public perception of AI. Front. Comput. Sci. 5:1113903. doi: 10.3389/fcomp.2023.1113903

Fast E, Horvitz E (2017). Long-Term Trends in the Public Perception of Artificial Intelligence, Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence AAAI-17

Mannino, I., Bell, L., Costa, E., Di Rosa, M., Fornetti, A., Franks, S., Iasillo, C., Maiden, N., Olesk, A., Pasotti, J., Renser, B., Roche, J., Schofield, B., Villa, R. and Zollo, F. (2021).

Supporting quality in science communication: insights from the QUEST project JCOM 20(03), A07. https://doi.org/10.22323/2.20030207

The Royal Society (2020). Communicating AI: the role of researchers in enabling public conversations about AI, Note of discussions at a Royal Society workshop.

Systematic review of strategies for science communication to mitigate mis- and disinformation

Christian Schuster, Andreas Scheu

The question of how contemporary societies can and should deal with mis- and disinformation is rapidly gaining importance (Adams et al., 2023; Borges Do Nascimento et al., 2022; Ecker et al., 2022). Particularly, the advancing changes in our media ecosystems and the current establishment of generative artificial intelligences (AIs) contribute to this development (Passanante et al., 2023; Powell et al., 2022; Zhao et al., 2023). Science communication efforts that address mis- and disinformation are increasingly understood as part of the solution to the infodemic challenge. This is not only indicated by the rapid growths of respective scientific evidence (Lazić & Žeželj, 2021; Passanante et al., 2023; Peng et al., 2023). It is also reflected by programs of German political parties, university development programs, and statements of political decision-makers in interviews (Anonymous, 2023).

Our contribution pursues the question of how science communication can and should respond to mis- and disinformation. A particular focus will be on the role of generative AI. On the one hand, AIs themselves can create and transport dis-/misinformation (De Angelis et al., 2023; Godulla et al., 2021; Spitale et al., 2023). On the other hand, artificial intelligences can also provide targeted information about current questions and uncertainties of recipients and thus counteract misinformation (Passanante et al., 2023; Spitale et al., 2023). Thus, generative AIs both challenge and offer new opportunities for science communication in dealing with dis- and misinformation.
In this context, we pose two questions:

RQ1: How can the research field "science communication to mitigate dis-/misinformation" be characterized and structured, and what role does the engagement with communicative AIs play?

RQ2: How can and should practical science communication deal with dis-/misinformation and what role does AI play in this respect?

Methods

From November 2023 to March 2024 we are conducting a systematic review of 188 interdisciplinary and international research articles. We searched the databases Communication and Mass Media Complete, PubMed, Scopus, and Web of Science. The search string is summarized in Figure 1.

Figure 1: General search string

((“science communication” OR “knowledge communication” OR “health communication” OR “risk communication” OR “climate communication” OR “climate change communication” OR “environment communication” OR “environmental communication” OR “technology communication”)
OR
(“science information” OR “health information” OR “risk information” OR “climate information” OR “climate change information” OR “environment information” OR “environmental information” OR “technology information”)
OR
(“communicating science” OR “communicating knowledge” OR “communicating health” OR “communicating risk” OR “communicating risks” OR “communicating climate” OR “communicating environment” OR “communicating environmental” OR “communicating technology” OR “communicating technological”)
OR
(“science dialog” OR “science dialogue” OR “science transfer”)
OR
(“public understanding” OR “public engagement with science” OR “public engagement with technology” OR “scientific literacy” OR “science and technology studies”))

AND

(“misinformation” OR “disinformation” OR “misleading information” OR “false information” OR “fake news” OR “fake science“ OR “false experts“ OR “false news“ OR “alternative facts“ OR “propaganda” OR "information disorder" OR "deception" OR “deep fakes” OR “conspiracy“ OR “debunk*“)

Note: Searches focused on titles, abstracts, and keywords. This presentation does not contain database specific syntax.

The systematic review follows the Cochrane approach (Higgins et al., 2019). Our review deviates in some points to meet the specific requirements of social science research, such as the relatively low standardization of published studies and the conceptual and methodological diversity in the interdisciplinary research field. Also, for resource reasons, we deviated from the consistent four-eye principle during screening and controlled screening on a sample basis (approx. 10% of the articles).

Of the recorded 1,978 hits, we selected 188 studies for analysis based on our selection criteria. The studies were evaluated using the review management software NACSOS[5] (Callaghan, 2021). For further analysis, only English and German peer-reviewed empirical studies that document their methodological approach were included. On a content level, articles were considered that discuss the possibilities and limits of science communication to mitigate dis- and misinformation.

Our evaluation strategy proceeds in two stages. In the first phase, we consider the entire literature (n = 188) based on a quantitative content analysis (Staender & Humprecht, 2023) and automated evaluation. Variables include first author, regional focus, method(s), and content focus. At the same time, topic modeling methods are used to identify clusters within the research field (Maier et al., 2018).

In a second phase, the selected studies are evaluated more in-depth qualitatively based on research categories (Scheu, 2018) with regard to RQ2. We selected studies that explicitly discuss options for science communication to mitigate mis- and disinformation. We differentiate preventive measures (e.g., proactive refutation of common misinterpretations), indirect preventive measures (e.g., promotion of resilience), and reactive measures (e.g., debunking). Our conference talk focuses on opportunities and challenges of generative artificial intelligences in this context.

Results – Outlook

The earliest studies in our corpus have been published in 2003; since the end of the 2010s, there has been a noticeable increase in publications. The focus of publications is on the topic of misinformation (n = 135), while disinformation is addressed in 28 studies. Thematically, the studies primarily deal with the communication of scientific evidence in the context of dis-/misinformation about societal health risks (mainly Covid, pandemic), and climate change. Methodologically, experiments (n = 46), reviews (n = 36), qualitative interviews (n = 25), and quantitative surveys (n = 22) were conducted.

Studies also discuss explicit options for science communication. Both direct preventive measures, such as the proactive communication of counterarguments to common misinterpretations (n = 89), and indirect preventive measures, for example, "prebunking" or "inoculation" strategies (n = 25), are examined, as well as reactive measures (e.g., debunking; n = 81). We differentiate the results with regard to dealing with dis-/misinformation in general and specifically in the context of generative AI.

The conference talk will discuss the state of research for concrete strategies of science communication in order to prevent and counter dis- and misinformation in the age of artificial intelligence (Schäfer, 2023).

References

Anonymous (2023). *Science Policy Strategies for Science Communication on the Level of German Federal States*. Unpublished research report.

Adams, Z., Osman, M., Bechlivanidis, C., & Meder, B. (2023). (Why) Is Misinformation a Problem? *Perspectives on Psychological Science*, *18*(6), 1436–1463. https://doi.org/10.1177/17456916221141344

Borges Do Nascimento, I. J., Beatriz Pizarro, A., Almeida, J., Azzopardi-Muscat, N., André Gonçalves, M., Björklund, M., & Novillo-Ortiz, D. (2022). Infodemics and health misinformation: A systematic review of reviews. *Bulletin of the World Health Organization*, *100*(9), 544–561. https://doi.org/10.2471/BLT.21.287654

[5] NLP-assisted classification, synthesis, and online screening: https://github.com/mcallaghan/tmv/wiki/Scoping-Documentation

Callaghan, M. (2021). Machine-learning-based evidence and attribution mapping of 100,000 climate impact studies. *Nature Climate Change*, *11*, 966–972.

De Angelis, L., Baglivo, F., Arzilli, G., Privitera, G. P., Ferragina, P., Tozzi, A. E., & Rizzo, C. (2023). ChatGPT and the Rise of Large Language Models: The New AI-Driven Infodemic Threat in Public Health. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4352931

Ecker, U. K. H., Lewandowsky, S., Cook, J., Schmid, P., Fazio, L. K., Brashier, N., Kendeou, P., Vraga, E. K., & Amazeen, M. A. (2022). The psychological drivers of misinformation belief and its resistance to correction. *Nature Reviews Psychology*, *1*(1), 13–29. https://doi.org/10.1038/s44159-021-00006-y

Godulla, A., Hoffmann, C. P., & Seibert, D. (2021). Dealing with deepfakes – an interdisciplinary examination of the state of research and implications for communication studies. *Studies in Communication and Media*, *10*(1), 72–96. https://doi.org/10.5771/2192-4007-2021-1-72

Higgins, J. P. T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., & Welch, V. A. (Hrsg.). (2019). *Cochrane Handbook for Systematic Reviews of Interventions* (1. Aufl.). Wiley. https://doi.org/10.1002/9781119536604

Lazić, A., & Žeželj, I. (2021). A systematic review of narrative interventions: Lessons for countering anti-vaccination conspiracy theories and misinformation. *Public Understanding of Science*, *30*(6), 644–670. https://doi.org/10.1177/09636625211011881

Maier, D., Waldherr, A., Miltner, P., Wiedemann, G., Niekler, A., Keinert, A., Pfetsch, B., Heyer, G., Reber, U., Häussler, T., Schmid-Petri, H., & Adam, S. (2018). Applying LDA Topic Modeling in Communication Research: Toward a Valid and Reliable Methodology. *Communication Methods and Measures*, *12*(2–3), 93–118. https://doi.org/10.1080/19312458.2018.1430754

Passanante, A., Pertwee, E., Lin, L., Lee, K. Y., Wu, J. T., & Larson, H. J. (2023). Conversational AI and Vaccine Communication: Systematic Review of the Evidence. *Journal of Medical Internet Research*, *25*, e42758. https://doi.org/10.2196/42758

Peng, W., Lim, S., & Meng, J. (2023). Persuasive strategies in online health misinformation: A systematic review. *Information, Communication & Society*, *26*(11), 2131–2148. https://doi.org/10.1080/1369118X.2022.2085615

Powell, L., Nour, R., Sleibi, R., Al Suwaidi, H., & Zary, N. (2022). *Democratizing the development of chatbots to improve public health: A feasibility study of COVID-19 misinformation (Preprint)* [Preprint]. JMIR Human Factors. https://doi.org/10.2196/preprints.43120

Schäfer, M. S. (2023). The Notorious GPT: Science communication in the age of artificial intelligence. *Journal of Science Communication*, *22*(02). https://doi.org/10.22323/2.22020402

Scheu, A. M. (Hrsg.). (2018). *Auswertung qualitativer Daten*. Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-18405-6

Spitale, G., Biller-Andorno, N., & Germani, F. (2023). AI model GPT-3 (dis)informs us better than humans. *Science Advances*, *9*(26), eadh1850. https://doi.org/10.1126/sciadv.adh1850

Staender, A., & Humprecht, E. (2023). Content Analysis in the Research Field of Disinformation. In F. Oehmer-Pedrazzi, S. H. Kessler, E. Humprecht, K. Sommer, & L. Castro (Hrsg.), *Standardisierte Inhaltsanalyse in der Kommunikationswissenschaft – Standardized Content Analysis in Communication Research* (S. 339–348). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-36179-2_29

Zhao, X., Chen, L., Jin, Y., & Zhang, X. (2023). Comparing button-based chatbots with webpages for presenting fact-checking results: A case study of health information. *Information Processing & Management*, *60*(2), 103203. https://doi.org/10.1016/j.ipm.2022.103203

Parallel Panel II: EXPLORING IMAGINARIES OF ARTIFICIAL INTELLIGENCE

Thursday, June 6, 11:20 – 12:40
Chair: Melanie Leidecker-Sandmann
Room: HG E 1.2

Generative AI in science communication research: The sociotechnical imaginaries informing in the fields' journal and publisher authorship policies

Michelle Riedlinger, Marina Joubert, Peta Mitchell

In early 2023, Springer Nature, the World Association for Medical Editors (or WAME) and the Committee on Publication Ethics (or COPE) released statements about using AI tools in academic publishing. Major academic journal publishers including Elsevier, Sage, and Taylor and Francis have also created policies that prohibit listing Generative AI tools, such as ChatGPT as an academic author, and require authors to document their use of these tools in the Methods or Acknowledgment sections of their manuscripts. Researchers and commentators predict that some of these policies, may have been hastily constructed or will be overturned as new developments emerge (Hufton, 2023).

As leaders in the technology engagement space, science communication researchers have important roles to play in guiding academic authorship debates and policies for the field. On the one hand, researchers recognize the relevance and potential of AI (Schäfer 2023), particularly "communicative AI" (Guzman & Lewis 2020: 79; Hepp et al. 2022) for producing novel outputs, shaping commentary for different audiences, and enabling interactive exchanges. However, the lack of transparency and explainability, systematic biases, data privacy and protection, maintaining the quality of scholarly work, and accountability and liability issues (STM, 2021; Wen and Wang, 2023) are critical considerations for science communication.

This study aims to inform debate on the impact of AI technologies on the broader science communication ecosystem by investigating the sociotechnical imaginaries embedded in the policies of journals and academic publishers serving science communication research communities. Science communication researchers publish in academic journals dedicated to the broad field of science communication, including 'Public Understanding of Science', 'Science Communication', 'JCOM', and 'International Journal of Science Education, Part B'. Other journals feature content relevant for science communication researchers, including those devoted to health and environmental communication, multi-disciplinary journals such as PLOS One and PNAS, and discipline-specific journals in fields such as climate change, health and conservation.

In this study, we are identifying relevant scholarly journals for the science communication by gathering relevant science communication research articles published in 2023. We are using academic database searches (including Scopus, DOAJ, and the Web of Science) and searching for all English-language research articles published in peer-reviewed journals since 1 January 2023 that contain on or more of the following terms in the title, abstract, keywords or body of the article: 'science communication'; 'communicating science', 'public engagement with science' or 'public understanding of science'. This search has delivered 2,520 items up to 3 December 2023. We will add new research articles up to 31 December 2023 so that we have data for a full calendar year.

After identifying the number of published articles for each academic journal and ranking the journals for relevance and quality according to the SCImago Journal & Country Rank (SJR), we will determine if these journals (and/or their overarching publisher) have a published authorship policy or guidelines that address AI authorship.

Taking inspiration from work focused on sociotechnical imaginaries (Jassanoff 2014), and contested AI social imaginaries (Cave, Dihal, and Dillon 2019) in particular, we will examine the prescriptions and proscriptions embedded in these policies and the ways that science communication authorship is imagined. We are particularly interested in authorship policy responses that move beyond the risk management framings of early position statements and policies, and those that might inform and generate rigorous debate, and more nuanced understandings within our field.

Important questions for discussion at the conference include: how approaches to the academy's "third mission" of science communication practice might reflect and reinforce these academic authorship policies and the imaginaries informing them (considering ethics, risk and trust issues, for example); how these Generative AI authorship policy imaginaries might support or discourage more dialogical research engagement at scale and encourage or discourage citizen or community science and participatory science communication efforts beyond communities with a keen interest in science.

References

Cave, S., Dihal, K., & Dillon, S. (2020). Introduction: Imagining AI. In S. Cave, K. Dihal, & S. Dillon (Eds.), AI Narratives: A History of Imaginative Thinking about Intelligent Machines (pp. 1-24)).

Guzman, A. L., & Lewis, S. C. (2020). Artificial intelligence and communication: A Human–Machine Communication research agenda. New Media & Society, 22(1), 70–86. https://doi.org/10.1177/1461444819858691

Hepp, A., Loosen, W., Dreyer, S., Jarke, J., Kannengießer, S., Katzenbach, C., ... & Schulz, W. (2023). ChatGPT, LaMDA, and the hype around communicative AI: The automation of communication as a field of research in media and communication studies. Human-Machine Communication, 6(1), 4. https://stars.library.ucf.edu/hmc/vol6/iss1/4/

Hufton, A. L. (2023). No artificial intelligence authors, for now. Patterns, 4(4). https://www.cell.com/patterns/pdf/S2666-3899(23)00071-5.pdf

Jasanoff, S. (2015). Future imperfect: Science, technology, and the imaginations of modernity. In J. H. Smith & W. T. Johnson (Eds.), Dreamscapes of modernity: Sociotechnical imaginaries and the fabrication of power (pp. 1-33).

Lund, B. D., Wang, T., Mannuru, N. R., Nie, B., Shimray, S., & Wang, Z. (2023). ChatGPT and a new academic reality: Artificial Intelligence-written research papers and the ethics of the large language models in scholarly publishing. JASIST, 74(5), 570-581.

Schäfer, M. S. (2023). The Notorious GPT: science communication in the age of artificial intelligence. Journal of Science Communication, 22(2), Y02. https://jcom.sissa.it/article/pubid/JCOM_2202_2023_Y02/

STM Association. (2021). AI ethics in scholarly communication: STM best practice principles for ethical, trustworthy and human-centric AI. https://www.stm- assoc.org/2021_04_29_STM_AI_White_Paper_April2021.pdf

Wen, J., & Wang, W. (2023). The future of ChatGPT in academic research and publishing: A commentary for clinical and translational medicine. Clinical and Translational Medicine, 13(3). doi: 10.1002/ctm2.1207

Negotiating AI as a sociotechnical phenomenon: Competing imaginaries of AI by stakeholders in the US, China, and Germany

Vanessa Richter, Christian Katzenbach, Jing Zeng

Utopian and dystopian visions regularly dominate the public discourse on artificial intelligence (AI) (Cave & Dihal 2019). These normative debates around AI repeatedly revolve around the relationship between humans and machines, but increasingly also around how AI (re)consolidates existing discrimination and social inequalities. In these debates, we need to understand AI not just as a taken-for-granted object of science communication, but more importantly as a notion and sociotechnical phenomenon that is an object of negotiation. While AI is now routinely treated as self-evident (Suchman 2023), it is still very much under formation as a sociotechnical phenomenon itself, with public perception and discursive framing having considerable influence.

This paper builds on the concept of imaginaries to study how AI is being negotiated between stakeholders in the US, China, and Germany – and thus to "trace its sources of power and to demystify its referents" (Suchman 2023, p. 1). The early sociological work on imaginaries has highlighted the role of perceptions, discourses, and future visions in the complex interactions and negotiations of co-constructing technological developments (Anderson, 1983; Taylor, 2003). More recent work on sociotechnical imaginaries (SI) (Jasanoff & Kim 2009, 2015) as "collectively held, institutionally stabilized, and publicly performed visions of desirable futures, animated by shared understandings of forms of social life and social order attainable through, and supportive of, advances in science and technology" (Jasanoff, 2015, p. 4) enables the reconstruction of the multiple, contested and often commodified (Mager & Katzenbach 2021) discursive negotiations between different actors in processes of technological development and its integration into society. In a more concrete form (Richter et al. 2023) it offers a constructive framework to question the role of different stakeholders in shaping imaginaries around AI as well as the often-contentious negotiation processes around different desirable future developments on AI innovation and application.

This paper addresses producers and proponents of AI-related communication, based on interviews with AI experts in industry, government, academia, media and civil society from three countries USA, Germany, and China. While a number of studies have analysed national, industrial, and political visions of digital media and automation (Felt & Öchsner, 2019; Mager,2017), there is a lack of research analysing the negotiation of imaginations on potential futures of AI between stakeholders in the actual field of AI development. Additionally, there is often a strong focus on national comparison when considering different perceptions because national perceptions are closely associated with funding and the regulation of new technologies such as AI. However, national perception is not a homogenous and monolithic point of view but rather a heterogeneous and controversial discursive construction process. Against this background, this paper examines the imaginaries of different stakeholders in the field - industry, government, academia, media and civil society - in three leading countries in AI development and debate: the USA, Germany, and China.

The data collection is based on semi-structured interviews (15-20 per country) with AI experts from the mentioned stakeholder groups for each country. The transcribed interviews were analysed employing both situational analysis (Clarke, 2019) and critical discourse analysis (Wodak, 2015) to map the relational development of imaginaries across stakeholder groups as well as major emerging imaginaries on AI. By considering the situational aspects of imaginary development, we are able to map controversial discourses on AI impacting future visions of the technology and its perception. The analysis highlights that the negotiation of AI imaginaries does

not necessarily take place in uniform ways across all stakeholder groups as (1) the US discourse can be mapped across several geographic AI centres, (2) while the German dataset revealed a focus on EU policy compliance and geographical distributions by stakeholder groups that led to localised AI imaginaries being foregrounded. Lastly, (3) the Chinese data emphasised a congruence with party policies minimising local specificities in the AI discourse. Moving beyond specific stakeholder groups to question the processes of national and international negotiation on future AI imaginaries reveals relevant intersections of imaginary building as well as the political and socio-technical agendas forming and disrupting current trajectories.

The analysis thus offers key results for understanding how AI as an object of (science) communication is actively negotiated between powerful stakeholders, based on data from theUS, China, and Germany. As a result, the paper makes visible how discourses and strategic activities of stakeholders condense into widely shared sociotechnical imaginaries and shape global and local technical and social developments. The debate around AI plays a decisive role in shaping the future design of AI systems and their integration into social sub-areas and individual everyday lifein times of profound automation.

References

Anderson, B. (1983). *Imagined communities. Reflections on the origin and spread of nationalism.* London: Verso.

Cave, S., & Dihal, K. (2019). Hopes and fears for intelligent machines in fiction and reality. NatureMachine Intelligence, 1(2), 74. https://doi.org/10.1038/s42256-019-0020-9

Clarke, A. E. (2019). Situating grounded theory and situational analysis in Interpretive QualitativeInquiry. *The SAGE Handbook of Current Developments in Grounded Theory*, 3–48. https://doi.org/10.4135/9781526436061.n3

Felt, U., & Öchsner, S. (2019). Reordering the "world of things": the sociotechnical imaginary of RFID tagging and new geographies of responsibility. *Science and EngineeringEthics* 25: 1425–1446.

Jasanoff, S., & Kim, S.-H. (2009). Containing the Atom: Sociotechnical Imaginaries andNuclear Power in the United States and South Korea. *Minerva, 47*(2), 119–146. https://doi.org/10.1007/s11024-009-9124-4.

Jasanoff, S., & Kim, S.-H. (Eds.) (2015). *Dreamscapes of Modernity: SociotechnicalImaginaries and the Fabrication of Power*. Chicago: University of Chicago Press.

Mager, A. (2017). Search engine imaginary: visions and values in the co-production ofsearch technology and Europe. *Social Studies of Science* 47(2): 240–262.

Mager, A., & Katzenbach, C. (2021). Future imaginaries in the making and governing of digitaltechnology: Multiple, contested, commodified. New Media & Society, 23(2), 223–236. https://doi.org/10.1177/1461444820929321

Richter, V., Katzenbach, C., & Schäfer, M. (2023). Imaginaries of artificial intelligence. In *Handbook ofCritical Studies of Artificial Intelligence* (pp. 209–223). Chapter, Edward Elgar Publishing.

Suchman, L. (2023). The uncontroversial 'thingness' of AI. *Big Data & Society*, 10(2). https://doi.org/10.1177/20539517231206794

Taylor, C. (2003). *Modern social imaginaries*. Durham: Duke University Press.

Wodak, R. (2015). Critical Discourse Analysis, Discourse-Historical Approach. In *The InternationalEncyclopedia of Language and Social Interaction* edited by Karen Tracey. Wiley Blackwell.

Visions of AI in the public eye: Comparing News Coverage in China, Germany, and the U.S. (2012-2021)

Jing Zeng, Daniela Mahl, Saba R. Brause, Mike S. Schäfer

Introduction

Artificial Intelligence (AI) plays an increasingly important role in today's societies, holding the potential to reshape entire sectors such as economics, education, and healthcare. As AI rapidly integrates into our daily lives, its potential risks and benefits are increasingly being discussed: While the technology's efficiency-enhancing potential has been strongly emphasized, AI-driven applications such as facial recognition have sparked public controversies. News media play a pivotal role in this regard: They raise awareness of the technologies' potential, societal implications, and ethical concerns, foster informed discussions, and help shape public attitudes toward and engagement with AI. Moreover, news coverage can drive public awareness, prompting government actions to align AI development with public values and interests. While the trajectories of discourses on AI vary across different cultural, political, and social contexts, previous research has predominantly focused on individual countries, primarily in the Global (English-speaking) North (Brause et al., forthcoming). Taking into account geo-political specificities, this study takes a cross-national comparative approach to map and analyze media constructions of AI through news coverage in China, Germany, and the United States (U.S.) over a ten-year period (2012-2021).

Theoretical Framework

This study builds on the concept of sociotechnical imaginaries, which refers to 'collectively imagined forms of social life and social order reflected in the design and fulfillment of nation-specific scientific and/or technological projects' (Jasanoff & Kim, 2009: 120). Sociotechnical imaginaries evolve over time in response to technological advancements, cultural shifts, and societal changes, underscoring the significant influence of publicly shared visions of communities on technological developments and the unpredictability of technology in shaping society. Using sociotechnical imaginaries as a theoretical lens allows us to comparatively assess the specificities of national and transnational discursive constructions of AI in a longitudinal research design. In addition, it allows us to assess who gets to shape these imaginaries of AI in the Chinese, German, and U.S. newspapers (cf. Gerhards & Schäfer, 2009). Guided by this theoretical framework, we propose the following research questions (RQ):

RQ1: How does the discursive construction of AI in the news media differ between China, Germany, and the U.S.? How does this change over time?
RQ2: To what extent do these three countries project different or similar imaginaries of AI through news reporting? Who contributes to the construction of these imaginaries?

Research Design

Drawing from a dataset comprising over 9,000 articles published in 15 leading national newspapers across China, Germany, and the U.S. from January 2012 to July 2021, this study employs computational content analysis alongside comprehensive qualitative investigations. To identify *cross-lingual topic patterns*, we performed multilingual topic modeling using BERTopic (Grootendorst, 2022). This technique exploits the capabilities of pre-trained transformer language models to generate document embeddings. To fine-tune the topic representation, Llama2—an open-source large language model developed by Meta—is used. More specifically, time series and sentiment analysis are used to compare the *salience* and *tensity* of public discourses across the three regions over time (**RQ1**). An in-depth qualitative comparison of publicly shared imaginaries of AI (**RQ2**) complements these analyses. This approach helps to contextualize and explain the distinct patterns in the discursive construction of AI in the three regions. The qualitative approach primarily entails a hybrid deductive-

inductive thematic analysis (Feredey & Muir-Cochrane, 2006).

Results

The study reveals considerable differences in the salience of topics across the three countries, with the most extensive coverage in China, followed by Germany, and the U.S. It also demonstrates that health-, policy-, and technology-related topics are most prominently discussed in China, while societal impacts and risks of AI are more prevalent in German and U.S. media (see Figure 1). Over time and across countries, news coverage about the societal impacts of AI has become less salient, while regulatory issues are taking up more space (see Figure 2). The sentiment analysis shows that U.S. media coverage of AI is generally more positive, while German media coverage of the technology is more negative. Especially with regard to the social impacts and risks of AI, there is comparatively negative coverage in Germany, while in the U.S., developments in the technology sector are discussed in a positive light. Chinese media are comparatively neutral on the identified topic groups.

The qualitative analysis demonstrates that even within topics, crucial differences in AI-related imaginaries can be found across countries. For example, in the healthcare context, AI is imagined to improve healthcare provision in all three countries. While U.S. coverage emphasized the health-related visions of private sector tech companies, Chinese newspapers focused more on speakers from the public sector, and German newspapers included the imaginaries of patients and NGOs.

Figure 1. Prevalence of topic groups

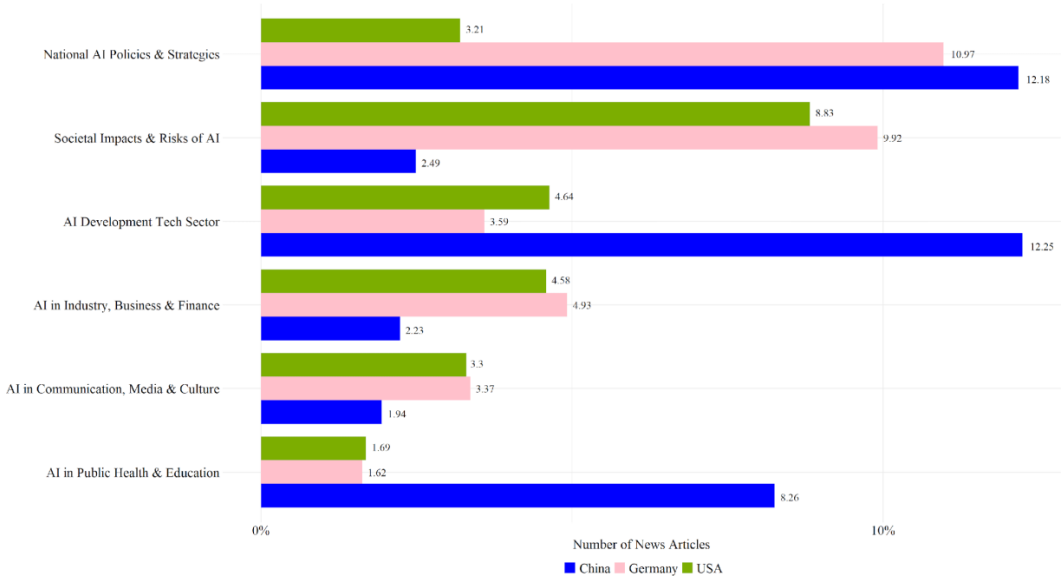
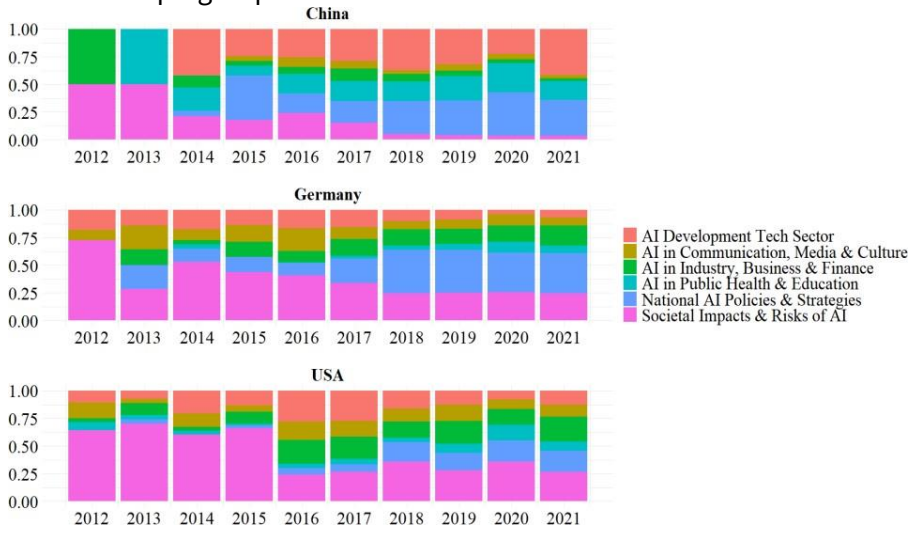


Figure 2. Evolution of topic groups over time



Contribution

This paper is one of the first to provide a comprehensive cross-lingual analysis of AI-related news discourses in China, Germany, and the U.S. Conceptually, the study contributes to advancing the theorization and operationalization of sociotechnical imaginaries in the context of AI. Methodologically, the study demonstrates how computational analysis and qualitative research can be utilized in conjunction to identify and make sense of discursive patterns in public communication technologies in a transnational context.

References

Brause, S. R., Zeng, J., Schäfer, M. S., & Katzenbach, C. (forthcoming). Media Representations of Artificial Intelligence, Surveying the Field. In Simon Lindgren (Ed.), Handbook of Critical Studies of Artificial Intelligence.

Fereday, J., & Muir-Cochrane, E. (2006). Demonstrating rigor using thematic analysis: A hybrid approach of inductive and deductive coding and theme development. International Journal of Qualitative Methods, 5(1), 80–92.

Gerhards, J., & Schäfer, M. S. (2009). Two normative models of science in the public sphere: human genome sequencing in German and US mass media. Public Understanding of Science, 18(4), 437–451.

Grootendorst, M. (2022). BERTopic: Neural topic modeling with a class-based TF-IDF procedure. arXiv preprint arXiv:2203.05794.

Jasanoff, S., & Kim, S.-H. (2009). Containing the Atom: Sociotechnical Imaginaries and Nuclear Power in the United States and South Korea. Minerva, 47(2), 119–146.

Restructuring social science communication in social media: An ethnographic study on the influence of algorithmic imaginaries in content creation

Clarissa E. Walter, Anne K. Krüger, Sascha Friesike

Theory and Research Question

Research in science communication is concerned with how recommendation algorithms on social media impact content visibility (Hoang, 2020) and thus influence content creation (Weingart, 2017). However, there is a lack of empirical research on how exactly recommendation algorithms affect content creation in science communication. This paper seeks to address this gap by offering empirical insights from an ethnographic study. In this paper we explore how "digital experts" perceive algorithms and how they translate "algorithmic imaginaries" (Bucher, 2017) into requirements that affect content creation. In this sense we draw on the analysis of "actual practices surrounding algorithmic technologies" (Christin, 2017, p. 1).

We argue that it is precisely the "sense-making" surrounding recommendation algorithms that profoundly affects content creation in science communication. In more detail, we examine the consulting practices of "digital experts" and delineate three requirements that they derived from algorithmic imaginaries: (1) Include diverse audiences, (2) Address user needs, (3) Persuade first, explain later.

Case

The article examines an 18-month long format development and production process of a YouTube channel communicating social science theories (Lewis et al., 2023).

The video channel and its content is co-produced by a German public broadcaster and a team of academic social scientists. The authors explored the collaborative content creation between the broadcaster's "digital experts" who claim authoritative knowledge on recommendation algorithms, and the academics, who assume responsibility for both editorial work and video moderation.

Data and Methods

In order to gain insights into the production process, the first author has conducted an ethnographic study from January 2022 to July 2023 while actively participating as one of three academics in the project. Such an ethnographic approach is particularly adept when examining work practices surrounding algorithmic technologies (Pinch & Bijker, 1984; Seaver 2017). During the fieldwork, the first author was able to capture the manifold stages of content creation and optimization, where imaginaries of algorithms played a pivotal role. Our data includes field notes from around 1200 hours of observation, including 36 full days of production in the broadcaster's studios as well as detailed documents of collaborative script writing, post-production, and a three-day workshop on digital content creation. After an initial round of inductive coding, our qualitative content analysis (Miles, 1994) focused on three themes: (a) algorithmic imaginaries, which Bucher (2017, p. 30) defines as ''ways of thinking about what algorithms are, what they should be and how they function'', (b) concluded requirements in content creation (c) frictions between digital experts and academics.

Findings

The prevailing imaginary of YouTube's recommendation algorithm among digital experts is that of a referee which decides on content visibility. They understand the referee's role on the platform in securing "user satisfaction" by selecting content that best meets the "users' needs". The narrative of "if you satisfy the user, the algorithm will reward you" therefore guides their advice on content creation. The relevance of this imaginary is illustrated through user statistics, which are provided by YouTube Analytics. To capture how this imaginary influences content creation, we have identified three requirements guiding digital experts' consulting practice.

Firstly, digital experts required science communicators to include diverse audiences. On the one hand, this means addressing users who already consume other science content on the platform. On the other hand, it means that the scientific nature of the language should be reduced so as not to exclude anyone. They recommended avoiding technical jargon, focusing on general interest topics, and creating attention-grabbing thumbnails. Or as they repeatedly summarised it: "You have to fill the church before you start preaching."
Secondly, they required science communicators to address user needs. In this case, it was argued that the question of how a particular theory works, is not a user need. Rather, a need comes more from everyday life. This requirement has changed content from explaining social science theories to explaining everyday problems with the help of social science theories.
Thirdly, digital experts required science communicators to persuade first, explain later. A promise, it was advised, keeps viewers watching a video for longer and consequently increases YouTube analytics' 'watch time' metric. This requirement has changed the production so that the beginning of each video is structured in such a way that you comprehend what you will understand when you have watched the video.
Taken together, “digital experts” translated these imaginaries of algorithmic evaluation criteria into requirements for content creation with the intention of achieving more visibility on the platform. The imaginaries were used as symbolic resources in the development process of the science communication content. While acknowledging the significant role of science within the YouTube platform, they advised to restructure the content according to their understanding of the recommendation algorithms (e.g., prioritising persuasion practices over scientific rigour). As assumed by Weingart (2017; 2020), recommendation algorithms indirectly influence the content of science communication, yet, as shown in this paper, particularly through “digital experts”, a new professional group that claims knowledge about opaque algorithms of social media platforms.

References

Bucher, T. (2017). The algorithmic imaginary: Exploring the ordinary affects of Facebook algorithms. Information, Communication & Society, 20(1), 30-44.

Bundesministerium für Bildung und Forschung (BMBF) (2021). Kompetenzaufbau Wissenschaftskommunikation. Bonn: BMBF.

Christin, A. (2017). Algorithms in practice: Comparing web journalism and criminal justice. Big Data & Society, 4(2), 2053951717718855.

Hoang, L. N. (2020). Science communication desperately needs more aligned recommendation algorithms. Frontiers in Communication, 5, 598454.

Lewis, J., Bartlett, A., Riesch, H., & Stephens, N. (2023). Why we need a public understanding of social science. Public understanding of science, 09636625221141862.

Miles, M. B., & Huberman, A. M. (1994). Qualitative data analysis: An expanded sourcebook. sage.

Pinch T. J. & Bijker W. E. (1984) The social construction of facts and artefacts. Or how the sociology of science and the sociology of technology might benefit each other. Social Studies of Science 14(3), 399–411.

Seaver, N. (2017). Algorithms as culture: Some tactics for the ethnography of algorithmic systems. Big data & society, 4(2), 2053951717738104.

Weingart, P. (2017). Wissenschaftskommunikation unter digitalen Bedingungen. Funktionen, Akteure und Probleme des Vertrauens. In Weingart, P., Wormer, H., Wenninger, A. & Hüttl, R. F. (Ed.). Perspektiven der Wissenschaftskommunikation im digitalen Zeitalter (29-59). Velbrück Wissenschaft.

Weingart, P. (2022). Trust or attention? Medialization of science revisited. Public Understanding of Science, 31(3), 288-296.

Parallel Panel II: Open panel: CLIMATE- AND HEALTH-RELATED SCIENCE COMMUNICATION

Thursday, June 6, 11:20 – 12:40
Chair: Elena Link
Room: HG E 1.1

Detecting manipulated visuals: A computational approach in the climate change discourse

Isaac Bravo, Katharina Prasse, Stefanie Walter, Margret Keuper

The emergence of Artificial Intelligence (AI) models is reshaping how scientists interact with this technology when conducting research. This has led to a growing interest in the role and impact of AI in the field of scientific communication over the last few years (Schäfer, 2023). The benefits of this technology that we can find as researchers are varied (Chian & Lee, 2023; Krishnan et al., 2023). Simultaneously, manipulated images pose a misinformation risk when viewers cannot determine the credibility of what they see (He, 2021). While computer science research in this field commonly uses images from "lab scenarios" to identify manipulated content, we focus on the polarized topic of climate change to explore how manipulated images shared on Twitter may contribute to polarizing debates between believers and sceptics (deniers) in anthropogenic climate change. This study contributes to two important aspects in the debate on science communication in the age of AI: a) How manipulated images are spreading on the social media landscape, and the resulting impact on public debates, and b) how methodological advances in machine learning and computer vision can help scientists to detect and analyse manipulated images.

Climate change is a global phenomenon which has received considerable media attention in recent decades (IPCC, 2022). Experts have recognized the urgency of addressing the impacts of this phenomenon and understanding how people perceive and engage with it (Falkenberg et al., 2022). In the digital media environment, the emergence of climate change information on social media, especially visual content, has changed how individuals understand this phenomenon and how it encourages collective action (Pearce et al., 2019). Despite the scientific evidence of the anthropocentric origin of climate change, contrarian voices still reject this reality and the related risks (though in many cases, these are minority opinions (Whitmarsh, 2011). The causes of such positions include disagreement between scientists (Patt, 2007) and people's attitudes and beliefs (Kahan et al., 2012).

Social media allows the circulation of opinions and manipulated content that may support or deny certain beliefs or facts. However, the emergence of image generative models has made it much easier for people to create ‘deep fake’ or synthetic images, and simultaneously, computer science research has started to improve their detection (Guan, 2022). Manipulated visual content can, for example, exaggerate or misrepresent climate change-related phenomena, mislead individuals, fuel scepticism regarding the veracity of climate change, and potentially erode trust in media and scientific institutions (Capstick & Pidgeon, 2014). Due to the limited development of computational studies, previous work mainly focused on analysing visual elements of climate change using qualitative approaches and small samples (Schäfer, 2020; Harb et al., 2020; Metag et al., 2016). Notably, there is a lack of studies on the prevalence and impact of manipulated content and how users interact with this type of visual content related to climate change on social media.

This study aims to answer the following research question: Do ‘real’ vs manipulated images about climate change on Twitter lead to different levels of engagement and interactions between believers and skeptics (deniers)? This study adopts a multimodal and computational approach combining

automated image and text analysis to examine more than 700,000 images, and replies shared by Twitter users in the year 2019 in the context of climate change. For the data collection, we sampled all tweets that included an image and the term "climate change" or the hashtag "#climatechange" in English. We compare methods detecting manipulated images, such as hash functions, to retrieve pairs of near-identical images, and analyse the manipulation that took place between the 'real' and the manipulated image within each pair. More precisely, we use the "Crop-Resistant Hash" proposed by Steinbach et al., where individual segments of the images are hashed and then compared to identify near-identical images. This method allows the matching even after major image manipulations. Then, we use different computational techniques such as topic modelling (BERTopic) and Latent Semantic Scaling (LSS) to analyze and classify comments between believers and sceptics.

Preliminary results reveal user differences in the distribution of engagement between manipulated and real images, as well as a concentration of user interactions around specific topics related mainly to the consequences of climate change. Furthermore, these differences concern not only the type of visual content engaged by deniers and believers but also how these users react to it. Here, we generally see that the believers exhibit more engagement than deniers when they are exposed to real images. While manipulated images only make up a small share of the total number of images shared on Twitter, they can lead to considerable user engagement and directly impact people's understanding of climate change.

This study makes a theoretical and methodological contribution to the field of science communication: From a theoretical perspective, the research delves into the framing strategies adopted by believers and sceptics (deniers) on Twitter, specifically comparing the use of manipulated and real images on a polarized climate change debate. The methodological contribution lies in using a computer science approach to detect manipulated images within the context of climate change-related data. By employing advanced detection techniques, the study also contributes to our understanding of the effectiveness of existing models in discerning between authentic and manipulated content in the climate change debate on social media platforms.

References

Boykoff, M., Aoyagi, M., Ballantyne, A., et al. (2022). World newspaper coverage of climate change or global warming, 2004-2022. Media and Climate Change Observatory Data Sets. Cooperative Institute for Research in Environmental Sciences, University of Colorado https://sciencepolicy.colorado.edu/icecaps/research/media coverage/world/index.html.

Capstick, S.B., & Pidgeon, N.F. (2014). What is climate change scepticism? Examination of the concept using a mixed methods study of the UK public. Global Environmental Change-human and Policy Dimensions, 24, 389-401.

Chiang, C., & Lee, H. (2023). Can Large Language Models Be an Alternative to Human Evaluations? Annual Meeting of the Association for Computational Linguistics.

Falkenberg, M., Galeazzi, A., Torricelli, M., Di Marco, N., Larosa, F., Sas, M., Mekacher, A.,

Pearce, W., Zollo, F., Quattrociocchi, W., & Baronchelli, A. (2022). Growing polarization around climate change on social media. Nature Climate Change, 12(12), 1114–1121. https://doi.org/10.1038/s41558-022-01527-x

Guan, J., Zhou, H., Hong, Z., Ding, E., Wang, J., Quan, C., & Zhao, Y. (2022). Delving into sequential patches for deepfake detection. Advances in Neural Information Processing Systems, 35, 4517-4530.

Harb, J. G. D., Ebeling, R., & Becker, K. (2020). A framework to analyze the emotional reactions to mass violent events on Twitter and influential factors. Information Processing & Management, 57(6), 102372. https://doi.org/10.1016/j.ipm.2020.102372

He, Y., Yu, N., Keuper, M., & Fritz, M. (2021). Beyond the Spectrum: Detecting Deepfakes via Re-Synthesis (arXiv:2105.14376). arXiv. https://doi.org/10.48550/arXiv.2105.14376

IPCC, 2022: Climate Change 2022: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [H.-O. Pörtner, D.C. Roberts, M. Tignor, E.S. Poloczanska, K. Mintenbeck, A. Alegría, M. Craig, S. Langsdorf, S. Löschke, V. Möller, A. Okem, B. Rama (eds.)]. Cambridge University Press. Cambridge University Press, Cambridge, UK and New York, NY, USA, 3056 pp., doi:10.1017/9781009325844.

Kahan, D., Peters, E., Wittlin, M., Slovic, P., Ouellette, L.L., Braman, D., & Mandel, G.N. (2012). The Polarizing Impact of Science Literacy and Numeracy on Perceived Climate Change Risks. Law & Psychology eJournal.

Krishnan, A., & Anoop, V.S. (2023). ClimateNLP: Analyzing Public Sentiment Towards Climate Change Using Natural Language Processing. ArXiv, abs/2310.08099.

Metag, J., Schäfer, M. S., Füchslin, T., Barsuhn, T., & Kleinen-von Königslöw, K. (2016). Perceptions of Climate Change Imagery: Evoked Salience and Self-Efficacy in Germany, Switzerland, and Austria. Science Communication, 38(2), 197–227. https://doi.org/10.1177/1075547016635181

Patt, A. (2007) Assessing model-based and conflict-based uncertainty. Global Environmental Change, 17, 37–46.

Pearce, W., Niederer, S., Ozkula, S., & Sanchez-Querubin, N. (2019). The social media life of climate change: Platforms, publics, and future imaginaries. Wiley Interdisciplinary Reviews: Climate Change, 10. https://doi.org/10.1002/wcc.569

Schäfer, M. S. (2023). The Notorious GPT: Science communication in the age of artificial intelligence. JCOM - Journal of Science Communication, 22(2), Y02. https://doi.org/10.22323/2.22020402

Schäfer, M. S. (2020). Introduction to visualizing climate change. In: Research Handbook on Communicating Climate Change (pp. 127-130). Edward Elgar Publishing.

Taskesen, E. (2022). undouble - Python library to detect (near-)identical images. (Version 1.2.0) [Computer software]. https://erdogant.github.io/undouble

Whitmarsh, L. (2011). Scepticism and uncertainty about climate change: Dimensions, determinants and change over time. Global Environmental Change-human and Policy Dimensions, 21, 690-700.

Politicization of science in German COVID-19 media coverage: Theoretical conceptualization and empirical evidence

Janise Brück, Julia Serong, Lars Guenther

In times of crisis, such as the COVID-19 pandemic, societies are most dependent on public trust in science, as science-based information provides explanation and guidance (Bromme et al., 2022). Simultaneously, public debates about COVID-19 have shown that scientific issues often have a political dimension (e.g., Post & Ramirez, 2018, p. 1151; Van Dooren & Noordegraaf, 2020, p. 614), and thus, the pandemic is assumed to have triggered a significant blurring of the boundaries between the scientific and political systems (Weingart, 2010, p. 157). The so-called politicization of science describes the process by which science acquires a political meaning and is used to pursue political goals (Schmid-Petri et al., 2022, p. 49). This includes both benefits (e.g., the exchange of resources in knowledge production) as well as threats to the science system: if scientific information are questioned or rejected by members of the public due to political ideologies, politicization can have a negative impact on public trust in science and its authority (e.g., Bolsen et al., 2019, p. 148; Gauchat, 2012, p. 170). Despite its considerable societal relevance, previous research has not adequately addressed politicization as a multifaceted phenomenon. Instead, it refers to different understandings and operationalize the construct divergently (e.g., Bolsen & Druckman, 2015; Hart et al., 2020; Schmidt, 2023). Therefore, this study aims to theoretically conceptualize politicization of science and to test the operationalization of the concept in a pilot study – a quantitative content analysis of German COVID-19 media coverage.

Against the background of Luhmann's systems theory (1984) and Habermas' theoretical reflections (1968) on science, politics, and the public sphere, the approach concentrates on the triangular relationship between science, politics, and the media as drivers of the politicization of science (Schmid-Petri et al., 2022, p. 49). The relationship is based on dependencies that are caused by resource exchange (e.g., funding in exchange for policy-relevant research) and promote a blurring of science and politics (Scholten & Verbeek, 2014, p. 189). Moreover, politics and science deal with similar (societal) issues which is why both scientific and political topics can occur in political or scientific debates (Bolsen & Palm, 2022, p. 86). However, political actors may also refer to scientific information for strategic purposes (e.g., to (de)legitimize political positions; Post & Ramirez, 2018, p. 1152), and scientists comment on political issues due to their assigned role (Pielke, 2007). Journalistic actors bring it all together by communicating scientific issues to the public via specific criteria and placing them in a political context (Fowler & Gollust, 2015, p. 157); hence, media reporting is the focus of this study.

Based on a comprehensive literature review of the triangular relationship, the aspects that point to a politicization of science in media coverage can be summarized in three overarching indicators: (1) The thematic blurring of politics and science means that scientific and political issues are linked by reference to (controversial) political angles, (political) story origins or interdependencies of scientific and political resources in scientific reporting (e.g., Schmid-Petri, 2017, p. 523; Schmidt, 2023, p. 35). Moreover, the extent to which different actors are addressed or comment on scientific or political issues can have an impact on the (2) (politicized) actor structure in the public (media) discourse (Leidecker-Sandmann & Lehmkuhl, 2022, p. 344). Central here is the (politicized) constellation of scientific and political actors in media reporting, based on the distinction of whether actors are mentioned or quoted (e.g., Chinn et al., 2020; Schmidt, 2023, p. 35). Finally, the (3) emphasis on science-related uncertainty is primarily expressed by political and journalistic actors, who (strategically) highlight the vulnerability of science to (de)legitimize political positions (Post & Ramirez, 2018, p. 1152) or to create news value (Fowler & Gollust, 2015, p. 158).

Considering the applicability of these three indicators and their relevance in crisis contexts, the empirically focused pilot study analyzed the extent to which science has been politicized in German COVID-19 reporting. Using a codebook developed for this purpose, a pre-registered, quantitative content analysis (ANONYMIZED) examined the indicators in COVID-19 related (offline/ online) media articles (N = 262) from journalistic quality media (FAZ, SZ, Die Zeit) regarding two selected time periods in the first and second pandemic wave in 2020 in Germany (t1: March 2nd – April 19th; t2: August 17th – October 4th).

The findings showed that, in terms of the indicators, there was no evidence of a significant politicization of science in the early phase of the pandemic; however, it is reasonable to assume a general political significance of science in COVID-19 media coverage. The data indicated a noticeable thematic blurring of politics and science as well as a political actor structure in the public (media) discourse right from the beginning of the pandemic. While the coupling between scientific and political actors was stronger in the media coverage of the first COVID-19 wave than in the second wave, solely political controversy and political story origins are more pronounced in later coverage and could be an indication of an increasing blurring. In contrast, the emphasis on science-related uncertainty was barely present in either time period.

Although our study focuses on a conceptualization of the politicization of science (in COVID-19 media coverage), the political relevance of science is not only prevalent in crises but also in general debates about scientific issues – including AI. Being aware of our limitations, we would like to discuss this study in the Open Panel at #aiscicomm24.

References

Bolsen, T., & Druckman, J. N. (2015). Counteracting the Politicization of Science. Journal of Communication, 65(5), 745-769. https://doi.org/10.1111/jcom.12171

Bolsen, T., & Palm, R. (2022). Chapter Five - Politicization and COVID-19 vaccine resistance in the U.S. In T. Bolsen, & R. Palm (Eds.), Progress in Molecular Biology and Translational Science (ed. 188, pp. 81-100). Academic Press. https://doi.org/10.1016/bs.pmbts.2021.10.002

Bolsen, T., Palm, R., & Kingsland, J. T. (2019). Counteracting Climate Science Politicization With Effective Frames and Imagery. Science Communication, 41(2), 147-171. https://doi.org/10.1177/1075547019834565

Bromme, R., Mede, N. G., Thomm, E., Kremer, B., & Ziegler, R. (2022). An anchor in troubled times: Trust in science before and within the COVID-19 pandemic. PLOS ONE, 17(2), 1-27. https://doi.org/10.1371/journal.pone.0262823

Chinn, S., Hart, P. S., & Soroka, S. (2020). Politicization and Polarization in Climate Change News Content, 1985-2017. Science Communication, 42(1), 112-129. https://doi.org/10.1177/1075547019900290

Fowler, E. F., & Gollust, S. E. (2015). The Content and Effect of Politicized Health Controversies. The ANNALS of the American Academy of Political and Social Science, 658(1), 155-171. https://doi.org/10.1177/0002716214555505

Gauchat, G. (2012). Politicization of Science in the Public Sphere: A Study of Public Trust in the United States, 1974 to 2010. American Sociological Review, 77(2), 167-187. http://www.jstor.org/stable/23102567

Habermas, J. (1968). Wissenschaft und Technik als Ideologie. Suhrkamp Verlag.

Hart, P. S., Chinn, S., & Soroka, S. (2020). Politicization and Polarization in COVID-19 News Coverage. Science Communication, 42(5), 679-697. https://doi.org/10.1177/1075547020950735

Leidecker-Sandmann, M., & Lehmkuhl, M. (2022). Politisierung oder Aufklärung? Analysen der Akteur: innen- und Aussagenstruktur in medialen Diskursen über gesundheitliche

Risikophänomene und die Rolle wissenschaftlicher Expert:innen. SCM Studies in Communication and Media, 11(3), 337-393. https://doi.org/10.5771/2192-4007-2022-3-337

Luhmann, N. (1984). Soziale Systeme. Grundriß einer allgemeinen Theorie. Suhrkamp Verlag.

Pielke, R. A. (2007). The Honest Broker: Making Sense of Science in Policy and Politics. Cambridge University Press. https://doi.org/10.1017/CBO9780511818110

Post, S., & Ramirez, N. (2018). Politicized Science Communication: Predicting Scientists' Acceptance of Overstatements by Their Knowledge Certainty, Media Perceptions, and Presumed Media Effects. Journalism & Mass Communication Quarterly, 95(4), 1150-1170. https://doi.org/10.1177/1077699018769668

Scheufele, D. A. (2014). Science communication as political communication. Proceedings of the National Academy of Sciences, 111(suppl. 4), 13585-13592. https://doi.org/10.1073/pnas.1317516111

Schmid-Petri, H. (2017). Politicization of science: how climate change skeptics use experts and scientific evidence in their online communication. Climatic Change, 145(3), 523-537. https://doi.org/10.1007/s10584-017-2112-z

Schmid-Petri, H., Bienzeisler, N., & Beseler, A. (2022). Chapter Three - Effects of politicization on the practice of science. In T. Bolsen, & R. Palm (Eds.), Progress in Molecular Biology and Translational Science (ed. 188, pp. 45-63). Academic Press. https://doi.org/10.1016/bs.pmbts.2021.11.005

Schmidt, H. (2023). Pandemics and Politics: Analyzing the politicization and polarization of pandemic- related reporting. Newspaper Research Journal, 44(1), 26-52. https://doi.org/10.1177/07395329221095850

Scholten, P., & Verbeek, S. (2014). Politicization and expertise: Changing research – policy dialogues on migrant integration in Europe. Science and Public Policy, 42(2), 188-200. https://doi.org/10.1093/scipol/scu040

Van Dooren, W., & Noordegraaf, M. (2020). Staging Science: Authoritativeness and Fragility of Models and Measurement in the COVID-19 Crisis. Public Administration Review, 80(4), 610-615. https://doi.org/10.1111/puar.13219

Weingart, P. (2010). Resonanz der Wissenschaft der Gesellschaft. In C. Büscher, & K. P. Japp (Eds.), Ökologische Auffilärung: 25 Jahre „Ökologische Kommunikation“ (pp. 157-172). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-92425-0_6

Science calls to action: Explainer videos as activist tool to promote sustainable streaming practices

Anna Schorn, Romina Behrend, Werner Wirth

When it comes to sustainable behavior, topics such as mobility, nutrition, or fashion are often discussed. However, a topic that has been largely overlooked by society are greenhouse gas (GHG) emissions caused by digital technology and the internet (e.g., The Shift Project, 2020). The entire digital technology sector already causes more GHG emissions than global aviation. More than half of these emissions originate from video streaming. By 2025, streaming could account for almost 6% of global GHGemissions—more than aviation and nearly as much as car traffic (The Shift Project, 2019). Therefore, achieving climate goals requires bringing this issue more into the focus of society and raising awareness about it.

A recent study indicates that an explainer video can play a decisive role in reducing GHG emissions through streaming (Seger et al., 2023). Despite the GHG emissions associated with providing such online videos, they serve as effective tools for various reasons. For example, videos can be integrated directly into the streaming platforms, addressing people where behavioral change is targeted. Moreover, explainer videos on scientific topics and climate change are popular and increasingly used on social media, meaning that they can reach a relevant audience (Allgaier, 2019; Koch & Bleisch, 2020). Furthermore, explainer videos can be highly persuasive due to design features such as storytelling, theuse of exemplars, an informal communication style, and the combination of voiceover with simple, meaningful animations (Schorn, 2022).

Simultaneously, many explainer videos pursue an agenda: YouTube channels that have taken up the cause of science communication (e.g., Kurzgesagt, Mai Lab, TED Ed) explain certain topics with a scientific basis, whereby they adopt an activist stance. The videos are characterized by the fact that information is not presented in a completely neutral way, and often there is a call-to-action at the end. Explainer videos are thus not only used to convey knowledge, but also to pursue persuasive purposes. However, previous research refers almost exclusively to learning effects and not to attitude and behavior changes (Schorn, 2022).

Therefore, the aim of this research is to investigate how science-based explainer videos can promote sustainable streaming practices. On the one hand, the study focuses on various call-to-actions (e.g., personal and social norm appeals) that complement the explanation. On the other hand, the study aims to enhance self-efficacy and behavioral control by providing concrete measures to reduce emissions caused by video streaming. A pilot study demonstrated that the topic is novel to many people and before being presented with relevant information, many were unaware of how they could prevent emissions from video streaming. Consequently, the main study will examine how specific measures to increase sufficiency and efficacy of video streaming can increase perceived behavioral control and self-efficacy, ultimately fostering the intention to change behavior.

This research is intended to provide insights into how scientific topics can be communicated in a way that not only expands recipients' knowledge but also stimulates them to question their own behavior, leading to a change in behavior towards more sustainable streaming practices. It is hypothesized that videos containing concrete measures for sustainable streaming will improve perceived (self-)efficacy and behavioral control and increase the intention to change behavior. Moreover, a call-to- action is expected to have a positive effect on behavioral intention. Without setting up a directed hypothesis, the study examines whether a personal norm appeal (emphasizing the moral obligation to act) or a social norm appeal (emphasizing that a shift in behavior is socially desirable and popular) is more effective (see OSF for manipulation).

In the main study, we will conduct a 3 (call-to-action: personal norm appeal vs. social norm appeal vs. control group) x 2 (concrete measures: present vs. absent) between-subjects-design to investigate how explainer videos can be optimized to be particularly persuasive and motivate more

sustainable streaming practices. The explainer videos will be created by a design agency (see OSF for storyboard). Participants will be recruited and compensated by a German market research institute (*N* =800). The study will be announced as a “study on video streaming”.
In the beginning, participants' streaming behavior will be queried without indicating that the study is about sustainability. After asking about platforms, viewing duration and devices, participants will be asked several questions that can provide information about the environmental impact of their streaming behavior (e.g., “I watch videos in the best possible quality”). Participants will then be randomly assigned to an experimental condition and asked to watch the respective video, before the questions on streaming behavior are asked again as dependent variables (e.g., “In the future, I plan to stream in a lower quality than before”). Individual users could either follow a sufficiency strategy by reducing their streaming duration, or increase the efficiency of video streaming by reducing the carbon impact per hour (e.g., by reducing the size of the screen or resolution). By asking about the current state beforehand, it will be possible to deduce the extent to which individuals intend to change their behavior in the future. Next, other variables such as reactance, perceived quality, self-efficacy, or behavioral control will be measured. 250 of the participants will be asked retrospective after a few weeks whether they have actually changed their behavior and implemented measures for more sustainable streaming.

The study was approved by the university’s ethics committee. The videos are currently being animated by the design agency. Subsequently, the study will be conducted so that the results can be presented in June.

References

Allgaier, J. (2019). Science and environmental communication on YouTube: Strategically distortedcommunications in online videos on climate change and climate engineering. *Frontiers inCommunication*, *4*, 36. https://doi.org/10.3389/fcomm.2019.00036

Koch, W., & Bleisch, N. (2020). Ergebnisse der ARD/ZDF-Onlinestudie 2020: Erneut starke Zuwächse beiOnlinevideo. *Media Perspektiven*, *9*, 482–500.

Schorn, A. (2022). Online explainer videos: Features, benefits, and effects. *Frontiers in Communication*, *7*,1034199. https://doi.org/10.3389/fcomm.2022.1034199

Seger, B. T., Burkhardt, J., Straub, F., Scherz, S., & Nieding, G. (2023). Reducing the individual carbonimpact of video streaming: A seven-week intervention using information, goal setting, and feedback. *Journal of Consumer Policy*, *46*(2), 137–153. https://doi.org/10.1007/s10603-023-09536-9

The Shift Project. (2019). *Climate Crisis: The Unsustainable Use of Online Video*.https://theshiftproject.org/en/lean-ict-2/ The Shift Project. (2020). *Expanding digital sufficiency: Executive summary*.

Supplementary Material on OSF:
https://osf.io/uf4rt/?view_only=e39fd7ea5f5e466b88f29f85d150a04a

"However I judge it, I think it's a gut feeling": Examining (un)trustworthiness cues in YouTube videos by real and feigned experts

Kaija Biermann, Monika Taddicken

Particularly in digital communication environments, misinformation on socio-scientific issues (e.g., Covid-19, climate change) can develop and spread rapidly (Freiling et al., 2023). Hence, epistemic authorities, including those of scientific experts, can be challenged, as social media platforms involve opening up professional roles to new actors (Neuberger et al., 2023). Feigned experts can present themselves to laypeople as experts without actually having the necessary expertise (Cook et al., 2017), making it difficult for laypeople to judge the expert status of a communicator (Hendriks et al., 2015). However, in circumstances where laypeople lack knowledge, which is typically the case for most scientific issues, they have to rely on trustworthy sources (Bromme & Gierth, 2021; Bromme & Goldman, 2014). Especially during the Covid-19 pandemic, laypeople were faced with the challenge of finding reliable information from such trustworthy sources in the multitude of available sources, as the pandemic was accompanied by an "infodemic" (WHO, 2020).

Although YouTube is one of the most widely used channels for information about science and research (WiD, 2021), research that deals with YouTube in the context of trust in science and scientific experts is scarce (e.g., Reif et al., 2020). Little is known about the complex mechanisms of recognizing real from feigned experts and the cues laypeople rely on to judge trustworthiness in audio-visual material. Our exploratory study aims to gain deeper insights into laypeople's subjective process of assessing trustworthiness of (feigned) scientificexperts in YouTube videos on Covid-19, hence, we ask:

RQ: On which cues rely laypeople to assess (feigned) scientific experts' trustworthiness in YouTube videos on Covid-19?

To answer our research question, we conducted 33 semi-structured interviews with laypeople[6] using the think-aloud-technique (Ericsson & Simon, 1998). As this study is part ofa larger preregistered research project, we were able to select the four most trustworthy rated video snippets of real and feigned experts on Covid-19 in a survey (n = 958) as stimulus material[7]. In order to put laypeople in the actual situation of evaluating trustworthiness, each interviewee were asked to think aloud when watching two of the four most trustworthy video snippets of real and feigned experts respectively. This approach is more reliable than asking laypeople which cues they think they use to determine trustworthiness (cf., Freiling, 2019). The thinking aloud part included training tasks to practice the method of thinking aloud, moreover, after assessing each video, the interviewees were asked again about the cues they used to assess trustworthiness. The interview guide included follow-up questions on cues thathad already been derived from the literature, such as references to the communicator's expertise or the importance of the channel, facial expressions, and gestures (Meinert & Krämer, 2022; Winter & Krämer, 2016). The data was analyzed using content-structuring qualitative content analysis according to Mayring (2015). In a deductive-inductive approach, overarching categories were first defined on the basis of the interview guide, subsequently, an inductive subcategorization was carried out.

A large number of different trustworthiness cues were identified, whereby the "experts" were rated differently by different interviewees in terms of their trustworthiness. Unsurprisingly, *familiarity of the "expert"* was frequently mentioned as trustworthiness cue, as did the *channel* and *format* (e.g., press conference). For all interviewees, *mentioning the title*, particularly in connection

[6] In order to obtain a picture as comprehensive as possible of trustworthiness cues, attention was paid to different age groups, educational levels, and occupational groups in addition to an equal distribution of gender in the sample.

[7] The total of 223 video snippets to be evaluated are based on a sample (n = 66) consisting of the most clicked YouTube videos (at least 10,000 views) on Covid-19 by real and feigned experts, which were determined in a multi-stage process.

with the *professional background*, was a decisive trustworthiness cue. Additionally, various trustworthiness cues were mentioned that relate to the *type of communication* (scientific jargon, comprehensibility). A large number of interviewees cited the *presentation of data and facts* as indicators of trustworthiness; "experts" were rated as less trustworthy when they expressed opinions from the perspective of the laypeople and when they commented on political issues that went beyond science. Additionally, a *calm voice, slow*, and *fluent speech* (communication style) were mentioned as trustworthiness cues. Non-verbal cues such as *friendly facial expression*, *straight posture*, and*gestures* were also frequently cited. *Outward appearance* also appears to be important for many interviewees when assessing the (feigned) scientific experts' trustworthiness. In addition to such visible cues, the interviewees also mentioned the *perceived closeness* to the "experts" and the *perceived sympathy*, while a perceived arrogance was rated as untrustworthy. Ultimately, interviewees also frequently cited their *gut feeling* as a decisive factor in assessing the trustworthiness.

Overall, our results indicate that a variety of different cues are used by laypeople to assess (feigned) experts' trustworthiness in digital communication environments, especially insituations in which the "experts" are unknown to them. At the conference, the identified trustworthiness cues will be put more in relation to each other and with real and feigned experts. As a limiting factor, it should be noted that these are ultimately only the perceived cues. It is unclear to what extent these expressed cues actually affect laypeople's trustworthiness assessments. However, based on the identified cues, our study offers a starting point to further examine laypeople's trustworthiness judgments in the context of (feigned) scientific experts on social media and ultimately contribute to empowering laypeople to distinguish between real and feigned experts.

References

Bromme, R., & Gierth, L. (2021). Rationality and the public understanding of science. In M.Knauff & W. Spohn (Eds.), *The handbook of rationality* (pp. 767–776). The MIT Press. https://doi.org/10.7551/mitpress/11252.003.0084
Bromme, R., & Goldman, S. R. (2014). The public's bounded understanding of science. *Educational Psychologist*, *49*(2), 59–69. https://doi.org/10.1080/00461520.2014.921572

Cook, J., Lewandowsky, S., & Ecker, U. K. H. (2017). Neutralizing misinformation throughinoculation: Exposing misleading argumentation techniques reduces their influence. *PLOS ONE*, *12*(5), e0175799. https://doi.org/10.1371/journal.pone.0175799

Ericsson, K. A., & Simon, H. A. (1998). How to study thinking in everyday life: Contrastingthink-aloud protocols with descriptions and explanations of thinking. *Mind, Culture, and Activity*, *5*(3), 178–186. https://doi.org/10.1207/s15327884mca0503_3

Freiling, I. (2019). Detecting misinformation in online social networks: A think-aloud studyon user strategies. *Studies in Communication and Media*, *8*(4), 471–496. https://doi.org/10.5771/2192-4007-2019-4-471

Freiling, I., Krause, N. M., Scheufele, D. A., & Brossard, D. (2023). Believing and sharingmisinformation, fact-checks, and accurate information on social media: The role ofanxiety during COVID-19. *New Media & Society*, 14614448211011451. https://doi.org/10.1177/14614448211011451

Hendriks, F., Kienhues, D., & Bromme, R. (2015). Measuring laypeople's trust in experts in adigital age: The Muenster Epistemic Trustworthiness Inventory (METI). *PLOS ONE*, *10*(10), e0139309. https://doi.org/10.1371/journal.pone.0139309

Mayring, P. (2015). Qualitative content analysis: Theoretical background and procedures. In A. Bikner-Ahsbahs, C. Knipping, & N. Presmeg (Eds.), *Advances in MathematicsEducation. Approaches to qualitative research in mathematics education* (pp. 365–380). Springer Netherlands. https://doi.org/10.1007/978-94-017-9181-6_13

Meinert, J., & Krämer, N. C. (2022). How the expertise heuristic accelerates decision-makingand credibility judgments in social media by means of effort reduction. *PLOS ONE*, *17*(3), e0264428. https://doi.org/10.1371/journal.pone.0264428

Neuberger, C., Bartsch, A., Fröhlich, R., Hanitzsch, T., Reinemann, C., & Schindler, J. (2023). The digital transformation of knowledge order: A model for the analysis of theepistemic crisis. *Annals of the International Communication Association*, 1–22. https://doi.org/10.1080/23808985.2023.2169950

Reif, A., Kneisel, T., Schafer, M., & Taddicken, M. (2020). Why are scientific experts perceived as trustworthy? Emotional assessment within TV and YouTube videos. *Media and Communication*, *8*(1), 191–205. https://doi.org/10.17645/mac.v8i1.2536

WHO. (2020). *An ad hoc WHO technical consultation managing the COVID-19 infodemic:call for action, 7-8 April 2020*. https://apps.who.int/iris/bitstream/handle/10665/334287/9789240010314-eng.pdf

WiD. (2021). *Wissenschaftsbarometer 2021*. https://www.wissenschaft-im-

dialog.de/projekte/wissenschaftsbarometer/wissenschaftsbarometer-2021/
Winter, S., & Krämer, N. (2016). Who's right: The author or the audience? Effects of user comments and ratings on the perception of online science articles. *Communications*,*41*(3), 339–360. https://www.degruyter.com/document/doi/10.1515/commun-2016- 0008/html

Parallel Panel III: PUBLIC PERCEPTIONS OF AI

Thursday, June 6, 14:00 – 15:00
Chair: Niels Mede
Room: HG E 1.2

More harm than good? Germans' attitudes towards generative AI in science communication

Bastian Kremer, Mike S. Schäfer, Liliann Fischer

Rationale and Research Questions
Generative artificial intelligence (AI) such as ChatGPT has emerged as an important part of contemporary information and communication ecosystems. With arguably the fastest technological rollout in history, it has the technological potential and the societalreach to profoundly influence individuals, organization and societal dynamics as a whole (e.g. Johnson 2022) – including science (Fecher et al. 2023) and science communication (Schäfer 2023).
Therefore, it is crucial to understand the attitudes and opinions of citizens towards generative AI and the factors shaping these attitudes. We ask:

> *RQ1:* What are Germans' attitudes towards generative AI, and how do they evaluate its potential risks and benefits in science communication?

With regards to explaining these attitudes, peoples' general attitudes towards science and research have been shown to affect their evaluation of specific technologies as well. For example, people whosee more risks of innovations science and research bring along rather than benefits might also be more critical towards new technologies such as generative AI (Méndez-Suárez et al. 2023). In order to investigate this, we examine the relation between peoples' general perceptions of science and research and their evaluations of generative AI as a source of scientific information. Explanatory factors include evaluations of the risks vs. benefits of science and research in general as well as trust in science. To investigate this, we assessed generalized trust in science as well as epistemic trust with its three dimensions: trust in the expertise, integrity, and benevolence of scientists (Hendriks et al. 2016). Laypeople generally cannot afford to invest the same amount of time and energy into researching specific scientific content – which is a similar starting situation to when people ask generative AI for help (Ferrario 2020). We, therefore, assume that evaluative attitudes like the perceived risks and benefits of science as research as well as the trust in science serve as relevant predictors of positive or negative evaluations of generative AI. In addition, we assume that peoples' sociodemographic characteristics – such as age, gender or education – are linked to evaluations of generative AI, as they have been shown to matter regarding evaluations of AI before (e.g. Neudert et al. 2020) and towards technologies in general (e.g. Frewer et al. 1998). Therefore, we ask:

> *RQ2:* How can attitudes towards generative AI in science communication be explained, and which role do general attitudes towards science and research as well as sociodemographic variables play in that regard?

Data and Method
Data from the 2023 German Science Barometer – an annual representative survey of the German population's attitudes towards science and research with 1,037 respondents (www.sciencebarometer.com) - are utilized. The survey captures various general attitudes towards science and research such as trust, perceived risks and benefits as well as the feeling of being informed. Additionally, it captures positive and negative attitudes towards "programs like ChatGPT" in eight items, particularly for the creation of science-related content, and their trustworthiness as

sources of scientific information. Descriptive statistics and multiple regression analyses are used.

Results

Germans are rather critical towards generative AI. While half of the respondents acknowledge that generative AI can be useful in explaining science in a simplified way, for example, the majority assumes that generative AI could contribute to the spread of science-related misinformation and finds it concerning that it “makes it more challenging to determine whether scientific content was authored by a human or a program". Accordingly, 44% say that they (rather) do not trust programs like ChatGPT as sources of scientific information. First multivariate analyses suggest that assessment is linked to general trust in science and differs between sociodemographicgroups. Further in-depth analyses using multiple regression will be conducted to follow up on the potential connections between these variables.

Relevance and Implications

The article contributes to the scholarly discussion on the relationship between general attitudes towards science and research and the perception of new technologies such as generative AI. It provides insights to researchers, policymakers, and technology developers onthe intersection of science and technology, especially in the context of emerging technologies like generative AI.

References

Fecher, B., Hebing, M., Laufer, M., Pohle, J. & Sofsky, F. (2023). Friend or Foe? Exploring the Implications of LargeLanguage Models on the Science System. https://arxiv.org/pdf/2306.09928 (30.11.2023)

Ferrario, A., Loi, M. & Viganò, E. (2020). In AI We Trust Incrementally: a Multi-layer Model of Trust to AnalyzeHuman-Artificial Intelligence Interactions. *Philos. Technol.* 33, 523–539. https://doi.org/10.1007/s13347- 019-00378-3

Frewer, L. J., Howard, C., & Shepherd, R. (1998). Understanding public attitudes to technology. Journal of RiskResearch, 1(3), 221-235.

Hendriks, F., Kienhues, D., & Bromme, R. (2016). Trust in science and the science of trust. In Blöbaum, B. (Eds.) Trust and communication in a digitized world: Models and concepts of trust research, 143-159

Johnson, A. (2022). Here’s What To Know About OpenAI’s ChatGPT—What It’s Disrupting And How To Use It.Forbes. Retrieved from https://www.forbes.com/sites/ariannajohnson/2022/12/07/heres-what-to-know-about-openais-chatgpt-what-its-disrupting-and-how-to-use-it/ (30.11.2023)

Méndez-Suárez, M., Monfort, A., Hervas-Oliver, J. (2023). Are you adopting artificial intelligence products? Social-demographic factors to explain customer acceptance. In European Research on Management andBusiness Economics, Volume 29, Issue 3, https://doi.org/10.1016/j.iedeen.2023.100223

Neudert, L. M., Knuutila, A., & Howard, P. N. (2020). Global attitudes towards AI, machine learning & automated decision making. Oxford: OII. https://philhoward.org/wp-content/uploads/2022/08/GlobalAttitudesTowardsAIMachineLearning2020.pdf

Schäfer, M. S. (2023). ‘The Notorious GPT: science communication in the age of artificial intelligence’. JCOM 22(02), Y02. https://doi.org/10.22323/2.22020402

AI talking science: Two experimental studies on the perception of large language models as a source of scientific information

Friederike Hendriks, Esther Greussing, Aike Horstmann, Bianca Nowak, Yannic Meier, Rainer Bromme

Since the launch of ChatGPT in 2022, Large Language Models (LLMs) are increasingly integrated into daily routines, acting as information intermediaries for a wide range of content, including science and health (Choudhury & Shamszare, 2023). These models generate text-based responses to user queries that are both coherent and contextually relevant, creating a simulation of understanding in their interactions with users, but not true understanding of the issue itself (Chavanayarn, 2023). Consequently, LLM-generated content often appears plausible, even though it might be factually incorrect (Spitale et al., 2023).

Besides the potential for generating incorrect information, certain standards of good science communication are not necessarily met by LLMs when providing information to users (Schäfer, 2023). Specifically, LLMs lack transparency in the process of generating content, such as about the disclosure of (reliable) source information, or on the selection of sources for training databases. A majority of Germans seem to be aware of the potential dangers to accuracy inherent in LLM: In the 2023 German Science Barometer 61% agree that technologysuch as ChatGPT might sometimes generate false information and 57% agree that such technology cannot evaluate information similar to humans (Wissenschaft im Dialog, 2023).

However, since LLM-generated content is presented in a clear, concise, easily understandable, and often highly personalized manner, users might be easily inclined to trust the presented information (Kaplan et al., 2023; Ruwe & Mayweg-Paus, 2023) in spite of the possible limits in accuracy; and instead of initiating further verification processes, such as consulting topic experts (c.f. easiness effect; Scharrer et al., 2012; 2017). Studies show that less use of jargon in scientific communication might increase the acceptance of messages through increased processing fluency (Bullock et al, 2019; Shulman et al., 2020). It is thus important to ask how users perceive and evaluate science content provided by LLMs, and whether they evaluate the credibility of information more critically, when explicitly made aware of flaws in LLMs' generation of scientific information (RQ1).

Furthermore, scientific information generated by LLMs might be perceived as even more accurate than information generated by humans via a rule of thumb such as the machine heuristic (Sundar & Kim, 2019), the idea that AI can generate better information than humans, as technology generated content is perceived as more objective, precise, and based on large amounts of data (Swiecki et al., 2022). As such, a second research question we pose is how an LLM as a source of science information is evaluated compared to a scientific source, and whether this evaluation is mediated by a machine heuristic (RQ2).

To investigate these research questions, we conducted two pre-registered experiments (total N ≈ 1.500; https://osf.io/saj6k?view_only=7912dd921d2e42579c4979e218b65287). In both experiments, participants are asked to envision a scenario in which a friend seeks advice on using sunscreen with nanoparticles. Subsequently, all participants receive information about nanoparticles in sunscreen and associated health risks, supported by scientific evidence. In our first between-person study, we test the effectiveness of a disclaimer warning about the uncertain quality of information provided by LLMs (vs. no disclaimer). This disclaimer is presented prior to participants receiving the information on nanoparticles in sunscreen via an LLM-based chatbot. In our second between-person study, we vary the source (LLM-based chatbot vs. human scientist) and the presentation of the information (static vs. dynamic). In both experiments, we are interested in the effect of the experimental manipulation on the ascribed objectivity and perceived trustworthiness of the source, the perceived credibility of the content presented, as well as in users'

strategies for coping with the division of cognitive labor (i.e., the effort expended in evaluating the content presented).

The first experiment was fielded in October 2023. The data was collected using the GapFish panel; the final sample consisted of N = 508 German-speaking adults. In this sample, AI-based chatbots were not often used (M = 2.68, SD = 1.82), and participants rated their own prior knowledge of AI-based chatbots to be small (M = 2.70, SD = 1.68); the attitude towards AI-based chatbots was rather neutral (M = 3.82, SD = 1.70); as was the attitude towards the topic (M = 3.61, SD = 1.48). All three items were rated on scales reaching from 1–7 (*little to high agreement*).

Contrary to our assumptions, the disclaimer warning about the uncertain quality of information provided by LLMs had no significant effects on any of the dependent variables (p> .05). This might be due to limited awareness of the information quality disclaimer (only 46% answered the manipulation check correctly, an additional 31% reported having to guess), and due to limited prior knowledge and attitudes towards AI-based chatbots and the scientific topic. For example (analyzing only the N = 394 who had answered the manipulation check correctly or had to guess), the expertise ascribed to the chatbot was predicted by participants' attitudes towards chatbots (β = .410, p < .001), their knowledge about chatbots (β = .095, p < .05), and their attitude towards nanoparticles in sunscreen (β = .237, p < .001). The second experiment (N ≈ 1000) is currently being finalized and will be fielded in December 2023.

Taken together, the results of the two experiments may inform practice about howusers could be supported in efficiently and successfully using generative AI for gatheringinformation about science-related issues, while being critically aware of limitations in accuracy and reliability.

References

Bullock, O. M., Colón Amill, D., Shulman, H. C., & Dixon, G. N. (2019). Jargon as a barrierto effective science communication: Evidence from metacognition. *Public Understanding of Science*, *28*(7), 845–853. https://doi.org/10.1177/0963662519865687

Choudhury, A. & Shamszare, H. (2023). Investigating the Impact of User Trust on theAdoption and Use of ChatGPT: Survey Analysis. *Journal of Medical Internet Research, 25*, e47184. https://doi.org/10.2196/47184

Chavanayarn, S. (2023). Navigating Ethical Complexities Through Epistemological Analysisof ChatGPT. *Bulletin of Science, Technology & Society*, 02704676231216355. https://doi.org/10.1177/02704676231216355

Kaplan, A. D., Kessler, T. T., Brill, J. C., & Hancock, P. A. (2023). Trust in ArtificialIntelligence: Meta-Analytic Findings. *Human Factors*, *65*(2), 337–359. https://doi.org/10.1177/00187208211013988

Ruwe, T., & Mayweg-Paus, E. (2023). "Your argumentation is good", says the AI vs humans – The role of feedback providers and personalised language for feedbackeffectiveness. *Computers and Education: Artificial Intelligence*, 100189.https://doi.org/10.1016/j.caeai.2023.100189

Scharrer, L., Bromme, R., Britt, M. A., & Stadtler, M. (2012). The seduction of easiness: Howscience depictions influence laypeople's reliance on their own evaluation of scientific information. *Learning and Instruction*, *22*(3), 231–243. https://doi.org/10.1016/j.learninstruc.2011.11.004

Scharrer, L., Rupieper, Y., Stadtler, M., & Bromme, R. (2017). When science becomes tooeasy: Science popularization inclines laypeople to underrate their dependence on experts. *Public Understanding of Science, 26*(8), 1003-1018. https://doi.org/10.1177/0963662516680311

Schäfer, M. S. (2023). The Notorious GPT: Science communication in the age of artificialintelligence. *Journal of Science Communication*, *22*(2), Y02. https://doi.org/10.22323/2.22020402

Shulman, H. C., Dixon, G. N., Bullock, O. M., & Colón Amill, D. (2020). The Effects of Jargon on Processing Fluency, Self-Perceptions, and Scientific Engagement. *Journalof Language and Social Psychology*. https://doi.org/10.1177/0261927X20902177

Spitale, G., Biller-Andorno, N., & Germani, F. (2023). AI model GPT-3 (dis)informs us betterthan humans. *Science Advances, 9*, eadh1850. https://doi.org/10.1126/sciadv.adh1850

Sundar, S. S., & Kim, J. (2019). Machine Heuristic: When We Trust Computers More than Humans with Our Personal Information. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–9. https://doi.org/10.1145/3290605.3300768

Swiecki, Z., Khosravi, H., Chen, G., Martinez-Maldonado, R., Lodge, J. M., Milligan, S.,Selwyn, N., & Gašević, D. (2022). Assessment in the age of artificial intelligence.*Computers and Education: Artificial Intelligence*, *3*, 100075. https://doi.org/10.1016/j.caeai.2022.100075

Wissenschaft im Dialog (2023). *Wissenschaftsbarometer 2023*. https://www.wissenschaft-im-dialog.de/fileadmin/user_upload/Projekte/Wissenschaftsbarometer/2023/WiD-Wissenschaftsbarometer2023_Broschuere_web.pdf

Predicting and describing the use of generative AI in science-related information search: Insights from a multinational survey

Esther Greussing, Lars Guenther, Ayelet Baram-Tsabari, Shakked Dabran-Zivan, Evelyn Jonas, Inbal Klein-Avraham, Monika Taddicken, Becca Beets, Dominique Brossard, Anwesha Chakraborty, Torben Esbo Agergaard, Antoinette Fage-Butler, Chun-Ju Huang, Kristian Hvidtfelt Nielsen, Siddharth Kankaria, Yin-Yueh Lo, Michelle Riedlinger, Hyunjin Jin Song

Generative Artificial Intelligence (GenAI), exemplified by large language models like ChatGPT, is entering mediated communication worldwide. In particular, GenAI increasingly serves as an information intermediary, offering content on a wide range of topics. An initial survey conducted among regular ChatGPT users in the US revealed that 36% of them utilize the model for information-seeking purposes, with trust in ChatGPT emerging as a significant factor (Choudhury & Shamszare, 2023). However, much remains to be learned regarding the factors that drive the regular useof such technology. Earlier research on the diffusion and acceptance of technologicalinnovations highlights the role of individuals' emotion and cognition in this respect (Xu et al., 2023). Specifically, attitudes and trust appear to positively predict behavioral intention (Kelly, Kaye, & Oviedo-Trespalacios, 2022). In the realm of GenAI, content generation unfolds in a distinct manner, introducing a risk of so-called hallucinations (De Angelis et al., 2023). Consequently, factual knowledge about AI might become a relevant factor influencing the use of this technology.

The potential of GenAI to facilitate easy access to information and customize content based on user characteristics could render the technology particularly suitable for searching for science-related information (Schäfer, 2023), but empirical evidence on this is presently lacking. Additionally, existing research tends to be specific to certain countries or regions, despite the global implications of GenAI on science communication. This study therefore asks:

RQ1: How do attitudes towards AI, knowledge about AI, trust in AI, and sociodemographics predict the regular use of GenAI technology?

RQ2: How do people report using GenAI for searching science-related information?

To address these research questions, from July to August 2023, we conducted an online survey gathering data from participants across seven countries/regions[8]: Australia, Denmark, Germany, Israel, South Korea, Taiwan, and the United States (*N*=4,321)[9]. In each country/region, our samples were representative of the respective populations in terms of age, gender, and education. We collected data on attitudes towards AI, factual knowledge about AI, trust in GenAI, and use of GenAI through online access panels in each country/region, with the questionnaire translated into the relevant primary language.

Regarding our first research question, we employed multilevel modeling, whereby respondents are nested within countries. We ran a logistic regression with random intercept, thereby controlling for potential country-level differences in the regular use of GenAI. Attitudes towards AI, factual knowledge about AI, trust in GenAI, and sociodemographics served as predictors. 18.1% (*n*=780) of our sample are regular users of GenAI technology. The results show that all predictor variables significantly increased the probability of being a regular user (*condR*^{2}=0.27). Specifically, respondents with positive attitudes towards AI (Odds ratio [OR] 1.65,95% CI 1.37-1.92, *p*<.001), higher trust in GenAI (OR 1.58, 95% CI 1.40-1.78, *p*<.001), and a more profound knowledge about AI (OR 1.08, 95% CI 1.03-1.13, *p*<.01) were more likely to use GenAI regularly. Moreover, the likelihood of regularGenAI use was associated with specific demographic factors. Being younger (OR 0.97, 95% CI 0.97-0.98, *p*<.001), male (OR 1.40, 95% CI 1.18-1.67, *p*<.001), and possessing a secondary education level (OR 0.52, 95% CI 0.42-0.64, *p*<.001, ref =higher education) all contributed to higher

[8] Given the unique status of Taiwan, in this manuscript, we refer to "countries/regions".

[9] The numbers of respondents per country/region are as follows: nAUS = 552, nDEN = 505, nGER = 566, nISR = 500, nKOR = 642, nTWN = 504, nUSA = 1.052.

odds of regular GenAI use.

Regarding our second research question, asking how people use GenAI-systems for searching science-related information, we focus on ChatGPT as the most commonly accessible model. Descriptive analyses showed that while Germany and Denmark have a relatively high percentage of respondents who have never used ChatGPT before (73.1% and 69.8%, respectively), Taiwan stood out with the highest proportion of regular users (25.6%), followed by South Korea (17.3%) and Israel (15.5%), indicating diverse adoption patterns in different countries/regions. Hence, ChatGPT's potential as a tool to search for science-related information is important. In Taiwan, 54.2% of the respondents have utilized ChatGPT for this purpose, compared to 8.1% in Denmark. Our data suggests that across all countries/regions, people who report using ChatGPT for science-related information searching tend to be younger, exhibit higher factual knowledge about both AI and GenAI, and have greater trust in GenAI compared to people who do not use this technology for this purpose. Also, users expressed satisfaction with the science-related information they have received from ChatGPT, as well as confidence in finding what they need.

Overall, our findings show that despite the novelty of ChatGPT, a significant proportion of the respondents reported using it for science-related information search. Adoption patterns, however, varied across countries/regions. The GenAI users we identified displayed distinct characteristics, suggesting that, currently, a specific segment of the population relies on these systems. This holds true for both general use of GenAI technology and, particularly, for its application in seeking science-related information. As such, this study encapsulates a distinctive moment in the global adoption of GenAI as a technology that simultaneously generates and disseminates science-related information. The comparative element further advances current understandings of the factors that influence the use of GenAI.

References

Choudhury, A. & Shamszare, H. (2023). Investigating the Impact of User Trust on theAdoption and Use of ChatGPT: Survey Analysis. *Journal of Medical Internet Research, 25*, e47184. https://doi.org/10.2196/47184

De Angelis, L., Baglivo, F., Arzilli, G., Privitera, G. P., Ferragina, P., Tozzi, A. E., & Rizzo, C. (2023). ChatGPT and the rise of large language models: the new AI-driven infodemic threat in public health. *Frontiers in Public Health, 11*, 1166120. https://doi.org/10.3389/fpubh.2023.1166120

Kelly, S., Kaye, S. A., & Oviedo-Trespalacios, O. (2022). What factors contribute toacceptance of artificial intelligence? A systematic review. *Telematics and Informatics, 77*, 101925. https://doi.org/10.1016/j.tele.2022.101925

Schäfer, M. S. (2023). The Notorious GPT: science communication in the age ofartificial intelligence. *Journal of Science Communication, 22*(2), Y02.

Xu, S., Kee, K. F., Li, W., Yamamoto, M., & Riggs, R. E. (2023). Examining the Diffusion of Innovations from a Dynamic, Differential-Effects Perspective: ALongitudinal Study on AI Adoption Among Employees. *Communication Research*, *0*(0). https://doi.org/10.1177/00936502231191832

Parallel Panel IV: OPPORTUNITIES OF AI IMAGERY

Thursday, June 6, 15:20 – 16:20
Chair: Sabrina H. Kessler
Room: HG E 1.2

An AI-based social media generator: STEM research artificially communicated

Elisabeth Jurack, Julia Gantenberg, Justus Henke, T.Y. Branch, Ingo Siegert

Science communication faces the challenge of bringing highly complex topics closer to different target groups (e.g. laypeople such as pupils or people with a lower level of education), choosing an appealing form of presentation and making scientific explanations understandable. The media landscape has changed considerably within the last decade, whereby scientific content is not only communicated via traditional text formats and lectures, but increasingly also via internet-based formats such as social media (Bik and Goldstein, 2013; Metag and Schäfer, 2017). The preparation of quality scientific content for social media is very time-consuming, but does not necessarily require the expertise of researchers. The generative models that have emerged since the advent of ChatGPT in November 2022 offer an opportunity to improve both the workload for researchers and the quality of science communication content (Dwivedi et al., 2021). So far, the use of AI-based technologies in science communication has been underexplored (Schäfer, 2023).

Our project proposes to use Artificial intelligence (AI) to explore automation opportunities in order to create credible science-content for Instagram. The AI will source information from a curated database of texts grounded in scientific content from STEM specialist publications, press releases, interviews etc. We will investigate the extent to which the process of generating scientific information using AI-generated posts for various channels and target groups on Instagram would be suitable for general use by researchers in science communication and researchers themselves. This involves investigating how usable social media is, and how targeted scientific information can be converted into high-quality posts on Instagram using AI.

The project will use a combination of quantitative and qualitative indicators permitting an in-depth analysis and evaluation of the AI-powered software's performance and acceptance. Pre-exposure to the posts, we will use qualitative analyses to seek insights into prevailing views of the target groups about AI-generated scientific information. Post-exposure to the posts, we have planned a survey on the topic of scientific literacy. Quantitatively, the factual (empirical) dimensions as well as the social, political and ethical content of the dataset will be examined to better understand the information landscape available to laypeople.

Additionally, our socio-reflexive philosophical analysis will use the dataset of the generated social media posts to analyse for the presence of epistemic and non-epistemic values in science. The suitability of AI-powered contributions for researchers who need to communicate science will be assessed based on how these values align with ethical questions about usability and responsibility. I.e.: How and to whom will the public attribute the discoveryof scientific knowledge? Who should be responsible for errors in the information? Does the impression of science change? Does the impression of researchers change? These questions and our research intend to contribute a practical dimension to broader discussions of trust in science and how to broker the relationship between science and society.

This abstract is part of a research proposal that will be reviewed by the end of 2023. Following a positive evaluation, the project will be launched in January 2024, whereby initial results could be presented at the Department of Communication and Media Research Annual Conference

of the “Science Communication” Division (DGPuK) in addition to the methodology and framing of research questions.

References

Bick, H. and Goldstein, M. (2013) An Introduction to Social Media for Scientists. BLOS Biology. https://doi.org/10.1371/journal.pbio.1001535

Dwivedi et al. (2021) Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. International Journalof Information Management. https://doi.org/10.1016/j.ijinfomgt.2019.08.002

Metag, J., & Schäfer, M. S. (2017). Hochschulen zwischen Social Media-Spezialisten undOnline-Verweigerern. *Studies in Communication | Media*, *6*(2), 160–195. https://doi.org/10.5771/2192-4007-2017-2-160

Schäfer, M. S. (2023). The Notorious GPT: Science communication in the age of artificialintelligence. *Journal of Science Communication*, *22*(2), Y02. https://doi.org/10.22323/2.22020402

AI Avatars in science communication: When Einstein and Curie resurrect

Jasmin Baake, Josephine Schmitt

The emergence of AI tools in science communication offers a new realm of possibilities, enhancing the efficiency and creativity in communicating scientific topics (De Angelis et al., 2023; Schäfer, 2023). One of the early examples of how synthetically generated videos used for science communication can be found on TikTok, where Albert Einstein, Marie Curie and other deceased scientists are brought back to life (see Figure 1).

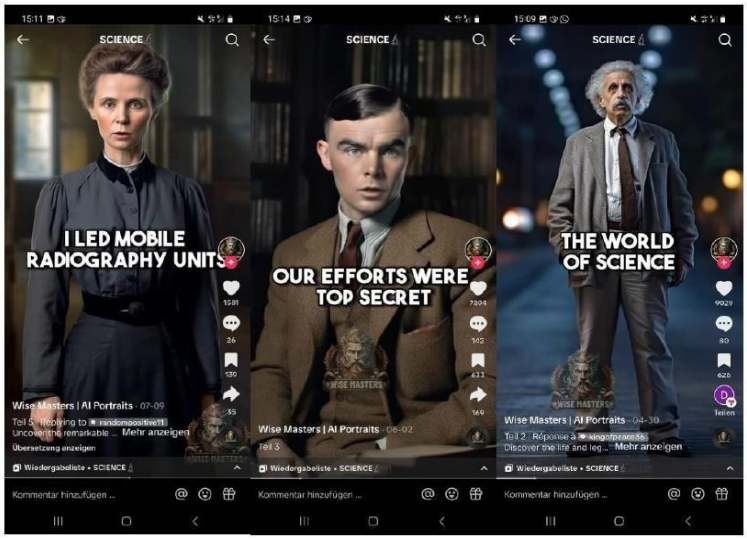


Figure 1 Video avatars of Marie Curie, Nikola Tesla and Albert Einstein on the TikTok channel @wisemasters.

With the help of AI-based video generators, such knowledge-sharing and engaging video avatars can be created. While these tools offer innovative storytelling methods, concerns arise regarding viewer skepticism due to the involvement of AI in content creation (Vaccari & Chadwick, 2020; WDR Innovation Hub, 2021). Additionally, the misuse of synthetic videos, as evident in instances of Deepfake manipulations, highlights challenges and public trust issues associated with synthetic media (Gräfe, 2022; Metzger & Schneider, 2022).

For science communication trust is a critical variable. Trust in science is central to evidence-based decision making in society and politics. Acceptance of scientific research is essential to ensure its integrity, independence and resources (Neuberger et al., 2022). In view of the growing hostility towards science, AI-generated content could offer more scope for attack (e.g., for conspiracy theorists) than human-generated/presented communication and thus impair trust in science. Against this background we wonder: How must AI-generated science communication content be designed to be trustworthy?

The role of presentation style for trust

The trust attribution in AI-generated content varies greatly depending on the application sector and is influenced by contextual factors and presentation styles (Glikson & Woolley, 2020; Kieslich et al., 2021). The impact of synthetic formats of science communication by recipients can only be guessed at so far. Studies on AI in journalism indicate that recipients are more critical of AI-generated information than information created by humans (Longoni et al., 2022). Research on intelligent, human-like avatars and algorithmic decision-making processes also indicates synthetic content is not necessarily trusted by users (Thaler et al., 2020).

Anthropomorphism and gender representation seem to play a central role for the perception of the protagonist. Studies show that highly realistic avatars might trigger discomfort (the so-called Uncanny Valley effect) and affect trust in the information source (Ho & MacDorman, 2010; Schwind et al., 2018). Conversely, stylised avatars could facilitate better parasocial interactions and

credibility (Di Natale et al., 2021).
Furthermore, the gender, e.g., represented by science communicators, is associated with stereotypes that have an impact on the evaluation of scientists (König & Jucks, 2019). For example, there is a tendency to perceive male scientists as "highly competent but less warm-hearted", while their female colleagues are often seen as "less competent but more warm-hearted"(Reif et al., 2020, S. 193). The attributed warmth in turn plays a role in the human likeness attributed to the avatars (Ho & MacDorman, 2010).

Against this background we ask:

RQ1: What influence does the *human likeness* of AI-generated video avatars communicating about science have on the perceived trustworthiness of the scientists portrayed?

RQ2: What influence does the *gender* of AI-generated video avatars communicating about science have on the perceived trustworthiness of the scientists portrayed?

To achieve meaningful results, research on AI-generated avatars must consider the interaction of different attributes and not just individual characteristics that apply to human communicators. Therefore, we are further interested in the interaction of human likeness and gender of the AI-generated video avatars (*RQ3*).

Method

This work-in-progress outlines a 2x2 experimental design focusing on the perceived trustworthiness of AI-generated video avatars based on anthropomorphism levels and gender representation (see Figure 2). Specifically, it investigates the influence of these factors on viewers' trust perceptions using the dimensions of the Muenster Epistemic Trustworthiness Inventory (METI): Expertise, Benevolence, Integrity (Hendriks et al., 2015). In the between-subjects design, participants will watch AI-generated video avatars that present their scientific findings in the context of gene therapy for red-green colour blindness. The preregistered study is currently undergoing a pretest with 500 participants from an online open access panel (https://aspredicted.org/WCF_8VT).

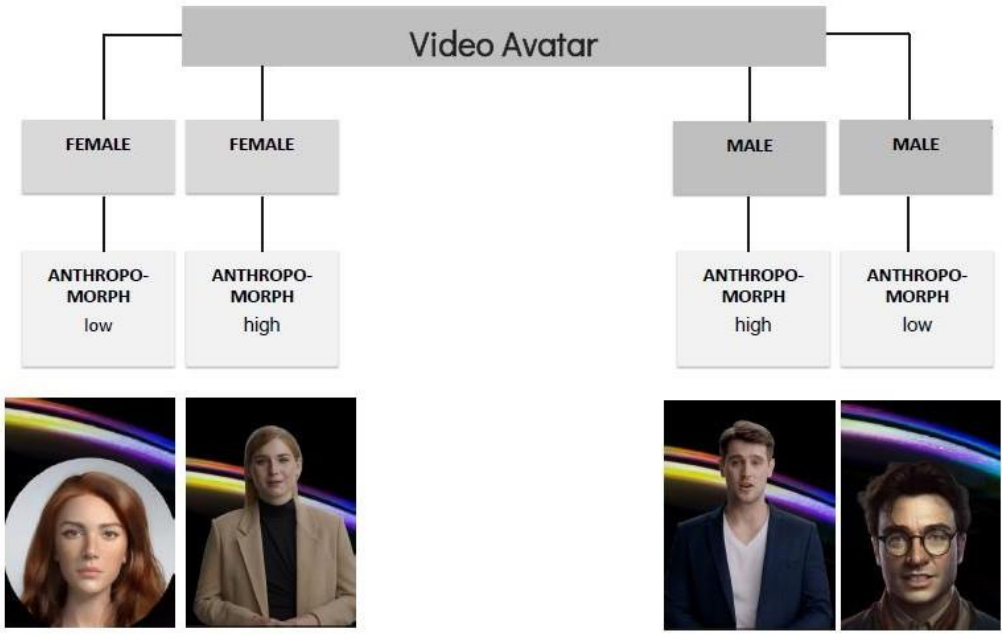


Figure 2 Overview of the four experimental conditions and corresponding avatars

Outlook

This study aims to shed light on the interrelation between AI-generated video avatars, anthropomorphism, and gender representation on perceived trustworthiness of science

communicators. Preliminary questions emerge regarding the discernible effects on trust dimensions and the perceived characteristics associated with highly anthropomorphic versus stylized avatars.

The ongoing pretest aims to refine the methods and prepare for further research. Methods and results will be discussed in more detail in the presentation. Further questions to be discussed in the plenary session would be: What degree of anthropomorphism is the right one for communicators of scientific topics if the avatar was generated using AI? Does a synthetic cartoon avatar or a human-like avatar (in good/bad quality) inspire more trust? Which dimensions of trust are relevant when evaluating AI video avatars? Are the three trust dimensions equally relevant when assessing the trustworthiness of AI video avatars? And in perspective: What happens when the synthetic content is no longer distinguishable from real video material? Has the Uncanny Valley then been crossed?

References

Burton, J. W., Stein, M.-K., & Jensen, T. B. (2020). A systematic review of algorithm aversion in augmented decision making. *Journal of Behavioral Decision Making*, *33*(2), 220–239. https://doi.org/10.1002/bdm.2155

De Angelis, L., Baglivo, F., Arzilli, G., Privitera, G. P., Ferragina, P., Tozzi, A. E., & Rizzo, C. (2023). *ChatGPT and the Rise of Large Language Models: The New AI-Driven Info-demic Threat in Public Health* (SSRN Scholarly Paper 4352931). https://doi.org/10.2139/ssrn.4352931

Di Natale, A. F., Triberti, S., Sibilla, F., Imperato, C., Villani, D., Mancini, T., & Riva, G. (2021). Behind a Digital Mask: Users' Subjective Experience of Animated Characters and Its Effect on Source Credibility. *Interacting with Computers*, *33*(5), 499–510.https://doi.org/10.1093/iwc/iwab030

Glikson, E., & Woolley, A. W. (2020). Human Trust in Artificial Intelligence: Review of Em- pirical Research. *Academy of Management Annals*, *14*(2), 627–660.https://doi.org/10.5465/annals.2018.0057

Gräfe, H.-C. (2022, März 30). *Deepfakes | Journalistikon*. https://journalistikon.de/deepfakes/ Heesen, J. (2022, Juli 30). *KI | Journalistikon*. https://journalistikon.de/ki/

Hendriks, F., Kienhues, D., & Bromme, R. (2015). Measuring Laypeople's Trust in Experts in a Digital Age: The Muenster Epistemic Trustworthiness Inventory (METI). *PLOSONE*, *10*(10), e0139309. https://doi.org/10.1371/journal.pone.0139309

Ho, C.-C., & MacDorman, K. F. (2010). Revisiting the uncanny valley theory: Developing and validating an alternative to the Godspeed indices. *Computers in Human Behavior*, *26*(6), 1508–1518. https://doi.org/10.1016/j.chb.2010.05.015

Kieslich, K., Dosenovic, P., Starke, C., Lünich, M., & Marcinkowski, F. (2021). *Artificial Intelligence in Journalism. How does the public perceive the impact of artificial intelligence on the future of journalism?* (Factsheet 4; MeMo:KI | Meinungsmonitor Künstliche Intelligenz). https://www.researchgate.net/profile/Kimon-Kieslich/publication/349882876_Artificial_Intelligence_in_Journalism_How_does_the_public_perceive_the_impact_of_artificial_intelligence_on_the_future_of_journalism/links/6045f606299bf1e07862c530/Artificial-Intelligence-in-Journalism-How-does-the-public-perceive-the-impact-of-artificial-intelligence-on-the-future-of-journa- lism.pdf

König, L., & Jucks, R. (2019). When do information seekers trust scientific information? Insights from recipients' evaluations of online video lectures. *International Journal of Educational Technology in Higher Education*, *16*(1), 1.https://doi.org/10.1186/s41239-019-0132-7

Longoni, C., Fradkin, A., Cian, L., & Pennycook, G. (2022). News from Generative Artificial Intelligence Is Believed Less. *2022 ACM Conference on Fairness, Accountability, and Transparency*, 97–106. https://doi.org/10.1145/3531146.3533077

Metzger, & Schneider. (2022, März 18). *Wie Deepfakes im Ukraine-Krieg genutzt werden*. https://www.zdf.de/uri/cbb730bf-b39c-4ecb-bee5-351dc2c96d32

Neuberger, C., Weingart, P., Schildhauer, T., Fähnrich, B., Wormer, H., Jarren, O., Passoth, J.- H., & Wagner, G. G. (2022). *Gute Wissenschaftskommunikation in der digitalen Welt: Politische, ökonomische, technische und regulatorische Rahmenbedingungen ihrer Qualitätssicherung*. Berlin-Brandenburgische Akademie der Wissenschaften.

Reif, A., Kneisel, T., Schäfer, M., & Taddicken, M. (2020). Why Are Scientific Experts Perceived as Trustworthy? Emotional Assessment within TV and YouTube Videos. *Media & Communication*, *8*(1), 191–205. https://doi.org/10.17645/mac.v8i1.2536

Schäfer, M. S. (2023). The Notorious GPT: Science communication in the age of artificial intelligence. *Journal of Science Communication*, *22*(2), Y02 https://doi.org/10.22323/2.22020402

Schwind, V., Wolf, K., & Henze, N. (2018). Avoiding the uncanny valley in virtual character design. *Interactions*, *25*(5), 45–49. https://doi.org/10.1145/3236673

Song, S. W., & Shin, M. (2022). Uncanny Valley Effects on Chatbot Trust, Purchase Intention, and Adoption Intention in the Context of E-Commerce: The Moderating Role of Ava- tar Familiarity. *International Journal of Human–Computer Interaction*, *0*(0), 1–16.https://doi.org/10.1080/10447318.2022.2121038

Thaler, M., Schlogl, S., & Groth, A. (2020). Agent vs. Avatar: Comparing Embodied Conversational Agents Concerning Characteristics of the Uncanny Valley. *2020 IEEE Inter- national Conference on Human-Machine Systems (ICHMS)*, 1–6.https://doi.org/10.1109/ICHMS49158.2020.9209539

Vaccari, C., & Chadwick, A. (2020). Deepfakes and Disinformation: Exploring the Impact of Synthetic Political Video on Deception, Uncertainty, and Trust in News. *Social Media+ Society*, *6*(1), 2056305120903408. https://doi.org/10.1177/2056305120903408

WDR Innovation Hub. (2021). *Synthetische Medien—Zukünfte der Medienproduktion mit Künstlicher Intelligenz: Ein Zukunftsreport des WDR Innovation Hub.* WDR Innovation Hub.

Can AI-generated imagery be used to communicate future climate scenarios?

Josephine Ewoma

The climate crisis, along with other environmental issues such as biodiversity loss, continue to persist. To address this, climate communicators have spent decades exploring the most effective modes to communicate the complex issues associated with the climate crisis to publics, in an attempt to promote pro-environmental behaviours and support policy decisions (Comfort & Park, 2018; Markowitz & Guckian, 2018). However, despite the progress made in this field, the impacts of the climate crisis are becoming realised in real time, exemplified by 2023 being 'the warmest year on record' (National Oceanic and Atmospheric Administration (NOAA), 2024; Carbon Brief, 2023). Yet, there remains doubts about the severity of the climate crisis amongst the general public (Pew Research Centre, 2023). At the same time, new technologies offer an array of opportunities for communicators to reach new audiences and garner public support for climate initiatives. Artificial intelligence (AI), has been heralded as an industry disruptor. One such example is the large language model (LLM) ChatGPT, which gained over one million users in its first week, highlighting public interest in these new technologies (Schäfer, 2023; OpenAI, 2023). Generative AI, as well as translating text, can generate audio, mimic voices (*VALL-E*) and generate imagery (*Midjourney, Stable Diffusion*).

AI-generated imagery is an especially interesting avenue of this new technology. Visual imagery has been suggested to be an essential facet of effective communication (O'Neill et al., 2013), in the context of the climate crisis visual imagery has the ability to "make visible" the impacts of the climate crisis (O'Neill & Nicholson-Cole, 2009). Common iconography of the climate crisis feature melting ice caps, polar bears and more broadly extreme weather events (Manzo, 2010; O'Neill et al., 2013; Smith & Joffe, 2009). However, research has identified that climate visuals rarely include people (Wang et al., 2018), despite the argument that without people climate visuals fail to promote salience as they may lead to individuals feeling removed from the issue (Corner, Webster & Teriete, 2015; Braasch, 2013).
AI-generated imagery could extend the range of image types used to communicate the impacts of the climate crisis. Further, AI-generated imagery has the ability to insert people into future climate scenarios, which has the potential to engage publics. Recently, AI filters have increased in popularity on social media platforms, such as TikTok. The "aged" filter on TikTok, which allows users to watch themselves age up to 50 years in real time has featured in over nine million videos on the platform (The Economic Times, 2023). The popularity of this filter, and other AI-assisted filters and apps, demonstrate public interest in reimaging themselves and their realities through AI. Therefore, it is worth climate communicators exploring whether this interest can be extended to informing and engaging publics with climate science, mitigation behaviours and inciting action.

As climate crisis impacts become more pervasive it is imperative that communicators engage with new technologies to reach a wider audience, and incite action towards climate goals. However, with new technology ethical considerations must be discussed. For example, AI algorithms have been found to show bias, due to the nature of training data used (Chen et al., 2023). Moreover, AI-imagery potentially catastrophizing the climate crisis could have impacts on users mental and emotional states. Further, AI predictions may be inaccurate leading to the spread of misinformation, which must be considered by researchers and communicators alike.

This study explores the ways in which AI-generated images can be used for climate communications, whether AI-generated images could be an effective communication tool and the ethical considerations associated with utilising AI-generated imagery as a communication strategy.

Parallel Panel IV: QUALITATIVE RESEARCH ON AI PERCEPTIONS

Thursday, June 6, 15:20 – 16:20
Chair: Julia Metag
Room: HG E 1.1

How do laypeople assess their trust in LLM-based chatbots when they seek science-related information? Results from a qualitative interview study using a hybrid trust approach

Evelyn Jonas, Esther Greussing, Monika Taddicken

Based on large language models (LLMs), chatbots like ChatGPT or Bing Chat can simplify complex content for laypeople and are therefore discussed as emerging intermediaries for health and science-related information (Ma & Hou, 2023; Schäfer, 2023). However, concerns regarding an AI-driven potential deterioration of information quality (Jungherr & Schroeder, 2023) reintroduce the question of who – or what – laypeople *trust* when navigating complex scientific issues within digital information environments (Bromme & Kienhues, 2014). This study explores trust in LLM-based chatbots as intermediaries for science-related information.

Trust is defined as a multidimensional relationship variable between a trustor and a trustee embedded in a situation of uncertainty and risk (Reif, 2021), traditionally drawing a conceptual distinction between technology trust (i.e., human trust in technologies) and interpersonal trust. The former includes the dimensions functionality, reliability and helpfulness (Mcknight et al., 2011), understood as reasons the trustor trusts the trustee. However, as users tend to perceive AI-based conversational interfaces in a human-like manner (Reeves & Nass, 1996; Ma & Huo, 2023), traditional notions of trust in technologies fall short to address trust in LLM-based chatbots, leading to a combination with concepts of interpersonal trust (Choung et al., 2022; Weidmüller, 2022). Considering the context of science communication, our study adopts this hybrid trust approach, extending technology trust with dimensions of epistemic trust, namely expertise, integrity, and benevolence (Hendriks et al., 2015; Mayer et al., 1995).

Moreover, AI-based applications encompass various layers, e.g., the application itself or the developers and the company behind it (Solomon & Wash, 2014). These can vary in their salience during interactions (Guzman, 2019) and thereby influence users trust assessment. Yet, it is uncertain whether users recognize different layers of the LLM-based trustee at all and if so, whether they associate them with all or only certain dimensions of technology and epistemic trust.

> **RQ1:** Which layers of LLM-based chatbots are users aware of while assessing their trust in the context of searching for science-related information?

Finally, trust is considered to be domain-specific (Mayer et al., 1995). This implies that the type of information presented by LLM-based chatbots can shape trustors' expectations towards the chatbot and influence the characteristics deemed relevant for assessing trustworthiness. Consequently, for topics characterised by greater individual involvement, such as health-related issues, interpersonal trust dimensions may be more relevant to the human trustor than for topics with no direct individual implications (Weidmüller, 2022).

> **RQ2:** How relevant is the type of information for user's trust assessment in LLM-based chatbots?

Method

We address these research questions with qualitative interviews. From June to August 2023, two researchers interviewed *n* = 34 people from the German general public (52.9 % female, mean age = 38.3 years (*SD* = 17.4), 58.8 % with prior chatbots experience) focusing on their perception and evaluation of LLM-based chatbots (⌀-duration: 2 hours). During this they used ChatGPT and Bing Chat to obtain scientific information on the topics of sustainable aviation (low personal involvement) and juice cleansing (high personal involvement). The verbatim transcribed interviews were analysed using flexible coding (Deterding & Waters, 2021).

Results

(RQ1) When assessing trust, participants not only referred to the LLM-based chatbots themselves, but also considered other layers, i.e., users, developers or companies. The last two appear to be salient in evaluating integrity (adherence to social and scientific standards and norms) and benevolence (concern for users' well-being). For example, regarding chatbots' adherence to scientific standards, interviewee 13 noted an underlying entity that dictates source prioritisation: "With Bing, I had the impression that the sources displayed were selected reasonably [...]. There is probably also a prioritisation behind it, [...] which is designed in the system, which has something to do with [...] scientific standards". Those findings suggest that it is not sufficient to consider trust in the application itself decoupled from trust in the underlying entities, turning LLM-based chatbots into complex objects of trust.

(RQ2) The interviews indicate a varying relevance of certain trust dimensions based on the nature of the topic. Regarding juice cleansing, some participants found ChatGPT's advice to consult professional doctors more personal and as an expression of interest in human well-being (benevolence), enhancing its trustworthiness. For sustainable aviation, perceived by some participants as a politicised issue, the main expectation was to receive a neutral, balanced and unemotional presentation of the facts. In this context, expectations related to benevolence were rarely addressed. It can be assumed that LLM-based chatbots as information intermediaries have to fulfil different, topic-sensitive functions, presumably enabled by their ample space of possible utilisations offered to the users.

This study offers valuable insights into user trust in LLM-based chatbots as intermediaries for science-related information, deepening our understanding of human-AI-communication in information search processes and stressing the importance to consider the information context and the trustee's complexity in future research on trust. Expanding technology trust dimensions with epistemic trust dimensions appears to be gainful, as it reveals that users perceive chatbots as multi-layered trustees, linking dimensions like benevolence to both technical systems and their developers/companies. Limitations include a controlled laboratory environment as well as the analysis of verbalized behaviour not fully reflecting natural use cases.

References

Bromme, R. & Kienhues, D. (2014). Wissenschaftsverständnis und Wissenschaftskommunikation. In T. Seidel & A. Krapp (Hrsg.), *Pädagogische Psychologie* (S. 55–81). Beltz.

Choung, H., David, P. & Ross, A. (2023). Trust in AI and Its Role in the Acceptance of AI Technologies. *International Journal of Human–Computer Interaction*, *39*(9), 1727–1739. https://doi.org/10.1080/10447318.2022.2050543

Deterding, N. M. & Waters, M. C. (2021). Flexible Coding of In-depth Interviews: A Twenty-first- century Approach. *Sociological Methods & Research*, *50*(2), 708–739. https://doi.org/10.1177/0049124118799377

Guzman, A. L. (2019). Voices in and of the machine: Source orientation toward mobile virtual assistants. *Computers in Human Behavior*, *90*, 343–350. https://doi.org/10.1016/j.chb.2018.08.009

Hendriks, F., Kienhues, D. & Bromme, R. (2015). Measuring Laypeople's Trust in Experts in a Dig- ital Age: The Muenster Epistemic Trustworthiness Inventory (METI). *PloS one*, *10*(10), e0139309. https://doi.org/10.1371/journal.pone.0139309

Jungherr, A. & Schroeder, R. (2023). Artificial intelligence and the public arena. *Communication Theory*, *33*(2-3), 164–173. https://doi.org/10.1093/ct/qtad006

Ma, X. & Huo, Y. (2023). Are users willing to embrace ChatGPT? Exploring the factors on the acceptance of chatbots from the perspective of AIDUA framework. *Technology in Society*, *75*, 102362.

https://doi.org/10.1016/j.techsoc.2023.102362
Mayer, R. C., Davis, J. H. & Schoorman, F. D. (1995). An Integrative Model of Organizational Trust. *The Academy of Management Review*, *20*(3), 709–734.
Mcknight, D. H., Carter, M., Thatcher, J. B. & Clay, P. F. (2011). Trust in a specific technology. *ACM Transactions on Management Information Systems*, *2*(2), 1–25. https://doi.org/10.1145/1985347.1985353
Reeves, B. & Nass, C. (1996). *The media equation: How people treat computers, television, and new media like real people and places*. Cambridge University Press.
Reif, A. (2021). Mehr Raum für Vertrauen? Potenzielle Veränderungen des Vertrauens in Wis- senschaft durch partizipative Online-Umgebungen. In T. Döbler, C. Pentzold & C. Katzenbach (Hrsg.), *Neue Schriften zur Online-Forschung: Band 16. Räume digitaler Kommunikation: Lokalität - Imagination - Virtualisierung* (1. Auflage, S. 210–243). Herbert von Halem Verlag.
Schäfer, M. S. (2023). The Notorious GPT: science communication in the age of artificial intelligence. *Journal of Science Communication*, *22*(02). https://doi.org/10.22323/2.22020402
Solomon, J. & Wash, R. (2014). Human- What Interaction? Understanding User Source Orientation. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *58*(1), 422–426. https://doi.org/10.1177/1541931214581088
Weidmüller, L. (2022). Human, Hybrid, or Machine? Exploring the Trustworthiness of Voice- Based Assistants. *Human-Machine Communication*, *4*, 85–110. https://doi.org/10.30658/hmc.4.5

‘Not for me’: An exploratory study into inclusive science communication about artificial intelligence in the Netherlands

Anne M. Dijkstra, Pien Spanjaard

Artificial intelligence (AI) applications, such as chatbots and recommendation systems, have become well-established tools which are integrated in many daily activities, while other applications as self-driving cars and social robots are developing fast. However, its development and use are not without consequences. AI is already radically impacting daily life, creating benefits but also raising concerns, for instance, about algorithmic discrimination, privacy and responsibility. AI operates in a social context, therefore, it is essential to consider how it affects citizens (cf Foulds et al, 2020). Science communication and engagement activities are core to these processes, however, these also have proven to structurally exclude and underrepresent minority groups of citizens (Dawson, 2018) even though these citizens are at the highest risk of experiencing negative consequences of AI applications.

Our study aimed to gain insight into those underrepresented citizens’ views and experiences with inclusion on the topic of AI in relation to science communication and engagement activities. Citizens from neighbourhoods with a higher average of inhabitants from a lower socio-economic-status (SES) in the city of Enschede in the Netherlands were recruited at the local shopping mall and asked to participate in semi-structured interviews. In total, 19 citizens ranging in age from 22 to 63 years and from different backgrounds participated with the majority having completed a lower or middle level of education. Participants’ perceptions of AI, their behaviour regarding information about AI, their intent and motivations to participate in science communication activities about AI, and their wishes, needs and recommendations for future activities were collected. The interviews were transcribed and analysed thematically using a combination of both deductive and inductive coding.

Participants gave detailed and nuanced accounts of their attitudes towards AI and their experiences with and perceptions of science communication and engagement activities about AI. Although they considered AI an important topic and held strong opinions about, in particular, the risks; a paradox existed regarding their information and engagement levels. Few participants who expressed a high level of interest in AI also regularly or occasionally consumed information about AI. They did so mainly through the mainstream media, internet, entertainment media or peers. However, the majority of the participants did not engage with any information about the topic. When inquired, lack of innate interest, mental health, age or perceived lack of influence were given as motivations.

Additionally, all but a few interviewees supported the idea of science communication activities on AI, while them majority signaled low intentions to join. The participants experienced a large gap between themselves and regular science communication-audiences. As said by Romero-Rodriquez et al. (2021), binary views, also held by the participants, create borders between groups. Barriers were material (financial, logistic) and social-emotional (feelings of alienation and shame, concentration issues, lack of mental space, age). The participants suggested more accessible, practical, welcoming and pro-active initiatives.

Overall, the participants were more knowledgeable and interested in AI than both researchers and even the participants themselves had anticipated. The majority of the interviewees were not opposed to engaging in science communication and engagement activities about AI. However, many barriers prevented participation. The findings contribute to dismantling the beliefs that citizens are ignorant and unfit to engage in activities and discussions about science and technology, due to a lack of understanding or interest. To create equal opportunities for all citizens to engage in AI

activities, exclusionary structures should be removed to make the activities better accessible and more inclusive.

References

Dawson, E. (2018). Reimagining publics and (non) participation: Exploring exclusion from science communication through the experiences of low-income, minority ethnic groups. Public Understanding of Science, 27(7), 772-786. https://doi.org/10.1177/0963662517750072

Islam, R., Keya, K. N., Pan, S., Sarwate, A. D., & Foulds, J. R. (2023). Differential Fairness: An Intersectional Framework for Fair AI. Entropy, 25(4), 660. https://www.mdpi.com/1099- 4300/25/4/660

Romero-Rodriguez, L. M., Civila, S., & Aguaded, I. (2021). Otherness as a form of intersubjective social exclusion. Journal of Information, Communication and Ethics in Society, 19(1), 20-37. https://doi.org/10.1108/JICES-11-2019-0130

How issues travel across social conversation: The case of AI in Italy

Massimiano Bucchi, Eliana Fattorini

The paper analyses public perception and discussion of AI within the framework of science communication as the social conversation around science (Bucchi and Trench, 2021).
The idea of science communication as social conversation around science is an inclusive definition that encompasses a wide range of formats and configurations, a spectrum with different degrees of formalization and purposiveness; from the more hierarchical to the more participatory; from closed to open formats.

Frameworks of the social conversation around science

Base model	Dissemination					Dialogue				Participation				
Sci-comm applications	Deficit	Defence	Promotion	Popularisation	Outreach	Engagement	Consultation	Interactivity	Deliberation	Chat	Play	Co-creation	Film & fiction	Art-science
Aspects of science	Findings: finished knowledge					Issues: applications and implications of knowledge				Processes: interpreting and (re-) constructing knowledge				
Public uses	Information, awareness, learning					Questioning, opinion, discussion				Sharing, creating, enjoyment, critique				
Social perspectives	Science literacy: scientism, technocracy					Science in society: Mode-2, post-normal, post-academic				Society in science: civic science, citizen science				
Orientation	Purposive Hierarchical Formal Closed									Non-purposive Participatory Informal Open				

In terms of communication models, the spectrum ranges from dissemination to dialogue to participation. Science communication issues can travel across this spectrum in different ways and directions: an issue can initially be introduced in terms of dissemination and then move towards more participatory and even conflictual configurations (e.g. nuclear energy, GMOs). Vice versa, issues can be introduced and fostered into the social conversation by citizen mobilization (e.g. specific health or environmental issues). Research activities and results can inspire – but also be inspired by - fictional content (e.g. robotics, space exploration).

The case of AI is particularly interesting in this respect (Brause, Zeng, Schäfer & Katzenbach, 2023; Schäfer & Metag, 2021). Recent data from Observa Science in Society monitor (a regular monitor of public perception and attitudes to science and technology that has been in place since 2003), for example, show that citizens recognize their general lack of information on AI, asking at the same time for rigid regulation. The issue is moving fluidly across the spectrum of social conversation, from a dissemination-like configuration (e.g. “what is AI?”) to a more dialogic (“what are the implications for work and the economy?”) and even controversial configuration (contested futures; regulation and governance; e.g. in connection with the temporary ban of ChatGPT in Italy in the spring of 2023, the London international policy meeting and the temporary firing of OpenAI CEO in November 2023).

Drawing upon quantitative data of public perception (another data collection is expected in Spring 2024) and qualitative analysis of media narratives (with a special focus on visual elements), the paper will focus on what we can learn from this case in terms of our conceptual understanding of science communication.

References

Brause, S. R., Zeng, J., Schäfer, M. S. & Katzenbach, C. (2023). Media Representations of Artificial Intelligence. In S. Lindgren (Ed.), *Handbook of Critical Studies of Artificial Intelligence*. Edward Elgar Publishing.

Bucchi, M. and Trench, B. (2021). Rethinking science communication as the social conversation around science JCOM 20(03), Y01. https://doi.org/10.22323/2.20030401

Schäfer, M. S. & Metag, J. (2021). Audiences of science communication between pluralisation, fragmentation and polarization In M. Bucchi & B. Trench (Eds.), *Routledge handbook of public communication of science and technology* (3rd ed., pp. 291–304). New York, NY, U.S.A.: Routledge

Panel V: COMMUNICATING WITH AI IN SCIENCE JOURNALISM AND SCIENCE COMMUNICATION

Friday, June 7, 09:45 – 10:45
Chair: Michelle Riedlinger
Room: HG E 1.2

The future of science communication - which role plays generative AI? A Delphi study with communicators and scientists

Josephine B. Schmitt, Matthias Begenat, Sandra Kero, Jasmin Baake

Looking at the (highly probable) future of science communication, it is no longer just about new formats that are distributed via digital platforms, but also about the production of synthetic media. Synthetic media refers to media products such as texts (including translations into all kinds of languages), videos, images and audio that are modified or generated using AI (Heesen et al., 2023; WDR Innovation Hub, 2021). Combined with each other, the tools can save a lot of resources (time, budget, labor). They allow for more extensive and versatile content production and new types of formats that would not be possible without their use. Large amounts of data can be processed and presented, complex phenomena can be introduced to the audience through interactive and immersive experiences. The disruptive potential of generative AI becomes even clearer when new developments that are already available one year after the release of ChatGPT are also considered (such as using ChatGPT through voice input and output, self-customized versions of ChatGPT for specific purposes, integration of AI in social media or new devices). The dynamics of progress will probably show that by the time of the conference, many of the applications mentioned in the submissions will already be outdated.

Until now, the structural change in scientific communication brought about by digitalization has been described along various dimensions: Pluralization and professionalization of actors, dynamization of processes, globalization of communication spaces, personalization, rationalization and economizationof communication (Vowe, 2016; Weingart, 2017). In many cases, this is accompanied by concerns about a loss of quality and trust (Boothby et al., 2021; Rubin et al., 2022). In the future, AI will have a direct impact on the targets of scientific literacy and trust and will present science communication with challenges that have not yet been considered (for a scenario for the development of universities and translation services, Seemann, 2023).

Synthetic content often impresses with its completeness, accuracy and, especially in the case of images and videos, its realistic representation. At the same time, however, there is something artificial about synthetic media that is often (still) easy to recognize and therefore unsettling (Schwind et al., 2018). The use of AI also has direct implications for professional ethics and quality standards in science communication. A possible loss of quality in communication due to improper or careless use could lead to an erosion of trust in science.

In various areas of society, such as journalism, the benefits and risks of the new AI tools are already being discussed (Heesen et al. 2023). In science communication, the specialist public and practitioners still seem largely unprepared. The current guidelines for science communication from various institutions do not (yet) reflect the possibilities outlined above (Deutscher Rat für Public Relations, 2022; Wissenschaft im Dialog & Bundesverband Hochschulkommunikation, 2016). Similarly, the "Perspectives for action in science communication" of #FACTORYWISSKOMM does

not contain any further explanations on this topic, only the keywords "artificial intelligence" and "automation" (FactoryWisskomm, 2021). We would like to fill this gap. And we would also like to interlink operative science communication and scientific research.

The question regarding the use of AI in science communication is not whether AI generated content isused in science communication, but how and under what ethical and legal framework conditions.

Method

Whether and how AI applications are part of the professional practice of science communicators and what use and impact assumptions they have and predict for synthetic content will be recorded by two Delphi studies with science communicators and researchers with a focus on science communication and AI. These studies will take place over the course of two years to reflect the dynamics of the field. In at least two written, standardized survey rounds and discussions in focus groups, well-founded assessments are collected. Based on the statements of the other group (researchers or practitioners) the experts can correct or expand their statements. The aim of the survey rounds is to record consensus and dissent in judgments (Döring & Bortz, 2016). Based on the experts' assessments, potential fields of application and the feasibility of AI in content production can be identified. At the same time, challenges in practice can be recognized. In this way, the necessary skills of communicators can be determined in order to produce quality-assured content.

Until the conference we will have finished the first round of the Delphi study with science communicators and researchers. We will discuss first results.

References

Boothby, C., Murray, D., Waggy, A. P., Tsou, A., & Sugimoto, C. R. (2021). Credibility of scientific information on social media: Variation by platform, genre and presence of formal credibilitycues. *Quantitative Science Studies*, *2*(3), 845–863. https://doi.org/10.1162/qss_a_00151

Deutscher Rat für Public Relations. (2022, June). *DRPR Richtlinie zur Wissenschafts-PR*. Drpr- Online.De. https://drpr-online.de/kodizes-2/ratsrichtlinien/wissenschaftskommunikation/

Döring, N., & Bortz, J. (2016). Datenerhebung. In N. Döring & J. Bortz (Eds.), *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften* (pp. 321–577).Springer https://doi.org/10.1007/978-3-642-41089-5_10

FactoryWisskomm. (2021). *Handlungsperspektiven für die Wissenschaftskommunikation*. https://www.bmbf.de/bmbf/shareddocs/downloads/files/factorywisskommpublikation.pdf?__blob=publicationFile&v=2

Heesen, J., Bieber, C., Lauber-Rönsberg, A., & Neuberger, C. (2023). *Künstliche Intelligenz im Journalismus: Potenziale und Herausforderungen für Medienschaffende* (p. 33 pages)[Application/pdf]. Lernende Systeme – Die Plattform für Künstliche Intelligenz. https://doi.org/10.48669/PLS_2023-1

Rubin, A., Brondi, S., & Pellegrini, G. (2022). Should I trust or should I go? How people perceive and assess the quality of science communication to avoid fake news. *Quality & Quantity*. https://doi.org/10.1007/s11135-022-01569-5

Schwind, V., Wolf, K., & Henze, N. (2018). Avoiding the uncanny valley in virtual character design. *Interactions*, *25*(5), 45–49. https://doi.org/10.1145/3236673

Seemann, M. (2023). Künstliche Intelligenz, Large Language Models, ChatGPT und die Arbeitswelt der Zukunft. Hans-Böckler-Stiftung. https://doi.org/10.5281/zenodo.8416853

Vowe, G. (2016). Wissenschaftskommunikation 2.0? *Publizistik*, *61*(1), 51–72.https://doi.org/10.1007/s11616-015-0249-1

WDR Innovation Hub. (2021). *Synthetische Medien—Zukünfte der Medienproduktion mit KünstlicherIntelligenz: Ein Zukunftsreport des WDR Innovation Hub*.

Weingart, P. (2017). Wissenschaftskommunikation unter digitalen Bedingungen. Funktionen, Akteureund Probleme des Vertrauens. In *Perspektiven der Wissenschaftskommunikation im digitalenZeitalter.* (pp. 31–59). Velbrück Wissenschaft.

Wissenschaft im Dialog, & Bundesverband Hochschulkommunikation. (2016, February). *Leitlinien zurguten Wissenschafts-PR*. https://www.wissenschaft-im-dialog.de/fileadmin/user_upload/Trends_und_Themen/Dokumente/Leitlinien gute-Wissenschafts-PR_final.pdf

University communication in the age of AI: First insights into the use and perspectives of generative AI tools

Justus Henke

This study examines recent developments in higher education communication in the age of artificial intelligence, focusing on the impact of generative AI tools such as ChatGPT. The increasing proliferation of artificial intelligence (AI) is significantly influencing the higher education landscape, not only in research and teaching but also in communication and organizational development. Generative AI tools will likely play an important role in the future of science communication, yet there is currently little research in this area (Schäfer, 2023). Higher education communication has already transformed significantly through digitalization, now utilizing numerous additional communication channels (Neuberger et al., 2021). Social media, for example, enables direct dialogue between universities and their target groups, but the future role of generative AI in this area remains unclear. Despite their potential to improve the efficiency and reach of communication, cautious use is currently advised (Fecher et al., 2023; Stokel-Walker & Van Noorden, 2023).

Drawing on concepts from Human-Machine-Communication (Guzman, 2018), Socio-technical Systems (Bijker et al., 2012; Orlikowski, 1992) and Technology-Acceptance-Model (Davis, 1986), this study derives the question: What are the expectations and perceived capabilities of generative AI tools among communication professionals in German universities, and how do these perceptions influence their acceptance and use of such technologies? These topics were the subject of an online survey conducted among the press offices and communication departments of all German universities in May 2023. A follow-up survey is planned for spring 2024.

Key findings of the survey indicate that while many universities have already experimented with ChatGPT, only a minority (22%) regularly use such chatbots in their communication work. There is a prevailing view that AI tools mainly offer efficiency gains, but not significant quality improvements. Other chatbots with integrated web search (e.g., Bing Chat) or document analysis (e.g., ChatPDF) are well-known, but their usage is low. The highest usage frequency is seen in AI-supported translation and language correction tools (e.g., DeepL, Grammarly) with 73% regular usage. Most respondents expressed mixed feelings about their satisfaction with these tools, with only a small number expressing high satisfaction or dissatisfaction. Most have not yet seen significant improvements in efficiency or changes in work practices as a result of using AI tools. However, the expectation for AI tools is predominantly time saving in content creation. Surprisingly, functions related to personalized communication and quality improvement of communication were deemed less important. Challenges in using AI tools include technical difficulties, suboptimal application, lack of adaptability, and insufficient training opportunities, as mentioned by respondents.

The survey findings align with existing literature and present a complex view of AI in university communications, highlighting critical challenges such as data protection (Arthur et al., 2023; Ninaus & Sailer, 2022; Zawacki-Richter et al., 2019) and the difficulty in distinguishing fact from fiction (Rawte et al., 2023; Zhang et al., 2023). Ethical considerations draw significant attention, aligning with concerns in academic discourse (Azaria et al., 2023; Dutta, 2023; Fecher et al., 2023), underscoring the urgent need for ethical guidelines. While few fear job losses, highlighting concerns over the digital divide and role displacement (Munoriyarwa et al., 2023; Noain-Sánchez, 2022; Peña-Fernández et al., 2023; Ray, 2023), the potential of AI to automate tasks and enhance communication efficiency is confirmed (Matz et al., 2023; Parycek et al., 2023). Yet, there is room for greater emphasis on personalized communication and quality improvement beyond mere efficiency gains.

The findings suggest a need for strategic alignment (Volk & Zerfass, 2020) in university

communications, emphasizing adaptive policies for navigating AI complexities. This approach calls for a balance between strategic planning and flexibility to address emerging issues in AI adoption. Not only are challenges to be overcome, but the potentials, especially those that have received little attention, need to be more prominently considered: Generative AI has the potential to not only increase efficiency but also significantly enhance the quality of communication.

References

Arthur, L., Costello, J., Hardy, J., O'Brien, W., Rea, J., Rees, G., & Ganev, G. (2023). On the Challenges of Deploying Privacy-Preserving Synthetic Data in the Enterprise. https://doi.org/10.48550/ARXIV.2307.04208

Azaria, A., Azoulay, R., & Reches, S. (2023). ChatGPT is a Remarkable Tool For Experts. https://doi.org/10.48550/ARXIV.2306.03102

Bijker, W. E., Hughes, T. P., & Pinch, T. (Hrsg.). (2012). The social construction of technological systems: New directions in the sociology and history of technology (Anniversary ed). MIT Press.

Davis, F. D. (1986). A technology acceptance model for empirically testing new end-user information systems: Theory and results [Phd, Massachusetts Institute of Technology]. http://hdl.handle.net/1721.1/15192

Dutta. (2023). The Ethics of Artificial Intelligence in Legal Decision Making: An Empirical Study. Psychology and Education, 55(01). https://doi.org/10.48047/pne.2018.55.1.38

Fecher, B., Hebing, M., Laufer, M., Pohle, J., & Sofsky, F. (2023). Friend or foe? Exploring the implications of large language models on the science system. AI & SOCIETY. https://doi.org/10.1007/s00146-023-01791-1

Guzman, A. L. (2018). Human-Machine Communication. https://www.peterlang.com/document/1055458

Matz, S., Teeny, J., Vaid, S. S., Peters, H., Harari, G. M., & Cerf, M. (2023). The Potential of Generative AI for Personalized Persuasion at Scale [Preprint]. PsyArXiv. https://doi.org/10.31234/osf.io/rn97c

Munoriyarwa, A., Chiumbu, S., & Motsaathebe, G. (2023). Artificial Intelligence Practices in Everyday News Production: The Case of South Africa's Mainstream Newsrooms. Journalism Practice, 17(7), 1374–1392. https://doi.org/10.1080/17512786.2021.1984976

Neuberger, C., Weingart, P., Fähnrich, B., Fecher, B., Schäfer, M. S., Schmid-Petri, H., & Wagner, G. G. (2021). Der digitale Wandel der Wissenschaftskommunikation. Wissenschaftspolitik im Dialog, 63.

Ninaus, M., & Sailer, M. (2022). Zwischen Mensch und Maschine: Künstliche Intelligenz zur Förderung von Lernprozessen. Lernen und Lernstörungen, 11(4), 2235-0977/a000386. https://doi.org/10.1024/2235-0977/a000386

Noain-Sánchez, A. (2022). Addressing the Impact of Artificial Intelligence on Journalism: The perception of experts, journalists and academics. Communication & Society, 35(3), 105–121. https://doi.org/10.15581/003.35.3.105-121

Orlikowski, W. J. (1992). The Duality of Technology: Rethinking the Concept of Technology in Organizations. Organization Science, 3(3), 398–427. JSTOR.

Parycek, P., Schmid, V., & Novak, A.-S. (2023). Artificial Intelligence (AI) and Automation in Administrative Procedures: Potentials, Limitations, and Framework Conditions. Journal of the Knowledge Economy. https://doi.org/10.1007/s13132-023- 01433-3

Peña-Fernández, S., Meso-Ayerdi, K., Larrondo-Ureta, A., & Díaz-Noci, J. (2023). Without journalists, there is no journalism: The social dimension of generative artificial intelligence in the media. El Profesional de la información, e320227. https://doi.org/10.3145/epi.2023.mar.27

Rawte, V., Sheth, A., & Das, A. (2023). A Survey of Hallucination in Large Foundation Models. https://doi.org/10.48550/ARXIV.2309.05922

Ray, P. P. (2023). ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. Internet of Things and Cyber-Physical Systems, 3, 121–154. https://doi.org/10.1016/j.iotcps.2023.04.003

Schäfer, M. S. (2023). The Notorious GPT: Science communication in the age of artificial intelligence. Journal of Science Communication, 22(2), Y02. https://doi.org/10.22323/2.22020402

Stokel-Walker, C., & Van Noorden, R. (2023). What ChatGPT and generative AI mean for science. Nature, 614(7947), 214–216. https://doi.org/10.1038/d41586-023-00340-6

Volk, S. C., & Zerfass, A. (2020). Alignment: Explicating a key concept in strategic communication. Future directions of strategic communication, 105–123.

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators? International Journal of Educational Technology in Higher Education, 16(1), 39. https://doi.org/10.1186/s41239-019-0171-0

Zhang, S., Heck, P. R., Meyer, M. N., Chabris, C. F., Goldstein, D. G., & Hofman, J. M. (2023). An illusion of predictability in scientific results: Even experts confuse inferential uncertainty and outcome variability. Proceedings of the National Academy of Sciences, 120(33), e2302491120. https://doi.org/10.1073/pnas.2302491120

Meet my new colleague, ChatGPT: How German science journalists perceive and use generative Artificial Intelligence in the selection, production, and distribution of news

Lars Guenther, Jessica Kunert, Bernhard Goodwin

"Artificial intelligence and the future of journalism go hand-in-hand" (Dhiman, 2023, p. 6) – this is what research on the topic often proclaims. Indeed, new developments regarding *(generative) Artificial Intelligence (AI)* have either been heralded as a savior or feared as the death of journalism (e.g., Borchardt, 2022; Broussard et al., 2019; Deuze & Beckett, 2022; Hansen et al., 2017; Moran & Shaikh, 2022; Soto-Sanfiel et al., 2022). Among the discussed benefits of AI are increased efficiency, reduced costs, and enhanced user experience (e.g., Beckett, 2019; Dhiman, 2023; Kunert, 2020; Stray,2019). Nevertheless, there is also a strong focus on the risks, such as the increasing lack of autonomy, e.g., when it comes to dependency from AI companies (e.g., "infrastructure capture", Simon, 2022; see also Borchardt, 2022; Cools et al., 2021) as well as questions of accuracy, bias, transparency, and accountability (e.g., Beckett, 2019; Broussard et al., 2019; Dhiman, 2023; Hansen et al., 2017). Only a few media companies have started to put their ideas into AI guidelines or an AI strategy (e.g., Beckett,2019; Hansen et al., 2017). These developments concern the whole process of making news, from selection, to production, to distribution (e.g., Beckett, 2019; Deuze & Beckett, 2022; Dhiman, 2023; Moran & Shaikh, 2022; Pavlik, 2023; Simon, 2022). Often, the pressing question is which tasks will be given to (generative) AI and which will remain with human journalists (e.g., Moran & Shaikh, 2022).

While this concerns journalism overall, it also concerns *science journalism*, which is further been seen as to be in crisis (e.g., Guenther, 2019; Maiden et al., 2023; Schäfer, 2023). For a long time, scholars have not expected that science journalism would be a suitable journalistic beat for generative AI (e.g., due to complexities and scientific uncertainty); however, this seems to be changing (e.g., Tatalovic, 2018). Research so far has focused on AI tool development for science journalists (e.g., Vadapalli et al., 2018), how they assess AI tools (Maiden et al., 2023), or how audiences evaluate AI authorship for science journalistic articles (e.g., Lermann Henestrosa, Greving,& Kimmerle, 2023). However, how science journalists themselves perceive generative AI and how they use it as part of their practices has yet to be addressed. A research endeavor like this falls into the "communication with AI" avenue for which Schäfer (2023) called scholars to work on. Insights generated by this research could shed light on the question if generative AI will likely expand on the crisis of science journalism or if it could potentially counteract it.

We make use of the *Technology Acceptance Model* (Davis, 1989; see also Soto-Sanfiel et al., 2022 for a similar approach), including its extensions and connections to the *Theory of Reasoned Action* and the *Theory of Planned Behavior* (Fishbein & Ajzen, 2010). According to these theories, the individual use of a new technology (such as generative AI) in a work context is affected by components such as the perceived usefulness of the technology, the perceived ease of use, attitudes towards using the technology, (injunctive and descriptive) subjective norms, and control beliefs. In this study, we adopt central elements of these theoretical approaches, to apply them to (science) journalistic practices. More specifically, we want to explore the actual behavior of German science journalists with regards to using generative AI to select, produce, and distribute news, the attitudes they have, the chances and risks they see (including ethical components), the social settings and norms in which this happens, as well as the future developments they think are likely. Based on that, the central research question (RQ) of this paper is:

How do German science journalists perceive and use generative Artificial Intelligence when selecting, producing, and distributing news?

To answer this RQ, we are currently conducting semi-structured interviews with 30 German science journalists. To include various perspectives, we selected journalists from leading media organizations such as public and private TV, (online) daily and weekly newspapers/news magazines, radio, science magazines, and podcasts. Selected individuals were approached and asked to take part in an (online) interview, led by trained interviewers. The interview guide covers all aspects of the theoretical models (Davis, 1989, Fishbein & Ajzen, 2010), extended by current research insights (e.g., Broussard et al., 2019; Beckett, 2019; Deuze & Beckett, 2022; Dhiman, 2023; Hansen et al., 2017; Moran & Shaikh, 2022; Pavlik, 2023; Simon, 2022; Stray, 2019). More specifically, we ask how the targeted science journalists are using generative AI when selecting, producing, and distributing news, what AI tools they use, what benefits they see, what risks they see (including autonomy, accuracy, bias, transparency, and accountability issues), if the company they work for has AI guidelines or an AI strategy, how their colleagues are using it, general attitudes towards generative AI, and future implications of generative AI for science journalism.

While we cannot present findings yet, we can emphasize that by June 2024, when the conference will take place, we will have applied qualitative content analysis to the interview transcripts and will be able to answer the RQ, and hence, how German science journalists perceive and use generative AI when selecting, producing, and distributing news. Given that (science) journalism is moving towards integration with generative AI, the insights of this study will shed light on the question if generative AI fuels the crisis of science journalism or has the potential to counteract it.

References

Beckett, C. (2019). New powers, new responsibilities. A global survey of journalism and artificial intelligence. Retrieved from: https://drive.google.com/file/d/1utmAMCmd4rfJHrUfLLfSJ- clpFTjyef1/view

Broussard, M., Diakopoulos, N., Guzman, A. L., Abebe, R., Dupagne, M., & Chuan, C.-H. (2019). Artificial Intelligence and journalism. Journalism & Mass Communication Quarterly, 96(3), 673– 695. https://doi.org/10.1177/1077699019859901

Borchardt, A. (2022). Go, robots, go! The value and challenges of Artificial Intelligence for local journalism. Digital Journalism, 10(10), 1919–1924. https://doi.org/10.1080/21670811.2022.2149584

Cools, H., Van Gorp, B., & Opgenhaffen, M. (2021). When algorithms recommend what's new(s): New dynamics of decision-making and autonomy in newsgathering. Media & Communication, 9(4), 198–207. https://doi.org/10.17645/mac.v9i4.4173

Davis, F. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly, 13(3), 319–340. https://doi.org/10.2307/249008

Deuze, M., & Beckett, C. (2022). Imagination, algorithms and news: Developing AI literacy for journalism. Digital Journalism, 10(10), 1913–1918. https://doi.org/10.1080/21670811.2022.2119152

Dhiman, B. (2023). Does Artificial Intelligence help journalists: A boon or bane? Preprint. https://doi.org/10.20944/preprints202303.0428.v1

Fishbein, M., & Ajzen, I. (2010). Predicting and changing behavior: The Reasoned Action Approach. New York: Psychology Press.

Guenther, L. (2019). Science journalism. In H. Ornebring (Ed.), Oxford Encyclopedia of Journalism Studies. New York: Oxford University Press.

Hansen, M., Roca-Sales, M., Keegan, J., & King, G. (2017). Artificial Intelligence: Practice and implications for journalism. Retrieved from: https://academiccommons.columbia.edu/doi/10.7916/D8X92PRD

Kunert, J. (2020). Automation in sports reporting: Strategies of data providers, software providers, and media outlets. Media and Communication, 8(3), 5–15. https://doi.org/10.17645/mac.v8i3.2996

Lermann Henestrosa, A., Greving, H., & Kimmerle, J. (2023). Automated journalism: The effects of AI authorship and evaluative information on the perception of a science journalism article. Computers in Human Behavior (online before print). https://doi.org/10.1016/j.chb.2022.107445

Maiden, N., Zachos, K., Franks, S., Nyre, L., & Linden, C.-G. (2023). Automating Science journalism tasks: Emerging opportunities. Journalism Practice (online before print). https://doi.org/10.1080/17512786.2023.2226116

Moran, R. E., & Shaikh, S. J. (2022). Robots in the news and newsrooms: Unpacking meta-journalistic discourse on the use of Artificial Intelligence in journalism. Digital Journalism, 10(10), 1756–1774. https://doi.org/10.1080/21670811.2022.2085129

Pavlik, J. V. (2023). Collaborating with ChatGPT: Considering the implications of generative Artificial Intelligence for journalism and media education. Journalism & Mass Communication Educator, 78(1), 84–93. https://doi.org/10.1177/1077695822114957

Schäfer, M. S. (2023). The notorious GPT: Science communication in the age of artificial intelligence. Journal of Science Communication, 22(02), Y02. https://doi.org/10.22323/2.22020402

Simon, F. M. (2022). Uneasy bedfellows: AI in the news, platform companies and the issue of journalistic autonomy. Digital Journalism, 10(10), 1832–1854. https://doi.org/10.1080/21670811.2022.2063150

Soto-Sanfiel, M. T., Ibiti, A., Machado, M., Marín Ochoa, B. E., Mendoza Michilot, M., Rosell Arce, C. G., & Angulo-Brunet, A. (2022). In search of the Global South: Assessing attitudes of Latin American journalists to Artificial Intelligence in journalism. Journalism Studies, 23(10), 1197–1224. https://doi.org/10.1080/1461670X.2022.2075786

Stray, J. (2019). Making Artificial Intelligence work for investigative journalism. Digital Journalism, 7(8), 1076–1097. https://doi.org/10.1080/21670811.2019.1630289

Tatalovic, M. (2018). AI writing bots are about to revolutionise science journalism: We must shape how this is done. Journal of Science Communication, 17(01), E. https://doi.org/10.22323/2.17010501

Vadapalli, R., Syed, B., Prabhu, N., Srinivasan, B. V., Varma, V. (2018). When science journalism meets artificial intelligence: An interactive demonstration. Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing (System Demonstrations), 163–168.

Parallel Panel VI: CHATGPT'S ROLE IN SCIENCE COMMUNICATION

Friday, June 7, 11:05 – 12:25
Chair: Anne M. Dijkstra
Room: HG E 1.2

"Chat GPT, is the influenza vaccination useful?" Comparing perceived argument strength and correctness of pro-vaccination-arguments from AI and scientific experts

Selina A. Beckmann, Elena Link, Marko Bachl

Artificial intelligence (AI) has become increasingly important and has attracted much public attention with the emergence of publicly available large language models (LLMs) such as ChatGPT (e.g. Schäfer, 2023). Such innovative technologies allow not only humans but also machines to produce content that can be used to inform or persuade (Palmer & Spirling, 2023) for example about the preventive health care measure of vaccination. As vaccine hesitancy is among the ten most crucial threats to global health (World Health Organization, 2019), this study examines the ability of LLMs compared to health experts from science to provide convincing arguments. We focus on the influenza vaccination, as this established vaccination is recommended annually for a large proportion of the German population (Robert Koch-Institut, 2023).

Previous research shows that the quality of AI-generated text is rated better than human-generated text when the source of the information is unknown, for example in the context of health recommendations such as vaccination against Covid 19 (Karinshak et al., 2023) or folic acid intake during pregnancy (Lim & Schmälzle, 2023). However, extant research revealed that the labelling as AI or LLM influences individuals' assessment as well as its persuasive impact. Research across various fields such as health communication, robot journalism, or marketing (e.g., Graefe et al., 2018; Palmer & Spirling, 2023) result in heterogenous results. Whereas in the US and the Netherlands labelling a message as algorithmically generated leads to lower quality ratings (Graefe et al., 2018; Waddell, 2018; Zheng et al., 2018), findings from China and South Korea show that algorithmically generated texts are rated more positively than human-generated texts (Jung et al., 2017; Zheng et al., 2018). Building on this state of research, we focus on Germany and analyse the perception of different arguments about influenza vaccination generated by AI or scientific experts. In a first study, we investigate which arguments are perceived as strong or weak (RQ 1). Further, we test whether arguments generated by AI are perceived as qualitatively better than those generated by scientific experts. Based on the current state of research, it is assumed that arguments generated by AI are rated to possess a higher argument strength (H1a) as well as correctness (H1b) than arguments written by scientific experts. In a second study, the labelling of the argument will be supplemented.

Method
A two-study design is planned. In this abstract, we focus on the first experimental study. To analyse the research question and test the hypotheses, an online survey applying a within-person experimental design (N = 294; 48.6% female, age: M = 56.73, SD = 14.81) recruited via an online access panel representative for a region in south Germany was conducted in November 2023. The respondents were asked to rate the quality of eight statements about influenza vaccination. Four of these arguments were generated using ChatGPT. Four everyday prompts were used to generate the AI arguments (e.g. "Is the influenza vaccination useful?"). ChatGPT was instructed to extract the strongest arguments from the responses, which in turn identified four recurring themes (see Table

1). One argument per theme was randomly selected for quality assessment. For the arguments of the scientific experts, the websites of the Robert Koch-Institute and the Federal Centre for Health Education were searched for arguments that corresponded thematically to the four AI arguments to establish comparability of content. Argument quality was assessed by combining perceived argument strength (PAS) (Zhao et al., 2011) and correctness (Kohring & Matthes, 2004), which includes most aspects of quality analysed in previous studies. RQ 1 was answered by descriptive analysis, RM-ANOVAs were conducted to test the hypotheses 1a/b.

Results of study 1 and discussion

Regarding RQ 1, we found that participants gave higher ratings to correctness than to PAS for almost all arguments (see Table 1). The assessments of the arguments seemed to be related to their thematic background. For PAS, the strongest argument is the scientific experts about the reduction of individual risks through vaccination, whereas the same AI argument received the lowest rating. The strength of the argument about protection against serious illness was also rated positive independent of the AI or expert version. Nevertheless, both experts and AI generated rather convincing arguments. Concerning H1, the results of the RM-ANOVAs showed that both the evaluation of PAS ($F(1,250) = 65.65$, $p < .001$, $\eta p2 = .208$, $f = .26$) and correctness ($F(1,268) = 25.35$, $p < .001$, $\eta p2 = .086$, $f = .09$) depend on the source of the argument. The PAS and correctness of arguments from scientific experts were significantly higher than that of arguments generated by AI (see Table 1). Thus, H1a and H1b need to be rejected.

Table 1: Descriptive Results of Argument Strength and Correctness

Argument Theme	PAS		Correctness	
	Argument by		Argument by	
	Experts	AI	Experts	AI
	M (SD)	M (SD)	M (SD)	M (SD)
Protection against serious illness	3.68 (0.90)	3.69 (0.93)	3.86 (1.00)	3.98 (1.00)
Protection of the community	3.55 (1.01)	3.24 (0.99)	3.73 (1.13)	3.55 (1.13)
Protection of the health care system	3.39 (1.03)	3.38 (1.02)	3.66 (1.11)	3.74 (1.06)
Reduction of individual risks	3.70 (0.92)	3.11 (1.07)	3.99 (1.00)	3.44 (1.17)
Overall	3.59 (0.80)	3.35 (0.84)	3.81 (0.91)	3.68 (0.94)
N = 294, Results of two RM-ANOVAs				

To sum up, the study does not confirm previous research on quality perceptions when the source remains unknown. While the strongest arguments come from both, experts and AI, the overall results show that experts' arguments are rated better in terms of PAS and correctness. One possible reason for this could be the methodology used to generate the AI arguments, which differs from other studies training ChatGPT in advance. In our case, we aimed to simulate everyday use as closely as possible.

Building on previous research, the second study will investigate whether the labelling of the source of the text has an impact on individuals' assessment. In addition, the mechanisms behind the quality ratings of texts with different labels, in particular innovativeness and trust (Jung et al., 2017) will be investigated in more detail. The presentation will cover results of both studies.

References

Graefe, A., Haim, M., Haarmann, B., & Brosius, H.-B. (2018). Readers' perception of computer- generated news: Credibility, expertise, and readability. Journalism, 19(5), 595–610. https://doi.org/10.1177/1464884916641269

Jung, J., Song, H., Kim, Y., Im, H., & Oh, S. (2017). Intrusion of software robots into journalism: The public's and journalists' perceptions of news written by algorithms and human journalists. Computers in Human Behavior, 71, 291–298. https://doi.org/10.1016/j.chb.2017.02.022

Karinshak, E., Liu, S. X., Park, J. S., & Hancock, J. T. (2023). Working With AI to Persuade: Examining a Large Language Model's Ability to Generate Pro-Vaccination Messages. Proceedings of the ACM on Human-Computer Interaction, 7(CSCW1), 1–29. https://doi.org/10.1145/3579592

Kohring, M., & Matthes, J. (2004). Revision und Validierung einer Skala zur Erfassung von Vertrauen in Journalismus. Medien & Kommunikationswissenschaft, 52(3), 377–385. https://doi.org/10.5771/1615-634x-2004-3-377

Lim, S., & Schmälzle, R. (2023). Artificial intelligence for health message generation: an empirical study using a large language model (LLM) and prompt engineering. Frontiers in Communication, 8, Article 1129082. https://doi.org/10.3389/fcomm.2023.1129082

Palmer, A. K., & Spirling, A. (2023). Large Language Models Can Argue in Convincing and Novel Ways About Politics.: Evidence from Experiments and Human Judgement.

Robert Koch-Institut (Ed.). (2023). Epidemiologisches Bulletin: Empfehlungen der Ständigen Impfkommission beim Robert Koch-Institut 2023.

Schäfer, M. S. (2023). The Notorious GPT: science communication in the age of artificial intelligence. Journal of Science Communication, 22(02). https://doi.org/10.22323/2.22020402

Waddell, T. F. (2018). A Robot Wrote This? How perceived machine authorship affects news credibility. Digital Journalism, 6(2), 236–255. https://doi.org/10.1080/21670811.2017.1384319

World Health Organization (Ed.). (2019). Ten Threats to Global Health in 2019. https://www.who.int/news-room/spotlight/ten-threats-to-global-health-in-2019

Zhao, X., Strasser, A., Cappella, J. N., Lerman, C., & Fishbein, M. (2011). A Measure of Perceived Argument Strength: Reliability and Validity. Communication Methods and Measures, 5(1), 48–75. https://doi.org/10.1080/19312458.2010.547822

Zheng, Y., Zhong, B., & Yang, F. (2018). When algorithms meet journalism: The user perception to automated news in a cross-cultural context. Computers in Human Behavior, 86, 266–275. https://doi.org/10.1016/j.chb.2018.04.046

How Generative AI Imagines and Communicates Science: Interviewing ChatGPT from the perspective of different audience segments

Sophia C. Volk, Mike S. Schäfer, Damiano Lombardi, Daniela Mahl, Xiaoyue Yan

Generative AI tools like ChatGPT have already, and will increasingly, become important sources of information about a wide range of topics. This also includes science (Fecher et al., 2023). In Germany, 64% of the population has heard of ChatGPT and 50% appreciate that it "allows complex scientific and research issues to be explained in a very simplified way" (WiD, 2023).

Conceptual Framework

Even though research on generative AI has mushroomed in recent months, its role in science communication is still underresearched (Schäfer, 2023). Therefore, analyzing how ChatGPT portrays science is important – especially given that many scholars assume that AI tools contain biases (Corless, 2023; Dwivedi et al., 2023; Hosseini et al., 2023), potentially reproducing "the same old trivialities and stereotypes" (Teubner et al., 2023, p. 99). Yet, since the underlying data of the model is proprietary, science-related biases potentially included in ChatGPT are not known. Hence, we ask:

- *RQ1: How does ChatGPT present science and related issues?*

Given that ChatGPT takes information provided in user profiles into account when responding to queries (OpenAI, 2023), it is plausible to assume that different audiences will receive different answers (Chen et al., 2022). However, since the specific training of ChatGPT is opaque, it remains unclear to what extent answers are targeted to different users. The possibility that different audience segments with varying perceptions of science – from supportive to critical (Schäfer et al., 2018) – may receive different (potentially biased) answers by ChatGPT evokes concern. We ask:

- *RQ2: Inhowfar do responses differ across user profiles?*

Finally, since prior analyses show differences in performance between the variants GPT-3.5 (free) and GPT-4 (paid) with regard to science-related issues (Bulian et al., 2023), this study aims to compare them. We ask:

- *RQ3: Inhowfar do responses differ between GPT-3.5 and GPT-4?*

Method

We first developed a set of "interview questions" to be answered by ChatGPT, structured into four blocks: (1) science and scientists, (2) science and the public, (3) scientific misbehavior, (4) controversial scientific topics. Second, we set up five user profiles for GPT-3.5 and GPT-4, respectively. One user profile description remained blank — representative of the likely usage by most users. The remaining four user profiles contained descriptions that aligned with the four audience segments of science communication proposed by Schäfer et al. (2018): *Sciencephiles, Critically Interested, Passive Supporters, Disengaged*. For each profile, we added statements like "I have a high interest in science" to reflect different attitudes toward science. Then we conducted *N* = 40 "interviews" with ChatGPT: eight per profile, four per GPT version. We used qualitative content analysis in MAXQDA to reconstruct how ChatGPT imagines science. Additionally, we used computational content analysis to identify differences in words and sentences per interview, lexical readability and diversity, and academic vocabulary.

Results

RQ1: The answers generated by ChatGPT predominantly associate science with STEM fields and a positivist-empiricist approach. ChatGPT, overall, takes a positive stance toward science, stating that science "can generally be trusted as a reliable and systematic method", but also mentioning that "it

is not infallible". The relationship between science and the public is described as important, but science communication is imagined primarily as a one-way transfer of knowledge. In quantitative terms, ChatGPT's responses are very difficult to read with an average Flesch Reading Ease Score (FRES) of 20.3.

RQ2: While there are strong similarities in substance across user profiles, the qualitative analysis reveals differences in the degree of user-personalized answers. These are more pronounced for the *Disengaged* and *Critically Interested*. For example, for the latter segment, ChatGPT generates answers stating that "your interests and beliefs about science make you well-suited for a nuanced understanding of this issue". The quantitative analysis reveals minor differences: Figure 1 shows that,on average, the blank profile receives longer responses than the four user profiles. Responses for the *Sciencephiles* are on average the most difficult to understand (M = 19.9) and contain more academic vocabulary (M = 61.3%). In contrast, the answers for the *Disengaged* are the easiest to read (M = 21) and contain least academic vocabulary (M = 60%). References in the responses are most frequently provided for the *Sciencephiles* (n = 147) and least for the *Disengaged* (n = 119).

RQ3: The comparison of the GPT variants shows pronounced differences in the degree of personalization: GPT-4 takes the user information provided more into account compared to GPT-3.5. For GPT-4, the strongest form of personalization is found for the *Critically Interested*, followed by *Sciencephiles,* giving answers beginning with e.g., "as someone who values being informed about science, you...". Answers are least targeted to the *Disengaged*. The quantitative analysis reveals that GPT-4 responses are shorter on average (Fig 1) and easier to read (four profiles: M = 23.5) compared to GPT-3.5 (four profiles: M = 17.3). GPT-4 responses also contain less academic vocabulary than GPT-3.5. GPT-4 provides more references overall (n = 416) than GPT-3.5 (n = 248) and does so more frequently even when not explicitly asked (n = 116 vs. n = 13).

Conclusion

This is the first study to analyze how ChatGPT portrays science and targets answers to different audiences, which could potentially have detrimental effects on those with skeptical attitudes towards science. Of course, this study is limited by analyzing only one AI tool and prompting in English, so future research is needed to compare the agency enacted by other AI tools and differences across languages, as well as real human-AI interactions and trust in AI tools.

References

Bulian, J., Schäfer, M. S., Amini, A., Lam, H., Ciaramita, M., Gaiarin, B., Huebscher, M. C., Buck, C., Mede, N., Leippold, M., & Strauss, N. (2023). *Assessing Large Language Models on ClimateInformation* [Online]. Available at http://arxiv.org/pdf/2310.02932v1

Chen, K., Shao, A., Burapacheep, J., &Li, Y. (2022). *How GPT-3 responds to different publics on climate change and Black Lives Matter: A critical appraisal of equity in conversational AI*. Available at http://arxiv.org/pdf/2209.13627v2

Corless, V. (2023, February 21). ChatGPT is making waves in the scientific literature. *Advanced Science News*. Retrieved December 3, 2023, from https://www.advancedsciencenews.com/where-and-how-should-chatgpt-be-used-in-the- scientific-literature/

Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koohang, A., Raghavan, V., Ahuja, M., Albanna, H., Albashrawi, M. A., Al-Busaidi, A. S., Balakrishnan, J.,Barlette, Y., Basu, S., Bose, I., Brooks, L., Buhalis, D., … Wright, R. (2023). Opinion Paper: "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management*, *71*, 102642. https://doi.org/10.1016/j.ijinfomgt.2023.102642

Fecher, B., Hebing, M., Laufer, M., Pohle, J., & Sofsky, F. (2023). *Friend or Foe? Exploring theImplications of Large Language Models on the Science.* https://doi.org/10.48550/arXiv.2306.09928

Flesch, R. (1948). A new readability yardstick. *Journal of Applied Psychology, 32*(3), 221-233. https://doi.org/10.1037/h0057532

Hosseini, M., Resnik, D. B., & Holmes, K. (2023). The ethics of disclosing the use of artificial intelligence tools in writing scholarly manuscripts. *Research Ethics*, *19*(4), 449–465.https://doi.org/10.1177/17470161231180449

Koch, C., Saner, M., Schäfer, M. S., Herrmann-Giovanelli, I., & Metag, J. (2019). "Space means Science, unless it's about Star Wars": A qualitative assessment of science communicationaudience segments. *Public Understanding of*

Science, *29*(2), 157-175. https://doi.org/10.1177/0963662519881938

OpenAI (2023). *Custom instructions for ChatGPT*. Retrieved December 3, 2023, fromhttps://openai.com/blog/custom-instructions-for-chatgpt

Schäfer, M. S. (2023). The Notorious GPT: science communication in the age of artificial intelligence. *Journal of Science Communication*, *22*(2), Y02.

Schäfer, M. S., Füchslin, T., Metag, J., Kristiansen, S., & Rauchfleisch, A. (2018). The different audiences of science communication: A segmentation analysis of the Swiss population's perceptions of science and their information and media use patterns. *Public Understandingof Science*, *27*(7), 836-856. https://doi.org/10.1177/0963662517752886

Teubner, T., Flath, C. M., Weinhardt, C., van der Aalst, W., & Hinz, O. (2023). Welcome to the Era of ChatGPT et al. *Business & Information Systems Engineering*, *65*(2), 95–101. https://doi.org/10.1007/s12599-023-00795-x

Wissenschaft im Dialog/Kantar (2023). *Wissenschaftsbarometer 2023*. https://www.wissenschaft-im-dialog.de/projekte/wissenschaftsbarometer/wissenschaftsbarometer-2023/

How well can ChatGPT replace human coders in quantitative content analysis? A case study

Clarissa Elisabeth Hohenwalde, Melanie Leidecker-Sandmann, Nikolai Promies, Markus Lehmkuhl

Introduction

The forthcoming annual conference of the Science Communication Division of DGPuK explores the impact of generative AI on science communication research. This paper contributes to the discourse regarding "the impact of AI on the methodology and methods of science communication research" by addressing the question:

> *RQ: How well can ChatGPT replace human coders in the quantitative content analysis task of identifying and categorizing actor groups?*

In communication science, content analysis is one of the most commonly used methods. Traditionally, it entails training human coders to classify texts through an extensive and costly iterative process based on detailed codebooks (Brosius et al., 2022). Automated coding with large language models (LLMs) such as ChatGPT could offer a cost-effective alternative, making studies more affordable and scalable for larger data samples, as well as enabling real-time analysis.

Our focus lies on the identification and categorization of actor groups in media coverage. Aligned with theoretical frameworks like Habermas (1996), which stresses the need for inclusive public debates on relevant issues, we leverage ChatGPT to classify actor groups such as "science", "advocacy", "politics" and "other". This assists in addressing questions such as who participates in public discourse, ensuring diverse representation, and understanding the evolution of visible actor groups over time.

State of Research

While scholars have suggested potential use cases for LLMs (Argyle et al., 2023; Binz & Schulz, 2023; Stokel-Walker & Van Noorden, 2023), their application in social science research remains limited. Our study aims to address this gap. Pioneer studies in other disciplines have shown promising results in text annotation (Gilardi et al., 2023; Ornstein et al., 2022; Törnberg, 2023; Wu et al., 2023; Zambrano et al., 2023). However, a literature review by Ollion et al. (2023) suggests that while LLMs like ChatGPT often match human performance, their effectiveness in coding tasks is partial and varies based on material, language, and prompt.

Method

We analyze a German media sample of science related articles, previously coded using named entity recognition (NER) combined with manual coding (Buz et al., 2022). We compare Chat-GPT's performance against these codes in four settings (figure 1 and chapter 6.2): First, we employ detailed codebooks, initially designed for human research assistants, to prompt ChatGPT (version gpt-3.5-turbo) in a zero-shot setting for identifying named entities and assigning them to actor groups. In a second approach, we optimize prompts by using few-shot learning principles and supplying the model with category keywords instead of exhaustive definitions. Thirdly, we integrate ChatGPT into a NERC pipeline, involving NER application, article contextualization, and optimized prompt coding. In the last setup, we assess whether the newer gpt-4-turbo version enhances results obtained by the NERC pipeline with gpt-3.5-turbo.

The optimized prompt's development phase involved a small subsample, while the final analysis used a distinct sample of 200 articles.

For settings 1, 2, and 3, we employed the gpt-3.5-turbo version, and for setting 4, we utilized gpt-4-turbo. To ensure a relatively deterministic result, we set temperature = 0 and topp = 0.5.

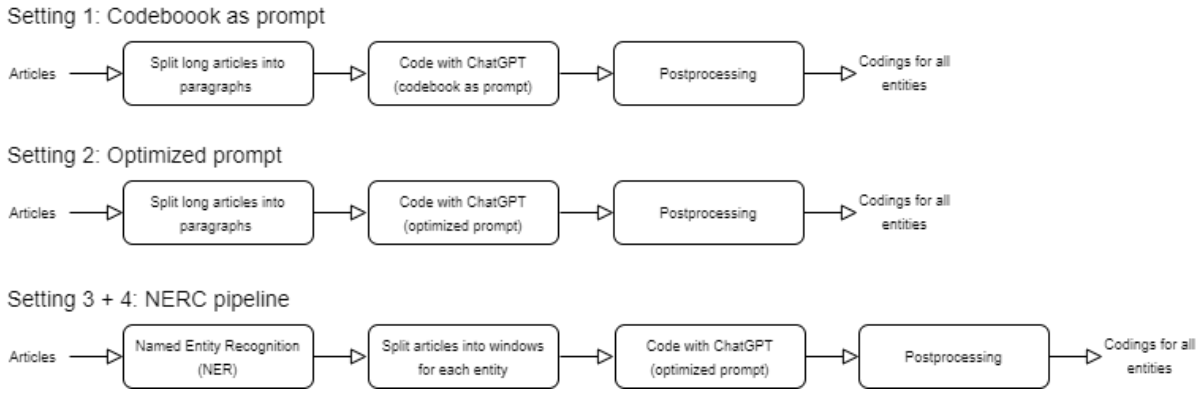


Figure 1: Procedures used for comparison of ChatGPT's performance

Results and discussion

In a multi-class classification problem with imbalanced class distribution, we can evaluate the classifier performance using the F-1 score. The proposed NERC pipeline (table 1) exhibits the best performance, further enhanced by employing the gpt-4-turbo model, which benefits from an expanded vocabulary, aiding in recognizing linguistic nuances.

	Codebook	Optimized prompt	NERC pipeline (gpt-3.5-turbo)	NERC pipeline (gpt-4-turbo)
Science	.49	.50	.90	.94
Advocacy	.00	.41	.77	.84
Politics	.47	.42	.91	.91
Others	.08	.09	.56	.63

Table 1: F1-scores for different settings by class

While "science" and "politics" are accurately identified, challenges arise in classifying "advocacy" and "other" actor groups. Upon closer examination of the confusion matrices, we find that advocacy actors are most commonly misclassified as science (table 2), primarily due to scientists in business being categorized as "science", despite the codebook considering them influenced by partial interests. Since LLMs operate based on the statistical probability of words occurring within a similar context and more science than advocacy related terms co-occur with these actors, teaching ChatGPT to discriminate between scientists in academia and scientist in business remains an issue.

		PREDICTED			
		Science	Advocacy	Politics	Other
EXPECTED	Science	236	3	2	1
	Advocacy	14	53	0	0
	Politics	2	1	68	3
	Other	8	5	1	15

Table 2: Confusion matrix for the NERC pipeline (gpt-4-turbo), displaying the number of entities for each category

Considering F-1 scores, it is crucial to acknowledge that ChatGPT's coding is benchmarked against manual coding, treated as the gold standard. However, the gold standard itself is not error-free, with an intercoder reliability of αKrippendorf = .77 indicating moderate human coder agreement.

Conclusions

In summary, ChatGPT proves effective in substituting human coders for tasks like classifying scientific or political actors. This could prove particularly beneficial when handling extensive or

multilingual materials requiring timely coding. For more intricate tasks and for the coding of advocacy and other groups, we recommend using LLMs as supportive tools for human coders, expediting and reducing content analysis costs.

Our findings indicate that traditional codebook construction methods may be insufficient for LLM classification tasks. Researchers are advised to adopt prompt engineering techniques, such as providing examples, breaking tasks into subtasks, and constructing specific training pipelines. Further research is essential to identify optimal settings for ChatGPT use and determine pipeline models and prompting strategies that align with the needs of science communication research.

References

Argyle, L. P., Busby, E. C., Fulda, N., Gubler, J. R., Rytting, C., & Wingate, D. (2023). Out of one, many: Using language models to simulate human samples. Political Analysis, 31(3), 337–351.

Binz, M., & Schulz, E. (2023). Using cognitive psychology to understand GPT-3. Proceedings of the National Academy of Sciences, 120(6), e2218523120.

Brosius, H.-B., Haas, A., & Unkel, J. (2022). Methoden der empirischen Kommunikations- forschung (8th ed.). Wiesbaden: Springer.

Buz, C., Promies, N., Kohler, S., & Lehmkuhl, M. (2022). Validierung von ner-verfahren zur automatisierten identifikation von akteuren in deutschsprachigen journalistischen texten. SCM Studies in Communication and Media, 10(4), 590–627.

Gilardi, F., Alizadeh, M., & Kubli, M. (2023). ChatGPT outperforms crowd-workers for text-annotation tasks. arXiv preprint arXiv:2303.15056.

Habermas, J. (1996). Between facts and norms: Contributions to a discourse theory of law and democracy. John Wiley & Sons.

Ollion, E., Shen, R., Macanovic, A., & Chatelain, A. (2023). ChatGPT for text annotation? mind the hype! SocArXiv preprint.

Ornstein, J. T., Blasingame, E. N., & Truscott, J. S. (2022). How to train your stochastic parrot: Large language models for political texts (Tech. Rep.). Working Paper.

Stokel-Walker, C., & Van Noorden, R. (2023). What ChatGPT and generative AI mean for science. Nature, 614(7947), 214–216.

Törnberg, P. (2023). ChatGPT-4 outperforms experts and crowd workers in annotating political twitter messages with zero-shot learning. arXiv preprint arXiv:2304.06588.

Wu, P. Y., Tucker, J. A., Nagler, J., & Messing, S. (2023). Large language models can be used to estimate the ideologies of politicians in a zero-shot learning setting. arXiv preprint arXiv:2303.12057.

Zambrano, A. F., Liu, X., Barany, A., Baker, R. S., Kim, J., & Nasiar, N. (2023). From nCoder to ChatGPT: From automated coding to refining human coding. In International conference on quantitative ethnography (pp. 470–485).

The impact of transparency: A qualitative investigation of LLM-based chatbots in science-related information search

Esther Greussing, Evelyn Jonas, Monika Taddicken

Since the introduction of ChatGPT, generative AI has emerged as a transformative force, shaping how users navigate and access information – including information on science and health (Choudhury & Shamszare, 2023). Chatbots based on large language models (LLMs) can make scientific topics understandable for diverse audiences (Schäfer, 2023), but concerns exist about their potential to convincingly present inaccurate information (Spitale et al., 2023), fostering repeated calls for transparency in these systems (Andrada, Clowes, & Smart, 2023).

While ample research exists on explainable AI in advice and decision-making (e.g., Jiang, Kahai, & Yang, 2022), our current understanding of how users perceive transparency in LLM-based chatbots that provide science-related information remains limited. Given the pivotal role of transparency in science communication (Olesk et al., 2021), this study compares user engagement with two prominent chatbots, Open AI's ChatGPT (version 3.5) and Microsoft's Bing Chat. This comparison is instrumental for uncovering nuanced differences in the way individuals use and perceive LLM-based chatbots in the context of science-related information search, considering the distinct features they offer. Most notably, while ChatGPT is opaque, Bing Chat indicates transparency by displaying the sources of the content presented through linked footnotes, and by visibly translating user queries into search phrases.

This study is guided by two research questions:
RQ1: How do users perceive transparency of LLM-based chatbots when seeking information on science-related topics?
RQ2: How does the transparency of LLM-based chatbots impact users' evaluation of the information presented?

The research questions are addressed through qualitative interviews with N=34 participants from the general public (52.9 % female, mean age=38.3 years (SD=17.4), 58.8 % have prior experiences with chatbots). As part of the interview, each participant engaged first with ChatGPT, then with Bing Chat to search for information on two science-related issues: sustainable aviation and juice cleanse. The interviews were conducted by the authors from June to August 2023 on-site at a German university (average duration approx. two hours). Verbatim transcriptions of the interviews were analyzed using a flexible coding approach (Deterding & Waters, 2021).

Our findings shows that in general, participants enjoyed conversing with LLM-based chatbots for basic, factual information. Addressing our research questions, a multifaceted perception of transparency unfolded. Most prominently, participants reflected on the unclear nature of the datasets generative AI operates on, and wished to receive source references for the information generated by ChatGPT. Notably, however, for some participants, only the visibility of sources on Bing Chat prompted them to reflect on the unclear informational foundation of ChatGPT. While Bing Chat did provide source references, participants became aware that these sources span a wide spectrum – from journalistic outlets and Wikipedia to corporate websites. For some participants, this spectrum raised doubts about the credibility of the content presented, and they did not consider Bing Chat useful for research on scientific topics anymore.

Another facet of transparency surfaced when participants noted that Bing Chat displays the translation of user input into a search phrase, resembling the functionality of a search engine. While some participants felt confirmed by this, believing that the chatbots "gather something from the internet" (interviewee 20), others felt that their own phrasing of a prompt loses significance, leading them to question the added value of the dialogical aspect of AI in science-related information search.

Transparency regarding the uncertainties of what is known about an issue was another facet that emerged. Participants emphasized that the chatbots do not make decisions or offer recommendations but present the advantages and disadvantages of a given subject. In particular, ChatGPT was discussed for its ability to generate comprehensive responses that weigh different perspectives, even without explicit prompting. Some participants viewed this positively, recognizing that they might have missed arguments when using a conventional search engine, particularly on complex or ambiguous issues. For them, the overview of advantages and disadvantages enhanced the perceived transparency of the system. On the other hand, Bing Chat was discussed for its concise responses and for clearly stating that information on certain topics is not available – also enhancing its perceived transparency. Notably, while ChatGPT's cohesive responses that include aspects that were not explicitly requested by the user were appreciated, there was concern that it may create an illusion of greater knowledge than is actually available. For some participants, however, Bing Chat's concise answers contained too few clues to obtain the desired information through well- chosen follow-up prompts.

In essence, this study emphasizes the critical importance of the multifaceted perception and role of transparency in generative AI when it comes to retrieving science- related information. Our analysis shows that differences in transparency among LLM-based chatbots emerge as noteworthy factors influencing user engagement. Most notably, regarding source information, it appears that ChatGPT's lack of transparency might not diminish its convincing power; rather, it seems to enhance it, contributing to a more effortless and compelling experience. Moreover, our data indicate that overall, our participants hold a rather competent view on generative AI for searching for science-related information, yet this may also reflect an awareness of societal expectations potentially inherent in a face-to-face interview.

References

Andrada, G., Clowes, R. W., & Smart, P. R. (2023). Varieties of transparency: Exploring agency within AI systems. AI & Society, 38(4), 1321-1331. https://doi.org/10.1007/s00146-021-01326-6

Choudhury, A. & Shamszare, H. (2023). Investigating the Impact of User Trust on the Adoption and Use of ChatGPT: Survey Analysis. Journal of Medical Internet Research, 25, e47184. https://doi.org/10.2196/47184

Deterding, N. M. & Waters, M. C. (2021). Flexible Coding of In-depth Interviews: A Twenty- first-century Approach. Sociological Methods & Research, 50(2), 708–739. https://doi.org/10.1177/0049124118799377

Jiang, J., Kahai, S., & Yang, M. (2022). Who needs explanation and when? Juggling explainable AI and user epistemic uncertainty. International Journal of Human-Computer Studies, 165, 102839. https://doi.org/10.1016/j.ijhcs.2022.102839

Olesk, A., Renser, B., Bell, L., Fornetti, A., Franks, S., Mannino, I., ... & Zollo, F. (2021). Quality indicators for science communication: results from a collaborative concept mapping exercise. Journal of Science Communication, 20(03). https://doi.org/10.22323/2.20030206

Schäfer, M. S. (2023). The Notorious GPT: science communication in the age of artificial intelligence. Journal of Science Communication, 22(2), Y02. https://doi.org/10.22323/2.22020402

Spitale, G., Biller-Andorno, N., & Germani, F. (2023). AI model GPT-3 (dis)informs us better than humans. ScienceAdvances, 9(26), eadh1850. https://doi.org/10.1126/sciadv.adh1850

Parallel Panel VI: Open panel: SHAPING PUBLIC DISCOURSE: SCIENTISTS AS COMMUNICATORS

Friday, June 7, 11:05 – 12:25
Chair: Marina Joubert
Room: HG E 1.1

Centralized vs. decentralized science communication in universities: Differences in the professional role identities of university communicators in various organizational contexts

Lennart Banse, Kaija Biermann, Monika Taddicken

In the evolving science communication (scicom) landscape, direct interaction between science and the public has become more important. Scientific organizations play a crucial roleat the interface between science and society (Schwetje et al., 2020). The decline in science journalism resources enhances the relevance of actors in Higher Education and Research Institutions (HERIs) in the broader scicom system (Barel-Ben David et al., 2020). To fulfill their scicom function, HERIs are bolstering their communication structures and staffing (Fürst et al., 2022). This development prompts the question of how communicative roles are distributed within HERIs, particularly as new actors like individual researchers or research departments are encouraged to engage in scicom.

Initial studies have quantitatively explored the role conceptions of HERI communicators in central press offices (e.g., Volk et al., 2023). However, there's a gap in understanding their roles in internal and external mediation activities across organizations andsystems (e.g., science, HERIs, society), commonly referred to as "boundary spanning" (Rödder, 2020). Particularly, research on professional communicators at decentralized units like faculties and research centers is scarce (Ojeda-Romano et al., 2022). With their stronger focus on scicom activities, apparently surpassing those in central press offices (Entradas et al.,2023), it's crucial to examine how they perceive their roles and interactions with other communicators within HERIs. Hence, we ask:

RQ: How do the role conceptions of central and decentral communicators at German HERIs differ in regard to their basic functions in scicom?

From April to July 2023, we conducted guided interviews with 29 HERI communicators from all HERIs in a large German federal state. Employing purposive sampling (Etikan et al., 2016), we ensured representation from each HERI (N = 10), with a focus on both central (n = 16) and decentral (n = 13) communicators at different career stages and positions. The sample included decentral communicators from various scicom structures like faculties, research centers, Clusters of Excellence, and medical school clinics.

The guide focused on professional self-perception, communication goals, scicom practices, autonomy, and factors influencing work and role understanding, aligning with a theoretical role framework on subjective role meanings (Kaplan & Garner, 2017). We used thematic analysis (Braun & Clarke, 2006) in MAXQDA, employing a deductive-inductive approach. Initial categories were established from the interview guide, followed by inductive subcategory formation. Two coders coded the main categories, regularly reviewing and refining subcategories. One coder then applied the final codebook to the entire dataset.

Our analysis shows that the interviewees have similarities in their basic role conceptions. They argue that they are not strictly science communicators in the sense of scientific content creators. Instead, a majority see their role as a structural boundary role in transmitting existing

information between science, scientific organizations, and publics. Most interviewees perceive themselves in certain sub-roles of this boundary function such as mediators between internal and external stakeholders, translators between scientific and public language and logics, and enablers supporting others in communicating directly with publics.

There are also differences between central and decentral communicators in the concrete implementation of these structural-functional roles. Central communicators align closely with the entire university and its leadership, with stronger connections to the public and media systems. This group divides into leadership, focusing on relationships with the university management as well as basic strategic planning at university level, and communication specialists, who see themselves in functionally differentiated expert roles for specific academic subject areas or communication channels. Decentral communicators identify less with the entire university, having closer ties to their sub-organizations. They form stronger networks within the scientific system but have less direct contact with public and media, often restricted by guidelines of the central communication office. Typically, as the sole communicator in their sub-organization, they assume a generalist role, handling a wide array of tasks, balancing service-provider functions and strategic advising roles in one person.

Contrary to prevalent research (Entradas et al., 2023), our study found no clear linkage between orientation on "pure" scicom ("Type 1" in Weingart & Joubert, 2019) vs. marketing/reputation communication ("Type 2") based on a central/decentral division. Patterns emerged when examining specific decentral structures: communicators in Clusters of Excellence were more aligned with scicom, while those in faculties and medical school clinics tended more towards marketing. Together with the fact that some central communicators do not see themselves in marketing roles at all, this indicates that the distinction of central vs. decentral does not necessarily determine a "pure" scicom orientation. Rather, specific organizational contexts seem to shape HERI communicators' role conceptions.

Overall, our study provides insights into how communicators at HERIs interpret their roles at the interface of science, organizational structures, and society. By elucidating decentral communicators' role conceptions in-depth for the first time, our research fills a significant research gap in scicom research. Contrary to current research, we found that the central/decentral divide appears to be too simplistic. Thus, our study offers a refined understanding of scicom from scientific organizations to the public. Future research should further explore the relationship between professional roles and organizational contexts, potentially guiding improvements in scicom within HERIs.

References

Barel-Ben David Y, Garty ES, Baram-Tsabari A (2020) Can scientists fill the science journalism void? Online public engagement with science stories authored by scientists. PLoS ONE, 15(1): e0222250. https://doi.org/10.1371/journal.pone.0222250

Braun, V. & Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Entradas, M., Marcinkowski, F., Bauer, M. W. & Pellegrini, G. (2023). University central offices are moving away from doing towards facilitating scicom: a European cross- comparison. PLOS ONE, 18(10), e0290504. https://doi.org/10.1371/journal.pone.0290504

Etikan, I., Musa, S. A., Alkassim, R. S., & others. (2016). Comparison of convenience sampling and purposive sampling. American Journal of Theoretical and Applied Statistics, 5(1)

Fürst, S., Volk, S. C., Schäfer, M. S., Vogler, D. & Sörensen, I. (2022). Assessing Changes in the public communication of higher education institutions: A survey of leaders of Swiss universities and colleges. Studies in communication sciences. https://doi.org/10.24434/j.scoms.2022.03.3489

Ojeda-Romano, G., Fernández-Marcial, V., Wilkinson, C. & Stengler, A. E. (2021). Organisational Forms of Scicom: the UK and Spanish European Higher Education systems as paradigms. Higher Education, 84(4), 801–825. https://doi.org/10.1007/s10734-021-00801-9

Rödder, S. (2020). Organisation matters: Towards an organisational sociology of science communication. Journal of Communication Management, 24(3), 169–188. https://doi.org/10.1108/jcom-06-2019-0093

Schwetje, T., Hauser, C., Böschen, S., & Leßmöllmann, A. (2020). Communicating science in higher education and research institutions. Journal of Communication Management, 24(3), 189–205. https://doi.org/10.1108/jcom-

06-2019-0094

Volk, S. C., Vogler, D., Fürst, S., Schäfer, M. S. & Sörensen, I. (2023). Role Conceptions of University Communicators: A segmentation analysis of communication practitioners in higher education institutions. Public Relations Review, 49(4), 102339. https://doi.org/10.1016/j.pubrev.2023.102339

Weingart, P. & Joubert, M. (2019). The conflation of motives of science communication — causes, consequences, remedies. JCOM, journal of science communication, 18(03), Y01. https://doi.org/10.22323/2.18030401

Between brokerage and advocacy – the role of organisations as science communicators in the COVID-19 pandemic

Simone Rödder, Anna-Lena Oltersdorf

The recent "organisational turn" (Schäfer and Fähnrich 2020) in science communication research acknowledges that science communication demands are typically dealt with at the organisational level and that organisations shape the social status of science and knowledge to a degree hitherto understudied (Rödder 2020). A case in point is the COVID-19 pandemic where in addition to numerous visible scientists (Joubert et al. 2023), several organisations took on central roles in communicating the science of the pandemic. Chiefly among them – in the German context – were the German National Academy of Sciences Leopoldina and the Science Media Centre Germany (SMC).

The Leopoldina engages in public and policy debates by bringing together the views of various experts from different disciplines. Since the beginning of the pandemic, the organisation has published ten so-called "ad-hoc" statements and one regular statement with recommendations, andpartly also demands, for political measures in the fight against COVID-19. Politics and the media have taken up and, in some cases, controversially discussed these statements. The SMC is a rather new addition to the science communication landscapes of many countries aiming to "inject the voice of science" into public and policy debates (Rödder 2015). It provides free 24/7 news services to journalists who report about scientific topics and is also instrumental in finding relevant experts. During the COVID-19 outbreak, the SMC Germany has established several new formats, including virtual press briefings where journalists could interact with scientific experts and published written summaries of new scientific publications including short assessments by relevant scientific experts.

Both organisations thus operate at the boundaries between science, politics and the media, and their contributions to science communication during the COVID-19 pandemic resonated broadly, such as in the news media. Moreover, both organisational actors possibly decouple the communication of and trust in science from individual scientists and unite it under the roof of their respective organisations. This calls attention to these organisations' strengths in crisis communication as well as possible (unintended) consequences, such as the definition of criteria applied to expert selection, dealing with various, and possibly conflicting, disciplinary knowledge bases as well as the handling of more than one scientific consensus that may result from debates in different expert committees. In academic research, however, these organisational actors are understudied so far, with a few current exceptions for the SMC Germany (Rödder 2020; Broer 2020; Broer and Pröschel 2022) and the Leopoldina (Beck and Nardmann 2021). The relevance of this research is emphasized by controversial public debates especially regarding the role of the Leopoldina in the context of the COVID-19 crisis (e.g., Hirschi 2021). We aim to contribute to the growing field of organisational science communication studies by asking how exactly the Leopoldina and the SMC Germany engaged in science communication during the COVID-19 crisis and what difference they made in the public debate from the perspective of staff members, researchers, and media professionals.

To address this research question, we combine three methodological approaches:

1. First, we qualitatively analyse the media presence of the Leopoldina in the period from March 2020 to February 2022 in the print and online editions of six influential German-language media (BILD, WELT, SZ, SPIEGEL, ZEIT and FAZ). We focus on the question of how the Leopoldina and its statements were presented in the media – e.g., concerning the role they played for pandemic policies.
2. Second, the content of the Leopoldina's statements as well as the products, which the SMC Germany offered to journalists during the pandemic, are qualitatively analysed – e.g., regarding the presentation of scientific evidence.
3. Third, we conduct problem-oriented expert interviews with researchers, media

professionals and staff members involved with one or both organisations, focused on the requirements of science communication during the pandemic.

This mixed-methods empirical study aims to gain insights into the conditions that facilitated or impeded science communication about COVID-19. Moreover, it aims to compare organisational science communication with the communication by individual researchers who are visible in the media and active in scientific policy advice.

Our initial results indicate how existing typologies on the role of individual scientists in politics could be adapted to organisational actors. Speaking with Pielke (2007), it seems almost impossible for organisational actors to act as *honest brokers* – the ideal type in his typology – in a situation of crisis. This is due to the temporal restrictions in a crisis, calling for fast political decisions while the science is uncertain, as well as the numerous requirements connected to the role of *honest broker*. By combining results from the media and document analysis with the interview data in the further course of the study, we strive to assess how the establishment of organisational actors affects the fundamental communicative challenges between science, media and politics.

In our talk, we will present initial results of our study and place them in the context of debates on challenges of science communication in the COVID-19 pandemic specifically. Thereby, we strive to contribute to the timely as well as challenging scholarly reflection of the COVID-19 crisis. More generally, we will discuss the results in the context of debates on the role of organisations in science communication and thereby contribute to this field of research.

References

Beck, Silke; Nardmann, Julian (2021): Wissenschaftliche Rückendeckung für politische Alternativlosigkeit? Kontroversen um Expertisen in der deutschen Corona-Politik. In Sebastian M. Büttner, Thomas Laux (Eds.): Umstrittene Expertise. Zur Wissensproblematik der Politik. 1. Auflage.Baden-Baden: Nomos (Leviathan. Sonderband, 38), pp. 187–214.

Broer, Irene (2020): Rapid reaction: ethnographic insights into the Science Media Center and itsresponse to the COVID-19 outbreak. In *Journal of Science Communication* 19 (05), A08.

Broer, Irene; Pröschel, Louisa (2022): Knowledge broker, trust broker, value broker: The roles of theScience Media Center during the COVID-19 pandemic. In *Studies in Communication Sciences* 22 (1), pp. 101–118.

Hirschi, Caspar (2021): Expertise in der Krise. Zur Totalisierung der Expertenrolle in der Euro-, Klima-und Coronakrise. In Sebastian M. Büttner, Thomas Laux (Eds.): Umstrittene Expertise. Zur Wissensproblematik der Politik. 1. Auflage. Baden-Baden: Nomos (Leviathan. Sonderband, 38), pp. 159–186.

Joubert, Marina; Guenther, Lars; Metcalfe, Jenni; Riedlinger, Michelle; Chakraborty, Anwesha;Gascoigne, Toss et al. (2023): `Pandem-icons' — exploring the characteristics of highly visible scientists during the Covid-19 pandemic. In *Journal of Science Communication* 22 (01), A04.

Pielke, Roger A. (2007): The honest broker. Making sense of science in policy and politics.Cambridge: Cambridge University Press.

Rödder, Simone (2015): Science Media Centres and public policy. In *Science and Public Policy* 42 (3),pp. 387–400.

Rödder, Simone (2020): Organisation matters: towards an organisational sociology of sciencecommunication. In *Journal of Communication Management* 24 (3), pp. 169–188.

Schäfer, Mike S.; Fähnrich, Birte (2020): Communicating science in organizational contexts: towardan “organizational turn” in science communication research. In *Journal of Communication Management* 24 (3), pp. 137–154.

From labs to politics: A mixed-methods study on researchers' participation in political debates

Nils Bienzeisler, Senja Post

Challenges, such as the loss of biodiversity and the COVID-19 pandemic, demonstrate how difficult it is to align science with politics. This alignment becomes particularly complex when uncertainty and urgency intersect (Scheufele, 2014). During the pandemic, the engagement of scientists has provoked pivotal debates about the proper involvement of researchers in political decision-making (e.g., Bogner, 2021). A central claim is that scientists have transformed into political actors harboring their own ambitions in political controversies.To the best of our knowledge, there has been limited exploration of this claim, with only a few exceptions (e.g., Alinejad & van Dijck, 2022; Besley & Nisbet, 2013; Post & Ramirez, 2018). Therefore, the purpose of this study is to explore if scientists dissociate themselves from political decision making or neglect this distinction. In this pursuit, we employ a mixed-methods design combining surveys and qualitative interviews with scientists engaged in pandemic and biodiversity research.

Background

Scientists are judged as especially trustworthy when evaluating matters (Kotcher et al., 2017). This leads researchers to distinguish themselves from other actors and requires them to embrace normative assumptions, such as being disengaged, impartial, and knowledgeable (Jasanoff, 1994, p. 12). Researchers who view their knowledge as irrelevant are probably less likely to share their expertise. Furthermore, to earn trust, they should aim for independence and practice restraint (cf. Lupia, 2013). However, these expectations conflict with scientists becoming visible political actors, challenging the traditional boundaries between scientific inquiry and political advocacy (Biermann et al., 2023). This contradiction prompts us to explore how scientists navigate their roles in science and politics. If our assumptions hold true, researchers could be differentiated into latent groups based on their dissociation from political controversies.

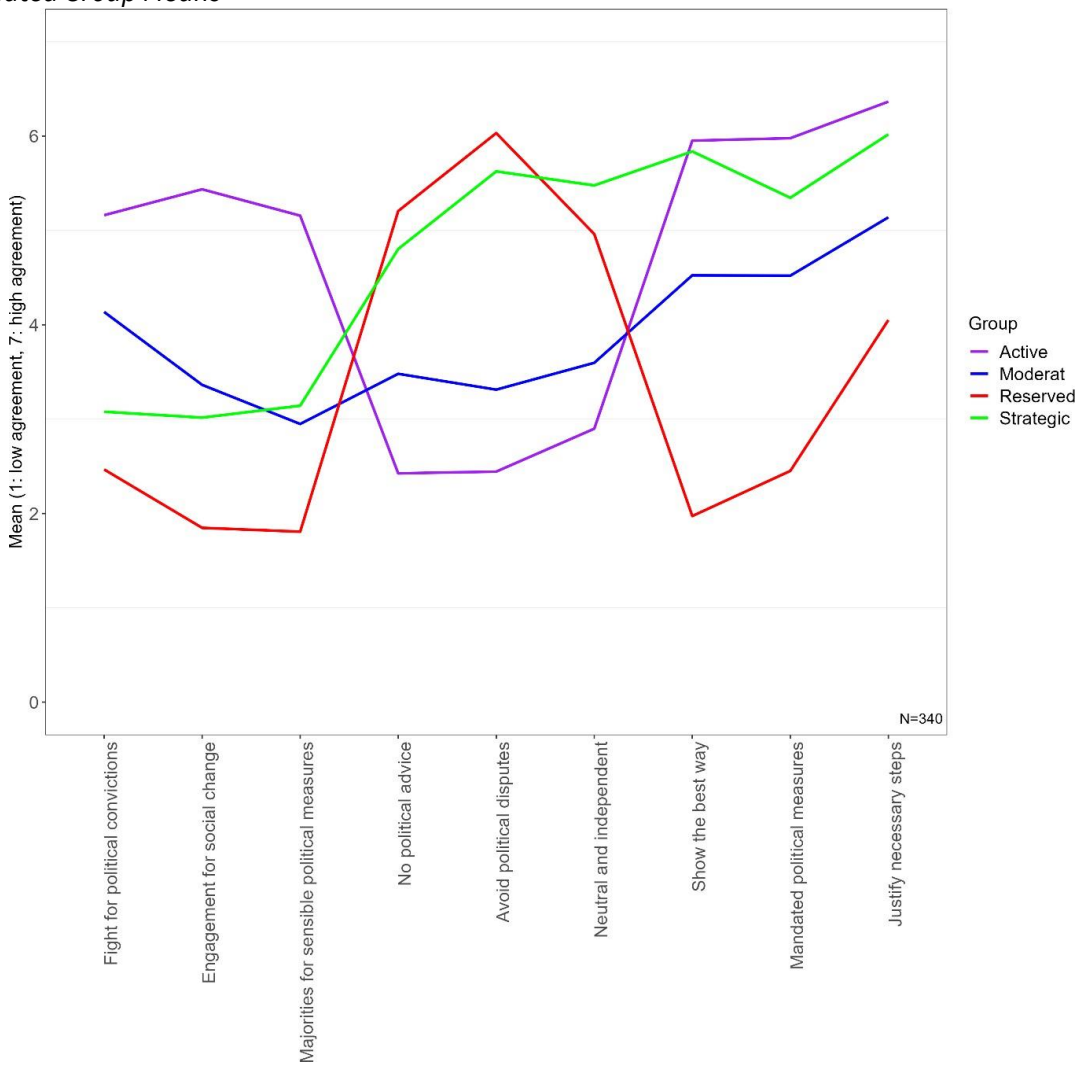
Methods

We surveyed N = 349 scientists from Germany who published medical (RR = 12.6%) or ecological papers (RR = 28.7%). Participants were asked to define their concept of expertise, rating a) their involvement b) neutrality, and c) epistemic authority in political debates. Subsequently, we interviewed a subsample of n = 24 participants that agreed to a follow-up interview in depth. The interviews were coded to align with the survey dimensions. We applied a latent profile analysis (LPA) with model specifications validated by sensitivity analyses and theoretical alignment evaluation, ensuring robust group identification. Last, we linked interviewees with class assignments.

Results

The initial qualitative results show a diverse range of perspectives among scientists in pandemic and biodiversity research. Many viewed neutralities as key to good science. They generally agreed on using their expertise to guide societal debates. However, there was less agreement on directly engaging in political activism, indicating a tendency to maintain a distance from political controversies.

Figure 1:
Estimated Group Means



Note: The figure displays estimated group means for involvement (first three items), neutrality(middle three), and authority (last three); item texts in the graph are abbreviated.

The LPA shows that scientists can be grouped based on their concept of expertise. The model differentiating four groups of scientists was chosen as optimal (AIC = 11204.40, *BIC* = 11388.19, Entropy ≥ 0.80) separating a moderate group (36.5%) showing average agreement to items measuring involvement, neutrality, and epistemic authority, a reserved group (7.6%) preferring non-involvement in political controversies, a strategic group (30.9%) high in agreement on epistemic authority and neutrality, and an active group (25.0%) showing strong involvement (cf. Figure 1). A multinomial regression analysis elucidated membership to groups with distinct characteristics: the active group was more progressive and saw advocacy as their task, the strategic group comprised individuals with higher seniority, and the reserved group was predominantly female (cf. Table 1).

Table 1:
Multinomial Logistic Regression of Group Membership

	Reserved	Strategic	Active
Progressivity	.821	.996	1.645**
External Efficacy	1.052	.931	1.201
Internal Efficacy	.790	1.146	1.175
Task to Influence	.809	1.068	1.455**
Task to Research	1.132	1.513*	1.164
Seniority	.563	3.154***	1.860
Female	2.527*	1.695	.586
Biodiversity Research	1.104	.568	.778

Note: N = 340, Nagelkerke Pseudo R^2 = .244. The table displays Exp(B) estimates of a multinomial logistic regression model for group membership. The reference category is the moderate profile; *** $p < .001$, ** $p < .01$, * $p < .05$.

The subsequent qualitative analysis corroborated the group characteristics identified in the LPA, with one example being the active group, whose members were more often engaged in political actions and more likely to utilize their authority for advocacy than others, aligning with their profile in the quantitative analysis. Additionally, the five interviewed members of the active group perceived activism as a viable option for enacting change.

Conclusion

The study demonstrates that scientists exhibit varied approaches to engaging with political debates. Scientists fulfill various latent roles in political controversies, ranging from reserved observers to activist voices. The insights gained highlight the importance of recognizing and fostering scientists' awareness of their potential impact on political decision-making. Practically, this calls for enhanced training and guidelines to help scientists navigate their involvement in political debates, ensuring their contributions are both impactful and ethically sound.

References

Alinejad, D., & van Dijck, J. (2022). Climate communication: How researchers navigatebetween scientific truth and media publics. *Communication and the Public*, 205704732211386. https://doi.org/10.1177/20570473221138612

Besley, J. C., & Nisbet, M. C. (2013). How scientists view the public, the media and thepolitical process. *Public Understanding of Science*, *22*(6), 644–659. https://doi.org/10.1177/0963662511418743

Biermann, K., Peters, N., & Taddicken, M. (2023). "You Can Do Better Than That!": Tweeting Scientists Addressing Politics on Climate Change and Covid-19. *Media andCommunication*, *11*(1). https://doi.org/10.17645/mac.v11i1.5961

Bogner, A. (2021). *Die Epistemisierung des Politischen: Wie die Macht des Wissens dieDemokratie gefährdet. (Was bedeutet das alles?): Nr. 14083*. Reclam.

Jasanoff, S. (1994). *The Fifth Branch: Science Advisers as Policymakers*. Harvard University Press.

Kotcher, J. E., Myers, T. A., Vraga, E. K., Stenhouse, N., & Maibach, E. W. (2017). Does engagement in advocacy hurt the credibility of scientists? Results from a randomizednational survey experiment. *Environmental Communication*, *11*(3), 415–429. https://doi.org/10.1080/17524032.2016.1275736

Lupia, A. (2013). Communicating science in politicized environments. *Proceedings of the National Academy of Sciences of the United States of America*, *110 Suppl 3*, 14048–14054. https://doi.org/10.1073/pnas.1212726110

Post, S., & Ramirez, N. (2018). Politicized science communication: Predicting scientists' acceptance of overstatements by their knowledge certainty, media perceptions, and presumed media effects. *Journalism & Mass Communication Quarterly*, *95*(4), 1150–1170. https://doi.org/10.1177/1077699018769668

Scheufele, D. A. (2014). Science communication as political communication. *Proceedings of the National Academy of Sciences of the United States of America*, *111*, 13585–13592.https://doi.org/10.1073/pnas.1317516111

Witnessing online harassment against scientists: Effects on scientists and public perceptions of science

Jana Laura Egelhofer, Christina Seeger, Alice Binder

Scientists who publicly speak out are increasingly affected by hate and harassment (Global Witness, 2023; Nogrady, 2021). Such incidents are growing more frequent, particularly on social media, which are increasingly vital for digital science communication (Neuberger et al.,2021). Despite the growing concerns surrounding the heightened hostility directed towards scientists, there exists a dearth of systematic research on its effects. Initial investigations have shed light on the detrimental consequences for the psychological well-being of affected scientists (Global Witness, 2023; Gosse et al., 2021; Nogrady, 2021). However, the amplified *visibility* of harassment through social media may also yield consequences – for (1) scientists who are not directly affected and (2) public perceptions of scientists. Yet, to the best of our knowledge, there is no study to date testing the effects of *witnessing* harassment against scientists. Therefore, we report the findings of two preregistered experiments investigating the impact of observing online harassment against scientists.

Hypotheses Study 1: Effects on scientists

As mentioned, previous survey and interview studies indicate that *direct* experiences with harassment, i.e., being targeted themselves, have adverse implications for scientists' emotional well-being and may instigate a 'chilling effect,' deterring them from participating in future public outreach (Global Witness, 2023; Nogrady, 2021). Intriguingly, concerns have been raised that *indirectly* experiencing harassment, i.e., witnessing fellow scientists being harassed online, could yield similar effects on observing scientists (Nogrady, 2021). Consequently, we examine whether exposure to online harassment of scientists leads to negative emotions **(H1.1),** an increased perceived risk of harassment for oneself **(H1.2),** and a 'chilling effect' on their willingness to engage in future science communication **(H1.3).** Moreover, affected scientists fear negative effects on their reputation and trustworthiness (Global Witness, 2023, p. 5). Therefore, we explore whether exposure to online harassment has a negative effect on the perceived trustworthiness of the harassed scientists **(RQ1.1).**

Hypotheses Study 2: Effects on public perceptions of scientists

Drawing from existing research that highlights the detrimental impact of uncivil and hostile comments on the trust and credibility of news media and their coverage, known as the ‚nasty effect' (Anderson et al., 2014), we posit that online harassment of scientists may exert a similar influence on trust in scientists and the information they provide. Therefore, we test whether exposure to harassment has a negative effect on the perceived trustworthiness of harassed scientists **(H2.1)** as well as the credibility of a scientific claim made by them **(H2.2).**

Moreover, concerns have been expressed regarding the potential negative impact of growing hostility towards scientists on public perceptions of scientists in general (Valiverronen & Saikkonen, 2021, p. 9). Thus, we also investigate whether exposure to harassment has negative effects on general trust in scientists **(H2.3)**. Additionally, we take into account the moderating role of science-related populism (Mede & Schäfer, 2020). Given the negative relationship between science-related populist attitudes and trust in science (Eberl et al., 2023), we anticipate that these attitudes will moderate the effects in H2.1-3 (**H2.4**).

Method

To test our hypotheses, we conducted two preregistered experimental studies that followed the same experimental design. In both studies, participants were exposed to two social media postings by two different scientists who shared information about a media interview. In the experimental conditions, these postings were accompanied by two user comments each, including harassment

targeting both the scientists' credibility and physical attractiveness. In the control conditions, participants saw only the original postings without comments.
The first study, focusing on the consequences for scientists, was conducted with a sample of German and Austrian scholars from various disciplines (*N* = 860). The second study, focusing on the consequences for public perceptions of scientists, was conducted with a representative sample of German citizens (*N* = 1,307).

Results
Study 1 (Scientists): Preliminary results indicate that exposure to harassment leads to stronger negative emotions (H1.1) and heightened perceived risk (H1.2). However, it has no effect on willingness to engage in outreach (H1.3). Furthermore, we find a very small positive effect on the trustworthiness of harassed scientists (RQ1.1).
Study 2 (Public): Initial results suggest that exposure to harassment has a negative effect on citizens' perceived trust in the harassed scientists (H2.1). However, there is no effect on the perceived credibility of a scientific claim made by these scientists (H2.2), and no effect on trusting scientists in general (H2.3). Regarding H2.4, science-related populist attitudes moderate the effect on trust in the harassed scientist, though no significant interaction effect is observed for claim credibility and general trust in scientists.

Conclusion
Our findings have several implications for scientists, science communication, and public perceptions of scientists. First, we show that not only direct but also *indirect* experiences with harassment have consequences for scientists' well-being (negative emotions, heightened risk perception). Encouragingly, our study does not reveal a 'chilling effect' of witnessing harassment on science communication. Interestingly, scholars perceive harassed scientists as slightly more trustworthy. However, when turning to the general public, there are indications of a 'nasty effect' of harassment on the perceived trust of targeted scientists, though no discernable effects on general trust in scientists.

References
Anderson, A.A., Brossard, D., Scheufele, D.A., Xenos, M.A., & Ladwig, P. (2014). The "NastyEffect:" Online Incivility and Risk Perceptions of Emerging Technologies: Crudecomments and concern. Journal of Computer-Mediated Communication, 19(3), 373–387. https://doi.org/10.1111/jcc4.12009

Eberl, J.-M., Huber, R. A., Mede, N. G., & Greussing, E. (2023). Populist attitudes towards politics and science: How do they differ? Political Research Exchange, 5(1), 2159847. https://doi.org/10.1080/2474736X.2022.2159847

Global Witness. (2023). Global Hating. How online abuse of climate scientists harms climate action. https:///en/campaigns/digital-threats/global-hating/

Gosse, C., Veletsianos, G., Hodson, J., Houlden, S., Dousay, T. A., Lowenthal, P. R., & Hall, N.(2021). The hidden costs of connectivity: Nature and effects of scholars' online harassment. Learning, Media and Technology, 46(3), 264–280. https://doi.org/10.1080/17439884.2021.1878218

Mede, N. G., & Schäfer, M. S. (2020). Science-related populism: Conceptualizing populist demands toward science. Public Understanding of Science, 29(5), 473–491. https://doi.org/10.1177/0963662520924259

Neuberger, C., Weingart, P., Fähnrich, B., Fecher, B., Schäfer, M. S., Schmid-Petri, H., & Wagner, G. G. (2021). *Der digitale Wandel der Wissenschaftskommunikation*.

Nogrady, B. (2021). 'I hope you die': How the COVID pandemic unleashed attacks on scientists. *Nature*, *598*(7880), 250–253. https://doi.org/10.1038/d41586-021-02741-x

Valiverronen, E., & Saikkonen, S. (2021). Science communicators intimidated: Researchers' freedom of expression and the rise of authoritarian populism. Journal of Science Communication, 20(04), A08. https://doi.org/10.22323/2.20040208