



OPERAS

open scholarly communication in the european
research area for social sciences and humanities

Fostering Trust in the Digital Age

**Report from the TrustOn2024 international
workshop**

**at the Royal Library of Belgium (KBR)
Brussels 26–27 June 2024**

& session at the Science Summit UNGA79

online 11 September 2024



TRUST ON 2024
TACKLING DISINFORMATION

Let's build a human-centric and trustworthy digital future!

Collectively gathered by OPERAS RI



Table of Contents

Foreword by Marc Vanholsbeeck, Belspo	3
1. Introduction	5
2. Track Reports	8
2.1. Infrastructure Track	10
“A People-Centered Approach to Combating Misinformation” by Mei Lin Fung	13
“Open Infrastructures for Open Science” by Chris de Loof	15
“Cultivating Trust in the Digital Information Ecosystem” by Amir Banifatemi	17
“Online Disinformation from a Freedom of Expression Perspective in the European Union” by Alejandra Michel	20
“(Re)informing and engaging readers in the Digital Age for a (francophone) knowledge platform” by Thomas Parisot	22
“MetaBelgica: A Trusted Reference Database for Cultural Heritage Entities” by Hannes Lowagie and Sven Lieber	24
“Intermediaries, Crafted by Trustees” by Dr. Jörg Lehmann	26
Key Recommendations from the Infrastructure Track	28
2.2. Science Track	29
“Trust and Disinformation in the Digital Age: The Role of Collaboration in Science Communication” by Julija Baniukevic	30
“Re-defining the ecosystem of trust in science: the role of stewards of trust in changing research environments” by Agata Gurzawska	32
“Responsible science communication for trustworthy science: First insights from the IANUS project” by Hendrik Berghäuser	35
“Fact or fiction: The battle against scientific disinformation in the social media arena” by Panagiotis Monachelis	37
“Guidelines on the Ethical Use of Artificial Intelligence in Education and Research” by Reda Cimmperman	40
“Science hoaxes, diminishing trust” by Isabel Mendoza-Poudereux and Carolina Moreno-Castro	41
“Building trust in science through science: brief recommendations” by Nelson Ferreira	43
“Public and mediator perspectives on responsible science communication and participation: Implications for navigating online science (mis)information” by Tine Ravn	45
“Changing the communication paradigm - from influencing to enabling public debate” by Mario Scharfbillig	49
Key recommendations from the Science Track	55
2.3. Mediation Track	57
“Data Curation as a Mechanism for Increasing Trust in Research” by Evelien Dhollander and Kevin Leonard	58
“Multilingual information in the age of AI: findings and perspectives” by Susanna Fiorini	60

“Rebuilding Trust in Online Information: The Role of Scholarly Communities” by Pierre Mounier	62
“Increasing confidence in science: how open science fights disinformation” by Véronique Stoll	64
“Fake news : need for a mix of technology, human sciences and international trust bodies to face borderless rapidly developing AI” by Matei Mancas	66
“Against unmonitored AI-generated content, journalists as gatekeepers” by Mathilde Dorcadie	68
“FROM TRUSTWORTHY DIGITAL TO INCORRUPTIBLE DIGITAL/AI?” by Guigone Camus and Christophe Calvin	70
“JPEG Trust: an international standard for establishing trust in digital media” by Frederik Temmermans	73
“Net Neutrality for a better Electronic Democracy” by Valentina Tirloni	75
“The GoTriple platform: mediation to access scientific results” by Sy Holsinger	78
“Building Digital Trust for a Human-Centered Internet” by Gaël Van Weyenbergh	80
Key Recommendations from the Mediation Track	88
3. TrustOn2024 at the Science Summit	92
4. Conclusion	93
Annex	96
Biographies	96

Foreword by Marc Vanholsbeeck, Belspo

Ladies and gentlemen,

According to UNESCO, by 2023, 60% of the world's population, or 4.75 billion people, will have used social networks to express themselves, keep informed and ask questions. They will enjoy a digital agora of expression and protest, building and fostering social relationships and identities.

On the other hand, at the beginning of 2023, an Ifop survey targeted young French individuals ([La mésinformation scientifique des jeunes à l'heure des réseaux sociaux - Fondation Jean-Jaurès](https://www.jean-jaures.org/fr/la-mésinformation-scientifique-des-jeunes-a-lheure-des-reseaux-sociaux) [jean-jaures.org]). The survey assessed how susceptible individuals were to scientific inaccuracies and examined the influence of their interaction with social media platforms. The results indicate a significant presence of conspiracy theory supporters and a growing inclination toward irrational beliefs and misinformation, particularly among frequent users of social media.

Social networks easily become bubbles of isolation, cocoons of disinformation, online ecosystems that sometimes lead to conspiracy and violence.

I believe much of our work during these 2 days will emerge from this apparent paradox. We are facing today an unprecedented mix of incredible opportunities and threats. There is hope for greater equity and inclusion in the way we share and communicate knowledge, empowering citizens to contribute to the public debates on which our democracies rely. At the same time, disinformation, misinformation and so-called “alternative facts” saturate our cognitive horizons.

Belspo – the Belgian Science Policy Office - is glad to support such an ambitious workshop that will combine consideration for science, mediation and governance, cross-fertilizing interactions between those, and providing concrete pieces of recommendation.

We are certainly looking forward to the outcomes of this conference, which is the last one to be placed under the auspices of the Belgian Presidency of the Council of the EU.

Today we can already look back to these 6 months of intense work and exchanges. Knowledge valorisation and the reinforcement of R&I in society have been two of our main priorities. More specifically, two of our Presidency events directly relate to our concerns of today. At the occasion of the Belgian EOSC tripartite in April 2024, we had the philosopher Sabina Leonelli as a keynote speaker. She emphasised that the current focus, in Open Science policies, on sharing data and materials has become, to some extent, an obstacle to the promotion of equitable and human-centric open science. Leonelli therefore suggested an alternative, process-oriented view of open science that seeks to establish “judicious

connections among systems of practice.” Probably Leonelli’s provocative thoughts can also be an inspiration for us, during the 2 coming days.

More recently, in Tervuren, Belspo organised with the Commission and the Belgian federated entities an important conference about the role of SSAH into evidence-informed policy making. According to the participants to this conference, there is a need to go from “ego-systems” to true “ecosystems” of scientific expertise, and to create “safe spaces” into which policy makers, researchers and citizens could discuss together topics for which scientific evidence is needed.

At the level of BELSPO, we are now implementing new “Science4Policy” funding programmes, and we will do our best to integrate the outcomes of those rich discussions and exchanges that the Presidency both stimulated and facilitated.

Ladies and gentlemen, dear participants to TrustOn2024 workshops, we at Belspo do strongly believe that strengthening access to free and reliable information, and particularly free and open access to research-based knowledge, also contributes to strengthening freedom of thought and expression.

To put it with Hannah Arendt,

“Facts inform opinions, and opinions, inspired by different interests and passions, can differ widely and still be legitimate as long as they respect factual truth. Freedom of opinion is a farce unless factual information is guaranteed and the facts themselves are not in dispute”

(Arendt, 1968, 238).

I thank you for your attention.¹

¹ This inspiring introduction speech was given by Marc Vanholsbeeck at the beginning of the TrustOn workshop on 26 June 2024.

1. Introduction

In today's rapidly evolving digital landscape, the challenges of disinformation and the erosion of public trust have reached critical levels. The rise of digital platforms and artificial intelligence, while offering tremendous potential, has also enabled the proliferation of false and misleading information on an unprecedented scale. This report, featuring 27 papers from the [TrustOn2024 Workshop](#) and the **Science Summit at the 79th United Nations General Assembly session “Fostering Trust in the Digital Age”**, seeks to address these pressing issues. Both events brought together experts to explore strategies to combat disinformation and strengthen trust in science and digital ecosystems.

The fight against disinformation is not new. Misinformation has historically threatened the stability of societies, but today's digital tools amplify its reach and impact. With just a few clicks, harmful content can spread across the globe, influencing opinions and decisions, often with dangerous consequences. As **Audrey Azoulay**, Director-General of UNESCO, noted in her recent address, disinformation and hate speech are major threats to social cohesion and peace. The UNESCO action plan, launched in late 2023, and the **Global Principles for Information Integrity**², unveiled by the **UN Secretary-General**, reflect the urgency of coordinated global efforts to address these threats.

OPERAS, as a research infrastructure dedicated to open scholarly communication in the social sciences and humanities, plays a vital role in this fight. Our focus on the quality of scholarly content, trust-based services, and knowledge transfer equips us to tackle disinformation head-on. By fostering reliable access to accurate information, we contribute to building a foundation of trust that is essential not only for academia but for the broader public sphere.

Disinformation and the United Nations Sustainable Development Goals (SDGs)

Addressing disinformation is also crucial for achieving the **United Nations Sustainable Development Goals (SDGs)**. Trustworthy and inclusive digital environments are key to ensuring equitable access to quality education (**SDG 4**), promoting gender equality (**SDG 5**), reducing inequalities (**SDG 10**), and building peaceful and inclusive societies (**SDG 16**). Moreover, global partnerships for sustainable development (**SDG 17**) depend on our ability to foster trust and integrity in the digital realm.

As highlighted during the Science Summit UNGA79, creating a trustworthy digital infrastructure and promoting ethical AI development are essential to navigating the complexities of today's information landscape. **Kenya's** innovative approach to

² See: <https://www.un.org/en/information-integrity/global-principles>

combating misinformation and hate speech during the 2022 Presidential Elections serves as an inspiring example. By establishing an early-warning system capable of monitoring and countering harmful content in real time, Kenya demonstrated how coordinated, tech-driven initiatives can promote peace and safeguard democracy.

A Call for Collective Action

The disinformation crisis is not an isolated issue—it is a collective challenge that requires global solutions. Whether addressing climate change, public health, or social justice, we will not succeed if disinformation continues to undermine public trust and democratic processes. As this report illustrates, tackling disinformation requires innovative solutions, participatory governance, a human-centered approach and a community to support it.

Through the diverse range of perspectives presented in this report, we aim to contribute to the global dialogue on fostering trust in the digital age. The insights gathered here are intended to inspire concrete actions and foster international collaboration. Together, we can build a more resilient and equitable digital ecosystem, one that empowers individuals, upholds the truth, and supports sustainable development for all.

We would like to express our sincere gratitude to all the contributors for sharing their abstracts and for the significant and important work you are doing around trust. Your perspectives have added significantly to the understanding and quality of this report. Thank you for your continued efforts, commitment and dedication to this important topic.

Defining: Information Disorder, Disinformation and Mis-information

As we are presenting a report with insights from different experts in related themes involving trust and disinformation, in this section we state the definitions that we consider while speaking about disinformation.

In Claire Wardle and Hussein Derakhshan's report published in 2017 by the Council of Europe, there are mainly three different types of information inside the bigger idea of "information disorder":

- **Dis-information.** Information that is false and deliberately created to harm a person, social group, organization or country.
- **Mis-information.** Information that is false, but not created with the intention of causing harm.
- **Mal-information.** Information that is based on reality, used to inflict harm on a person, organization or country. (Wardle, Derakhshan, 2017, p. 20).

From this differentiation we state that every time we speak about disinformation, there is intention to harm involved in it. On the other side, misinformation also involves wrong content, although not produced with the intention to harm. These are important differentiations. Although we state them here, it is possible that in individual texts from experts these concepts are not that clearly stated. That's why it is important for us to have them already clarified in this first section.

Also following the work of Wardle and Derakshan, we avoid using the term “fake news”. The term is inadequate to grasp the complex context of information disorder, while it is also being inadequately appropriated by politicians to refer to coverage they find disagreeable (Wardle, Derakshan, 2017, p. 5).

2. Track Reports

In addressing disinformation and building trust in the digital age, the Report, consisting mainly of abstracts from the **TrustOn2024 Workshop**, is organised into three distinct yet complementary tracks: **infrastructure**, **science**, and **mediation**. Each track offers a unique lens through which to tackle these critical issues, recognizing that solutions must be multi-faceted and interconnected to effectively combat disinformation.

1. Infrastructure: Building a trustworthy foundation

The **Infrastructure** track focuses on the foundational elements that enable the digital ecosystem to function effectively and securely. This includes the development of digital infrastructures that support the transparent and trustworthy dissemination of information. Research infrastructures, such as those provided by **OPERAS**, play a crucial role by offering reliable, open, and accessible scholarly content. The goal is to design infrastructure that can mitigate the spread of disinformation by ensuring the quality and reliability of the data and content it supports.

In this context, infrastructure does not merely refer to physical servers or software but extends to the protocols and governance structures that underpin these systems. **Data cooperatives**, **ethical AI development**, and **digital literacy** initiatives fall under this category, all working towards creating systems that foster trust and reduce vulnerabilities to misinformation. These infrastructures provide the tools and frameworks needed to identify, prevent, and correct the spread of false information. Such work is critical for the long-term sustainability of the digital ecosystem, as it strengthens the very systems through which we access and share knowledge.

2. Science: Ensuring credible and transparent knowledge

The **Science** track emphasises the importance of research in building trust. Trust in science is crucial, particularly as disinformation increasingly targets scientific facts and findings. Research institutions, universities, and research communities must not only produce credible, peer-reviewed content but also take an active role in ensuring information integrity and preventing the dissemination of false information related to research.

In this track, the focus is on fostering open, transparent, and ethical research practices that make scientific findings more accessible and understandable to the public. Efforts to ensure that science is communicated accurately and clearly to diverse audiences are critical in promoting public trust. The integration of **open access publishing**, **open data sharing**, and **collaborative research** plays a pivotal role in this track, ensuring that academic knowledge is disseminated in ways that

are transparent and free from manipulation. By aligning scholarly communication with robust ethical standards and integrity, this track works towards countering the growing influence of disinformation on public perception of science.

3. Mediation: Addressing disinformation through communication and engagement

The **Mediation** track examines how we engage with the public and mediate the flow of information. This track focuses on the role of media, communication professionals, and educators in addressing disinformation and building trust. **Mediation** is not just about countering false information; it is about creating inclusive environments where people can critically engage with information and recognize the difference between trustworthy sources and misleading content.

This track draws attention to the **role of journalists, media literacy programs, and public information campaigns** as tools for building trust in the digital space. By equipping citizens with the skills to critically evaluate information and engage in constructive dialogue, mediation helps to build a more informed and resilient society. In this context, the work of fact-checking organisations, media outlets, and public awareness campaigns becomes essential in shaping a more reliable and trustworthy information ecosystem.

A holistic approach to tackling disinformation

Each track—**infrastructure, science, and mediation**—offers a different perspective on how to address disinformation and build trust in the digital age. The **Infrastructure** track builds the technical foundation to ensure the quality and accessibility of information. The **Science** track focuses on making scientific knowledge credible and transparent, while the **Mediation** track addresses how we communicate and engage with the public to foster critical thinking and resilience to misinformation.

These three approaches are interconnected and must work together to create a comprehensive strategy to address disinformation. Whether through building ethical infrastructures, promoting transparent scholarly communication, or empowering citizens with the tools to discern fact from fiction, each perspective contributes to the shared goal of fostering trust in the digital age.

By synthesising insights from these diverse tracks, this report aims to offer actionable solutions and foster collaboration among stakeholders across sectors. Together, we can develop a more informed, resilient, and trustworthy digital ecosystem, crucial for addressing not only disinformation but also the broader challenges facing societies today and have an impact on how society deals with these challenges.

2.1. Infrastructure Track

Mei Lin Fung, chair of the *People Centered Internet*, gave a presentation titled "**A People-Centered Approach to Combating Misinformation**". In it, the dual nature of AI was highlighted, showing its potential for both positive and harmful use, such as spreading misinformation. Fung emphasised the importance of a people-centered approach to combat misinformation by engaging all individuals, not just elites, and expanding participation opportunities. She discussed the challenges of misinformation's widespread impact and the lack of coordination and common standards. Recommendations included empowering citizens through a digital bill of rights (and a corresponding digital asset registry) and fostering collaborative governance models with regional regulatory sandboxes. The discussion also touched on future AI regulation and the need for adaptive, feedback-driven regulatory networks.

In his presentation "**Open Infrastructures for Open Science**", **Chris De Loof** from **BELSPo** discussed the technical aspects of digital public infrastructure (DPI), comparing it to public utilities like electricity and roads. He described DPI as a robust framework offering digital services to both citizens and private companies, promoted by the United Nations to advance Sustainable Development Goals. The architecture of DPI includes application layers, middleware, and physical infrastructure. De Loof highlighted challenges such as the privatisation of public spaces and issues with tech concentration. Recommendations for enhancing DPI included improving security, interoperability, training, transparency, and accountability. For research, he suggested creating digital commons, investing in hyperscale computing facilities, and supporting public alternatives in collaboration with private entities.

In his presentation, "**Cultivating Trust in the Digital Information Ecosystem**," **Amir Banifatemi** from **AI Commons** addressed the critical need to build public trust in the age of generative AI. He highlighted the current crisis of trust, citing Ipsos data from 2023 that shows social media as the leading source of information, where disinformation and fake news are prevalent. Banifatemi noted that misinformation poses significant global risks alongside extreme weather. He discussed the misuse of AI in spreading disinformation through deep fakes and manipulated content, emphasising the need for collective action. Current solutions include AI detection tools, content authentication, public education, and legal measures. Banifatemi proposed a Global Challenge to foster international collaboration, involving phases for identifying ideas, prototyping, and scaling solutions, with prizes to incentivize open-source results.

"**Online Disinformation from a Freedom of Expression Perspective in the European Union**," was the title of **Alejandra Michel's** talk, from **CRIDS/University of Namur**. She discussed the Digital Services Act (DSA), which aims to ensure online trust by imposing due diligence obligations on intermediary service

providers, including content moderation. She addressed the challenge of balancing the fight against illegal and harmful online content with the protection of freedom of expression. Michel explained that while the DSA distinguishes between illegal content and harmful but legal content, the regulatory approach must vary accordingly. She emphasised the need for appropriate measures that address content-specific risks for illegal content and content-neutral strategies for harmful disinformation, warning against censorship as a means to build trust.

Thomas Parisot from *cairn.info* stated in his presentation "**(Re)informing and engaging readers in the Digital Age for a (francophone) knowledge platform**" that the challenge of scientific platforms is gradually giving way to the need to establish a relationship of trust and quality with the readership. With this in mind, Cairn.info, a platform dedicated to the dissemination of French-language scientific publications, has recently upgraded its service, with the introduction of video interviews with renowned French-speaking scientific authors, accompanied by reading lists and dossiers written to help as many people as possible discover the state of the art. By balancing technological advancement with respect for the human authorship behind scientific claims, Cairn.info seeks to deepen trust in science without compromising its inherent complexity and debate.

In their presentation "**MetaBelgica: A Trusted Reference Database for Cultural Heritage Entities**," **Hannes Lowagie and Sven Lieber** from *KBR* discussed the challenges of ensuring trustworthiness in cultural heritage institutions' data, addressing issues like copyright disputes, provenance, and historical research. He emphasised the importance of credibility, transparency, and accuracy in earning global trust in internet infrastructures. Lowagie and Lieber recommended investing in libraries and cultural heritage institutions and highlighted MetaBelgica as a use-case for shared entity management using WikiBase. This approach involves metadata, domain, and tech experts, manual curation, community engagement, and alignment with international standards. They concluded that leveraging established resources is preferable to outsourcing or reinventing solutions.

"**Intermediaries, Crafted by Trustees: Datasheets for Digital Cultural Heritage**," a presentation by **Jörg Lehmann** from the *Berlin State Library*, emphasised the role of cultural heritage institutions in developing and applying AI to digital heritage. He discussed the importance of trust in these institutions, which are mission-driven rather than profit-focused. Lehmann highlighted the need for heritage institutions to educate the public about machine learning and biases in datasets, explaining how large language models can produce inaccurate results due to their probabilistic nature. He introduced the concept of datasheets and model cards as tools for documenting machine learning models, noting the absence of templates for digital heritage datasets until recently. Lehmann called for standardisation in documenting dataset creation and use, including metadata on biases and motivations. He proposed creating a web app for metadata documentation and stressed the value of authentic, human-created text for future machine learning models, emphasising the importance of libraries in digitising collections. He also

pointed out that because the amount of born-digital text today is increasingly generative, authentic human-created text will be valuable for machine learning models. Therefore, the importance of libraries and digitising their collections cannot be overstated.³

³ The introduction of the infrastructure track was prepared by Julie M. Birkholz & Amanda Hemmons, KBR.

“A People-Centered Approach to Combating Misinformation”

by Mei Lin Fung

Principles and Strategies

1. Engagement as a Tool for Resilience:

Engaging communities in public discussions equips them with the tools and knowledge needed to discern truth from falsehood. This engagement can take various forms, such as community discussions, digital literacy workshops, and social media campaigns that promote accurate information and debunk myths. As people engage, they share their insights and learn from each other, creating a collective intelligence that strengthens the community's resilience to misinformation. By encouraging people to add their own voices and perspectives, they develop critical thinking skills and become more vigilant against misinformation. This grassroots involvement not only enhances individual media literacy but also fosters a community-wide defense mechanism against fake news.

2. Building Trust through Transparency and Participation:

Trust is fundamental in combating misinformation. Governments and institutions need to earn public trust through transparent and participatory practices. By involving citizens in the decision-making process and providing clear, accessible information, trust is built and maintained.

3. Historical Lessons from Medicine and Transportation:

Learning from the historical evolution of sectors like medicine and transportation, which have successfully expanded to serve diverse populations, provides actionable insights for today's digital challenges. These sectors thrived through regulatory frameworks, public trust, and widespread community engagement. Similarly, a coordinated, people-centered approach can enhance the effectiveness of digital governance and the fight against misinformation.

Recommendations

1. Collaborative Governance Models:

Foster collaborative governance models that involve multiple stakeholders, including governments, tech companies, and civil society organizations to build trust and enhance citizen engagement. Regional digital cross-sector regulatory sandboxes can be established to test and refine strategies for combating misinformation while ensuring interoperability and shared protocols.

2. Digital Bill of Rights:

Establish a Digital Bill of Rights to give citizens ownership and a stake in the digital realm. This can include rights to accurate information, privacy protections, and mechanisms to report and rectify misinformation.

3. Set up Regional Networks to learn faster together:

Encourage regional cooperation to address misinformation, recognizing that fake news often crosses borders. A coordinated network for responding to common threats and learning in parallel can allow a successful response to be quickly replicated broadly.

4. Continuous Feedback and Improvement:

Create feedback loops where community input is continuously sought and used to improve misinformation countermeasures. This ensures that the strategies remain relevant and effective in an ever-changing digital landscape.

5. Digital Literacy Programs:

Implement community-based digital literacy programs that educate people on recognizing and reporting misinformation. These programs should focus on critical thinking skills, understanding the sources of information, and the impact of fake news.

6. Engage with Technology:

Utilize AI and other advanced technologies to bring in more involvement by people to detect and combat misinformation. However, this must be done transparently and ethically, with clear labels and explanations of how these technologies work to avoid mistrust and misuse.

[Link](#) to TrustOn2024 Presentation.

References

1. Stein, J., Fung, M., Bertalosa, M., 2024 Improving Global Governance: Data Cooperatives for Global Cooperation. Italy G7, Think7 published policy brief.
<https://think7.org/improving-global-governance-data-cooperatives-for-global-cooperation/>
2. Fung, M, Bussolli, P., Stein, J., 2024 A Gateway to Finance for MSMEs Paving a Practical Path to Prosperity. Italy G7, Think7 published policy brief.
<https://think7.org/cbdcs-a-gateway-to-finance-for-msmes-paving-a-practical-path-to-prosperity/>
3. Fung, M., Tobing, D., Bertolaso, M. Potluri, V., 2023, People-Centered Science and Digital Transformation: A Practical Proposal for the G7 and G20 Japan G7, Think7 published policy brief
<https://think7.org/people-centered-science-and-digital-transformation-a-practical-proposal-for-the-g7-and-g20/>
4. Fung, M., Stein, J., Flynn, S. 2024, Implementing Financial Resilience by Digital Asset Registries in CBDCs. Unpublished paper ([link](#)) Implementing Financial Resilience by Digital Asset Registries in CBDCs.
5. Singh, T., Fung, M., Murray, E., Lacavaro, A., Gamser M., Omwenga. B., Ndemo, B., Kharas, H. 2022, Global Public-Private Digital Utilities for MSME Recovery and Transition German G7, Think7 published policy brief
<https://think7.org/global-public-private-digital-utilities-for-msme-recovery-and-transition/>
6. Beirne, J., Bilotta N. Fung, M., Singh T. 2022, Central Bank Digital Currencies: Governance, Interoperability, and Inclusive Growth German G7, Think7 published policy brief
<https://think7.org/central-bank-digital-currencies-governance-interoperability-and-inclusive-growth/>

Full list of publications includes additional G7 Germany and Italy Think7 policy

briefs ([link](#)).

“Open Infrastructures for Open Science” by Chris de Loof

In the modern digital era, the concept of Open Infrastructures has become increasingly significant. These infrastructures are essential for promoting transparency, collaboration, and innovation within the scientific community and beyond. Open Science, which advocates for the free sharing of research outputs and methodologies, relies heavily on robust digital public infrastructures (DPIs). These infrastructures are not only pivotal for scientific advancement but also for fostering trust between citizens and governments. The TrustOn2024 conference, held in Brussels, underscored the importance of trust, governance, and digital public infrastructures, making it an ideal platform to discuss these themes.

Public Goods and Digital Public Infrastructures: Traditionally, public goods such as electricity and roads have been characterized by their fair use, regulation, and the balance between public and private sectors. Similarly, digital public infrastructures (DPIs) are designed to provide essential digital services to citizens and private entities. These infrastructures are promoted by the United Nations as high-impact initiatives to accelerate the achievement of Sustainable Development Goals. DPIs drive inclusion, innovation, and respect for rights and fundamental freedoms, all governed by a comprehensive legal and policy framework. By building trust between citizens and governments, DPIs enable the delivery of citizen-centric services while ensuring security, privacy, and accessibility.

Building Trust through Digital Public Infrastructures: Trust is a cornerstone in the relationship between citizens and governments. DPIs play a crucial role in building this trust by ensuring that digital services are secure, private, and accessible. Successful implementations of DPIs have demonstrated their potential to enhance public trust significantly. For instance, the establishment of secure and interoperable systems that provide seamless user experiences can greatly improve public confidence in digital services. Moreover, the inclusive approach to training and education ensures that all citizens can benefit from these infrastructures, further strengthening trust.

Architecture of Digital Public Infrastructures: The architecture of DPIs is multi-layered, comprising several critical components. The DPI layer, which resembles a Common European Data Space, forms the foundation. This is followed by the applications layer, which includes services related to health, agriculture, education, and other innovative sectors. Middleware, identification, and data exchange layers ensure interoperability and seamless data flow. The physical infrastructure, encompassing connectivity, power, and data centers, supports the entire system. However, managing public spaces by private companies can pose challenges, as seen in the comparison between Twitter/X and Mastodon. Discussions on net neutrality and AI market concentration highlight the need for careful governance of these infrastructures.

Requirements for Effective Digital Public Infrastructures: For DPIs to be effective, several requirements must be met. Security measures such as data protection, encryption, and authentication are paramount. Interoperability between different systems ensures a seamless user experience. An inclusive approach to training and education maximizes the benefits of DPIs for all citizens. Transparency and collaboration are also crucial, with open frameworks and accountability mechanisms in place to ensure that DPIs operate fairly and effectively.

Open Science Infrastructures: Open Science Infrastructures aim to create a digital commons for research, functioning as public digital infrastructures. These infrastructures involve building hyperscale computing facilities for research, with national investments providing global access. Public-private partnerships are essential in developing and managing these infrastructures, ensuring that public interests remain at the forefront. Security measures, such as those provided by Geant and EDUGAIN, and interoperability initiatives like the European Open Science Cloud (EOSC), are critical components. An inclusive approach to training, transparency, and collaboration further enhances the effectiveness of Open Science Infrastructures.

Governance and Regulation: Effective governance and regulation are vital for the success of DPIs. Comprehensive rules and standardization ensure that these infrastructures operate smoothly and fairly. Data governance legislation, including the General Data Protection Regulation (GDPR), provides a robust framework for protecting citizens' data. The regulation of big tech companies is also necessary to prevent market monopolies and ensure fair competition. Clear procurement and funding rules, along with the establishment of authoritative bodies for oversight, further strengthen the governance of DPIs.

Recommendations for European e-Infrastructures: To stimulate public digital alternatives to big tech, a coherent European Commission strategy and funding for e-Infrastructures are essential. Recognizing and enforcing existing e-Infrastructures, such as national research and education networks (NREN), will enhance their role in the digital commons for research. Broader support for community-based open-source software and increased investment in cybersecurity are also recommended. These measures will ensure that European e-Infrastructures remain key players in the global digital landscape.

Conclusion: Open Infrastructures for Open Science are crucial for achieving a trusted digital future. By building robust, secure, and inclusive digital public infrastructures, we can foster greater trust between citizens and governments. Continued investment and collaboration in these infrastructures are essential for promoting transparency, innovation, and sustainable development.

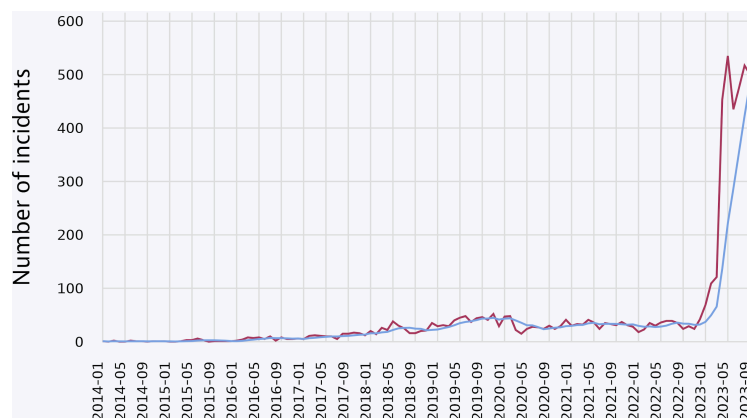
“Cultivating Trust in the Digital Information Ecosystem” by Amir Banifatemi

Introduction

In an era dominated by digital technologies, the way information is created, shared, and consumed has undergone a profound transformation. While advancements in artificial intelligence (AI) and digital platforms have enabled unprecedented access to knowledge, they have also introduced significant challenges. Trust in digital information has eroded as society grapples with the proliferation of disinformation, deepfakes, and manipulated content. According to a [2023 Pew Research](#) study, 64% of adults worldwide report difficulty distinguishing between accurate and misleading information online. The consequences of these trends affect not only individuals but also public institutions, private organizations, and the broader social fabric.

Generative artificial intelligence systems are revolutionizing how information is created, shared, and consumed. These technologies significantly enhance the speed and diversity of information flow across platforms. However, they also have the potential for misuse, such as campaigns of disinformation, the creation of deep fakes, and other manipulated content. These misuses carry severe negative consequences for individuals and society. Additionally, generative AI systems can be exploited for cyberattacks, especially through social engineering tactics, further exacerbating the risks associated with their misuse.

The [OECD AI Incident Monitor](#) shows a drastic increase in the number of AI incidents related to mis- and disinformation since 2022 (see below):



This issue transcends national borders and requires global cooperation for complementary policy and technology solutions. Moreover, it transcends technical approaches, necessitating an interdisciplinary approach, as techno-solutionism alone will not suffice.

Addressing these challenges requires a dual approach that integrates policy innovation, technological solutions, and an emphasis on leveraging data responsibly. This synergy can create actionable frameworks to rebuild trust in the digital information ecosystem, ensuring that technological progress aligns with

societal needs while incentivizing positive behavior in the use and governance of data.

The Current State of Trust in Digital Information

The rise of generative AI and advanced content creation tools has transformed the landscape of information dissemination. However, this transformation has been accompanied by a growing mistrust in digital content. Generative AI enables the creation of highly realistic manipulated content, fueling disinformation campaigns that undermine public trust in institutions, media, and governance. Studies show that deepfake videos and AI-manipulated content are shared 70% more frequently than authentic content, amplifying their societal impact. Existing systems for verifying content authenticity struggle to keep pace with these advancements, and global regulatory responses remain fragmented and inconsistent. The lack of robust frameworks to address disinformation at scale has further eroded institutional credibility and public confidence.

The rise of generative AI and advanced content creation tools has transformed the landscape of information dissemination. However, this transformation has been accompanied by a growing mistrust in digital content. Generative AI enables the creation of highly realistic manipulated content, fueling disinformation campaigns that undermine public trust in institutions, media, and governance. Deepfake videos and AI-manipulated content present a significant challenge to trust in digital ecosystems. A study by the Center for Strategic and International Studies (CSIS) found that participants correctly identified AI-generated content only about 51.2% of the time, equivalent to a coin toss. Furthermore, a 2022 Royal Society report highlighted that even with warning labels, detection accuracy for deepfakes improved only marginally, from 10.7% to 21.6%. These findings underscore the difficulty of distinguishing deepfakes from authentic content, amplifying their societal impact (source: CSIS, Royal Society). Existing systems for verifying content authenticity struggle to keep pace with these advancements, and global regulatory responses remain fragmented and inconsistent. The lack of robust frameworks to address disinformation at scale has further eroded institutional credibility and public confidence.

At the same time, the management and use of data play a critical role in influencing trust. Data is often weaponized to mislead or manipulate, exacerbating the challenges of content verification and authenticity. Without ethical guidelines and transparent mechanisms for data governance, the potential for misuse grows, further undermining trust. The [2024 Edelman Trust Barometer](#) reveals a key paradox—while rapid innovation has the potential to advance prosperity, 61% of respondents emphasize that collaboration between business and government is essential for ensuring innovation is implemented in a trustworthy manner. Furthermore, 65% of respondents believe that businesses must take the lead in ensuring that innovation is safe, beneficial, and accessible. This highlights the need to align data-driven innovations with societal values and to incentivize the responsible use of data through well-designed frameworks and policies.

The Role of Policy and Technology Collaboration

Building trust in the digital information ecosystem requires a concerted effort to align policy and technology. Policymakers must establish governance structures that promote ethical behavior and transparency, while technologists must design solutions that operationalize these policies effectively. Collaboration between these communities can lead to innovative frameworks that are both actionable and scalable.

One critical area for collaboration is the development of integrated policy-technology solutions that address the dual challenges of misinformation and data misuse. For example, technical tools such as blockchain can provide provenance tracking for digital content, while policy mandates can establish standards for labeling and verifying information. Public-private partnerships can amplify these efforts by pooling resources and expertise from governments, technology companies, and civil society organizations.

To foster positive behavior, incentive structures must be embedded into these frameworks. Encouraging open-source contributions, rewarding transparency in AI systems, and offering recognition for ethical data practices can drive innovation in the right direction. Additionally, knowledge-sharing platforms that provide access to best practices, policy templates, and technical tools can accelerate the adoption of trustworthy solutions across regions and sectors.

Data as a Cornerstone of Trust

Data is both the foundation of innovation and a potential source of mistrust. To rebuild trust in digital ecosystems, it is essential to establish mechanisms that ensure data integrity, privacy, and ethical use. Policies must mandate transparency in how data is collected, processed, and shared, while technological solutions should enforce these policies through secure architectures and accountability mechanisms.

For example, AI models trained on biased or unverified datasets can amplify misinformation, highlighting the importance of curating high-quality, representative datasets. A [McKinsey report](#) from 2023 estimates that organizations leveraging transparent and ethical data practices see a 25% improvement in user trust and engagement. Incentives can play a crucial role in this area by rewarding organizations that prioritize ethical data practices and contribute to open data initiatives. These efforts not only enhance trust but also promote innovation by creating a shared pool of resources that benefits the broader community.

Toward a Trusted Digital Information Ecosystem

The need to cultivate trust in the digital information ecosystem has never been more urgent. Collaborative efforts between policy and technology communities can provide innovative pathways to address disinformation, empower users, and

foster a healthier digital space. By integrating ethical data practices and incentivizing positive behavior, these efforts can create a virtuous cycle of trust and innovation.

One initiative aiming to drive this change is the [Global Trust Challenge](#), developed in partnership with organizations like the OECD and IEEE. The IEEE Global Initiative on Trust Technology is an example of how collaboration can help develop and scale impactful solutions (source: [IEEE Trust Technology](#)). Additionally, partnerships highlighted by the OECD, IEEE, and UNESCO demonstrate the global effort to address AI-related trust issues through multi-stakeholder collaboration ([OECD Global Challenge Partners](#), [UNESCO Call for Global Challenge Partners](#)). By bringing together policymakers, technologists, researchers, and practitioners, this challenge fosters collaboration to develop scalable and impactful solutions. Participants work on integrating policy frameworks with technical tools to address trust issues at their roots, with a strong emphasis on data integrity and transparency. The challenge emphasizes diversity and inclusivity, ensuring solutions are adaptable across different cultural and geographical contexts.

The Global Trust Challenge exemplifies the potential of coordinated global efforts to build a trustworthy information ecosystem. By supporting and participating in initiatives like these, we can ensure that digital technologies serve as a force for good, empowering societies while upholding the integrity of information.

“Online Disinformation from a Freedom of Expression Perspective in the European Union” by Alejandra Michel⁴

The aim of the presentation was to examine the EU legislator's legal response to online disinformation, while drawing attention to the principles of freedom of expression (“FoE”). At the end of 2022, the European Union adopted the Digital Services Act (“DSA”)⁵. The regulation is truly a multifaceted instrument dealing with many aspects, but one of its major objectives is to ensure online trust⁶. To this end, the DSA requires intermediary service providers to comply with a set of tailored due diligence obligations that notably relate to content moderation. While these requirements contribute mainly to the fight against the dissemination of illegal content (i.e., a content or an activity that does not comply with current legislation, be it national or EU law), they can also, to a lesser extent, tackle harmful content (i.e. a complete legal content that could nonetheless cause harm). The EU legislator is setting this fight against illegal and harmful online content within a framework that effectively respects and protects human rights.

⁴ Senior Researcher in digital law and Head of the “Media Law” research unit at the Research Centre in Information, Law and Society (CRIDS/NaDI), University of Namur.

⁵ Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act), *O.J.*, L 277, 27 October 2022 (hereinafter “DSA”).

⁶ See notably DSA, rec. 9.

This context thus leads to the following question: “What (regulatory) method should be used to fight online disinformation while ensuring freedom of expression?”. To solve it, the presentation highlighted three key elements, which constitute the various steps of the reasoning:

1. The **qualification of online disinformation regarding the scale of FoE’ protection;**
2. The **establishment of measures that comply with FoE’ principles;** and
3. The **place of the fight against online disinformation within the DSA.**

The first step is the **qualification of online disinformation and the assessment of the relationship with freedom of expression.** Commitments to act against online disinformation are often expressed in general terms, without insisting on or without considering nuances in qualification. It is however essential to differentiate between harmful disinformation and illegal disinformation. While in most cases disinformation constitutes a harmful content under EU law, it can sometimes be an illegal one when a national provision renders a certain type of disinformation illegal, or when the disinformation operates through a content that is per se illegal (for example, a content that disinforms with the aim of inciting the commission of a terrorist act). This assessment must always be made on a case-by-case basis.

The second step is the **determination of the appropriate** measures to act against disinformation. In a human rights approach, the very nature of the content must be taken into account when determining the measures to apply. Thus, the dichotomy between illegal and harmful content is highly important for the determination of measures: illegal content does not benefit from the protection by FoE, whereas (legal) harmful content does. This means that we cannot act in the same way against disinformation when it is harmful and when it is illegal: while outright removal is acceptable for illegal disinformation, more flexible measures will have to be applied for harmful disinformation in order to comply with the three-steps-test⁷.

The third step examines **the place of the fight against online disinformation within the DSA.** On this point, there is some confusion and controversy, mainly due to the very recent nature of the instrument. Although no binding provision of the DSA explicitly mention disinformation, the recitals refer on several occasions to disinformation campaigns, misleading information or intentional manipulation of the service⁸. One of the key DSA mechanisms that applies to disinformation is the obligation imposed on very large online platforms (VLOPs) and very large search engines (VLSEs) to “assess” and “mitigate” systemic risks arising from the design, operation and use of their services⁹. Online disinformation can in practice encounter each of the four categories of systemic risk listed in the regulation. It

⁷ Communication from the Commission to the European Parliament, the Council, the European economic and social committee and the Committee of the regions, ‘Tackling online disinformation: a European Approach’, 26 April 2018, COM(2018) 236 final, pp. 1 and 4. Restrictions on freedom of expression must be lawful, pursue a legitimate aim and be necessary in a democratic society. Regarding the three-steps-test, restrictions to freedom of expression must be lawful, pursue a legitimate aim and be necessary in a democratic society.

⁸ See mainly DSA, rec. 2, 9, 69, 83, 84, 88, 95 and 104.

⁹ DSA, art. 34 and 35.

could constitute (i) a risk linked to the dissemination of illegal content, or a harmful content likely to have a negative effect on (ii) electoral processes, (iii) public health, or on (iv) the fundamental right to information or on the exercise of another fundamental right that could be impacted by the receipt of false information. While some have expressed fears for public debate with such provisions, M. Husovec recently highlighted the essential difference of scope between the assessment obligation and the mitigation obligation. According to him, while the extent of the risk assessment obligation is very broad (article 34 DSA), VLOPs and VLSEs “do not have to act upon all risks [...] and when they do act, they are not required to act against all risks in the same way” (article 35)¹⁰. This brings us back to FoE principles and the need to take into account the very nature of the content when determining the appropriate measures. According to M. Husovec, if the disinformation content is an illegal one, the EU Commission may require content-specific mitigation measures. On the contrary, if it is a harmful content, the EU Commission can only go for content-neutral mitigation measures (i.e., measures that are the same for the whole service irrespective of content) such as users’ empowerment¹¹.

To conclude, the regulatory approach to fight the online dissemination of illegal and harmful content constitutes only a small part of the solution but helps to ensure online trust. Building trust also requires getting the right measures and respecting human rights. We need to clean up illegalities online to ensure trust, but we cannot censor legitimate expressions, even if they are disturbing, because they are at the heart of any democratic society.

“(Re)informing and engaging readers in the Digital Age for a (francophone) knowledge platform” by Thomas Parisot

From the development of access to scientific publications, the challenge of scientific platforms is gradually giving way to the need to establish a relationship of trust and quality with the readership. Increasingly multi-channel communication - platforms, applications, social networks - and increasingly multi-formats - text, images, video, data - require a rethinking of the quality reading experience, in favour of quality, readable and original information.

With this in mind, Cairn.info, a platform dedicated to the dissemination of French-language scientific publications, has recently upgraded its service, with the introduction of video interviews with renowned French-speaking scientific authors, accompanied by reading lists and dossiers written to help as many people as possible discover the state of the art¹². The combination of academic and journalistic know-how is used here to facilitate the appropriation of texts, not only

¹⁰ M. HUSOVEC, “The Digital Service Act’s Red Line: What the Commission Can and Cannot Do About Disinformation”, pp. 4-5.

¹¹ *Ibidem*.

¹² <https://shs.cairn.info/rencontres?lang=en>

to popularize their subject matter, but also to stimulate scientific debate with embodied, well-argued positions, beyond the highly formatted production of scientific publications.

These formal innovations go hand in hand with an ambitious strategy of international promotion and translation of the French-language publications we distribute¹³: translation of interfaces, enrichment of English and Spanish metadata sets for the journals and books we distribute, subtitling of scientific videos, etc. The development of partnerships with more than 2,000 higher education and research institutions in 70 countries is also an essential part of this internationalization of knowledge in a variety of cultural and intellectual contexts, particularly in the humanities and social sciences, our core area of expertise¹⁴.

The issue of cultural barriers should also not be underestimated. More and more video interviews are being organized with authors from the South¹⁵ (Tunisia, Morocco, Senegal, etc.), in order to decentralize points of view and diversify analyses on issues that are directly relevant to the concerns of civil society: geopolitics, economic development, the environment, health, and so on. More broadly speaking, the growing acceptance of publications from academic institutions in the South is one of the keys to a genuine relationship of trust and interest with readers outside of Europe.

In addition, a dissemination tool must itself integrate artificial intelligence, and this is also what we have started to do, for example with RAG (Retrieval Augmented Generation)¹⁶. We are also deploying devices that highlight the attribution of analyses to authors and that genuinely invite the reading of synthesized sources. This is an essential point for us, as our platforms are spaces where research habits and access to knowledge are structured, for confirmed researchers as well as more or less advanced students.

Confidence in a response engine should not come at the cost of erasing the polemical nature of scientific debate or the revocable nature of certain proposals, even if they are based on scientifically valid arguments at a given time. Notwithstanding the importance of consolidating trust in scientific platforms and avoiding fueling the prevailing relativism, we believe that it is both useful and important in building trust to show that scientific claims also have their limits, and that delegation to a tool, however sophisticated, cannot be an absolute solution, and that there are flesh-and-blood authors behind every claim.

¹³ <https://shs.cairn.info/?lang=en> / <https://shs.cairn.info/?lang=es>

¹⁴ <https://apropos.cairn.info/en/mission/international>

¹⁵ Ex. : <https://droit.cairn.info/entre-discours-fondateur-de-droit>

¹⁶ <https://apropos.cairn.info/en/mission/ai>

“MetaBelgica: A Trusted Reference Database for Cultural Heritage Entities” by Hannes Lowagie and Sven Lieber

In a world where disinformation threatens the integrity of historical, cultural, and scholarly records, the role of accurate, transparent, and reliable data is more crucial than ever. At the Royal Library of Belgium (KBR), we are working with other Belgian federal scientific institutes to create MetaBelgica¹⁷, an official reference database that will provide trustworthy data on cultural heritage institutions.

Disinformation and Heritage Institutions

Heritage institutions hold vast amounts of data, ranging from historical records to artifacts of immense cultural significance. This data that is usually curated by professionals within the institutions, provides a certain degree of trustworthiness. And this trustworthiness is needed, because unreliable or manipulated data (on the web) can lead to serious problems, including:

- Copyright Disputes: Conflicting claims over ownership and usage rights of cultural materials.
- Nationality Debates: Misrepresentation of an author's or artist's nationality, leading to distorted historical narratives.
- Provenance Issues: Disputed origins or authenticity of artifacts, impacting their historical and cultural value.
- Research Challenges: Inaccuracies in the historical data of people, organizations, and objects can mislead researchers and scholars.

The trustworthiness of the data stems from the fact that heritage institutions ensure that their data is credible, accurate, and immune to disinformation. The credibility of heritage institutions is paramount because their collections often form the foundation of historical research and public knowledge. When data is questioned, it undermines trust in not only the institution but also the historical, cultural, and legal truths that these collections represent. For example, contested copyright or disputed provenance can lead to expensive legal battles or misinformed academic studies. Thus, it is critical for these institutions to safeguard their data against manipulation, errors, or bias, ensuring it is seen as a reliable source by both the public and academic communities. To earn global trust and confidence in internet-based data infrastructures, institutions must focus on the following principles:

- Credibility and Reliability: The data should be managed by experts with proven competence in handling cultural heritage. A long history of data curation and professional expertise lends authority to the data provided.
- Transparency and Accountability: Open data practices and transparent methodologies build user confidence. Institutions should make their data management processes visible and subject to external scrutiny to ensure

¹⁷ Information about the MetaBelgica project: <https://www.kbr.be/en/projects/metabelgica/>

their accountability.

- Integrity and Accuracy: Strict governance practices, regular audits, and manual curation help maintain the integrity and accuracy of the data, which is crucial in combating misinformation.
- Consistency and FAIR-ness: The data must adhere to international standards and best practices for data management, ensuring that it is Findable, Accessible, Interoperable, and Reusable (FAIR). This consistency across platforms and systems allows for seamless use and re-use of the data worldwide.

Libraries and cultural heritage institutions already possess much of the necessary infrastructure and expertise to manage vast amounts of data. By investing in these institutions and their personnel, it is possible to build a trusted, high-quality data infrastructure. The key lies in updating their expertise and fostering a synergy between content knowledge and modern data management practices. This approach not only builds upon an established foundation but also enhances the credibility and reliability of the data being managed.

The MetaBelgica Solution

MetaBelgica improves the status quo by combining data from different institutions and therefore increases the quality by data enrichment and duplicate detection. It is a collaborative project involving multiple Federal Scientific Institutes (FSIs) in Belgium, aimed at creating a unified reference database for cultural heritage entities. Within this project, running from 2023 to 2026, we make the combined data accessible according to the FAIR principles. This ensures that the data is not only accurate but also compatible with international standards, ensuring its global relevance and usability. The types of entities it covers include persons, organizations, time/events, and locations.

Why MetaBelgica is Trustworthy

MetaBelgica tackles disinformation by adhering to the principles of credibility, transparency, integrity, and consistency:

1. **Credibility and Reliability:** The collaborative nature of MetaBelgica, combined with the long history of data management and curation of the participating member institutions, enhances its credibility. The involvement of multiple FSIs ensures that data is carefully curated and verified by experts, making it a trusted resource.

2. **Transparency and Accountability:** MetaBelgica follows open science practices, which promote transparency in its methodologies and data management practices. Resources related to the project are openly shared on platforms such as GitHub and Zenodo, making them easily accessible to anyone interested. By embracing open-source principles, the project aligns with values of openness and long-term sustainability.

3. **Integrity and Accuracy:** MetaBelgica will implement strict data governance practices, ensuring that all data is curated and verified by professionals. Despite the rise of automated data curation tools, MetaBelgica relies heavily on manual oversight by experienced professionals. This hands-on approach ensures that the data is carefully vetted for accuracy, making it a more reliable source of information. The project also provides verifiable references for its data, which enhances the trustworthiness of the information.

4. **Consistency and FAIR-ness:** MetaBelgica assigns persistent identifiers to entities, ensuring that they are consistently referenced over time. Its alignment with international standards, such as RDF and ISNI, ensures that its data can be easily reused and shared across platforms, further increasing its value on a global scale.

Conclusion

MetaBelgica aims to combat disinformation by providing a reliable, high-quality dataset curated by professionals of the public sector. Through collaboration, human curation, and adherence to international data standards, it strives to become the definitive source for Belgian cultural heritage data. Instead of outsourcing data management to commercial entities or reinventing the wheel, MetaBelgica builds on the solid foundation of existing resources, ensuring that our cultural heritage remains trustworthy and accessible for future generations.

“Intermediaries, Crafted by Trustees” by Dr. Jörg Lehmann

Cultural heritage institutions (CHIs) are trusted institutions, and they should prolong the trust placed on them by their users into the infrastructures wherever they provide openly accessible datasets, as well as via educational activities.

Even in the (dis-)information age, CHIs are trusted institutions specialised in information retrieval. They are supported by the public, have reliably fulfilled expectations for centuries and represent consistency. In their role as institutions serving public education, they need to mediate between people overly enthusiastic of the capabilities of AI applications and luddites aiming to abandon this technology all-at-once, an opposition creating mistrust in AI. Explaining core fallacies of AI enthusiasts (like the incapability of Large Language Models to distinguish whether a statement is true or false), attenuating fears of job loss (supposing that humans will be replaced by machines) by explaining that the far more likely future will bring man-machine-co-operations, and disseminating a sound estimation of what AI applications are capable of should become a core activity of CHIs as well as further trusted institutions active in the educational sector.

Cultural heritage practitioners should directly exchange with the machine learning community in order to explain the specificities of digital cultural

heritage and thus build up mutual trust.

There are no datasets without biases (“statistical distortions”). Although representative data sets have been created in the recent past, they also reflect social biases. Cultural heritage datasets are equally characterised by biases. They are clearly evident due to the historical distance, thus reflecting the positionality of the dataset producer, and they should be documented in accompanying datasheets (“instruction leaflets”) functioning as **intermediaries**, because machine learning models trained on such datasets scale these biases. They can therefore lead to harms for social groups, for example through a disadvantageous, stereotypical or discriminatory representation of such populations. CHIs need to produce datasheets accompanying the datasets published by them or similar forms of documentation to explain the inherent biases to dataset consumers and thus to create trust. These datasheets should be machine-readable, interoperable with catalogues and digital infrastructures, reliable and referenceable. Moreover, cultural heritage practitioners need, in their roles as **trustees**, curators of datasets and authors of datasheets, to mediate between the domains of cultural heritage and machine learning on face-to-face occasions like conferences, seminars, and workshops.

Cultural heritage institutions should ensure the authenticity of man-made contents by using international standards covering both integrity at the file level and authenticity at the content level to build trust in data infrastructures.

Content produced by generative AI applications will soon lead to a true crisis of representation. Without the use of analytic tools, users will no longer be able to distinguish synthetically generated texts, images, audio-visual materials, news, notifications, reports etc. from man-made content. This forms the core dilemma of the disinformation age. However, companies developing AI applications are interested in using authentic, man-made content and want to avoid using synthetic data for the training of AI applications in order to avoid model collapse. CHIs as holders of man-made content from publishing houses, photographers, archived original audio-visual materials as well as curated metadata need therefore to ensure the authenticity of the digital assets which they feed into infrastructures by using e.g. the International Standard Content Code ISCC (ISO Norm ISO 24138).

Key Recommendations from the Infrastructure Track

Prepared by Julie M. Birkholz & Amanda Hemmons, KBR

Fighting misinformation and disinformation is a civic act. It is democratically urgent, it requires many actors to accomplish, it needs to be a global effort, and it also requires an agreed upon definition of trust (among other terms). The objective of OPERAS TrustOn2024 workshop was to evaluate the Internet's trust dimensions through diverse partnerships, create a multi-stakeholder model for responsible Digital Public Infrastructures, and promote best practices for maintaining online content quality amid AI-driven disinformation.

Infrastructure can be many forms including technical (physical), social, governance, legal, political, cultural, economic. And the speakers reflected these diverse forms. The infrastructure recommendations include:

- Establishing a digital bill of rights and a digital asset registry to safeguard and manage digital resources;
- Reinforcing regional networks to develop adaptive regulation sandboxes to facilitate compliance with legislation and allow for experimental approaches in living labs. Emphasising multidisciplinary collaboration, the proposals called for engagement from National Research and Education Networks (NREN), heritage institutions, publishing houses, government organisations, and (big) tech companies.
- Developing clear rules of participation to ensure transparency in funding, and standardising practices through the adoption of international standards and common protocols.
- Defining terms and establishing common terminology for effective communication and implementation between stakeholders.

2.2. Science Track

The Science track at the TrustOn2024 event focused on mis- and disinformation in the context of **science**. Specifically, this track explored (digital) mis- and disinformation by focusing on fostering **trust in science**. It derived from the science-based experience of four EU-funded Horizon Europe Projects, namely [COALESCE](#), [IANUS](#), [POIESIS](#) and [VERITY](#). The track took a practice-oriented approach to develop policy recommendations on addressing mis- and disinformation through enhancing trust in science. The track comprised three sessions discussing the topic from three angles: (1) relevant **audiences** for mobilisation of science communication; (2) **participatory methods** for navigating online (mis)information; (3) collaboration between actors to enhance trust in science - **Stewards of Trust**. To better understand the topic, the session was supported by science-based presentations and discussions around the challenges of trust in science.

The scepticism surrounding the COVID-19 vaccine rollout echoed the hostility to climate science we have seen in recent years. Science is vulnerable to mis- and disinformation. Nevertheless, the pandemic and increased exposure to science and scientists have highlighted the role of science in fighting disease around the world and, ultimately, the role of science in people's lives (Wellcome Global Monitor, 2021). Science helps to understand and tackle grand contemporary challenges and assist with achieving the Sustainable Development Goals (SDGs) in relation to, for instance, global health, climate change, sustainability, and inequalities. However, scientific advancements are not enough if they are not accompanied by sufficient consideration of societal perceptions, public support and trust in science. While trust in science and scientists around the globe is relatively high (European Commission, 2021; Wellcome Global Monitor, 2020), the fact that people, in general, trust science, does not mean that they follow science-based recommendations on specific issues (Ratner & Riis, 2014). Trust in science is not constant, it is a changeable and dependable phenomenon. The hesitancy or resistance to follow science-based recommendations is caused by several factors, including (1) the nature of science; (2) the interdependence of science, politics and economy; (3) changing research environments; (4) the changing ecosystem of trust.¹⁸

¹⁸ The Introduction of the science track was written by dr. Charlotte Bruns, Erasmus University Rotterdam, EU-funded IANUS and COALESCE projects, dr. Agata Gurzawska, Trilateral Research IE, Coordinator of the EU-funded VERITY project and dr. Tine Ravn, Aarhus University, Coordinator of the EU-funded POIESIS project.

“Trust and Disinformation in the Digital Age: The Role of Collaboration in Science Communication” by Julija Baniukevic

In today's digital age, the landscape of trust in science and the dissemination of reliable information has become increasingly complex, particularly in the wake of crises such as the COVID-19 pandemic. This multifaceted collaboration is essential not only to combat disinformation but also to enhance public trust in scientific discourse. The pandemic served as a critical case study in understanding the dynamics of trust and disinformation, revealing how quickly misinformation can spread and how deeply it can influence public perceptions of science and health [1]; [2]; [3]; [4]

Analysis of media panel in Lithuania [5]. A thought-provoking discussion in Lithuania was organised by the Research Council of Lithuania, addressing the pivotal question “Do You Trust in Science?” This dialogue revealed the challenges faced in science communication, where the reliability of scientific information was scrutinised amidst an overwhelming flood of data. Participants included four focus groups: science communicators, science journalists, scientists, and students, who engaged in discussions that underscored the importance of critical thinking and media literacy in navigating the digital information landscape. Key observations drawn from the first media panels conducted as part of the IANUS project provide valuable insights into the complex dynamics of trust in science within contemporary society, particularly from a Lithuanian perspective.

Main findings [5]: One primary theme was the role of artificial intelligence (AI) in shaping scientific communication. While AI can assist researchers in drafting articles and analysing data, it also raises significant concerns regarding the authenticity and reliability of information, especially in non-peer-reviewed contexts. Participants noted that the proliferation of AI-generated content complicates the ability of readers to discern credible sources from misleading narratives, creating a landscape where misinformation can thrive. This challenge is exacerbated by the rapid spread of misinformation online, often outpacing efforts to correct false information and leaving the public vulnerable to confusion.

Another critical observation made during the discussions was the necessity of clearly communicating the boundaries of scientific expertise. Participants emphasised that scientists must be aware of their limitations and refrain from commenting outside their areas of expertise to maintain the integrity of scientific communication and foster public trust. This principle was particularly evident during the pandemic, where experts in epidemiology provided essential insights into virus transmission and public health measures, while voices from unrelated fields sometimes muddied the waters with opinions that lacked a scientific basis. The discussions highlighted the importance of ensuring that scientific

communication is both accurate and aligned with the expertise of the individuals sharing it.

The responsibility of science journalists was also a focal point of the discussions. Journalists play a vital role in bridging the gap between scientific communities and the public; however, they must adhere to ethical standards that prioritise accuracy and context. Misrepresentation of scientific findings can exacerbate public distrust and contribute to the spread of disinformation. Participants noted that collaborative efforts between journalists and scientists are essential to ensure that complex scientific concepts are accurately conveyed to diverse audiences. For instance, when journalists collaborate closely with scientists, they can produce articles that not only explain the science but also contextualize the findings within broader societal issues, thereby enhancing public understanding and engagement.

Public literacy regarding scientific processes emerged as another crucial factor in fostering trust. Many individuals lack a comprehensive understanding of how science operates, which can lead to skepticism toward scientific findings. During the pandemic, misinformation overshadowed factual data, often fueling public fear and resistance to health guidelines. This gap in understanding necessitates improved science communication strategies that not only inform but also engage the public, making scientific processes more accessible and relatable. Educational programs that demystify scientific inquiry and promote critical engagement can empower individuals to make informed decisions regarding their health and well-being.

Additionally, the role of social media in disseminating information was acknowledged as a significant factor influencing public perceptions. Platforms such as Twitter, Facebook, and Instagram have become primary sources of news for many individuals, yet they often prioritise engagement over accuracy, facilitating the rapid spread of sensationalised or misleading information. The panels underscored the need for social media platforms to take responsibility for curbing misinformation. Collaborative efforts between tech companies and scientific communities could lead to the development of tools that help users identify credible sources and distinguish facts from misinformation. For example, initiatives could include fact-checking features or algorithms designed to promote scientifically accurate content.

In conclusion, the interplay between trust and disinformation in the digital age necessitates a concerted effort among various stakeholders, including scientists, policymakers, educators, and communicators. Collaborative frameworks that unite these groups are essential for effective science communication, particularly during crises. By fostering critical thinking, clarifying the boundaries of expertise, ensuring ethical journalism, enhancing public literacy, and addressing the psychological dimensions of information processing, we can build an ethical framework for knowledge sharing.

References:

1. Case Study „Lessons Learned From Dear Pandemic, a Social Media–Based Science Communication Project Targeting the COVID-19 Infodemic“. Sandra S. Albrecht, et al. Public Health Reports Volume 137, Issue 3, May/June 2022, Pages 449-456.
<https://journals.sagepub.com/doi/epub/10.1177/00333549221076544>
2. Infodemic: disinformation and media literacy in the context of COVID-19. J. Posetti et al. Internet Sectoral Overview Number, 3 September, 2021 Year 13. Website: <chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://nic.br/media/docs/publicacoes/6/20210923161126/internet_sectoral_overview_year-13_n3_infodemic.pdf>
3. „Who Speaks for Science?“ Douglas Allchin. Science & Education (2022) 31:1475–1492. Website: <https://link.springer.com/article/10.1007/s1191-021-00257-4>
4. Model uncertainty, political contestation, and public trust in science: Evidence from the COVID-19 pandemic. S.E. Kreps and D. L. Kriner, Sci. Adv. 2020; 6 : eabd4563 21 October 2020. Website: <chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/<https://www.science.org/doi/pdf/10.1126/sciadv.abd4563>>
5. Insights from media panel in Lithuania, lanus project, 2024. Website: <https://trustinscience.eu/newsroom/news/insights-from-the-media-panel-discussion-in-lithuania/>

“Re-defining the ecosystem of trust in science: the role of stewards of trust in changing research environments” by Agata Gurzawska

Science, technology and innovation offer a pathway to achieving Sustainable Development Goals (SDGs) and fostering the transformation towards sustainable societies and economies. However, scientific advancements must be accompanied by consideration of societal perceptions, public support and trust in science. Ensuring public trust must start with acknowledging that research and innovation (R&I) are inherently connected to society (Mazzucato, 2019). Trust in science is shaped by broader historical, social, economic and political contexts. It is an integral aspect of democratic governance since trust relationships are the building blocks of social cooperation (Warren, 2018). Although there is a popular narrative about a crisis of trust in science, survey data suggests otherwise (European Commission, 2021; Wellcome Global Monitor, 2020). However, the fact that people, in general, trust science, does not mean that they follow science-based recommendations on specific issues (Ratner & Riis, 2014).

Trust in science is a nuanced and complex phenomenon based on relationships between trusters (a diverse group of members of the public) and trustees (individual or collective science actors). This relationship is increasingly challenged by the fact that the ecosystem of trust in science is changing. Scientific research and research environments have evolved over the decades (von Schomberg, 2024). This transformation involves the internationalisation and globalisation of R&I and multi-actor scientific projects; the multi-, inter- and transdisciplinary character of research; and the increasing use of AI-based technologies for research. While trust in science is greater than trust in politics and the economy (Bromme et al., 2022), nowadays, science is inevitably intertwined with both (Funk et al., 2019; Melo-Martin, 2021; Rathenau Institute, 2021). This entanglement exacerbates power relations and affects trust in science. The ‘economisation’ and ‘privatisation of knowledge’, where

science focuses on the competitive character of research and innovation (R&I) may erode an essential basis of public trust in science: the conviction that science serves the common good (Peters, 2015). Those shifts may challenge the imperative for science, research and innovation to benefit individuals and societies and uphold ethical standards, academic freedom and human rights.

The ecosystem of trust in science is changing, and trust is moving into the hands of the many. The traditional trust ecosystem was primarily composed of institutions and professionals who were held publicly accountable (van Dijck, 2021) and fostered not only 'trust in science' but 'trust within science' (Hendriks, et al., 2016:146). Today, this system has expanded to include e.g. companies, public-private partnerships, collaborative multi-actor and multi-sector research consortia, influencers, citizen science participants and civil society organisations (CSOs), as well as social media platforms, knowledge and education platforms (such as Wikipedia and Massive Open Online Courses, MOOCs), and generative artificial intelligence (AI) chatbots. In this way, the traditional model of linear information flow, governed by established stakeholders such as experts, institutional filters, and gatekeepers (van Dijk & Alinejad, 2020), is undergoing a significant transformation. This transformation has also been inscribed in the reshaping process from 'institutional trust' to new forms of 'distributed trust'. Such reorganisation of the trust ecosystem has been enabled by the internet and high-tech platforms significantly influenced by the digital revolution (Botsman, 2017; Thunert, 2021). Thus, trust in science is a complex social phenomenon (Origgi, 2022) connected to the reputation of scientists as experts, the scientific community, scientific institutions and the science system as such.

In this new landscape, trust is co-governed by 'new' stakeholders, who act as Stewards of Trust with a critical role and responsibility in nurturing trust in science. Trust in science extends beyond just the research itself. Effective communication, transparent practices, and equitable distribution of benefits are equally important. In this regard, trust is not merely about scientific results but how they are presented and how the public engages with them. Therefore, the roles of Stewards of Trust span various domains, including science production, education, communication, policy, funding and advocacy, implementation, oversight and protection as well as science-society facilitation and citizen science. Their roles and responsibilities emanate either from their mandate and mission or their *de facto* power and influence.

The VERITY research project's findings, integrating insights from surveys, fieldwork, and thematic analysis, provide a comprehensive understanding of the trust dynamics and highlight the need for developing collaborative relationships within the ecosystem of trust. It is essential that networking, communication, information sharing, and collaboration occur within the ecosystem to develop relationships between traditional and non-traditional Stewards of Trust and society. VERITY findings show a consensus that the science-society relationship needs to be nurtured and facilitated. There is a need for multiple stakeholders to safeguard trust in science and establish long-term and meaningful partnerships across important societal institutions, such as civil society and academic institutions, as a

way of amplifying voices in science and, by extension, trust in science. Open science approaches, public engagement via citizen science and citizen engagement, outreach activities and public involvement in decision-making should be continuously supported and promoted. Stewards of Trust and those functioning within the science ecosystem should credibly demonstrate that, and how, they are working for the public interest and values ensuring the equal right of all people to share and enjoy scientific progress and its applications (Article 27(1) UDHR; Article 15(1)(b) ICESCR). Because scientists and policymaking actors are interdependent, evidence-based governance, regulations and scientists' involvement in official governmental bodies are necessary. This calls for a coordinated effort from both parties. Transparency and clarity are crucial regarding roles, responsibilities, funding sources, and connections to private and political actors within the trust ecosystem.

In the VERITY project, we recognise the collaborative nature of science and trust and changing research environments. Therefore, we advocate for the individual responsibility of scientific actors, along with the shared responsibility of the scientific community for nurturing trust in science through building and strengthening inclusive and meaningful relationships within the science ecosystem.

Acknowledgements

Evren Yalaz, Orla Drummond, and Mariano Martín Zamorano Barrios from Trilateral Research Ireland contributed to this research and report.

The research leading to this report received funding from the European Union under grant agreement No. 101058623 (VERITY).

References

- Botsman, R. (2017). *Who can you trust? How technology brought us together and why it might drive us apart*. UK: Portfolio Penguin.
- Bromme R. & Hendriks F. (2022). "Trust in Science: When the public considers whom to trust - the example of COVID-19". In: Mayer R C and Mayer B M (eds) *A Research agenda for trust: Interdisciplinary perspectives*. Northampton, Massachusetts: Edward Elgar Publishing.
- European Commission. (2021). Special Eurobarometer 516: European citizens' knowledge and attitudes towards science and technology. Retrieved from <https://europa.eu/eurobarometer/surveys/detail/2237>.
- Funk, C., Hefferon, M., Kennedy, B., & Johnson, C. (2019). "Trust and mistrust in Americans' views of scientific experts". *Pew Research Center*, 2, 1-96., ISO 690.
- Hendriks, F.; Kienhues, D.; Bromme, R. (2016). "Trust in Science and the Science of Trust. In *Trust and Communication in a Digitized World: Models and Concepts of Trust Research*"; Blöbaum, B. (Ed.); Springer International Publishing: Cham, Switzerland; (pp. 143–159).
- Mazzucato, M. (2019). *Governing missions in the European Union. Independent Expert Report*.
- Melo-Martín, D. (2021). "Vaccine Hesitancy" by Maya J. Goldenberg: The Publics Are Not Ignorant but Distrustful. *Philosophy of Medicine*, 2(2), 1-7.
- Origgi, G. (2022). "The social indicators of the reputation of an expert." *Social Epistemology*, 36(5), 541-549.
- Peters, H. P. (2015). "Science dilemma: Between public trust and social relevance". *Euro Science*. Verfügbar unter <https://www.euroscientist.com/trust-in-science-as-compared-to-trust-in-economics-and-politics/>
- Rathenau Institute (2021). *Trust in science in the Netherlands* (2021 survey).
- Ratner, R. K., & Riis, J. (2014). "Communicating science-based recommendations with memorable and actionable guidelines". *Proceedings of the National Academy of Sciences*, 111, 13634-13641.
- Thunert, M. (2021). "Want trust in (scientific) experts and expertise?". *American Culture Studies*, 30, 61-85.

- van Dijck J, Alinejad D. (2020). "Social media and trust in scientific expertise: Debating the Covid-19 pandemic in the Netherlands". *Social media + Society* 6: 981057.
- van Dijck, J. (2021). "Governing trust in European platform societies: Introduction to the special issue". *European Journal of Communication*, 36(4), 323-333.
- von Schomberg, R. (2024). "Towards a new ethos of science or a reform of the institution of science?" In R. von Schomberg, C. Mitcham, S. Leonelli, F. Lukas, A. Nordmann, & M. Edwards-Schachter (Eds.), *Towards a New Ethos of Science or a Reform of the Institution of Science? Merton Revisited and the Prospects of Institutionalizing the Research Values of Openness and Mutual Responsiveness*.
- Warren ME (2018). Trust and democracy. In: Uslaner EM (ed.). *The Oxford Handbook of Social and Political Trust*. Oxford: Oxford University Press, (pp. 75–94).
- Wellcome Global Monitor, (2021), "How Covid-19 affected people's lives and their views about science". Available online: <https://wellcome.org/sites/default/files/2021-11/Wellcome-Global-Monitor-Covid.pdf>

“Responsible science communication for trustworthy science: First insights from the IANUS project” by Hendrik Berghäuser

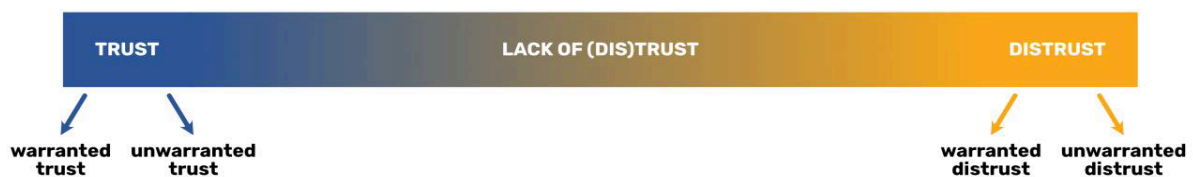
The research project “Inspiring and Anchoring Trust in Science, Research and Innovation” (**‘IANUS’**) funded by the European Commission¹⁹ strengthens warranted trust in science, research and innovation at a systemic level by encouraging participation in research as a co-creative and inclusive process, sensitive to societal values, concerns and needs. Value-driven research means that researchers are engaged and sensitive to societal needs and values, as well as dedicated to helping society deal with urgent global challenges. For this reason, the IANUS project delivers valuable insights for the Science Summit at the 79th United Nations General Assembly. The contribution from IANUS for this input is twofold: First, IANUS helps to **unpack the black box of “trust in science”** by developing a **conceptual frame** and second, can help to explain why in particular science communication is central to trust in science and to engage with different audiences about science and technology in times of crisis.

The **IANUS project has four objectives**: (1) Enable societal stakeholders to distinguish valid from unsubstantiated trust and healthy from unfounded distrust, (2) enable societal stakeholders to deal with the uncertainties, incompleteness and epistemic pluralism inherent in scientific knowledge, (3) enable researchers to foster trust in science through radical methodological transformation, making research inclusive, transparent and responsive to societal needs and concerns, lowering barriers between researchers and societal stakeholders, and (4) enable researchers to conduct relevant, engaged and value-driven research while foregoing partiality, ideological biases and conflicts of interests.

In general, **trust** (in science) is a **complex mental state** which involves vulnerability and risk-taking by consciously giving up (some) control and being ready to feel betrayal. Thus, trust acts as a mechanism that aids us in navigating through uncertain scenarios. It involves a conscious decision to take a leap of faith without

¹⁹ This project is receiving funding from the European Union's Horizon Europe research and innovation programme under grant agreement No. 101058158.

an assured safety net. **Public trust in science can mean several things:** trust in what scientists say (epistemic trust) and what they do, trust in a specific method (e.g. with regard to questions of reproducibility and replication of their work), trust in research results, trust in individual scientists or research organisations, as well as trust in science as a system. **Both trust and mistrust may be rational or irrational.** Perhaps we have no reason to trust, but we trust anyway, irrationally. This kind of trust is unwarranted. Alternatively, there may be reasons to trust, but we refuse them and distrust. This kind of distrust is unwarranted. Irrational trust or irrational mistrust means that we are not fully cognizant of risks; that we do not have sufficient knowledge of the trustee; that we do not fully take evidence or previous experience into account. Whereas irrational trust ignores the risks that something could betray trust, irrational mistrust overestimates the possibilities of betrayal. Both **trust and distrust have a gradual nature.** Mere lack of trust doesn't automatically translate to active distrust; it could mean a neutral stance of neither trust nor distrust. Absence of trust should not be equated with distrust. Rather, absence of trust describes the centre stage of this continuum between trust on the one hand and distrust on the other hand, where a trustor neither trusts nor distrusts a trusting object. Distrust or the absence of trust can be described as an attitude of healthy skepticism, as the following graphic shows:



Furthermore, **trust is a relationship**, as it involves a **positive attitude of the trustor towards the trustee.** As a trustor you trust (attitude) the trustee to do something (activity or commitment). Also, trust is influenced by many personal characteristics such as personal values, education, socio-economic status, world view, information sources but also previous experiences with the trustee.

Science communication in particular is often seen as a **proxy for trust in science**, as for many citizens, science is only accessed indirectly. Science communication therefore serves as an intermediary actor between science and society and it builds linkages to society through dialogical and participatory modes of interaction. However, in the IANUS project we identified various systemic **challenges**, both for the science communication system, but also within the science system as such:

Challenges within science comm. system	Challenges within the science system
<ul style="list-style-type: none"> • Must compete in an information-saturated digital media environment • Public is increasingly exposed to contradictory and inconsistent information 	<ul style="list-style-type: none"> • Pandemic of overpromising: Over-enthusiastic and sensational reporting about (preliminary) research findings leading to

<ul style="list-style-type: none"> • Fast spread of misinformation, often fueled by political agendas / actors • Increasing difficulties in addressing and engaging public • High funding pressure and closure of prestigious science-related magazines • Pressure to conform to sensationalism and political agenda • Paradox: Distilling distrust through dissemination 	<p>exaggerated expectations among citizens</p> <ul style="list-style-type: none"> • Researchers feel pressured by research funders to inflate the (potential) outcomes of their research • Politicisation and weaponisation of science (conflicts of interests) • Scientists increasingly confronted with hate speech and hostility
--	--

With regard to identified challenges we elaborated **first recommendations** that will be further validated with various actors of the science governance system:

- We call for more reflection by scientists and science communication initiators about their motives and desired outcomes
- If a Science Communication activity has a promotional or persuasive dimension, then it should be conducted ethically
- We call for more reflection about own biases and preconceptions
- We propose responsible promise management and epistemic humility
- We call for moving away from a rhetoric of deficit
- We call for resilience against editorial demands that compromise the integrity of science communication
- We see a need for strategies and training for scientists and science communication actors to deal with critique, hate speech and hostility.

“Fact or fiction: The battle against scientific disinformation in the social media arena” by Panagiotis Monachelis

The following work is a part of the project VERITY, a funded project by the EU that seeks to rebuild the relationship between science and society and focuses on re-defining the ecosystem of trust and stewards of trust and connecting them with effective methods of enhancing trust in science. VERITY is centred around three questions; What do people trust? Whom do people trust? How is trust built? VERITY attempts to address these questions through interdisciplinary research tasks aiming to develop a Protocol of Recommendations providing guidelines and methods for traditional and non-traditional ‘Stewards of Trust’ to enhance trust in science and facilitate science-society co-creation.

Among the various tasks, the analysis of data from Online Social Networks (OSNs) is included in order to examine the issue of trust in science and scientific disinformation through social media.

SOCIAL MEDIA IN NUMBERS

To realise the necessity of studying OSNs that have become a part of our daily life, let’s take a look at the following numbers:

- 59% of EU individuals using social networks in 2023. Countries with the highest use are Denmark (91%), Cyprus (83%) and Hungary (81%).
- 26% of people in the EU use social media platforms and blogs to access news. TV dominates with 75%.
- Online Social Networks and blogs are the second most used source of information (29%) about science and technology after TV (63%).
- Traditional media seem to be more trustworthy for the people in the EU (49%) in comparison with social media (14%) and other online news platforms (11%).
- In the last 5 months of the 2016 US presidential election, 25% of 30 million tweets spread either fake or extremely biased news.

Additionally, as the world's population from 2017 to 2023 increased from 7.8 to 8.04 billion, the number of social network users worldwide increased over the same period from 2.73 to 4.9 billion revealing that more than half of the world's population are on social media.

Furthermore, considering the three main characteristics of OSNs, a) the large volume of information that circulates within them, b) the variety of the overall content, and c) the velocity with which information circulates, the need for a settlement of trust in science through OSNs becomes more apparent.

Regarding the scientific disinformation through OSNs, some questions are raised:

- How can information sources be evaluated as trustworthy and untrustworthy?
- Who has the most influence on the public?
- What are the characteristics of disinformation messages?
- How do people react to disinformation? What feelings does disinformation evoke?

VERITY'S METHODOLOGY AND FINDINGS

VERITY seeks to address the above question based on specific methodology examining a dataset of posts from Twitter (X) platform related to COVID-19 vaccination. The methodology consists of the following stages:

- i) Evaluation of credibility in OSNs: One of the initial stages of the data processing was the evaluation of information credibility. The dataset was studied from the perspective of articles that were included in the posts from external sources, and then external sources could be evaluated as trustworthy and untrustworthy according to fact check journalistic platforms providing a crucial finding, **during the COVID-19 vaccination period, one in four messages on social networks came from an unreliable source.**
- ii) Analysis based on the numerical data of the posts, such as number of likes and reshares. This analysis provided results revealing the influential accounts and the influential posts. Examining the dataset, it is observed

that in general during the vaccination period untrustworthy messages with a great number of likes and reshares were at slightly lower levels than trustworthy ones. **It is noteworthy, however, that at the beginning of the vaccination period untrustworthy messages have significantly more likes and reshares.**

- iii) Social network analysis examining the network of the users: The social network analysis examined the network of the users who shared trustworthy and untrustworthy information revealing that **users engaging with trustworthy sources appear more likely to engage with information from multiple sources while users engaging with untrustworthy sources appear slightly more clustered around single sources.**
- iv) Analysis of the content of the messages: The mass analysis of all messages content can give a visualisation showing which users are highly engaged and which content is engaging users. **Specific conspiracy theories were detected and also accounts related to disinformation were revealed, inferring who and what concerns the public.**
- v) Deep learning techniques for sentiment and subjectivity analysis: AI techniques applied to the dataset can detect what causes negative or positive sentiment to the public and also identify the stance of users and detect possible controversial topics. **The analysis yielded significant results by identifying specific users and specific messages that elicited negative public sentiment associated with disinformation.**
- vi) Additional analysis of Eurobarometer data in the field of trust in science through social media conducted as well: Eurobarometer analysis emphasised on the demographics as OSNs lack this kind of data. The analysis identified some important findings such as **which countries people use social media as a source of science and technology news, younger people are more likely to choose social media to get science and technology news, and social media users have a higher tendency to believe in disinformation and conspiracy theories.**

RECOMMENDATIONS TO POLICYMAKERS

According to the findings some recommendations could be suggested to policymakers:

- Political ideology seems to be one of the major factors that influence people fostering scepticism towards science. Politicians must act in concert to strengthen trust in science leaving scientific issues away from political expediencies
- The high circulation of fake anti-scientific news on social networks reveals the urgent need to educate the public about the proper use of social networks by fostering scepticism. Training programs related to the proper use of social media could help the public to cultivate a critical mindset.
- Communities and groups of people may be identified from the social network analysis and the Eurobarometer analysis that can be more

vulnerable to disinformation. Policymakers can approach these groups with specific actions.

- Social media platforms should integrate mechanisms for making available the information of trustworthiness. EU Funded project EUNOMIA has addressed this issue providing the creation of a social network that prioritises the term “trust”.
- Disinformation is a particularly worrying phenomenon in times of crisis. Especially in the beginning of a crisis, debunking fake news is crucial. Policymakers need to have quick reflexes in critical times.
- Policymakers could effectively enhance trust in science through social media recruiting influential persons or organisations.

"Guidelines on the Ethical Use of Artificial Intelligence in Education and Research" by Reda Cimmperman

Ethical Use of Artificial Intelligence in Science and Studies

The digital age has transformed how information is disseminated and consumed, with Artificial Intelligence (AI) playing a pivotal role in this evolution. While AI offers remarkable potential in enhancing the efficiency and scope of scientific research and education, it also poses significant challenges, particularly concerning trust and disinformation. This abstract examines the ethical considerations around trust and disinformation in the digital age, as outlined in the ["Guidelines on the Ethical Use of Artificial Intelligence in Science and Studies"](#) by the Ombudsperson for Academic Ethics and Procedures of the Republic of Lithuania.

AI, especially generative AI, has revolutionised content creation, enabling the generation of text, images, and other media with minimal human intervention. However, this capability also makes AI a potent tool for spreading disinformation. AI can generate fake news, deepfakes, and other misleading content that can be easily distributed across digital platforms, undermining public trust in information and institutions. The guidelines highlight the necessity of recognizing and mitigating the risks associated with AI's potential to amplify disinformation.

Trust in AI-generated content is a central concern in the digital age. The guidelines emphasise that AI should be used transparently, with a clear distinction between human and AI-generated content. Academic institutions are encouraged to implement robust verification processes to ensure the reliability and authenticity of AI-generated outputs. This includes developing policies that require the clear identification, citation, and contextualization of AI-generated material in research and educational settings.

The ethical use of AI in education and research is paramount to maintaining trust in these institutions. The guidelines stress the importance of transparency, accountability, and the ethical deployment of AI technologies. Institutions are

urged to critically evaluate AI tools to ensure they do not contribute to the spread of disinformation or undermine trust in academic outputs. This includes a commitment to maintaining academic integrity, where the results generated by AI are rigorously assessed and validated against established scientific methods.

To combat the spread of disinformation, the guidelines recommend a multi-faceted approach:

- 1 *Critical Evaluation*: Researchers and educators are encouraged to critically assess AI-generated content, recognizing the limitations and potential biases inherent in AI systems. This includes verifying the sources and validity of AI outputs before integrating them into academic work.
- 2 *Education and Awareness*: Increasing AI literacy among students, researchers, and educators is essential for fostering an environment where AI tools are used responsibly. This involves training individuals to recognize and counteract disinformation, ensuring they can discern between credible and non-credible sources.
- 3 *Regulatory Compliance*: Institutions must ensure that their use of AI tools complies with legal frameworks designed to protect against the spread of disinformation. This includes adhering to the European Union's General Data Protection Regulation (GDPR) and other relevant laws that safeguard against the misuse of personal data and intellectual property in AI applications.

In an era where trust in digital content is increasingly fragile, the guidelines underscore the need for a cautious and principled approach to AI. Institutions are encouraged to foster a culture of transparency and ethical responsibility, where AI is seen as a tool to enhance, rather than replace, human judgement and creativity. This involves not only adhering to ethical standards but also actively participating in the broader societal effort to build and maintain trust in digital information.

The integration of AI into science and education presents both opportunities and challenges, particularly in addressing the issues of trust and disinformation. The guidelines provide a comprehensive framework for the ethical use of AI, emphasising the importance of transparency, accountability and critical evaluation. By adhering to these principles, academic institutions can help mitigate the risks associated with AI-generated disinformation and foster a digital environment where trust in information is preserved and strengthened.

“Science hoaxes, diminishing trust” by Isabel Mendoza-Poudereux and Carolina Moreno-Castro

The proliferation of science hoaxes and misinformation, particularly during periods of crisis, poses significant threats to public health and safety. The COVID-19 pandemic underscored the urgent need for more effective strategies to counter

misinformation, especially on private messaging platforms where verifying information is challenging. Ensuring public access to accurate, scientifically grounded information is crucial for addressing complex global challenges and restoring confidence in scientific processes. Without such efforts, the public remains vulnerable to harmful hoaxes that undermine informed decision-making and may endanger lives. While the public generally trusts scientists, this trust is not universal (Hoogeveen et al., 2022). Evidence indicates that trust in science is politically polarised, with conservatives displaying lower levels of trust than progressives (Azevedo & Jost, 2021; Altenmüller et al., 2024). In this context, the continued dissemination of disinformation further erodes confidence in science and scientific institutions. From science communication research, we addressed the following approaches:

Science-based information refers to knowledge derived from scientific methods and principles supported by empirical evidence, peer-reviewed research, and reproducible results. It provides a reliable foundation for addressing complex issues, making informed decisions, and guiding policies and practices in medicine, climate change, environmental science, engineering, and social sciences. Furthermore, science communication must consider potential conflicts of interest among the researchers involved in each study.

Misinformation, in contrast, consists of false information that individuals may believe to be accurate, whereas disinformation is the deliberate spread of false information for specific purposes. Since 2016, the term "fake news" has gained prominence, particularly during the U.S. presidential election, as a descriptor for such harmful messaging. Disinformation manifests in various forms, including memes, manipulated media, propaganda, and hoaxes, following the seven types of Mis- and Dis-information, based and extended on Wardle's taxonomy (2017).

A science hoax is a type of misinformation that deliberately propagates false or misleading claims under the guise of scientific authority and could promote harmful behaviours, such as rejecting vaccinations or endorsing unproven treatments. These hoaxes often involve exaggeration, fabrication, or distortion, exploiting scientific terminology to create an appearance of credibility while lacking legitimate scientific support. Some scholars argue that hoaxes, whether related to science or not, are fundamentally lies but with an additional element of sensationalism, resourcefulness, or outrageousness, which amplifies their lasting impact on the public.

The [ScienceFlows](#) research team at the University of Valencia analysed hoaxes circulated via WhatsApp between March and April 2020, shortly after Spain's first lockdown. Of the 2,353 messages analysed, 584 were identified as hoaxes, with 126 focused explicitly on COVID-19 prevention and cures. Often attributed to self-proclaimed medical professionals or anonymous sources, those hoaxes promoted natural substances like lemons and garlic or synthetic treatments such as antimalarial drugs. While those messages appeared to offer helpful advice, they frequently encouraged dangerous practices, such as gargling disinfectants, potentially delaying proper medical treatment. Psychological factors, such as risk

perception and confirmation bias, contributed to misinformation, as individuals were likely to believe information aligned with their pre-existing beliefs (Moreno-Castro et al., 2022).

The Role of COALESCE in Combating Science Misinformation

The [COALESCE Project](#), funded by the European Commission under Horizon Europe, aims to address declining trust in science by establishing a [European Competence Centre for Science Communication](#). This initiative seeks to generate, consolidate, and mainstream knowledge in science communication, creating sustainable solutions that build public trust and, in the long term, enhance the public's understanding of science.

In an era where misinformation and hoaxes spread quickly, the [European Competence Centre for Science Communication](#) will provide tools and information to restore public trust in science. This will be achieved through transparent communication, fostering trust outside crises, and linking science information to individuals' personal experiences and values. Establishing this Competence Centre is crucial for ensuring that science continues to play a central role in shaping European Union policies and addressing global challenges, ranging from pandemics to climate change. Through these efforts, [COALESCE](#) will address the issue of misinformation and offer a long-term strategy for rebuilding public trust in science, which is particularly important as we move into 2024 and beyond.

References:

- Altenmüller, M. S., Wingen, T., & Schulte, A. (2024). Explaining Polarised Trust in Scientists: A Political Stereotype-Approach. *Science Communication*, 46(1), 92-115. <https://doi.org/10.1177/10755470231221770>
- Azevedo F., Jost J. T. (2021). The ideological basis of antiscientific attitudes: Effects of authoritarianism, conservatism, religiosity, social dominance, and system justification. *Group Processes & Intergroup Relations*, 24(4), 518–549. <https://doi.org/10.1177/1368430221990104>
- Hoogeveen S., Haaf J. M., Bulbulia J. A., Ross R. M., McKay R., Altay S., Bendixen T., Berniūnas R., Cheshin A., Gentili C., Georgescu R., Gervais W. M., Hagel K., Kavanagh C., Levy N., Neely A., Qiu L., Rabelo A., Ramsay J. E., van Elk M. (2022). The Einstein effect provides global evidence for scientific source credibility effects and the influence of religiosity. *Nature Human Behaviour*, 6(4), Article 4. <https://doi.org/10.1038/s41562-021-01273-8>
- Moreno-Castro, C., Vengut-Climent, E., Cano-Orón, L., & Mendoza-Poudereux, I. (2022). Exploratory study of the hoaxes spread via WhatsApp in Spain to prevent and/or cure COVID-19. (2021). *Gaceta Sanitaria*, 35, 534-541. <https://doi.org/10.1016/j.gaceta.2020.07.008>
- Wardle, C. (February 16, 2017) Fake News. It's Complicated. First Draft News. Available at: <https://firstdraftnews.org/articles/fake-news-complicated/>.

“Building trust in science through science: brief recommendations” by Nelson Ferreira

- I. **Support researchers and RPO with training designed by RFO and RPO**
 1. **On Targeting.** It is often taken for granted that researchers, being skilled in their fields, are perfect communicators of their own science. Nevertheless, that is really

far from being true, as researchers are tendentiously focused on their expertise and not so much on the social context of research. Researchers do not necessarily know the impact or reception their research may have in different communities. In that sense, they must learn how to reach communities, but first, who really are those communities and why they must be reached out. Training should focus on:

- a) how to identify and engage with who are the immediate beneficiaries of their research;
- b) how to identify and engage with who could benefit from their research (as a positive externality);
- c) how to identify and engage with who could be sceptical or resistant to the research procedures or research results (and understand why);
- d) how to identify and engage with who could be a stakeholder in the research procedure, without benefiting from it directly; (e.g. policy makers, unrelated industries, etc)
- e) what kinds of misuses their research may have and by whom;
- f) workshops for designing research projects and calls for funding considering communities/groups identified in a. b. c. d. e.;

2. On citizen science. An adjusted and focused Performance of Citizen science can be an optimal method for building Trust in the long term inside communities. Small communities at a local level are clusters, especially in municipalities, favouring a chain of relations that can boost awareness, collaboration, interest and dissemination. Addressing Trust through massive channels is not effective for creating roots in society regarding science. If citizens are included in the process, those citizens are optimal disseminators, consumers and contributors for science and also arbiters and whistleblowers in general communities, regarding desinformation. Trust is easier to be built at a local level because of proximity, acknowledgement of peers and established networks from civil society. Training should focus on:

- a) How to actively integrate citizens in research beyond 'data collection tasks' (for researchers).
- b) Training citizens to be 'science ambassadors' in their communities.
- c) Engagement with civil society at local level.
- d) Reconsider social networks at rural level to foster scalability.

II. Make 'disinformation' and 'trust' topics of research to be funded by RFO

Funding of science must consider 'disinformation' and 'trust' critical dimensions of scientific developments and outputs. Otherwise, there is the risk of just contributing to very specific tangible aspects of technological and human development, while setting aside the 'knowledge prevalence' and long-term impact on social level. In that sense all constituents of society must be potential targets of scientific outputs, even those who do not want to listen. Inclusion instead

of marginalisation of those ignoring science must be a strategic goal for actively improving Trust on Science and reducing the impact of disinformation.

1. Calls for funding should consider proposals including a 'trust plan' or 'misinformation plan' in the same way they ask for a Data Management plan, among others.
2. There should be calls for funding addressing objectively 'Trust' and 'disinformation', covering diverse scientific areas.
3. Policy makers should be involved in 'close follow ups' of scientific initiatives for generating awareness to resources involved, restraints and potential of Research Objectives.
4. RPO organisations should be enrolled in Open knowledge sharing procedures with the Policy makers, directed to building know-how with political stakeholders for pushing science in the agenda for 'voting campaigns'.
5. Include procedures of 'citizen science', 'Trust Building', 'practices for addressing disinformation' and 'targeted communication plans' as relevant elements for CV evaluation of researchers.
6. Initiatives directed to demystification of science should establish a constructive approach and not antagonistic one addressing those who follow misinformation, disinformation or pseudoscience. I.e. those who do not want to listen to what science has to say must be heard so they can be included in their own process of education.

“Public and mediator perspectives on responsible science communication and participation: Implications for navigating online science (mis)information” by Tine Ravn

The POIESIS project (Probing the impact of integrity and integration on societal trust in science, <https://poiesis-project.eu/>) is a three-year research initiative (2022-2025) led by Aarhus University, with seven European partners. The other six project partners are: the London School of Economics and Political Science (UK), Wissenschaft im Dialog (Germany), the National Technical University of Athens (Greece), Centre National de la Recherche Scientifique (France), Iscte – Instituto Universitário de Lisboa (Portugal), and Agencia Estatal Consejo Superior de Investigaciones Científicas (Spain). POIESIS is funded by the European Commission (grant number 101057253).

The POIESIS project broadly examines the relationship between research integrity, societal integration, and trust in research. It focuses on the role of various institutions connected to research, innovation, communication, and funding in

fostering a research environment that supports societal trust in science. It is commonly assumed that increased integration of relevant social actors throughout all phases of the research process, along with responsible research practices, naturally leads to public trust in research. The POIESIS project explores these assumptions and their underlying mechanisms, aiming to shed light on the connection between research and dissemination practices and public perception of them. The overarching goal is to understand how and to what extent societal trust in science, research, and innovation is influenced by adherence to research integrity principles and by the inclusion of citizens and societal actors at various stages of the research process. In this regard, a key aspect of the project is to examine how scientific misconduct, questionable research practices, poor or absent communication, and/or misinformation affect public trust. The project aims to develop a set of recommendations 1) on how to address societal mistrust 2) on how to strengthen public engagement and responsible research practices and 3) on how to communicate science responsibly.

The project conducts a wide range of empirical studies involving different stakeholders, for instance deliberative workshops with citizens, expert interviews with researchers and science communicators, focus group interviews with institutional actors, and survey experiments with citizens. All studies are conducted in each of the seven partner countries, allowing for a vast and comparative data set.

At the time of presentation, the following studies were conducted or ongoing:

- Seven Public deliberative workshops across seven countries (n=169 citizens):
 - The study explores citizens' attitudes towards research integrity, research involvement, trust in research and research institutions, as well as research communication.
- 22 focus group interviews across seven countries (n=131 institutional actors):
 - The study focuses on how institutions can provide policies and procedures that enable researchers to act in ways that promote public trust in research. Additionally, the study examines how research integrity and public involvement are perceived.
- Ongoing study with 112 expert interviews with science mediators and researchers to explore complex 'chains of mediation' processes.

An overview of completed studies and results can be found at <https://poiesis-project.eu/deliverables/>

The overall topic of trust in research is central to our society and the significance of research. Scholars have argued that we are experiencing a crisis of trust, pointing to shifts in science-society relations, including the evolving social roles and functions of science (Saltelli & Funtowicz, 2017). Trust in research is a highly complex and multifaceted issue, shaped by relational and contextual factors. It depends on various influences, such as individual and societal factors, including lived experiences, morals, beliefs, trust in institutions, socioeconomic conditions, and knowledge of science, among others (Master & Resnik, 2013; Seethaler, 2019; Gallup,

2019).

Despite this perceived crisis, global surveys generally indicate a high degree of trust in research, while also revealing significant challenges with mistrust. While differences in trust and research institutions differ across countries, a global trend of concern over government interference in research appears to be a cross-country phenomenon (Gallup, 2019; Nature Editorial, 2024).

Public engagement with science and technology (PEST) has gained significant societal and academic attention, with growing recognition of the need for innovative, non-traditional PEST initiatives and platforms. In recent years, there has been a shift from one-way knowledge dissemination toward two-way interaction and dialogue, emphasising co-creation and openness throughout the research process (Burgess, 2014; Stilgoe, Lock and Wilsdon 2014). Nevertheless, challenges remain, such as 'deficit-like assumptions' in dialogues, unexplored impacts, and a need for alternative models of scientific governance (Burchell et al. 2009; Irwin 2014). Additionally, some have noted that certain public engagement activities may be outdated, discouraging, or conducted as mere 'tick-box exercises,' lacking meaningful integration into public services (Andersson 2014).

Results from public deliberations with citizens and focus group interviews with institutional actors indicate that while study participants generally support citizen involvement, this support comes with certain conditions. There is broad consensus that public integration and participation can bring society closer to science and scientists, reflecting the democratic nature of science by allowing scientific activities to be scrutinised and judged. However, opinions differ on the ideal extent of citizen involvement, with several participants noting that not all research projects are equally suited for such engagement (Dubois 2024; Entradas, Sousa and Yan, 2023).

Specifically, citizens voiced concerns about the lack of expertise among laypeople, which they feared could compromise data quality and validity. Many citizens emphasised that public involvement would be most appropriate at the early stages of the implementation phase rather than in the decision-making process, given the limited expertise of the general public. In general, citizens called for effective, substantive, and meaningful engagement, readily identifying and critiquing "tokenistic" or "window-dressing" participation aimed merely at legitimization (Arnstein, 1969), as opposed to genuine engagement based on substantive and normative rationales (Stirling, 2008).

The deliberative workshop study also found that citizens perceive science communication as a crucial factor for trust in science, even more so than research integrity. Additionally, the reputation and credibility of those communicating science (whether institutions or individuals) are essential, along with transparency, clarity, and accessibility of sources. Social media was generally viewed as an untrustworthy source of scientific information, while traditional, reputable media -

such as television, newspapers, and radio - were considered more reliable (Entradas, Sousa and Yan, 2023).

Institutional actors - such as those working in research funding, research support, science communication, citizen science, research evaluation, research administration, open science, and open access infrastructures - viewed measures to increase trust in science as dependent on several factors, including the following (Dubois, 2024):

- Scientists should be provided with the necessary training and resources to engage in sustainable and meaningful public engagement activities. Participatory processes that fail to deliver on their promises can generate public mistrust.
- The promotion of a culture of social integration requires the mobilisation of all actors in the chain of mediation.
- Focus on participation cannot replace a massive institutional communication strategy around scientific information on social media.
- New ways of consulting citizens at a local and regional level such as assemblies for citizens should be developed and considered.

In relation to the specific objectives of the TrustOn2024 workshop - to develop actionable strategies “for a more trusted digital future” and to “establish a multi-stakeholder collaboration model” that considers a holistic approach to all actors within complex science-society relations - the findings from the POIESIS studies could contribute valuable empirical insights. These insights and recommendations could help elucidate citizen and stakeholder perceptions of trust in science, along with identified concerns and challenges related to science communication, disinformation, and public engagement with science.

References:

- Andersson et al. (2014): From Fairy Tale to Reality - Dispelling the Myths around Citizen Engagement. RSA and Involve.
- Arnstein, Sherry R. A Ladder of Citizen Participation. AIP, Vol. 35, No. 4, July 1969, pp. 216-224.
- Burchell, Kevin, Sarah Franklin and Kerry Holden (2009): Public culture as professional science: final report of the ScoPE project – Scientists on public engagement: from communication to deliberation? September, BIOS, London School of Economics and Political Science.
- Burgess, Michael. 2014. “From ‘trust us’ to participatory governance: Deliberative publics and science policy”. *Public Understanding of Science*, 23:1, pp.48-52.
- Dubois, M. (2024). D.3.2: Focus Groups – Findings. Exploring Institutional Roles in Fostering Public Trust in Science. POIESIS. European Commission.
- Editorial Nature (2024). Making the most of trust in scientists. Jan, 31. *Nature* 626, 8. doi: <https://doi.org/10.1038/d41586-024-00238-x>
- Entradas, M, I. Sousa and F. Yan (2023) D2.2: Public Deliberative Workshops – Findings. POIESIS. European Commission.
- Gallup. 2019. Welcome Global Monitor - First Wave Findings. How does the world feel about science and health? Available at: <https://wellcome.org/sites/default/files/wellcome-global-monitor-2018.pdf>
- Irwin, Alan (2014): From deficit to democracy (re-visited). *Public Understanding of Science*, Vol. 23(1) 71–76.
- Master, Z., and D.B. Resnik. 2013. Hype and public trust in science. *Science and Engineering Ethics* 19

(2): 321–335. doi: 10.1007/s11948-011-9327-6. Epub 2011 Nov 2. PMID: 22045550; PMCID: PMC3954634.

Saltelli, A. and S. Funtowicz, What is science's crisis really about? *Futures*, 2017, vol 91, pages 5-11, ISSN 0016-3287, <https://doi.org/10.1016/j.futures.2017.05.010>.

Seethaler, S., J. H. Evans, C. Gere, and R.M. Rajagopalan. 2019. Science, Values, and Science Communication: Competencies for Pushing Beyond the Deficit Model. *Science Communication*, 41(3), 378–388. <https://doi.org/10.1177/1075547019847484>

Stilgoe, Jack, Simon J. Lock, James Wilsdon. 2014. "Why Should We Promote Public Engagement with Science?" *Public Understanding of Science*, 23 (1): 4-15.

Stirling, Andy. 2008. "'Opening up' and 'Closing Down': Power, Participation, and Pluralism in the Social Appraisal of Technology." *Science, Technology, & Human Values*, 33 (2): 262-294.

“Changing the communication paradigm - from influencing to enabling public debate” by Mario Scharfbillig²⁰

“Declining trust in governments and administrations worldwide has prompted the need to re-evaluate the role of public communication. As the information environment evolves, effective communication becomes increasingly challenging but crucial for promoting democracy. Our new report offers evidence-based insights and recommendations to help public communicators navigate new information ecosystems, fostering trust and strengthening democracies.”

Today's information environment could be described as chaotic and overwhelming. The term “drinking from the firehose” captures our everyday experience. Add algorithmic curation for attention maximization, ubiquitous news alerts, a seemingly incessant stream of push messages and a slew of dis- and misinformation casting doubt on everything we think we know and it's hardly surprising many people to cope or have trouble knowing what truly matters to them.

Enter the public communicator competing for the attention of citizens and wanting to communicate what their institution is doing and why it matters. They are trying to find a way to connect and cut through all the noise to provide often vital information. How can it be done in an effective way? We try to give an answer in our new report on [“Trustworthy Public Communication - How communicators can strengthen the future of democracies”](#).

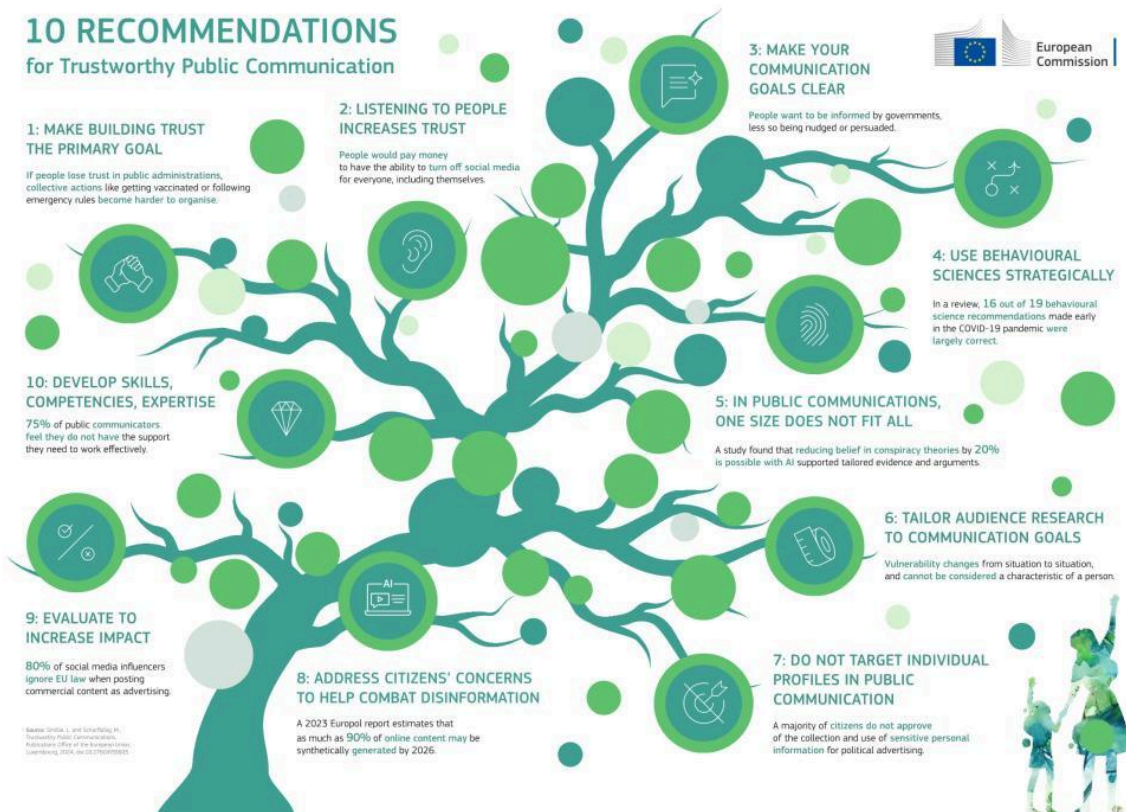
Public communication (i.e. Communication by public institutions, governments and administrations) is a very broad field and can mean anything from press releases, public speeches, and shiny reports to direct mail exchanges with citizens or journalists, public apps, legal notices, and forms people need to fill out. Each of those interactions with the public is an opportunity for communication. Each of those communications should also follow its own logic, responding to the needs of the audience and the channel employed. However, there are some fundamental principles we can distil from research on what the goals of communication should

²⁰ This paper has been published on the EC website: https://knowledge4policy.ec.europa.eu/blog/changing-communication-paradigm-influencing-enabling-public-debate_en and the full report is accessible at <https://publications.irc.ec.europa.eu/repository/handle/JRC137725>.

ideally be.

We have reviewed research from multiple disciplines, including behavioural sciences, public policy, political science, sociology, philosophy, and more, spoken to experts in those fields and public communicators on the front line to find the answers. The work also builds on our previous lessons from our Enlightenment 2.0 research programme where we interrogated [our political nature](#), dived deeper into peoples' [values and identities](#) and the [influence of technology on democracy](#).

There are a number of key messages in our report that are designed to help public communicators to communicate in a more effective and trustworthy way, providing a guiding vision for their work. The key question for public communicators is: what do we communicate for? What is the role of public communication representing public administrations and governments (and we exclude political party communication from this consideration)?



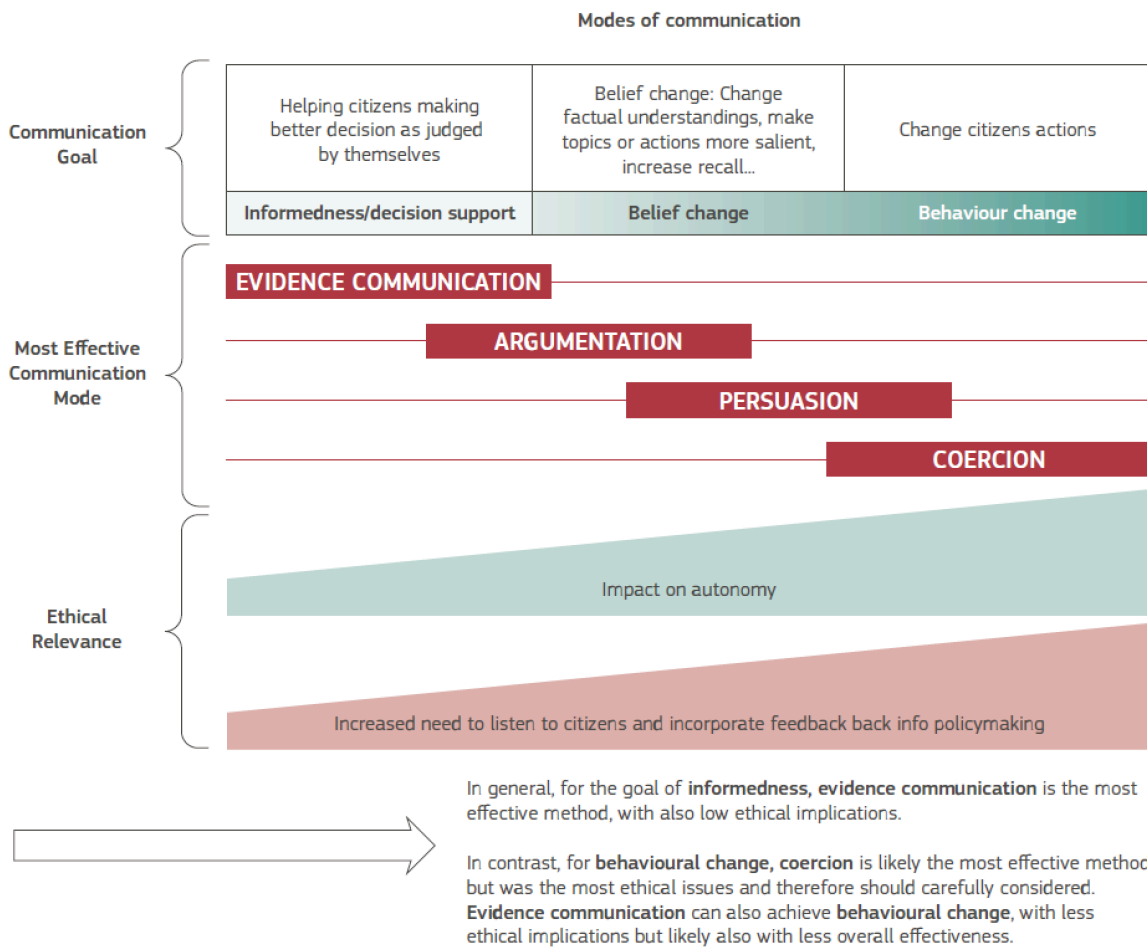
Our answer is that no matter the concrete goal, whether there is a need to inform, persuade, change or listen to the public, communication can only be successful when the communicator is seen as worthy to listen and to talk to. This is reflected in the first key message: “(1) **Building and retaining public trust in their public administration, through being trustworthy at all times, should be a public communicator’s primary goal.**” People are influenced by those they trust and judge what they read and hear in light of their relationship. Is this information coming from someone they can reasonably trust or is the sender known for shady

and unreliable behaviour or outright manipulation? If people [distrust the sender, then the information they receive from them will be discounted](#) heavily, no matter how powerful it is. Building and retaining trust as a source of information will therefore help achieve the mission of the public administration.

The best way to increase trust is to act in a trustworthy way at all times, as building trust is much more difficult than losing trust, something that is known as the “asymmetry” principle: A single case of maladministration and bad communication can destroy long built trust relations, as for example people can be [relatively insensitive to magnitudes](#) when it [comes to moral decision making](#). Thus “(2) **Public communicators should invest more in effective ways of listening to citizens to increase trust in their public administration and democracy.**”

Having established trust by being trustworthy, then we can come to the concrete goal of communication. Answering this question can be broadly divided into two paradigms, the marketer and the enabler. The marketer sees his task as focusing on getting the word out there, persuading and changing behaviour. In contrast, the enabler approaches communication as a facilitator of public dialogue and as providing information and arguments that allows citizens to make up their own mind.

There are merits and drawbacks to both paradigms, and usually public communication is neither purely one nor the other. We reflect this in our report, proposing a spectrum of communication goals and accompanying modes. The primary reflection that is needed by the public administration, which means policymakers together with communicators, is to decide when addressing a policy problem what the goal of the communication is. Is it to enable citizens to make their own decisions, is it to persuade or change behaviour or is it to gather more information from the audience because not enough is known about local or other specificities?



In line with this spectrum of goals, we also offer explanations for which modes of communication are most effective in reaching this goal, for example, evidence communication may serve “informedness” as a goal best. Importantly, we also highlight that the ethical implications and therefore the needs for justification of the communication (and related policies) increase along the spectrum. Governments and public administrations can be argued to have an obligation to change peoples’ behaviour in some areas, be it through persuasion, nudging or even coercion. For example, it is reasonable to expect persuasive action on reducing criminal behaviour, encouraging safe driving, improving vaccination rates against communicable diseases or simply to get citizens to pay their taxes. We offer a full chapter on behavioural insights, many of which can directly be applied to communication, referring to our key messages “(4) **If behaviour change is the communication goal, behavioural sciences should guide the selection of the most appropriate tools**”.

Behavioural sciences can play a dual role on the spectrum between marketer and enabler. On the one hand, the dominant voices from behavioural sciences are usually found on the marketer side, partially because behavioural sciences are focusing by definition on explaining and changing behaviour. Nudges are firmly following in this tradition, where established goals of the administration are sought to be better implemented by recognizing the true nature of human decision

making, rather than following some “rational” model of choice. These nudges are usually [broadly supported](#) in Europe, including nudges like decreasing obesity through food labelling, increasing organ donations through prompting active choice, reducing smoking through graphical images or increasing retirement savings using default saving plans. But as the popularity of these tools increases, so does the range of applications, not all of which may benefit from such large public support. It is essential therefore to also use the important insights from behavioural sciences to better understand citizens, be it in abstract through better understanding their [values](#), or concrete by [engaging with them to](#) define the goals of nudges. Alternatively, use behavioural sciences to [develop boosts](#) – promoting capabilities that citizens can use in multiple contexts.

Unfortunately, in reality a lot of public communication still appears to be following a different mode of “silence, silence, silence, press release: We achieved it all, ribbon cutting, silence.” Whether driven by fear of public repercussions to non-agreed and non-perfect policy solutions, a wish to not be seen as vulnerable or directionless or simply by the lack of resources, this is a potentially dangerous mode of communication. Dis- and misinformation thrives in information vacuums, filling vital gaps of legitimate public interest with wrong, harmful and hateful information that will harm the public administration more than demonstrating a more open, transparent and engaging mode of communication that is however less assertive. Our key message on this is “(8) **Public communicators should acknowledge public concerns pre-emptively, before policy solutions have been developed; this includes strategies to combat mis- and disinformation.**”

The most successful public communicators will be the ones successfully navigating the different goals and modes of communication, keeping in mind that they all have their different merits and pitfalls. For example, the marketer's methods can be highly effective in cutting through the noise and capturing attention in our oversaturated media landscape, but, it can sometimes come across as insincere or manipulative if not executed with care and ethical consideration. On the other hand, it is clear that the enabler, while fostering deeper engagement and understanding, faces challenges in scaling their impact or in engaging audiences who are used to more direct and assertive forms of communication, or not interested in engaging at all. The [new digital formats of online simultaneous and multilingual discussions](#) seem a particularly promising way to overcome this latter challenge though.

What now? If you are a public communicator and you are currently working on a piece, we have a practical test in the report, the TARES test (p. 21), that could be used by institutions to have a general ethical framework. But for now, just ask yourself: How democratic is your article? Does it enable public debate by informing, or is it designed to make you and your organisation look good? What would you want from the communication if you were “just” a citizen? Perhaps, we can build a reputation for being trustworthy organisations by demonstrating [meaningful listening to citizens](#) and stakeholders, respecting their input [together with input](#)

[from scientists and experts](#) to achieve sustainable positive impact? After all, the belief that public administrations care about citizens input is one of the [most powerful predictors of trust in the administration as shown in a large-scale survey](#) by the OECD.

One thing is clear though, considering the fast-paced change in the global information ecosystem. Most public administrators we talked to said that public communication today is not like anything they learned when they started their jobs, even for those who are not yet working for very long in their profession. There are only a few jobs that have undergone a more thorough transformation. Therefore, we recommend “(10) **New challenges require new skills, competences and centres of expertise to support the public communication profession**”.

Key recommendations from the Science Track

Prepared by dr. Charlotte Bruns, Erasmus University Rotterdam, EU-funded IANUS and COALESCE projects, dr. Agata Gurzawska, Trilateral Research IE, Coordinator of the EU-funded VERITY project and dr. Tine Ravn, Aarhus University, Coordinator of the EU-funded POIESIS project

To fight mis- and disinformation in science, a holistic approach is necessary to foster trust in science, where all actors in the ecosystem of trust collaborate. This involves collaboration at the macro, meso and micro levels to address the interconnected character of science, politics and economy and ensure the equal right of all people to share and enjoy scientific progress and its applications (Article 27(1) UDHR; Article 15(1)(b) ICESCR).

At the macro level, we strive for Global Governance solutions to support responsible global research and the role of policymakers including the UN in, first, COORDINATING misinformation and disinformation issues at the global level to ensure accountability in practice. This involves the regulation of social media and industry actors as main beneficiaries and actors in science production, funding, implementation and communication (monetisation of science). We see the Brussels effect of EU regulations, and we hope to see a similar impact in the future of the Digital Service Act, Digital Markets Act, Code of Practice on Dis- und Misinformation, Political Advertising Act, etc. Second, in ENSURING freedom of science through the protection of scientists, we need a global strategy for research-performing organisations (RPOs) (e.g. Universities) to protect researchers.

At the meso level, addressed towards research funding organisations (RFOs) and research RPOs, there is a need for DEVELOPING guidelines for responsible science communication that delineate the responsibilities of different actors and support citizens in evaluating trustworthy communications. They also need to ENCOURAGE more accessible communication of scientific results including reward and recognition systems, training, and resources.

Finally, at the micro level, with responsibilities of all actors engaged in the trust in science ecosystem, we advocate for (1) DIVERSIFYING the presentation of facts and values in narratives to better reflect the diversity of the population; (2) IMPLEMENTING incentives to foster innovative hybrid collaboration approaches that actively involve diverse publics and cultural groups, aiming to build sustainable PE activities; (3) COMMITTING to evidence-based practices and objectives, critically apply methods, and enhance science communication for various audiences.

Our research findings and results of the TrustOn2024 event emphasise the collaborative nature of science and trust. Science is a collaborative enterprise; it does not involve just scientists, but the entire science community and the public. Science requires inclusive and meaningful relationships in the ecosystem of trust in science and across all actors and communities.

References

- European Commission. (2021). Special Eurobarometer 516: European citizens' knowledge and attitudes towards science and technology. Retrieved from <https://europa.eu/eurobarometer/surveys/detail/2237>.
- Ratner, R. K., & Riis, J. (2014). "Communicating science-based recommendations with memorable and actionable guidelines". Proceedings of the National Academy of Sciences, 111, 13634-13641.
- Wellcome Global Monitor, (2021), "How Covid-19 affected people's lives and their views about science". Available online: <https://wellcome.org/sites/default/files/2021-11/Wellcome-Global-Monitor-Covid.pdf>

2.3. Mediation Track

The **Mediation** track of this report explores how effective communication, engagement strategies, and innovative tools can help restore trust in the digital ecosystem while combating the spread of disinformation. In this section, we delve into how diverse sectors—ranging from multilingual information efforts to scholarly communities and journalism—play a pivotal role in navigating the complexities of digital information.

Key contributions within this track examine how to **rebuild trust** in the information we encounter online. For example, **Evelien Dhollander and Kevin Leonard** discuss how the mere sharing of data is not sufficient for that data to be useable/trustable, and advocate for Data Curators to play an intermediary role in improving trust in data, while **Susanna Fiorini** examines challenges and solutions related to multilingual information in an age increasingly shaped by artificial intelligence. **Pierre Mounier** addresses the crucial role of scholarly communities in fostering reliable, open, and accessible information that combats misinformation.

Further, **Véronique Stoll** explores how **open science** practices are instrumental in enhancing public confidence in scientific knowledge, offering transparency and accountability as antidotes to misinformation. In contrast, **Matei Mancas** advocates for a multidisciplinary approach to tackling disinformation, highlighting the need for a combination of **technology, human sciences, and international trust bodies** to address the borderless and rapidly evolving nature of AI-driven disinformation.

The theme of **AI-generated content** is also central to this section. **Mathilde Dorcadie** stresses the importance of journalists as critical gatekeepers in an era dominated by unmonitored AI, while **Guigone Camus** and **Christophe Calvin** ask whether AI can ever be trusted without corruption. In a similar vein, **Frederik Temmermans** introduces a promising development in establishing trustworthiness in digital media through the creation of an **international standard** for verifying content, **JPEG Trust**.

Lastly, the track touches upon **net neutrality** and its role in supporting a more democratic and trustworthy online environment, as **Valentina Tirloni** suggests, and **Sy Holsinger** showcases how **the GoTriple platform** is helping mediate the access to scientific knowledge and build trust in scholarly research. **Gaël Van Weyenbergh** analyses **data cooperatives** as a transformative model for rebuilding digital trust.

These contributions highlight the essential role of **mediation** in shaping how information is created, shared, and evaluated in the digital age. By focusing on ethical practices, transparency, and critical engagement, this track provides valuable insights into the tools and strategies that can help restore trust in the digital ecosystem.²¹

²¹ The introduction of the mediation track was prepared by Suzanne Dumouchel, OPERAS.

“Data Curation as a Mechanism for Increasing Trust in Research” by Evelien Dhollander and Kevin Leonard

The principal function of sharing research data is to enable reuse of that data, be it to validate and reproduce reported findings or to provide a launching point from which new scientific questions can be posed and answered. Indeed, published data enhances the trust placed in the scientific claims which it supports (Rosman et al., 2022). Therefore, the free and open²² sharing of research data is a necessary pillar of good Open Science policy, however simply sharing data is not sufficient to ensure that said data can or will be used. Beyond simple data sharing, data and metadata quality must enter the conversation.

Often, documentation for published research data is limited or of poor quality, which makes it difficult for potential data re-users to find data which might be relevant to themselves – or to even interpret said data once found. Efforts like the FAIR data principles have brought awareness to these concerns about data publishing and have helped operationalize concepts like “findability” and “reusability” to guide practical implementations of data sharing standards (Wilkinson et al., 2016). However, the problem remains that researchers themselves experience increasing time-pressures and often lack the time and specific skills necessary to become familiar with these concepts or to ensure that their published data conform to the newest standards or best-practices of data sharing and documentation (Zuiderwijk et al., 2020). As a result, the reuse potential of the published data falls short of its potential.

We argue in favour of an increased professionalisation of – and research funding allocated towards – Data Curators: individuals whose job it is to remain current on data sharing and documentation practices and to assist practically in the FAIRification of published research data. We believe that investment in curation as an integral part of research data publishing can have the dual benefit of improving the utility of published research data, while reducing the already high administrative burdens for researchers.

Data Curation at Various Levels

Data curation can be implemented at various points within the research data lifecycle and can be focused on different granularities with respect to data and metadata, and these different varieties of curation come with distinct benefits.

When data curation is conducted as early as possible in the research process – namely, prior to the publication and sharing of that research data – curators can evaluate the data themselves. They can verify the completeness of the dataset prior to its publication and can perform reproducibility checks on the data and/or

²² Or at least “as open as possible, as closed as necessary” given legitimate (e.g., legal/ethical) reasons for access restrictions.

software. By performing these quality checks prior to publication, curators can identify scientific inaccuracies early in the publication process, thereby preventing them from entering the scholarly record. Furthermore, by providing edits for clarity, the curation process can also increase the interpretability and therefore reusability of research data. Thus, by increasing the clarity of published research data and reducing the frequency of data errors, data curators enhance the trustworthiness of published data and therefore the other scholarly outputs that those data support.

Alternatively, when curation occurs after the data are already published, the ability to modify the data themselves is diminished as updating the published data would likely require the use of data versioning on the data repository, which may be prohibitive to some researchers. However, for many data repositories, the metadata associated with published datasets can be freely edited. Curators can recommend changes be made to metadata fields such as the Title, Abstract, or Keywords – fields commonly indexed by various search engines – to increase the likelihood that the dataset is found when searched for with relevant search terms. They can also recommend broader changes to fields like the Abstract to include some of the re-use relevant documentation that might have been absent from the README or other documentation file (e.g., definition of variables, explanation of file contents, etc.). Lastly, the curator can also recommend the research make explicit links in the metadata between the dataset and other research outputs, such as published articles which the dataset supports or software that was used to process the data, thereby increasing the interconnectedness of the research.

Curation, Trust, and AI

Although concerns regarding trust in science have long been raised (Haerlin & Parr, 1999), the emergence of LLMs like Chat-GPT as research tools have exacerbated the potential for misinformation – and therefore mistrust – to proliferate. At present, the information provided by these AI tools is only as good as their training and input data, and therefore are subject to “illusions of objectivity” and “monocultures of knowing” insofar as their training data contain gaps or biases. As researchers and other consumers of scientific research offload more of their cognitive burden to AI tools when reviewing the existing scientific literature, it is increasingly important that we ensure that the scientific corpus is as complete and well-documented as possible, thoroughly described by machine-readable metadata. Here data curators can also play an integral role, preparing and documenting research data for not only human but also machine audiences to ensure that important scientific works don't fail to enter our collective consciousness just because they aren't FAIR enough.

References

- Haerlin, B., & Parr, D. (1999). How to restore public trust in science. *Nature*, 400(499).
<https://doi.org/10.1038/22867>
- Messeri, L., Crockett, M.J. Artificial intelligence and illusions of understanding in scientific research.

Nature 627, 49–58 (2024). <https://doi.org/10.1038/s41586-024-07146-0>

Rosman, T., Bosnjak, M., Silber, H., Koßmann, J., & Heycke, T. (2022). Open science and public trust in science: Results from two studies. *Public Understanding of Science*, 31(8).

<https://doi.org/10.1177/09636625221100686>

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boitn, J.-W., Santos, L. B. dS., ... Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3. <https://doi.org/10.1038/sdata.2016.18>

Zuiderwijk, A., Shinde, R., & Jeng, W. (2020). What drives and inhibits researchers to share and use open research data? A systematic literature review to analyze factors influencing open research data adoption. *PLOS ONE*, 15(9). <https://doi.org/10.1371/journal.pone.0239283>

“Multilingual information in the age of AI: findings and perspectives” by Susanna Fiorini

Early 2024, a group of researchers warned about the massive presence of unsupervised or barely-supervised machine-translated content on the Web²³. Yet translation is a linguistic and cultural mediation requiring advanced human skills, which cannot be fully replicated by machines without risks, especially in a sensitive context like information. Therefore, the current landscape could compromise access to reliable information in multiple languages, and especially in the low-resource ones, which are not much represented with genuine human-written content on the Internet and are often the primary communication languages of already marginalised communities. Authors also found evidence of a selection bias in the type of content translated into many languages, consistent with low quality English. This means that the original content that is typically submitted to machines for linguistic mediation is already inaccurate, so that the translated output can hardly be reliable. On top of that, these low-quality translations are then collected from the Web to build the datasets that are used to train the models underlying the neural machine translation engines and generative AI tools that will serve to translate more content. The vicious circle is closed: the machine-translated output published online can only become worse with time, and – again – especially in low resources languages.

The main consequences of such a massive, uncontrolled use of machine translation are the following:

- **Lack of traceability:** in the current context, it is generally difficult to find information about the origin of a translation – for example, determining if the translation is machine-generated, if a human was involved in the process and in this case, which level of supervision was ensured. Unlike clearly inadequate machine translations which are reasonably easy to recognise, the most recent AI technologies can produce very credible but not necessarily accurate output, which could contribute to misinformation or disinformation.
- **Lack of accountability:** if there is no information on the origin of the

²³ Brian Thompson et al. 2024. A Shocking Amount of the Web is Machine Translated: Insights from Multi-Way Parallelism. <https://doi.org/10.48550/arXiv.2401.05749>

translations published on the Web or social media, not only it is hard to determine who is accountable for the reliability of such translations, but also the essentially human knowledge which is the foundation of the translating process – including automated translation – becomes invisible and therefore loses value.

- **Lack of transparency:** without information on the nature and the origin of existing translations, the datasets used to train AI systems are by definition opaque. Unawareness of what data is used for training not only amplifies the “black-box effect” which makes predicting and interpreting the machine output more challenging, but also makes it virtually impossible to know if copyrighted, protected or sensitive material was collected.
- **Lack of representation:** data collection for AI training is basically performed on the Web via automated data collection techniques, which must be replicable on large scales and thus require substantial standardisation. While being core factors for efficient data collection, replicability and standardisation limit by their very nature the kind of data that can be collected: therefore, the datasets used for AI training are intrinsically biased, because their construction is affected by technical and legal constraints – for example, only content having given standardised formats and licences are collected – and also reflects real-world forces so that data is generally oversampled in favour of dominant cultures, genders, or languages.

The conclusion that emerges from such a scenario is not that AI should be rejected in the context of digital information, but rather used in an informed way within a more regulated and transparent framework. Potential measures that could help to identify trustful content, content producers and mediators are the following:

- **Promoting a transparent process for an informed use:** a trustful content producer should put users in the condition to know how the content was produced in order to make an informed use of it. To this aim, translated content could come with metadata indicating the origin of the translation, and any AI tools used to produce it. Thanks to this “set of instructions”, users could be able to activate and focus their critical thinking in order to keep using the translated content, but still be aware of any potential underlying issue, interest and bias.
- **Implementing a clear labelling system:** the metadata described in the previous point could come in a standardised format, or labelling system which could be quickly and easily recognised by users (for example, the “Made with AI” label introduced by some social media, or the author statements required by some academic journals). The labelling system could also be designed to give more visibility to the human who produced the content, whether with the aid of AI or not, in order to determine clear accountability for that content: in the case of translation, for example, the labelling system could systematically include the name of the translator – which is rarely the case today.
- **Reaffirming human (expert) added value:** making human contributions (more) visible could help to strengthen trust in human experts. The idea is not to stop using AI, but to trust human mediators to interpret and learn how to

critically use machine-generated content for (multilingual) information and knowledge dissemination.

“Rebuilding Trust in Online Information: The Role of Scholarly Communities” by Pierre Mounier

In an era characterised by the rapid proliferation of disinformation and mistrust in online content, scholarly communities have a unique role in rebuilding trust in information. The challenge of managing information quality, especially in the digital age, is a critical concern. The OPERAS Research Infrastructure, a dynamic and community-driven initiative, is dedicated to advancing open scholarly communication in the social sciences and humanities (SSH) across Europe. Through its services, OPERAS aims to make open science more accessible and impactful, addressing the fragmented landscape and distinctive challenges of SSH disciplines. By empowering scholars to innovate and collaborate, OPERAS fosters a culture of transparency, accountability, and trust in academic knowledge.

The Paradox of Disinformation

A striking paradox defines our digital era: while the internet has provided unprecedented access to information, it has also facilitated the rapid spread of disinformation. Academic institutions, which have played a central role in the development of the internet, are uniquely positioned to counteract this challenge.

Historically, the internet's core technologies, such as TCP/IP (Transmission Control Protocol/Internet Protocol), were developed within academic and research settings. TCP/IP, the foundation for all internet communication, emerged from research funded by the United States Department of Defense's ARPANET project. However, much of the intellectual work was carried out in academic institutions such as UCLA, Stanford, and MIT, where scholars collaborated to create an open and interoperable communication system. This openness and the collaborative spirit in which the internet was born originally aimed to facilitate the free exchange of ideas and knowledge.

Similarly, the World Wide Web, invented by Sir Tim Berners-Lee at CERN (the European Organization for Nuclear Research), was designed with academia in mind. The development of HTML (HyperText Markup Language) and the creation of the first web browser enabled researchers to share documents and data across the globe. The ethos of academic openness was embedded in the DNA of the web, with Berners-Lee's vision emphasising that the web should be universally accessible and free from commercial or governmental control. The web was created as an infrastructure for reliable, peer-reviewed knowledge-sharing, open to all and maintained through academic integrity.

However, despite these noble origins, the internet and the web have also become breeding grounds for disinformation. The tools that once enabled the democratic dissemination of knowledge now allow for the rapid spread of unverified and misleading content. This paradox is at the heart of the current crisis in online trust. The institutions that laid the foundations of the digital world must now play a key role in addressing the challenges of disinformation and restoring faith in online content.

Academia's long-standing tradition of trust in information, based on rigorous peer review and empirical validation, is more relevant than ever. The mechanisms that academia developed to ensure the credibility of knowledge—such as peer-reviewed publications and standards for evidence-based research—are essential to rebuilding trust in today's digital ecosystem. By reasserting the principles of transparency, accountability, and peer validation, academia can help mitigate the harmful effects of disinformation.

The Role of OPERAS in Rebuilding Trust

OPERAS, as a major research infrastructure for open scholarly communication, embodies this academic tradition of fostering trust in information. Through initiatives like the Peer Review Information Service for Monographs (PRISM) and the Diamond Open Access Standard (DOAS), OPERAS seeks to enhance the quality and credibility of scholarly communication in the digital age.

The Peer Review Information Service for Monographs (PRISM), provided by the Directory of Open Access Books (DOAB), reinforces the cornerstone of academic integrity by increasing transparency around peer review processes, particularly for monographs. By providing accessible information on how academic work is evaluated, PRISM helps ensure that the public can verify the credibility of the content.

Meanwhile, the Diamond Open Access Standard (DOAS), developed under the lead of the Spanish Foundation for Science and Technology (FECYT) in the context of the EC-funded DIAMAS project, extends this focus to the broader context of information production, promoting best practices across areas like funding, governance, editorial quality, and inclusion. DOAS helps publishers maintain high standards of transparency and ensures that scholarly publishing aligns with the principles of equity, diversity, and accountability.

The Contribution to the Information Quality Protocol

OPERAS is committed to contributing to the future development of an "Information Quality Protocol," a framework that will further define the standards for trustworthy information in the digital age. Through PRISM and DOAS, OPERAS offers practical solutions for enhancing the transparency, quality, and accessibility of scholarly communication. By fostering a culture of open science and rigorous peer review, OPERAS strengthens the infrastructure necessary to rebuild trust in online information.

Conclusion

In conclusion, scholarly communities, particularly those engaged in open access initiatives, have a critical role to play in rebuilding trust in online information. Academia, which helped lay the foundations of the internet and the web, must now address the challenge of disinformation by leveraging its long-standing principles of transparency and peer validation. Through initiatives like PRISM and DOAS, OPERAS is leading the way in creating a more transparent and accountable system for scholarly communication. As we continue to navigate the challenges of the digital age, these tools will be essential in ensuring that academic knowledge remains a trusted and reliable resource for researchers, policymakers, and the public at large.

“Increasing confidence in science: how open science fights disinformation” by Véronique Stoll

A common definition of open science is the free dissemination of the results, methods and products of scientific research. This means accessing, sharing and using information for more transparent and citizen-centred science: open access to scientific literature, open data for research data, open source for source code and even open educational resources for teaching materials. Science is seen as a common good to improve access to science for all, to enhance scientific collaboration, transparency and scientific integrity, and to improve the sustainability and reproducibility of research.

By providing transparent access to publications, data, tools and methods, Open Science plays a crucial role in combating disinformation and reducing mistrust between science and society. Transparency is a motivating factor to do things right and to prioritise scientific goals. Open science does not prevent fraud, but helps to detect it through two key points: the scientific process and peer reviewing.

Open access to scientific data and publications: Open science allows more people to access research results, raw data and scientific publications without financial barriers. This increases transparency and allows anyone to check the results, reducing suspicions of manipulation or bias.

Open research practices: By making research processes more transparent, including through open peer review and the publication of research protocols, it becomes easier to understand and follow scientific reasoning. The Open Science Framework (OSF) movement allows researchers to share their protocols, raw data and intermediate analyses. For example, pre-registration of studies on OSF improves transparency and helps reduce publication and confirmation bias.

Replicability of studies: By providing open access to data and methods, other scientists can replicate studies and verify results, strengthening the credibility of scientific discoveries.

Awareness: By making scientific knowledge more accessible, open science can better inform the public and educators. Open publication platforms and institutional repositories provide a wealth of reliable and verified information that is easily accessible to the general public.

Citizen engagement: Citizen science initiatives and participatory platforms enable non-scientists to participate in research projects, thereby increasing mutual understanding and acceptance of scientific processes. Open science strengthens the link between science and society: participatory research, capacity building and skills development for non-academic actors enable them to be better informed and thus engage as active participants. By enabling better dissemination of scientific information, open science is a weapon against disinformation and the spread of false information.

Transparency for the public: Increased transparency and the ability for the public to see and understand how scientific knowledge is produced and validated strengthens trust in science. Example: Open access scientific journals such as PLOS ONE and eLife publish studies online with transparent peer review, allowing the public to see not only the results but also the critical review process that validated them.

Fast response to false information: During the COVID-19 pandemic, initiatives such as Retraction Watch played a crucial role in quickly identifying flawed or fraudulent studies, which helped to correct misinformation disseminated in the media. The COVID crisis is a good example of "open science" practices, integrity lapses and disinformation. Data were exchanged between teams in a very short time. The COVID-19 gene was sequenced in a matter of weeks. Thousands of sequences are also regularly recorded and shared to track mutations (a process called "phylogeny"). All relevant journals have made their publications freely available at the request of policy-makers and scientists. It is reassuring to see that, in the event of a pandemic, global cooperation works and that all stakeholders are mobilised. Open science has been set in motion, knowledge has been circulated for scientific progress, and also to combat the rampant infodemic. This is a prime example of the successes and benefits of open science, which made it possible to work immediately on the development of treatments and vaccines on a global scale. Open science practices have demonstrated their importance in rapidly disseminating the knowledge needed to address the health crisis.

However, available studies suggest that the sharing of COVID-19 research data remained low, well below expectations, and there was no massive data sharing. Even during a crisis of this magnitude, barriers to data sharing within research communities remained significant. Unvalidated preprints were used by the scientific community and the public ; this could lead to fake news. More importantly, some public decisions were even made on the basis of questionable scientific articles ("Lancetgate"). The quality of the filtering of scientific communication has been questioned. At the level of the general public, the legitimacy of scientific journals has been attacked, even if these deviations remain

marginal compared to the total number of articles published, and this has fuelled widespread mistrust of science and various conspiracy theories.

The pandemic has also highlighted the need to provide information tailored to each target audience. Thus, it is not the "raw" availability of scientific articles by researchers that will restore public trust in science, but rather well-adapted, synthesised, well-argued and understandable reformulations that will strengthen the science-society transfer.

For the benefits of open science to truly contribute to the advancement of sustainability science, certain conditions need to be met, including

- A new approach to research evaluation, based not only on quantitative criteria but also on the intrinsic quality of scientific work, taking into account its diversity and including, among other factors, its societal impact;
- National and international policies, accompanied by resources, that support the development of a culture and practice of open science within scientific communities;
- Reducing the digital divide that disadvantages many regions of the world, especially countries in the Global South, in accessing and creating open data repositories and archives.
- Developing other practices such as sharing data in trusted repositories or peer communities.

We can also consider other publication processes that emphasise preprints and organised open peer review. For example, Peer Community In (PCI) implements a preprint labelling process within a thematic community, where a transparent evaluation is carried out leading to the 'recommendation' of an article. All information leading to the recommendation of an article is made public. The name of the recommender responsible for recommending the article, their editorial decisions and the text of the recommendation, the reviews and suggested corrections, and the authors' responses are available on the PCI thematic website. Successive versions of the preprint are freely available in open archives. Only the names of the reviewers can be withheld if they wish to remain anonymous.

“Fake news : need for a mix of technology, human sciences and international trust bodies to face borderless rapidly developing AI” by Matei Mancas

Fake news is often faced on a very specific facet of it. Either approaches are very technical and focused on only deep fakes (based on generative AI) for example, either it is too much based on knowledge databases and context or only on the education side. While all those approaches are good, they are not enough when used alone.

The AI4Debunk EU project (<https://ai4debunk.eu/>) intends to take into account several aspects of fake news debunking. There is a technical aspect purely focused on the data itself. However, multimodal data coherence (such as matching between images and text description in posts) and the context of knowledge bases (knowledge graphs of news already known as being fake) are also taken into account. All those AI-based cues together provide suggestions to human moderators which ultimately take the responsibility of the decision of classifying news as real or fake. The AI is supposed to provide a transparent set of reasons for the recommendations (such as links to similar fake news, areas in images, text, videos which are incoherent in the same post, etc.) that the human moderators validate or modify. As the project focuses on citizens, the transparency and the explanation of why the news is considered as fake is important.

In addition to those aspects, the project also intends to provide schools with tools to augment young people's awareness on fake news. This facet of the deep fake fight is crucial to avoid future generations having a too fuzzy border between real and fake in general.

The speed of AI development which can be applied to improve fake news is impressive. This is why, bodies which can make a link between the very long time needed to achieve new legislation and the very short time needed by the technical developments to be used by specialists and non-specialists are highly needed. On national or continental levels, bodies of trust from different university specialists need to rapidly let the legislator be aware of new possible AI-based methods which can be used for fake news and provide ideas on how to handle them based on existing legislation or provide hints for new legislation if needed.

Finally, trust is complex to acquire and easy to lose. This is why there is a need for trusted servers where the fake news debunk methods are run and fake news are stored as any corruption of this data can lead to a brutal loss of trust from the public. There is also a need for moderators which are independent from economic or political pressures at an international level. Those servers and human moderators are key to keep a high level of trust of the public in a new fake news debunk system.

As a conclusion, a technical debunk system needs to put together dynamically updated knowledge about fake news and analyse data incoherences or signs of data corruption in the signal. However, this is not enough alone, and other crucial points are needed :

- Its use in a game-based approach in schools to augment young people's awareness on fake news (national level).
- Bodies which can make a link between the very long time needed to achieve new legislation and the very short time needed by the technical developments to be used by specialists and non-specialists (national or continental levels).
- An independent and secured platform and consortium for moderation led by an international body (international level).

“Against unmonitored AI-generated content, journalists as gatekeepers” by Mathilde Dorcadie

I. AI-generated content : what are we talking about in the media?

> for some years now, media outlets have utilised AI to produce various types of content, primarily texts that compile extensive data ranges (e.g., sports competition results, financial market analytics). These tools are using what is called “natural language processing”, they are trained and kept under human control, are fed with selected databases, using logical and understandable processes.

> **AI tools can then offer important support to the work of journalists to produce content in a high-speed media environment.**

> Recently, a new kind of tools have emerged (and reached a larger public) which are able to process even more data by using the “deep learning” technology. Unlike traditional AI tools based on pre-encoded languages, these tools can create their own codes. They possess “artificial neurons” and can learn automatically, mainly by mimicking existing content.

> *What does that mean ?*

Generative AI is now being used to create a large number of **fake media websites**. For a relative low cost, one can create a fake media website. AI will not only produce articles, but it will also generate images and even fake authors, with fake profile pictures and fake biographies, sometimes complete with fictitious personalities. (e.g., a food critic sharing their tastes and best recipes). These websites mimic real news websites in layout, content, and writing style.

So when it comes to culinary portals or travel blogs, we could say it is a pity, but it is not that harmful (at the worst your pie will taste weird or your next city-trip may turn out very disappointing). Such websites primarily drive traffic for commercial purposes and sell advertising space.

But now you can imagine the damages if that technology is being used to create news websites that mislead the audience for partisan purposes, such as spreading disinformation before an election.

In the US, there are two parallel trends: the decline of local news outlets due to economic constraints and the proliferation of cheap, fake local news websites.

II. Real or Fake: for everyone it is becoming harder and harder to navigate among the information flow

Basically, generative AI is now able to counterfeit everything : from people’s voices to historical pictures, from best-sellers books to the music style of a specific artist.

One thing it can also do is **“fabricated sources”**. For instance, if you ask Chat-GPT to write a journalistic article it will likely produce a text quoting experts, referring to some studies, providing related data. Why? Because its goal is to produce a

credible-looking article. Not necessarily a reliable article. It has to “look like” an article.

These quotes, references and data might be correct (“copied” from elsewhere) but they might be entirely fabricated. Specialists call that phenomenon “**hallucination**”.

> *But as a reader, how can you be certain of what is correct and reliable, and what it is not?*

There is not a single answer to that question. There are plenty of solutions that could be implemented.

> First of all, **disinformation** is existing not only online, but also on TV and newspapers and it is spread by all kind of players including politicians, influencers and even scientists (as we have seen it during the pandemic or among climate-change deniers). For years now, journalists have combated fake news with **dedicated “fact-checking” team**. Spoiler: those efforts can only rectify part of the misinformation. And as you might know, the rectification has often less echo than the initial fake news. But still, despite this, that “fact-checking” job remains fundamental and must continue.

> What we are facing now, the current challenge, is that disinformation is no longer solely produced by human but mostly by machine in a larger scale and faster rate than ever before. It is a significant global challenge and it is **harming increasingly the functioning of democracies**.

On top of that, we can no longer fully trust tools like *Google News* anymore, because it is not designed to promote human-made content and demote IA-generated content. *Google News* puts side by side reliable and unreliable content and it is very problematic.

I will not even start to talk about social media platforms and the way they deal with that matter....

III. What can journalists do ? the role of “certifiers”.

The ethical challenges posed by AI are significant for us, journalists, as they are for many professions. We are just one part of the solution.

> Let’s start by saying : My profession is facing public trust decline. Not everywhere, not everyone of us. If we perform our job correctly, we remain among the most crucial players in this misinformation battlefield.

What we have in two strengths: firstly **we are human** and secondly we **have professional codes of conduct**.

> Obviously, computers cannot go on the ground, we, as journalists can do that and can bring back genuine and untold stories. I have personally done that, and I can’t imagine how any software would be able to bring the same kind of information, without going and see with my eyes people’s reality or experience.

As humans, we can think and understand topics in ways that AI cannot. The technology is called “intelligent” but it cannot grasp the essence of human experiences because it does not possess its own experiences. Consequently, AI cannot inherently possess morality or ethics. (It can be trained to have one maybe. It can only be programmed to simulate them).

> That is why, journalists must increasingly play **the role of “certifiers”** of information, basing their work on ethics and fairness, acting as « filters » for the constant flow of news production.

Journalists and media freedom organisations are already joining forces for an ethical AI in their environment. They formalise **guidelines and AI charters** for their outlets and beyond. International committees aim to come up with “a strong international reference” to preserve the quality of information and public trust in journalism. Because trust has always been the bedrock of journalism **and a key differentiator from other information sources.**

Quoting an expert:

Building rapport with readers and sources is “not something you can do from typing a prompt into ChatGPT”. (Felix Simon in Politico)

AI must remain under democratic control and this should be done through **transparency obligations.**

CONCLUSION

> AI is not the enemy; the enemy is the people who could misuse them on purpose or not. As researchers and reporters, we have to keep in mind not only the tools behind disinformation, but the intent, too. We have to remember the purpose of disinformation and propaganda : Some of it is financially motivated, some of it is used to build political will, and some of it is created to widen societal divisions.

> Less repetitive tasks, more time to go on the field and to investigate, more appreciation from the public for journalistic work: that sounds like the dream of many journalists. AI is able to help journalists, fact-checkers and investigators, as long as we follow the moto : “human decide, AI does”.

“FROM TRUSTWORTHY DIGITAL TO INCORRUPTIBLE DIGITAL/AI?” by Guigone Camus and Christophe Calvin

In an era characterised by an unprecedented proliferation of information provided by IA through digital technologies, disinformation and miscommunication represent major risks for each and every one of us’ future, whether our capacity to decode or evaluate what multimedia and technology say and show. Regardless of the domains of expertise — political, social, economic decision-support tools; business models; scientific and technological forecasting, simulation or progress —

AI has become a huge performance improvement gateway to enhance accuracy and increasing efficiency, by offering data-driven insights. Consequently, the superimposing role of AI raises the need for controlling a stable, trustworthy and resilient digital environment. The latter encompasses not only core AI but the entire ecosystem evolving around and for it: data, communication, privacy, etc.

* **Technologies for trustworthy digital unquestionably improved in the last decade.** Data privacy has improved. For instance, *Fully Homomorphic Encryption* (FHE) allows computations to be performed on encrypted data without first having to decrypt, thus eliminating the need for processing data in the clear, thereby preventing attacks on the data while it is being processed. *Synthetic data* are non-human-created data that mimic real-world data and that are created by a computational algorithms, and simulations based on generative artificial intelligence technologies. They can be used instead of personal data, in order to create similar data with the same statistically relevant information without exposing private and sensitive data. Transaction and communication privacy has also improved (*Virtual & Invisible Private Network* (VIPN));so did Data traceability (*Data tattoo*).

* **Although AI also improved, it shows persistent deficiencies.** Among them AI Hallucinations are the most known. These limitations non exhaustively cover : (a) *incorrect predictions* — inaccurate or unlikely data-based prediction such as weather forecasting nonsense — (b) *false positives* — AI identifies something as a problem/threat when it's not, such as legitimate transaction incorrectly flagged as fraudulent — (c) *false negatives* — AI fails to detect a real problem/threat such as a tumour on a scan — (d) *nonsensical outputs* — AI generates nonsensical outputs or unrelated to data inputs — (e) *factual errors* — AI states an info as a fact that is objectively incorrect in its training data — (f) *visual artefacts* — AI Image Generator creates unrealistic or contradictory elements such as proportions — (g) *biased or offensive outputs* — AI output generates unexpected or offensive content such as the denial of Holocaust. There are also Learning data biases such as skin tumour tracking programs based only on white skin learning databases. Distinction between Correlation and causation, which is still a major question in crucial scientific fields such as medical research, also challenges AI systems by requesting a large and trustworthy corpus of data in order to identify causal relationships.

* **Traceable gaps.** Some basic AI weaknesses are easily identifiable. This is the case in open technologies suffering from a lack of data-verification (Microsoft Chatbot (2016); ChatGPT) and open technologies with no quality measurement of training data. The excuse that the models are “completely open” is an illusion: understanding LLM models with trillions of parameters is an impossible task. Eventually, legal tools — e.g., European AI Act, March 13, 2024 — regulate the uses of AI. But they do not anticipate its misuses nor do they warn about biases, quantification and errors. Indeed, they keep silent as to the regulation of data-quality and data-quality measurement constraints. Besides, jointly elaborating legal global tools is also an illusion. The General Data Protection Regulation (GDPR),

which is a Europe-wide legal text, is far from being currently systematised in other countries, either for personal data collection or personal data based LLM creation.

*** Solutions imply a double commitment: technology and human beings must engage.** A “better AI” i.e., trustworthy, fair, ethical, just and robust, requires technical solutions. Bayesian approaches and their typical algorithms are used to deal with uncertainty in order to improve the efficiency of probabilistic models. Hybrid AI combines quantity — ML exploiting statistical models to analyse data — with quality — Symbolic AI, a semantic-based AI — in order to emphasise on the importance of meaning and logic and correct correlations either a posteriori or “live”. Explainable AI allows users to rely on a set of processes and methods that decode the functioning of an LLM, thus offering to apprehend and trust the results and output created by machine learning algorithms.

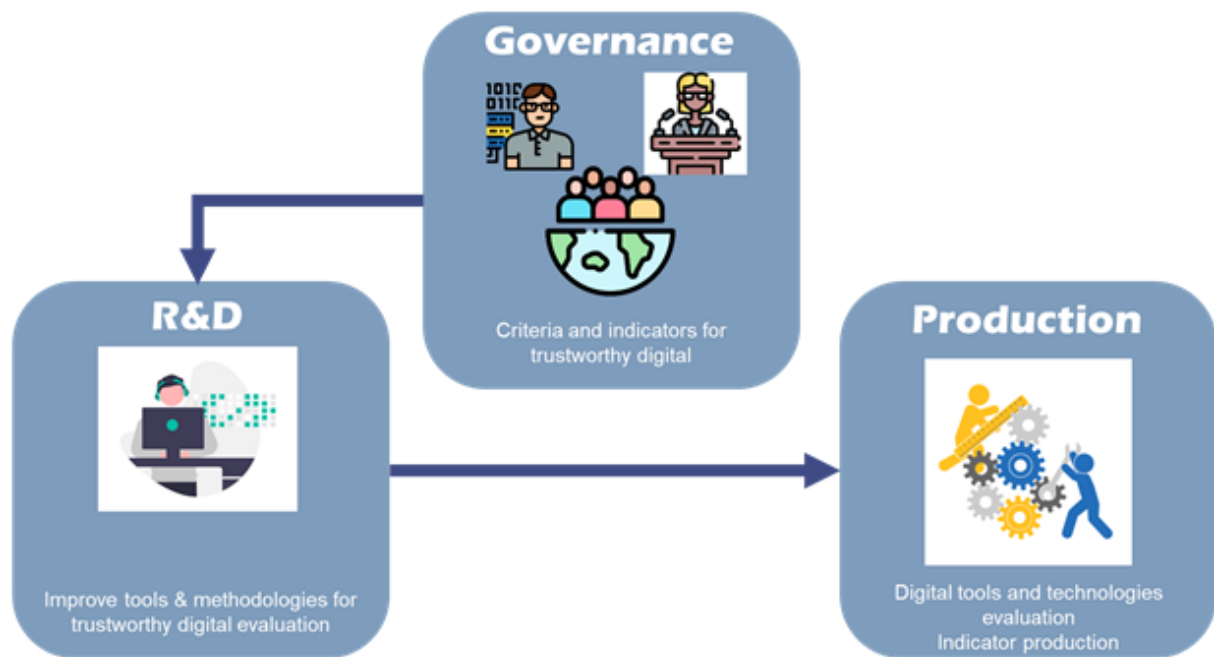
Simultaneously, human-beings must commit to delivering *and* acquiring a better and constantly growing digital and AI Literacy in order to protect ethical and equal values for each and every one of us, in the present time and in the future. Literacy is the human-centric way of reducing miscommunication, misinterpretation and disinformation (Brazil G20, 2024). Self-criticism towards digital advantages and limits is the key to provide not only personal but collective empowerment and autonomy, and even well-being — especially in the case of the use of AI for decision-making regarding inclusivity, inequality and eradicating poverty in the context of low- and middle-income countries. Then, in a pendulum movement, a large and various digital literate human collectivity can in return contribute to AI by refining, specifying and correcting AI frequent and numerous biases — socio-cultural, economical, racial, religious, gender, equality, linguistic, political, Human Rights biases....

*** Proposal: an AI trustworthy digital score.**

Although a global AI (i.e. an AI that would have been trained on unbiased data and data encompassing cultural worldwide subjectivities) is technically possible, it is not desirable for universal stays an illusionary and a controversial concept. Nevertheless, it is reasonable to assume that, here and today, ideas can be suggested in order to contribute to the actual favourable reflection ecosystem.

The concept of a **Trustworthy digital score** has emerged from the observation that the trust placed in digital technology mostly are highly human-feelings centred, thus, very subjective. This is not to say it has no value, but that the evaluation of digital and AI trust should be beget from a balanced mix of subjective and objective and quantifiable criteria. Ideally, the building of such a trustworthy score should rely on the intertwined relationship between 3 interactive strategic directions: (1) R&D (2) Governance (3) Production. The R&D requires the creation and/or the improvement effort for robust tools & methodologies to evaluate digital technologies (Generative AI, data and transaction privacy etc). Governance requires human resources, involving technology creators (R&D), civil society and deciders and stake-holders together, that should be capable of elaborating an openly shared

collectively built criteria of evaluation, in order to define what trust means for individual, collective, technological, political, social and cultural points of view. The R&D takes the responsibility to implement robust tools and methodologies that evaluate these criteria. Finally, after these 2 steps, Production is in charge of the evaluation of the digital tools that will provide a trust score as incorruptible as possible. These three components have to be of intergovernmental nature.



“JPEG Trust: an international standard for establishing trust in digital media” by Frederik Temmermans

Today’s images arise from many sources: digital cameras, film photo scans, photo editing software, artificial intelligence, and combinations thereof. These technological innovations allow us to produce novel new imagery and even new knowledge derived from our media assets, but at the same time, they can confuse and even deceive us. To take advantage of the beneficial outcomes of modern digital media production technologies while lessening the negative outcomes, we need to know if and how we can trust the media we encounter.

To facilitate global interoperable media asset authenticity, JPEG (ISO/IEC JTC 1/SC 29/WG 1) initiated the development of a new international standard: JPEG Trust. JPEG Trust arises from an exploration that started five years ago of requirements for addressing mis- and dis-information in digital media. JPEG Trust provides a comprehensive framework for individuals, organisations, and governing institutions interested in establishing an environment of trust for the media that they use and supporting trust in the media they share online. This framework addresses aspects

of provenance, authenticity, integrity, copyright, and identification of assets and stakeholders.

The JPEG Trust framework is built in compliance with JPEG standards as well as other widely adopted industry standards to ensure a smooth integration into existing digital media ecosystems. The framework can be integrated into ecosystems that use any of the JPEG family of standards. In addition, due to its generic nature, many aspects of the framework can also be applied to other image file formats or other media modalities such as video or audio.

Establishing Trust

Trustworthiness of media is subjective and dependent on context. JPEG Trust does not explicitly define trustworthiness but rather provides a framework and tools for individuals, organisations, and governing institutions to establish trust in accordance with the conditions they specify. For example, when a photograph of a damaged car is shared among family members, they would not question the veracity of the photograph. However, when this same photograph is shared with an insurance company, additional indicators of authenticity would likely be required. The context-dependent and often subjective nature of trust is accommodated by the JPEG Trust framework.

The JPEG Trust Framework

Currently JPEG Trust consists of one part, the Core Foundation (ISO/IEC 21617-1). This foundation handles three main areas: annotating provenance information, extracting and evaluating trust indicators, and handling privacy and security concerns. The following three sections elaborate upon these aspects in further detail.

Annotating provenance information

JPEG Trust standardises means to link media assets together with their associated provenance annotations in a tamper-evident manner. The presence or absence of this information provides the contextual information needed for the establishment of trust in the media asset. The model for storing and accessing media asset provenance information is aligned with the industry supported Coalition for Content Provenance and Authenticity (C2PA) specification²⁴. The C2PA specification defines the technical means for combining statements of fact together with a digital signature. Some of these facts include cryptographic bindings, information about human, non-human, and AI-based actors and the actions (creation, editing, type of editing etc.) that were performed in the creation or modification of the media asset. Hence, existing media assets that have C2PA-compliant provenance information are fully compatible with the JPEG Trust framework. JPEG Trust also adds additional provenance functionality such as signalling the extent of modifications.

²⁴ <https://c2pa.org/specifications/specifications/1.4/index.html>

Extracting and evaluating trust indicators

The JPEG Trust framework specifies how to extract an extensive array of Trust Indicators from any given media asset. These indicators can originate from the metadata, the media content or provenance information. Subsequently, specific conditions for trustworthiness, expressed in Trust Profiles, allow individuals, organisations, and governing institutions to evaluate relevant trust indicators according to the requirements for their specific usage scenarios. The resulting evaluation can be expressed in a Trust Report to make the information easily accessed and understood by the end user.

Handling privacy and security concerns

The JPEG Trust framework provides a mechanism to annotate media assets with information about their trustworthiness. In many scenarios, it is important that this information can be protected from exposure. JPEG Trust also provides the means to protect that information using means based on the provisions of JPEG Privacy and Security (ISO/IEC 19566-4). Privacy provisions allow for the protection of information about an image when appropriate.

Next steps

The JPEG Trust Part 1 - Core Foundation is expected to be published as an International Standard later in 2024. This first part is the starting point of the JPEG Trust framework that will evolve over time and be extended with additional functionalities in the future. These functionalities will include support for media tokenization such as declaration of authorship, ownership and terms of use as well as specifications to enable workflows that require watermarking.

“Net Neutrality for a better Electronic Democracy” by Valentina Tirloni

Democracy is the future of our free societies: we must keep it safe in order to keep alive our rights and freedoms and to live in stable and peaceful societies.

Information and Communication new technologies have had a deep and strong impact on human lives and activities since the digital paradigm was established as metaphysically relevant. Politics too is a field impacted by those new technologies that have radically changed the way citizens deal with policy, elections, information, and political debates. If traditional medias have a vertical approach to information (top down), nowadays, by using ICT, citizens can contribute to create, broadcast and exchange information through an horizontal way that cuts away any mediation. Citizens became “prod-users” of information so that a general populism could appear. The lack of mediation could be a risk for democracy.

The major disruption came with Web 2.0, when users changed their status and became producers (*produser*) and broadcasters or distributors of information. Information channels became horizontal, and the media gradually lost their legitimacy, authority and credibility. New "journalists" are appearing on web pages, and the code of ethics - though not mandatory or binding in some countries - is losing all validity. Digital readers no longer seem to have the protection, guidance or mediation they need to make sense of the far more complex technical and political content that *reductio ad unum* seems to present them with. This tendency to "bypass" the mediation of experts (regardless of whether they really are experts or not) is one of the current characteristics of communication and information, whether encouraged by private economic operators, defended by users or criticised by the "old" expert mediators.

The access to information, its truthfulness, its verification, its production, and dissemination are crucial aspects for a healthy democratic system. The proliferation of social media has led to the loss of the intermediation of traditional media, to a growing divide between opinions and individuals, changing drastically the Habermas' theory on public space that is nowadays closer to a mosaic of separate individuals. Citizens, whose main communicative flaw is that of being extremely fragmented, are finding new impetus in the new technologies of communication: they can now get together, mobilise; they can also make information (no matter whether true or false) and can animate, moderate political debates thanks to the creation of blogs or discussion threads. In political science terms, citizens are acquiring, and reinforcing their power of agenda setting and agenda building. With new ICT, the process of information has changed according to a large importance to technical means. If technical means are designed following specific technical biases, the risk is that no liberty is still possible on the actual Internet. In order to have a good democracy, we need informed citizens so that information is the fourth power: the one who controls information, has the power.

What is information? We can define it as a representation of what is happening. From a more philosophical point of view, we should consider that there are possible different ways of encoding reality, according to specific political ideas or economical visions. The possibility of discussing those different representations is the main core of politics since centuries ago when sophists were able to treat a topic from opposite perspectives.

Furthermore, in the last years, disinformation, fake news, and misinformation are very central in the social debate and political arenas: that is the essence of politics. There are other two important phenomena related to information that do count in democratic systems: silence and reticence. Those two are as dangerous as disinformation and fake news because not communicating an information means hiding part of what is happening to citizens. A partial representation of external word is a threat for citizens unable to have the full real framework for decision making.

New technologies are the future of our *modus operandi*: Techno-democracy is the process of digitalisation that new technologies could offer to increase citizens' participation and social engagement. Digital democracy is a political-social phenomenon which first emerged some fifteen years ago in separate and diverse national contexts that cannot be reduced to e-voting procedures.

The main technological tool today is the Internet, that is the crucial and metaphysical structure of our world we inhabit, and the infrastructure that makes possible political action and social debate. Internet is the technical infrastructure of the digital paradigm and also the condition of possibility for information and therefore democracy. In 1996 at Davos, John Perry Barlow imagined Internet as a truly free space for people to connect without external control from providers, according to the basic end-to-end principle. Nowadays, the economic model of the Internet introduces technical biases: if you want to obtain content, you must pay in order to have a fast and powerful connection. We are so far from the ideologically free idea of the Internet.

The risk is that this instrument is already deeply unequal and biased by an economic model that shows little concern for the protection of fundamental rights and freedoms. There is a staggering distance between today's Internet and its disruptive aurora, steeped in Californian counterculture: as long as the Internet is configured in this way, no form of safe digital democracy will be possible. If net neutrality is not protected, every bit of information will be merchandised. According to every national legal system, the way net neutrality is ruled can vary: therefore, there can be a strong difference between US and EU law systems in treating the idea of net neutrality. For instance, in EU, "the open Internet" does not forcedly imply that same Internet is also free.

Furthermore, users utilise tools that have been designed for pre-established uses, and which are rarely questioned or challenged by those same users. The design of a tool is always driven by a purpose, a strategic choice. An instrument that does not serve a purpose would be a pure art product. What is more, in the digital and electronic technological age, the question of the neutrality of technology is more easily resolved. Whereas in the past philosophical ethical dissertations about neutrality used ambiguous examples, such as knives and firearms, today's electronic products are more enlightening: is there such a thing as an algorithm that serves no purpose? Can I run software that does not have an output to produce? So, technics is always a political object and is never neutral because of its purposes.

Every technical object is always a political object: technics is a means and as such it is designed for a purpose. According to ethical enquiries, a purpose is always a political design.

The access, the production and broadcasting of information are very important and crucial parts of a healthy democracy.

“The GoTriple platform: mediation to access scientific results”

by Sy Holsinger

The GoTriple platform highlights the importance of accessible and open scientific resources, particularly in the Social Sciences and Humanities (SSH). Developed through the TRIPLE project, funded under the Horizon 2020 program, GoTriple aims to be a transformative tool for interdisciplinary research, facilitating the discovery, reuse and collaboration of scholarly SSH resources across Europe. It serves as an open science hub for SSH, providing multilingual access to millions of publications, datasets and project information. It addresses the need for a unified platform where researchers, educators and policymakers can interact with SSH content and with each other. By centralising verified, reliable resources, GoTriple also combats disinformation by giving users access to trustworthy information and fostering transparency.

Platform Features, Community Engagement, and Multilingualism

GoTriple’s key features includes a powerful search engine, data visualisation tools, a recommendation system and web annotation, enabling users to analyse, contextualise and navigate SSH content effectively. Researchers and authors create profiles that allow users to connect, reinforcing transparency by linking research to specific individuals, which reduces risks of misattribution and misuse—a key element in the fight against disinformation. Since its launch, GoTriple has experienced rapid growth, with monthly visitors increasing from 1,000 to more than 30,000 over the last eight months and 1,400+ data sources from 20 major (well-known) providers, giving users access to a broad array of curated SSH resources. By maintaining high-quality sources, GoTriple minimises exposure to unreliable information, directly addressing risks associated with misinformation.

The platform’s multilingual interface, available in nine languages, broadens accessibility for Europe’s linguistically diverse research community. Keywords are searchable in 11 languages, making it easier for users to locate reliable information across language barriers. By enhancing cross-linguistic search capabilities, GoTriple prevents the spread of disinformation that can result from translations or linguistic misunderstandings.

Innovative Services and AI Integration

GoTriple is experimenting with an AI-driven chatbot, designed as a “Research Assistant” to help users interact with resources more intuitively. This chatbot moves beyond basic keyword searches, offering users conversational interactions to find specific research answers. The platform’s AI features are limited to only the content within GoTriple and is being carefully designed to ensure transparency, by offering insights into the sources used in generating responses, which contributes to ethical AI practices and allows users to verify the authenticity of the information they receive. In a time where AI can often generate content without verifiable sources,

GoTriple's transparent approach to AI is directly tackling concerns about AI-driven disinformation.

Governance and Sustainability

A long-term focus on financial sustainability is also central to GoTriple's strategy. Although a pricing model is under discussion, GoTriple's management team continues to explore alternative business models that balance financial viability with the platform's mission of open access. Sustaining GoTriple's reliable, verified sources financially is a key factor in maintaining a dependable resource that users can trust, further reducing economic pressures that sometimes lead other sources to compromise on quality or open access, which can fuel disinformation.

The platform is structured around a shared service delivery model that benefits from its community-driven governance. By involving users in its governance, GoTriple minimises deviations from its commitment to openness and social responsibility, ensuring that the platform remains aligned with the values of the SSH community. This user-governance approach also creates a community standard for quality and reliability, where users are empowered to hold the platform accountable and collectively work against potential sources of misinformation.

Conclusion

GoTriple's commitment to **accessible research outputs** promotes the consumption of accurate information by making reliable resources available to a broad audience. The platform's design supports transparency by linking research to individual profiles, clarifying authorship and reinforcing content credibility. This visibility makes it harder for misinformation to flourish, as verified sources and their origins are clearly identifiable. **Multilingualism** plays a crucial role in GoTriple's accessibility strategy, ensuring that research materials are available across languages to reach a diverse audience. This multilingual approach reduces the potential for misinterpretation when research crosses linguistic boundaries—a key risk in the spread of disinformation.

Visualisation tools enhance user experience by providing context, fostering users' trust in scientific content and reducing the risk of oversimplified or misleading interpretations. The platform's **verified data providers** and sources ensure that GoTriple holds itself to maintaining high standards of accuracy and trustworthiness, building a robust defence against misinformation. **Community engagement** is also vital; researchers and users can directly connect with peers and experts, enabling them to clarify information, share insights and perform fact-checking collaboratively. This community-centred approach encourages a culture of ongoing feedback and improvement, which aligns GoTriple with the changing needs of its users and strengthens the platform's credibility as a defence against disinformation.

GoTriple also places a strong emphasis on **ethical AI** use. By clearly indicating the sources in AI-generated responses, the platform enables users to trace the origins

of information, reducing the likelihood of AI-driven misinformation. Documentation on data model tuning will be made available as much as possible, which helps users understand the processes shaping AI interactions on the platform. Although GoTriple's AI models are trained on content from reputable sources, the platform continues to address ethical concerns about data ownership, ensuring responsible AI development within an open-access framework.

In summary, GoTriple is advancing SSH research by combining accessibility, collaboration and ethical AI to create a comprehensive and trustworthy research tool. Through these efforts, GoTriple not only enhances SSH access but also provides a safeguard against misinformation that can erode trust in scientific research.

“Building Digital Trust for a Human-Centered Internet” by Gaël Van Weyenbergh

The trust, shared interests, and mutual accountability that underpin real-life social networks—enabling reciprocal help and cooperation—are largely absent in the digital realm. Instead, our digital networks provide a framework for large-scale communication, not for large-scale cooperation. This creates a paradox: while we can detect humanity's collective problems, identify solutions, and connect billions of minds for idea exchange, we lack the means to translate this potential into coordinated action. Internet pioneers have even suggested that the digital revolution has yet to truly begin.

To resolve this paradox, we propose learning from natural networks of cooperation, exemplified by the "small world phenomenon." Popularized by the concept of "six degrees of separation," this idea suggests that any two individuals on Earth are connected by a surprisingly short chain of intermediaries. These networks develop naturally, requiring no central supervision, and span diverse social groups, fostering resilience and trust across boundaries. Built on personal relationships shaped by trust and shared interests, they evolve through individuals' conscious or unconscious assessments of time, attention, and emotional investment.

Trust, however, is not certainty—it is a belief, inherently shaped by risk. In an age of increasing interconnectivity and interdependencies, where everything affects everything else, traditional risk-averse approaches—where trust does not align with a business model—are no longer sufficient. Shielding off "bad things" within clear group boundaries is inadequate in systems without clear edges, where solutions often emerge from unexpected places and relationships beyond formal authority. This calls for a risk-aware mindset that embraces uncertainty while adapting to the dynamics of boundaryless networks.

While the original ethos of the internet was defined by cooperation, commercial interests have shifted its focus to communication. Today's online social networks

prioritize engagement to drive profit, often shielding users from meaningful participation and undermining the dynamics that foster trust, shared interests, and mutual accountability. To rekindle the cooperative spirit of the internet, we suggest a dual approach. First, digital tools, along with their incentives and governance structures, must adapt to natural social dynamics rather than distort them. Second, community dynamics or social capital must be translated into the digital realm without eroding their essence. We propose innovating with a "few-to-few" architecture—small, tightly connected networks designed to support personal relationships defined by trust and shared accountability—over the "one-to-many" influencer model or the "many-to-many" anonymous model. By realigning digital networks with the principles of real-life cooperation, we can unlock the internet's untapped potential for large-scale collective action.

The Challenges of Digital Trust

The current digital landscape reveals a significant trust deficit. While social platforms promise to connect individuals globally, their reliance on profit-oriented algorithms often prioritizes engagement over authenticity. As a result, users find themselves in environments that erode trust rather than foster meaningful connections. Even newer experiments in decentralization associated with Web3 introduce complexities, such as tokenized interactions, that risk commodifying relationships instead of strengthening them (De Filippi & Loveluck, 2016).

This "trustless internet" is characterized by several paradoxes. For example, while digital platforms have enabled the identification of problems and resources, they have failed to facilitate self-coordination and collective action at scale, leaving critical global challenges unresolved (Shirky, 2008). Additionally, the internet's democratization of information has led to the proliferation of misinformation, echo chambers, and AI-generated disinformation, undermining public discourse (Lecheler & Egelhofer, 2022). Finally, although these platforms have connected individuals, they have also contributed to social atomization, weakening the sense of genuine community (Turkle, 2011).

Marginalized communities face unique challenges in this ecosystem. For example, Indigenous groups attempting to safeguard their cultural heritage often find their narratives overshadowed or distorted by dominant digital platforms driven by profit incentives. Algorithms may amplify stereotypes or reduce their rich histories to commodified snippets, eroding trust in these platforms as allies. Furthermore, the lack of community-led governance structures leaves these groups vulnerable to exploitation or exclusion, reinforcing systemic inequities rather than addressing them (Carroll et al., 2020).

Cumulatively, these paradoxes hinder the broader societal benefits that digital technologies aspire to deliver. Instead of fostering collaboration, innovation, and shared understanding, the current internet often exacerbates divisions and isolates individuals. Addressing these challenges requires a fundamental rethinking of the structures and incentives that underpin digital trust.

Building Blocks for a Trustworthy Digital Ecosystem

To restore trust, digital ecosystems must embrace human-centric designs that prioritize authentic interactions and equitable governance. First, a shift in network architecture is essential. Current models—characterized by vast, many-to-many, impersonal networks—should be replaced by smaller, community-oriented clusters that mirror offline social dynamics. For instance, capping the size of personal networks at Dunbar's number, the cognitive limit of approximately 150 stable relationships, fosters deeper connections and more meaningful engagement. Few-to-few networks reflect the "small worlds phenomenon," the natural network of cooperation that already exists within human society, as demonstrated by the six degrees of separation theory (Milgram, 1967).

Second, online platforms must redefine their incentive structures. Instead of metrics like follower counts or engagement rates, which encourage superficial interactions, digital environments should prioritize social capital. Social capital refers to the networks of relationships and the trust, reciprocity, and shared norms that facilitate collective action and cooperation (Granovetter, 1985). For example, emotionally resonant digital platforms foster trust by connecting individuals with shared experiences, values, or interests. By prioritizing authenticity through meaningful interactions and emphasizing reciprocity, these platforms enable users to cultivate deeper connections, grounded in mutual understanding and trust—ultimately enriching both personal relationships and the broader social fabric.

Second, online platforms must redefine their incentive structures. Instead of metrics like follower counts or engagement rates, which encourage superficial interactions, digital environments should prioritize social capital. Social capital refers to the networks of relationships and the trust, reciprocity, and shared norms that facilitate collective action and cooperation (Granovetter, 1985). For example, the community-driven digital platforms foster trust by connecting neighbors to exchange information, goods, and services within their local areas. By prioritizing authenticity through verification processes and emphasizing reciprocity, these help residents form meaningful connections while strengthening social capital within their neighborhoods. This demonstrates how localized digital communities can translate online interactions into tangible, real-world trust.

By aligning incentives with community-driven values, platforms can encourage the distribution of accountability, shifting power from algorithms to users (Bouncken, Gast, Kraus, & Bogers, 2015). Peer evaluation, shared values, and interdependent interests become central to maintaining trustworthiness, as networks of peers can supersede traditional game theory mechanics by fostering collaborative rather than purely competitive dynamics. In such systems, individuals are motivated to act ethically and cooperatively because their long-term success depends on mutual trust and accountability.

Finally, governance must transition from autocratic, top-down models to community-driven approaches. Data cooperatives, as emerging intermediaries, exemplify this shift by allowing users to collectively manage and benefit from their

data. These cooperatives ensure democratic decision-making and equitable distribution of value, addressing many of the trust-related challenges posed by traditional platforms (Bühler et al., 2023). This shift is not merely a matter of top-down or bottom-up governance but rather relies on a dynamic micro-macro feedback loop. These relationships enable the emergence of self-maintained structures where individual actions influence systemic outcomes, and systemic conditions, in turn, shape individual behavior. Such feedback loops create resilient ecosystems that balance centralized governance with local autonomy, fostering trust and accountability at all levels of interaction.

The Promise of Data Cooperatives

Data cooperatives represent a transformative model for rebuilding digital trust by offering an alternative to the centralized structures of Web2 and the overly commodified dynamics of Web3. A data cooperative is a collectively governed organization where members pool their data to achieve shared benefits, including enhanced privacy, equitable data usage, and fair distribution of value. Rooted in the principles of democratic governance, collective ownership, and fiduciary accountability, these cooperatives empower individuals and communities to reclaim agency over their digital interactions and data management.

Historically, cooperatives have played a pivotal role in balancing economic and social power during periods of rapid industrial and technological change. From trade unions to cooperative banks, they have demonstrated the capacity to distribute power and resources more equitably. Data cooperatives continue this tradition by addressing the challenges of the digital economy, offering a framework that prioritizes collective interests over profit-driven motives (Hardjono & Pentland, 2019).

What sets data cooperatives apart is their fiduciary representation. Unlike traditional Web2 platforms, which often exploit user data for profit, and Web3 models, which can commodify interactions through tokenization, data cooperatives act as stewards of member data. This fiduciary duty ensures that the cooperative operates transparently and in alignment with the collective good of its members. Accountability mechanisms within these cooperatives reinforce trust, as members have both visibility into and control over how their data is used.

Importantly, data cooperatives do not overshadow the community dynamics that underpin the small world phenomenon—the natural networks of trust and cooperation that exist within human society. While many traditional social platforms attempt to leverage these dynamics, their potential is often choked by the superior profit-driven interests of the platforms themselves. Data cooperatives, by contrast, enhance these community networks by preserving their organic nature and allowing trust to flourish without interference from external profit motives.

The Data Governance Act (DGA) introduces a favorable legal framework for the growth of data cooperatives within the European Union. By recognizing data cooperatives as intermediaries, the DGA provides support for their development. This framework legitimizes the role of cooperatives in the digital economy and sets the stage for their broader adoption. However, the DGA also presents challenges, such as navigating potential national deviations and integrating cooperatives into the broader EU data law landscape (Bayamlioglu, 2021).

The applications of data cooperatives are wide-ranging. For instance, they can facilitate urban policy planning by pooling data to create smarter, more sustainable cities (UNCTAD, 2024). In healthcare, data cooperatives enable secure, collective data sharing to advance medical research while safeguarding patient privacy. They also play a crucial role in combating disinformation by enforcing transparency and fostering diversified information flows, reducing the risks of algorithmic manipulation and echo chambers. Indigenous communities can significantly benefit from data cooperatives, using them to protect cultural heritage and manage natural resources while avoiding the pitfalls of digital colonialism (Carroll et al., 2020).

Despite their promise, data cooperatives face several challenges. Legal ambiguities, limited public awareness, and scalability issues hinder their widespread adoption. Addressing these obstacles requires strategic alliances with stakeholders, targeted educational campaigns to demystify the cooperative model, and innovative governance frameworks that can adapt to evolving digital landscapes (Mannan, Bietti, Etxeberria, & Wong, 2019).

By harmonizing with global regulatory trends such as the DGA, data cooperatives can provide a replicable model for digital trust worldwide, balancing local community needs with international governance frameworks. These structures offer a pathway for a more equitable and accountable digital ecosystem, demonstrating the transformative potential of collective action in the digital age.

Actionable Framework for Digital Trust

A scalable trust model must balance emotional resonance with digital scalability. Emotional resonance—the ability to foster genuine and meaningful connections—is essential because trust is inherently relational and thrives on shared values, mutual understanding, and empathy. Without this foundation, digital interactions risk becoming purely transactional, undermining the collaborative potential of online networks. At the same time, scalability is critical for addressing the complex, interconnected challenges of modern digital ecosystems. Achieving both requires innovative approaches that combine personal trust with systemic reach.

Leveraging trust transitivity, where trusted peers act as bridges to expand networks, is a key strategy for scaling trust without diluting its integrity. This approach mirrors natural social dynamics, enabling trust to propagate through interconnected

communities. Blockchain technology can further enhance transparency by providing immutable records of data provenance, reinforcing trust in decentralized systems (Szabo, 2017). For instance, blockchain's role in verifying supply chains highlights its capacity to enhance transparency without dictating decision-making. By providing a tamper-proof record of product origins and movements, blockchain ensures accountability while allowing human-led governance to manage broader strategic decisions.

To ensure accountability, peer-led governance systems are vital. These systems prioritize relational integrity by enabling communities to collectively manage trustworthiness. For example, capped personal networks—reflecting authentic relationships—can act as units of social capital or curated pools of opportunities one wants access to, creating an attractive force that incentivizes community-positive behavior. This dynamic encourages individuals to align their actions with shared values and mutual accountability, reinforcing trust and cooperation. By replacing superficial metrics such as follower counts, these networks ensure that digital interactions are rooted in meaningful connections (Edunov, Bhagat, & Filiz, 2016).

Future innovations must also address trust challenges associated with artificial intelligence. Ensuring that AI systems are transparent, unbiased, and culturally sensitive is critical for maintaining trust in increasingly automated digital ecosystems (Walz & Firth-Butterfield, 2019). These systems must operate within ethical boundaries, with mechanisms in place to prevent harm and promote fairness. Additionally, governance structures must scale without sacrificing local authenticity. Scale-independent architectures—designed to operate effectively across different levels of interaction—offer a promising solution. These architectures enable decentralized frameworks to integrate seamlessly with global coordination efforts, striking a balance that preserves trust across diverse contexts (Watts & Strogatz, 1998).

By combining emotionally resonant interactions, transparent technology, and adaptive governance, this framework lays the foundation for a trustworthy digital ecosystem. It demonstrates that trust is not merely a static attribute but a dynamic and scalable quality that can flourish in both personal relationships and systemic structures.

Conclusion

Trust is both a fragile and antifragile cornerstone of human interaction. It is fragile because it takes time to build yet can be destroyed in an instant through breaches or betrayals. At the same time, trust is antifragile—it has the potential to grow stronger when nurtured under conditions of mutual accountability and shared values. This duality makes trust a critical yet elusive foundation for digital ecosystems, as highlighted in the introduction: without trust, the internet is reduced to a tool for communication, unable to achieve its collaborative potential.

While digital platforms have often eroded trust by prioritizing scale and profit over relational integrity, emerging frameworks rooted in human-centric values and adaptive governance offer a path toward restoration. Data cooperatives exemplify this potential, transforming data governance into an equitable and community-driven process where individuals and communities retain control over their digital interactions and benefit from collective ownership (Bühler et al., 2023). By prioritizing transparency, accountability, and community empowerment, these cooperatives address many of the structural flaws in existing digital platforms.

Building a trustworthy internet requires collective action across disciplines—ethics, sociology, technology, and governance. It demands systems that balance emotional resonance with scalability, align incentives with community-driven values, and establish governance models that are both locally authentic and globally coordinated. By prioritizing trust, the internet can rekindle its original spirit and evolve into a collaborative, equitable ecosystem. This transformation would not only restore confidence in digital networks but also empower individuals and communities to navigate an increasingly interconnected world with agency, resilience, and shared purpose.

References

- Ashby, M. (2020). Ethical regulators and super-ethical systems. *Systems*, 8(4), 53. <https://doi.org/10.3390/systems8040053>
- Bayamlioglu, E. (2021). Data cooperative: A new intermediary on the horizon. KU Leuven, Center for IT and IP Law.
- Boldyreva, E. L., Grishina, N. Y., & Duisembina, Y. (2018). Cambridge Analytica: Ethics and online manipulation with decision-making process. In V. Chernyavskaya, & H. Kuße (Eds.), *Professional Culture of the Specialist of the Future, vol 51. European Proceedings of Social and Behavioural Sciences* (pp. 91–102). Future Academy. <https://doi.org/10.15405/epsbs.2018.12.02.10>
- Bouncken, R. B., Gast, J., Kraus, S., & Bogers, M. (2015). Coopetition: A systematic review, synthesis, and future research directions. *Review of Managerial Science*, 9(4), 577–601.
- Bühler, M. M., Calzada, I., Cane, I., Jelinek, T., Kapoor, A., Mannan, M., Mehta, S., Mookerje, V., Nübel, K., Pentland, A., Scholz, T., Siddarth, D., Tait, J., Vaitla, B., & Zhu, J. (2023). Unlocking the power of digital commons: Data cooperatives as a pathway for data sovereign, innovative and equitable digital communities. *Digital*, 3(3), 146–171.
- Carroll, S. R., Garba, I., Figueroa-Rodríguez, O. L., Holbrook, J., Lovett, R., Materechera, S., Parsons, M., Raseroka, K., Rodriguez-Lonebear, D., Rowe, R., Sara, R., Walker, J. D., Anderson, J., & Hudson, M. (2020). The CARE principles for Indigenous data governance. *Data Science Journal*, 19. <https://doi.org/10.5334/dsj-2020-043>
- De Filippi, P., & Loveluck, B. (2016). The invisible politics of Bitcoin: Governance crisis of a decentralized infrastructure. *Internet Policy Review*, 5(3). [https://doi.org/\[specific DOI\]](https://doi.org/[specific DOI]).
- Edunov, S., Bhagat, S., Moira, Diuk, C., & Onur Filiz, I. (2016, February 4). Three and a half degrees of separation. *Facebook Research*.
- Granovetter, M. (1973). The strength of weak ties. *American Journal of Sociology*, 78(6), 1360–1380.
- Granovetter, M. (1985). Economic action and social structure: The problem of embeddedness. *American Journal of Sociology*, 91(3), 481–510.
- Hardjono, T., & Pentland, A. (2019). Data cooperatives: Towards a foundation for decentralized personal data management. Cornell University.
- Lecheler, S., & Egelhofer, J. L. (2022). Disinformation, misinformation, and fake news: Understanding the supply side. In J. Strömbäck, S. Wikforss, K. Glüer, T. Lindholm, & H. Oscarsson (Eds.), *Knowledge resistance in high-choice information environments* (pp. 88–106). Routledge. <https://doi.org/10.4324/9781003111474>
- Milgram, S. (1967). The small-world problem. *Psychology Today*, 1(1), 61–67.

Shirky, C. (2008). *Here comes everybody: The power of organizing without organizations*. Penguin Press.

Szabo, N. (2017). Money, blockchains, and social scalability. Blog.

Turkle, S. (2011). *Alone together: Why we expect more from technology and less from each other*. Basic Books.

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature*, 393(6684), 440–442.

Key Recommendations from the Mediation Track

Prepared by Suzanne Dumouchel, OPERAS

The **TrustOn2024 event** addressed the complex issue of trust and disinformation, proposing a variety of solutions to rebuild trust in information systems and foster healthier digital communities. The discussions were centred around the importance of reliable sources, engagement within communities, technological design, and the broader information landscape. Below is a synthesis of the key discussion points, challenges, and actionable recommendations that emerged from the event.

Key Discussion Points

1. Trusted Sources & Data Cooperatives

Participants emphasised the need to promote **trusted sources** of information to counter disinformation. One proposed solution was the creation of **data cooperatives**, where communities collaboratively share, validate, and authenticate information, ensuring that it aligns with collective trust values.

2. Reciprocity & Community Engagement:

Trust was linked to **reciprocity** and deep community engagement. Building trust within digital spaces requires meaningful connections between community members. It was noted that **similar individuals collaborating** could foster strong relationships, but care should be taken to avoid excessive **homogeneity** that stifles diversity of thought.

3. Argumentation over Polarisation:

A key strategy to combat polarisation involves **encouraging nuanced argumentation** rather than reducing conversations to binary yes/no choices. Facilitating thoughtful and respectful **dialogue** can help reduce extremes and foster a more inclusive discussion space.

4. Tech Design and Incentives:

Designing technology that encourages **meaningful engagement** rather than simplistic feedback mechanisms (e.g., thumbs up/down) is essential. The focus should be on training tech designers to create systems that reward **thoughtful contributions** and also consider **ecological impacts** alongside user engagement.

5. Comprehensive Information Landscape:

Understanding the broader information ecosystem is critical, and **journalists** were identified as key players in maintaining this balance. The challenge is to mediate between **regulation** and **human behaviour**, as humans are naturally social beings, and their interaction with information is complex.

6. Mediation Levels:

The session discussed how information is mediated across various platforms and actors. Understanding the **different levels** of mediation and the roles of mediators (including platforms, journalists, and technology) is critical for addressing the spread of disinformation.

Main Challenges Identified:

1. Information Overload:

Participants acknowledged the overwhelming nature of **information overload**, where users are bombarded with too much data, leading to **fatigue** and reduced ability to critically assess the information.

2. Conscious Consumption:

Defining what constitutes **conscious consumption** of information is a major challenge. Guiding citizens through this complex landscape requires clear strategies for **distinguishing important** from less relevant information.

3. Rebuilding Trust:

The erosion of **trust** in information sources and media was highlighted as a major issue. Finding ways to **rebuild trust**—through transparency, accountability, and the promotion of credible sources—was considered essential.

4. Governance:

Effective **governance** of mediation and information-sharing processes is a central challenge. This includes exploring self-regulation versus external governance to ensure that mediation is done fairly and effectively.

Recommendations and Proposed Actions

1. Community-Based Groups:

Form **community groups** to define core values and act as mediators. These groups should build on **social capital** and local trust to moderate and validate information within the community.

2. Honesty and Lifelong Learning:

Encourage **honesty** in public discourse, such as rewarding corrections and publishing errata. Promoting a culture of **lifelong learning** is also crucial to empower individuals to navigate the evolving information landscape.

3. Self-Regulation & Legal Frameworks:

A balance between **self-regulation** and the creation of **legal frameworks** was suggested. While communities should have the autonomy to regulate information flows, overarching legal frameworks are needed to ensure accountability and protection against disinformation.

4. Pragmatic Information Selection:

Participants recommended a **pragmatic approach** to information consumption, teaching users how to select and prioritise information.

Training should also extend to those who design and manage platforms, emphasising empathy and user experience.

5. Education & Media Literacy:

Improving **digital literacy** is critical, especially for younger generations. This includes certifying experts, updating educational practices, and integrating **media literacy** into school curricula, ensuring both educators and students are equipped to critically assess information.

6. Ethical Frameworks:

Develop **ethical frameworks** that guide how digital interactions should occur, ensuring that all stakeholders—users, platforms, and regulators—adhere to standards that promote transparency and trust.

7. Multi-Dimensional Communication:

Recognize the importance of **multi-dimensional communication** by supporting diverse communication channels. Designs should accommodate a variety of interaction styles and foster richer dialogue rather than incentivizing simplistic exchanges.

8. Citizen Engagement:

Actively **involve citizens** in discussions about information mediation and governance. Communities should define their own values and roles, while legislative actions should support rather than undermine local efforts to build trust.

9. Transparency in Human-Machine Interactions:

Increase **transparency** in interactions between humans and digital systems, ensuring that algorithms and automated processes are understandable and accountable. This fosters critical thinking and reduces the potential for manipulation.

10. Inclusive Infrastructure:

Promote **inclusive infrastructure** and co-design processes that continuously evaluate the effectiveness of mediation efforts. Certification processes for technology that meets trust-building criteria were proposed to maintain accountability.

11. Diversity of Approaches:

Emphasise the need for **diverse approaches** to combat disinformation, recognizing that no single solution fits all contexts. Solutions must be adaptable to local cultures, values, and needs.

12. Disruption and Rebuilding Mediators:

Address the **disruption** of traditional mediation tools, such as professional journalism, and propose the creation of new mediators. This includes establishing new governance models to ensure effective oversight of digital

information flows.

13. Community Building and Knowledge Sharing:

Strengthen relationships between communities to enhance **knowledge sharing** and trust-building. Leveraging the inherent social nature of humans can help in restoring trust in mediated information.

14. Regulations and Community Values:

Regulations should be linked to **community-defined values**, ensuring that both governmental and community-based mechanisms align with local needs and foster trust rather than imposing top-down solutions.

15. Community-Controlled Frameworks:

Advocate for **community-controlled frameworks** that align with local values and ensure that mediation processes are transparent, democratic, and reflective of the communities they serve.

The TrustOn2024 event discussions highlighted the need for a **multi-faceted approach** to tackle disinformation, emphasising **community involvement, ethical frameworks, and technological design** that promotes trust and transparency. By fostering **reciprocity**, promoting **media literacy**, and involving citizens directly in the governance of information, a more resilient and trusted information ecosystem can be built. The recommendations propose **flexible, community-driven solutions** that are both pragmatic and adaptable, focusing on rebuilding trust from the ground up.

3. TrustOn2024 at the Science Summit

A follow-up of the event occurred in the context of the Science Summit of the 79th United Nations General Assembly in September 2024. OPERAS RI has chaired a session “Fostering Trust in the Digital Sphere: A Multi-Stakeholder Approach Introduction” on 11 September 2024 gathering 10 panelists that either contributed to the TrustOn2024 event or have been invited to bring another perspective.

This session was in line with the United Nations' vision for a digital future founded on trustworthiness. It emphasised a people-centred approach that goes beyond mere technological fixes.

The session aimed to delve into the multifaceted concept of trust within the digital sphere. It recognized that trust is not a singular entity but rather a complex interplay of factors including technological reliability, effective governance, fair regulation, unbiased mediation, and the advancements of science. By engaging a diverse array of stakeholders, the workshop sought to gain insights from various perspectives, ensuring a comprehensive understanding of trust dynamics.

Furthermore, the session sought to pioneer a collaborative approach to digital infrastructure development. Inspired by the ethos of community living labs, this approach emphasised the active involvement of stakeholders from different sectors. By fostering collaboration among civil society, academia, technology experts, youth advocates, policymakers, and others, the workshop aimed to co-create digital public infrastructures (DPIs) that are not only technologically robust but also socially responsible and ethically sound.

Moreover, the session acknowledged the critical importance of maintaining the quality and integrity of online content in an era dominated by AI technologies. With the proliferation of AI-driven disinformation campaigns, ensuring the credibility of online information has become paramount. Drawing from academic expertise in editorial practices, the session explored strategies to combat disinformation and uphold the standards of truthfulness and accuracy in digital spaces.

In summary, the session has represented a concerted effort to advance the United Nations' vision of a trusted digital future. By prioritising a human-centric approach, fostering multi-stakeholder collaboration, and promoting editorial integrity, it aimed to lay the groundwork for responsible and inclusive digital transformation. Through its practical recommendations, it started to pave the way for a more resilient and equitable digital landscape.

4. Conclusion

The TrustOn2024 workshop and the session at the Science Summit UNGA 79 marked a significant milestone in addressing the intricate dynamics of trust in the digital age, particularly concerning science and research. It needed a space to facilitate interaction, bring together diverse perspectives, interlink them, create new approaches, and contribute to building a supportive community. Bringing together experts from diverse fields, the workshop facilitated rich discussions that yielded several key outcomes and recommendations.

Three areas of achievement can be highlighted for a conclusion and a way forward. First, an **enhanced understanding of the dynamics of trust**: the efforts helped to deepen the understanding of how trust in science is evolving in the context of digital transformation, with presentations exploring the multifaceted nature of trust and considering factors such as technological reliability, governance and the role of AI in shaping public perceptions.

Second, the **interdisciplinary collaboration**: the first edition of the TrustOn2024 workshop and the following session at the Science Summit UNGA79 showed the potential for a collaborative environment, featuring stakeholders from academia, policy making, technology sector, and the civil society, all of them engaging in a meaningful dialogue. This interdisciplinary approach was pivotal in identifying common challenges and opportunities to foster trust in the digital space and also to strengthen trust in each other.

Finally, TrustOn2024 gave spotlight into **innovative strategies** for digital governance models. Many discussions emphasized the importance of community-driven frameworks and the potential blockchain technology in enhancing transparency, without compromising decision-making processes.

Besides the achievements, it is crucial to elaborate and retake the many recommendations offered by speakers and experts during the event to better grasp them in a path for the future:

1. **Promote Trusted Sources and Data Cooperatives:** TrustOn2024 participants underscored the necessity of promoting reliable sources of information and establishing data cooperatives. These cooperatives can serve as collaborative platforms for sharing, validating, and authenticating data, thereby enhancing collective trust.
2. **Enhance Public Engagement and Transparency:** Engaging the public through transparent and participatory practices was identified as crucial for building trust. Governments and institutions are encouraged to involve citizens in decision-making processes and provide clear and accessible information.
3. **Leverage Historical Lessons:** Drawing from the evolution of sectors like

medicine and transportation, TrustOn2024 experts recommended adopting regulatory frameworks and community engagement strategies that have proven effective in building trust.

4. **Implement Ethical AI Practices:** The integration of AI in research and communication should adhere to ethical standards that ensure transparency and accountability. Institutions are urged to critically evaluate AI tools to prevent the spread of misinformation.

The perspective of the **infrastructure track** showed that fighting misinformation and disinformation is a civic act and that it needs to be a global effort and an agreed upon definition of trust. Therefore the recommendations include:

- Establishing a digital bill of rights and a digital asses registry
- Reinforcing regional networks and create living labs as well as emphasising multidisciplinary collaboration
- Developing clear rules of participation and standardising practices
- Defining terms and establishing common terminology

Similarly, the **science track** identified the need for a holistic approach as necessary to foster trust in science, where all actors in the ecosystem of trust collaborate. This collaboration has to take place at the macro, meso and micro levels. At the macro level, global governance solutions are needed to COORDINATE misinformation and disinformation issues and to ENSURE freedom of science by protecting scientists. At the meso level, there is a need to DEVELOP guidelines for responsible science communication and ENCOURAGE more accessible communication of scientific results. Finally, at the micro level, an advocacy is needed to (1) DIVERSIFY the presentation of facts and values in narratives, (2) IMPLEMENT incentives to foster innovative hybrid collaboration approaches, (3) COMMIT to evidence-based practices and objectives.

In the **mediation track** the main challenges which were identified are: Information Overload, Conscious Consumption, Rebuilding Trust and Governance. Following discussions around these main challenges several actions and recommendations were identified, always considering a multi-faceted approach to tackle disinformation, emphasising community involvement, ethical frameworks, and technological design. The recommendations and proposed actions are:

- To form community-based groups to define core values and act as mediators
- Encourage honesty in public discourse and promoting lifelong learning
- Balance between self-regulation and the creation of legal frameworks
- Pragmatic approach to information consumption
- Improving digital literacy and integrating media literacy into school curricula
- Develop ethical frameworks
- Recognize the importance of multi-dimensional communication
- Actively involve citizens in discussions
- Increase transparency in interactions between humans and digital systems
- Promote inclusive infrastructure and co-design processes
- Emphasise the need for diverse approaches

- Address the disruption of traditional mediation tools and propose the creation of new mediators
- Strengthen relationships between communities to enhance knowledge sharing and trust-building
- Regulations should be linked to community-defined values
- Advocate for community-controlled frameworks that align with local values

The **first step** towards addressing trust in a multifaceted approach begins with the publication of this report. By bringing together key insights, research, and actionable recommendations, this document serves as a foundation for meaningful progress. We invite all stakeholders to engage with its findings and collaborate in building a human-centric and trustworthy digital future.

By considering several actions planned for 2025, from a policy brief to events, the journey towards fostering trust in the digital age for OPERAS is just beginning, and we invite you to join us in collectively taking action.

Annex

Biographies

Agata Gurzawska is Research Manager at Trilateral leading research group on ethics, human rights and emerging tech. She has worked on +15 EU-funded projects as a human rights and ethics expert. Currently, she is coordinating the Horizon Europe VERITY project on trust in science. Driven by curiosity and human rights values, she believes that strategic responsible innovation management (StRIM) can boost economic benefits and lead to socially desirable and ethically acceptable technology. **Science track**

Alejandra Michel is a senior researcher in digital law at the Research Centre in Information Law and Society (CRIDS/UNamur), where she is Head of the Media Law research unit. She also teaches media law, human rights in the digital environment and archives law and digital. Her research focuses on the regulation of online content moderation and the right to information. She is also a member of the Conseil de Déontologie Journalistique (French-speaking Belgium) on behalf of civil society. **Infrastructure track**

Amir Banifatemi is a seasoned technology executive and strategist with over 25 years of experience in AI and emerging technologies. He is an accomplished alliance builder, having managed private and public investment funds in the US and Europe. He co-founded the AI for Good Summit, fostering dialogues on human-centric AI for SDGs, and the AI Commons collective championing AI as a public empowerment tool. He co-chairs the Global Partnership on AI's Responsible AI group, shaping policies for ethical AI development and usage. Previously, as XPRIZE's Chief Innovation Officer, he oversaw global competitions advancing AI, robotics, digital health, and energy transition. **Infrastructure track**

Carolina Moreno-Castro is a Full Professor of Journalism at the University of Valencia and a member of its Institute on Social Welfare Policy (POLIBIENESTAR). She launched her research career in 1993 at the Institute for Advanced Social Studies (IESA-CSIC). Recognized for ensuring the societal impact of her work, she has collaborated extensively with organizations such as the Spanish Foundation for Science and Technology (FECYT) and the Organization of Ibero-American States (OEI). Her international experience includes research stays at prestigious institutions, including Cornell University, Harvard University, and Austral University. Currently, she leads the ScienceFlows group, where she coordinates European research initiatives, including Horizon 2020, Horizon Europe and Erasmus+ projects. Her primary research interests encompass science communication, risk communication, science journalism, and public perceptions of science and technology. **Science track**

As Senior Policy Advisor at the Belgian Science Policy Office (BELSPO), **Chris de Loof** works on digital strategies for Scientific Research and Culture Heritage. Chris is a longtime promoter of open and FAIR data and open science. He has an interest for innovative, cross-cutting collaborations between cultural heritage professionals and researchers, mostly those in the fields of the Humanities and Social Sciences. Chris is a delegate to the Commission expert group on the common European Data Space for Cultural Heritage (CEDCHE). He is also the Chair of the General Assembly of DARIAH, the European Research Infrastructures for digital humanities and arts, and Belgian representative in the European Research Infrastructure for Heritage Science (E-RIHS). In the past Chris has been involved in several successful EU funded projects, often in a managerial role. Chris studied Business Information Management at the KULeuven focusing on Information Technology, process- and data management. **Infrastructure track**

Christophe Calvin is CEA Senior Fellow and Director of Research at the Fundamental Research of CEA. He is currently the assistant to the director in charge of HPC, numerical simulation and digital health. He is in charge of the implementation of the institutional policy on scientific and technical data. **Mediation track**

Evelien Dholander and **Kevin Leonard** are data curators at Ghent University. As curators, they work directly with researchers to improve the FAIRness of their published data. They also work more broadly, both within Ghent University and at the Flemish and Belgian level, to improve the findability and accessibility of published research data. **Mediation track**

Frederik Temmermans is guest professor in multimedia at the Department of Electronics and Informatics (ETRO), associated with the Vrije Universiteit Brussel and imec. His research focuses on media privacy, security, authenticity, and integrity. He has been involved in various research projects in the medical, mobile, and cultural domains. Frederik is an active member of the JPEG standardisation committee (ISO/IEC JTC1/SC29/WG1) where he currently chairs the JPEG Systems and Integration subgroup and leads the JPEG Trust (ISO/IEC 21617) standardisation activities. Frederik is also co-founder of the VUB spin-off company Universum Digitalis. **Mediation track**

As a co-founder of Meoh.io, **Gaël Van Weyenbergh** is deeply engaged in advancing discussions on digital cooperation and on new digital intermediaries like data cooperatives that merge cooperative principles with the new asset class of data. Gael frequently shares his expertise and insights at various forums, striving to craft a more vibrant, equitable, and inclusive digital future so that the benefits of the digital transformation are shared by all. **Mediation track**

Guigone Camus is a Scientific and Technical Data Policy Officer at the French Alternative Energies and Atomic Energy Commission (CEA). As an anthropologist, her field of expertise also covers the Pacific area, specifically Kiribati. She is involved

in research projects related to the effects of global warming on oceans and on the livelihood of island communities. **Mediation track**

Hannes Lowagie is head of the catalographic section in KBR. He is currently working on the transition to RDA, Linked Open Data and the use of AI-technologies in the cataloguing process. He is Vice-Chair of EURIG and supervisor of MetaBelgica project that will be presented. He has a PhD in (Medieval) History (University of Ghent, 2012). **Infrastructure track**

Hendrik Berghäuser is a project manager in the Competence Center Policy and Society at Fraunhofer ISI in Karlsruhe. Since 2013, he has been working on numerous research projects and evaluations on science, technology and innovation policies. His research focus is on knowledge transfer between science and society, responsible research and institutional change within research organisations. Hendrik holds a PhD in Political Science and Economics from the German University for Administrative Science Speyer. **Science track**

Isabel Mendoza-Poudereux has worked in research since 2005, working both on the technical side and in communication and management roles. She has a PhD in Biotechnology and a Master's in Business Innovation. Currently, she is a member of the ScienceFlows Group at the University of Valencia, currently working on the COALESCE project, while also serving as the Chief Communications Officer for The Global Plant Council. **Science track**

Jörg Lehmann is a post-doctoral researcher employed at the Staatsbibliothek zu Berlin – Berlin State Library in the project "Human.Machine.Culture —Artificial Intelligence for Digital Cultural Heritage. Before he joined the Berlin State Library, he worked as a researcher at several universities, mostly in the fields of digital humanities and data science. At Berlin State Library, his current work centres on data provision and curation for AI, thus preparing data sheets and model cards as well as a field manual for the documentation of ethical, legal and social issues in cultural heritage datasets. **Infrastructure track**

Julija Baniukevic is a professional serving as the IANUS project coordinator on behalf of the Research Council of Lithuania, with a focus on enhancing trust in science. With a background as a European Commission expert-evaluator and experience as a liaison officer at the Lithuanian RDI Liaison Office in Brussels, Julija excels in bridging connections between Lithuanian scientists, innovators, and EU policymakers. Her growing influence as a LinkedIn influencer reflects her ability to engage a wide audience, while her role as a JRC trainer of Evidence for policy demonstrates her expertise in training researchers to effectively inform public policy design and navigate the intricate relationships between scientists, policymakers, and communicators. **Science track**

Mario Scharfbillig uses behavioural insights to improve evidence-informed policymaking and democratic processes in the EU. He is working at the Joint

Research Centre of the European Commission, leading the Enlightenment 2.0 research programme. He is member of the advisory board to the Democracy & Belonging Forum at the UC Berkeley and the Science for Policy Working Group at the European Geoscience Union. He received his PhD in economics from the University of Mainz, where he specialized in research on behavioural economics, public policy and corporate governance. **Science track**

Matei Mancas holds a PhD in applied sciences from the University Of Mons (UMONS), Belgium. He works on human visual attention modelling since his doctoral thesis and he is focusing on AI applied to humans. Dr. Mancas is part of the EU project AI4DENUK (<https://ai4debunk.eu/>) which focuses on the use of both social sciences and technologies to combat fake news. UMONS is involved in the technical part of the project. He has published several dozen of articles in peer reviewed international conferences and journals, and edited a book entitled "From Human Attention to Computational Attention: A Multidisciplinary Approach" published by Springer Series in Cognitive and Neural Systems. **Mediation track**

Mathilde Dorcadie is a French freelance journalist based in Brussels. Since 2018, she has worked as co-editor at Equal Times (www.equaltimes.org). She has been reporting on social and economic issues in Europe and beyond. Previously, she worked as a foreign correspondent in Brazil from 2014 to 2018 for French-speaking media outlets. **Mediation track**

Mei Lin Fung, Co-Chair of People Centered Internet which she co-founded with Vint Cerf. An early pioneer of CRM, working with Tom Siebel and Marc Benioff at Oracle, she also worked at Shell and Intel. Mei Lin studied Finance at MIT under two future Nobel Economics winners. She served as Socio Technical lead for the US government Federal Health Futures, and was finalist in the GSA Citizen Engagement GEAR challenge and chairs the IEEE Technical Committee on Sustainability. She organised "Program for the Future" 2008 with Douglas Engelbart and "40th Anniversary of the Internet" (www.tcpip40.com) in 2014 with Vint Cerf. Fellow of Hasso-Plattner-Institute. Mei Lin curates the People Centered Internet's Digital Cooperation and Diplomacy network and co-authored G7 policy contributions. She initiated and chaired the 3-day series of 60 speakers, including ITU Secretary General Doreen Bogdan Martin, Bob Kahn and Vint Cerf the fathers of the Internet as part of 10 panels on Digital Cooperation, Governance and Regulations at the UN Science Summit in Sept 2023. She organised with the IEEE History and David Gonzalez, IEEE Ambassador, the 50th Anniversary of the Internet. **Infrastructure track**

Nelson H. da Silva Ferreira is an invited researcher and open science officer at the University of Coimbra. His research and publications encompass a range of fields, including ancient linguistics, ancient Sumerian and Roman agriculture, the history of medicine, and open science infrastructures. Moreover, he teaches courses in Classical Philology and the History of the Ancient World, and engages in collaborative endeavours with other academic institutions and museums with the

objective of disseminating and transferring academic knowledge on SSH fields to the wider community. His research has primarily focused on the anthropological impact of economic dependence on agricultural production in ancient, unrelated contexts. This has led to an integration of the history of ancient agriculture with the history of medical texts as a source of knowledge regarding consumption and production habits, particularly in the Roman Empire. Recently, Nelson H. S. Ferreira has also been engaged in the study of pseudoscience and disinformation as a phenomenon, European projects on open science infrastructures and heritage and citizen science, and he is the project manager of CONVIVIUM (ID 101178921 — HORIZON-CL2-2024-HERITAGE-01). **Science track**

Panagiotis Monachelis belongs to the Laboratory Teaching Staff of the Department of Electrical & Electronic Engineering of the University of West Attica in Athens. He has two master's degrees in data communication and artificial intelligence and is a PhD candidate in the field of social media data visualisation. His research focuses on Social Network Analysis involving artificial intelligence techniques. He is involved in the EU-funded VERITY project investigating public's trust in science and has previously been involved in the EU-funded EUNOMIA project investigating trustworthiness on social media while has been participating in other projects in communication and dissemination strategies. **Science track**

Pierre Mounier is the co-coordinator of OPERAS Research Infrastructure with Suzanne Dumouchel. He supports cooperation between OPERAS members and contributes to the strategic roadmap of the infrastructure. Mounier is trained in classical studies and social anthropology. He is affiliated to the École des Hautes Études en Sciences Sociales (EHESS); deputy director of OpenEdition, the French national infrastructure dedicated to open scholarly communication in the SSH, and co-director of the Directory of Open Access Books (DOAB) with Niels Stern. He regularly publishes on digital humanities and open science topics, and more largely on the social and political impact of Information and Communication Technology (ICT). **Mediation track**

Reda Cimmerman is the Ombudswoman for Academic Ethics and Procedures of the Republic of Lithuania, holding a doctoral degree in biomedical sciences from Vilnius University and multiple degrees in microbiology, biochemistry, and management. Over 15 years at the Research Council of Lithuania, she has served as a Board Member, Scientific Secretary, and Head of the Science Foundation Secretariat. She has led European Commission-funded projects in academic ethics and open science, and initiatives for gender equality in science. She actively engages in various science and policy committees, including serving as Vice-Chair of the Lithuanian National UNESCO Commission. **Science track**

Susanna Fiorini is a translator and consultant in translation technology for research institutions and international organisations. Since 2020, she has been coordinating the Translations and Open Science project, funded by the French National Fund for Open Science, with the aim to explore the opportunities offered by translation

technologies to foster multilingualism in scholarly communication. **Mediation track**

Sven Lieber is a data manager in the research and innovation department at the Royal Library of Belgium (KBR). He contributes to making the data about Belgian cultural heritage available to researchers and the public by increasing data quality and improving data processing workflows. He mainly works on the projects MetaBelgica and BELTRANS. See <https://sven-lieber.org/en/> **Infrastructure track**

Sy Holsinger is the Chief Technology Officer (CTO) at the OPERAS AISBL. He is in charge of the technical vision and service strategy, coordinating the distributed teams behind the solutions and ensuring the delivery, sustainability, efficiency and effectiveness of the OPERAS portfolio of services. Sy has almost 20 years of experience in EU-funded projects related to the development, implementation and innovation of e-Infrastructures supporting research. In addition, he is a certified expert, trainer and auditor in both FitSM (Service Management) and ISO/IEC 27001 (Information Security) standards, and volunteers as the Co-chair of the FitSM Committee to evolve the standard. **Mediation track**

Deputy Managing Director of the Cairn.info platform, **Thomas Parisot** oversees the platform's institutional and editorial relations, and is also actively involved in the work of the Syndicat National de l'Édition (SNE) and the Groupement Français des Industries de l'Information (GF2I) in France. **Infrastructure track**

Tine Ravn is a Senior Researcher at the Danish Centre for Studies in Research and Research Policy within the Department of Political Science at Aarhus University. Her work broadly concerns the relationship between science and society, particularly focusing on public trust in science and research integrity, the social and ethical aspects of emerging biomedical technologies, and public engagement with science and society. She is currently acting as co-PI on a research project examining public trust in science in relation to research integrity and integration (POIESIS, supported by Horizon Europe). In addition, she is co-leading a work package aimed at understanding the ethical implications and public perceptions of organoid technologies (HYBRIDA, supported by Horizon 2020). **Science Track**

Valentina Tirloni is associate professor at Université Côte d'Azur in Nice. She graduated in Philosophy and Law at Pavia University. She also holds a PhD in Political Philosophy, a Professorial Thesis in Communication Studies and the professional habilitation as lawyer. Her research work deals with the anthropology and Ethics of Technics to investigate how new technologies have impacted human life, human body (transhumanism) and social life. In particular, she focused on the political impacts of technological communication means and electronic democracy. New technologies can be ambiguous in political Communication. **Mediation track**

Véronique Stoll is General Curator of Libraries and Director of the Paris Observatory

Library. As co-chair of the Research Data Working Group of the French Open Science Committee, she works in particular on data management and dissemination policies. **Mediation track**

OPERAS

open scholarly communication in the european
research area for social sciences and humanities

OPERAS AISBL

Quai aux Briques 76
1000 Brussels, Belgium
operas-eu.org

Contact:

Scientific Coordination: Suzanne Dumouchel, Coordinator (Partnerships),
suzanne.dumouchel@operas-eu.org

Communication: communication@operas-eu.org

DOI

[10.5281/zenodo.14621383](https://doi.org/10.5281/zenodo.14621383)



Editors:

Suzanne Dumouchel, OPERAS Coordinator (Partnerships)

Marlen Töpfer, OPERAS Communication Manager

Lorena Caliman, OPERAS Content and Communications Specialist

Carol Delmazo, OPERAS Service Marketing and Community Outreach Officer