



A Toolbox to Deal with Misinformation in Disaster Risk Management

By Rosa Vicari and Nadejda Komentatova (IIASA), with contributions from Irina Dallo (ETHZ), Abraham Yosipof (CLB), Camille de Carolis (EMSC), Carmit Rapaport (CLB),



Introduction & Background



Context & Objectives



Eight Stages to Tackle Misinformation



INTRODUCTION & BACKGROUND



- ❖ Recent literature on social **media misinformation** emphasize the importance of integrating **multidisciplinary research** :
 - ❖ Psychological principles,
 - ❖ AI tools and platform-based solutions,
 - ❖ and community-driven trust ratings.

- ❖ **There is currently no general framework to comprehensively address and manage misinformation on social media related to both anthropogenic and natural disasters and hazards.**

- ❖ Research focusing on **misinformation related to hazards and disasters** highlights that addressing fake news requires a combination of :
 - ❖ Fact-based refutation,
 - ❖ Transparent content moderation,
 - ❖ Authoritative sources,
 - ❖ Comprehensive educational strategies,
 - ❖ And advanced technological tools.

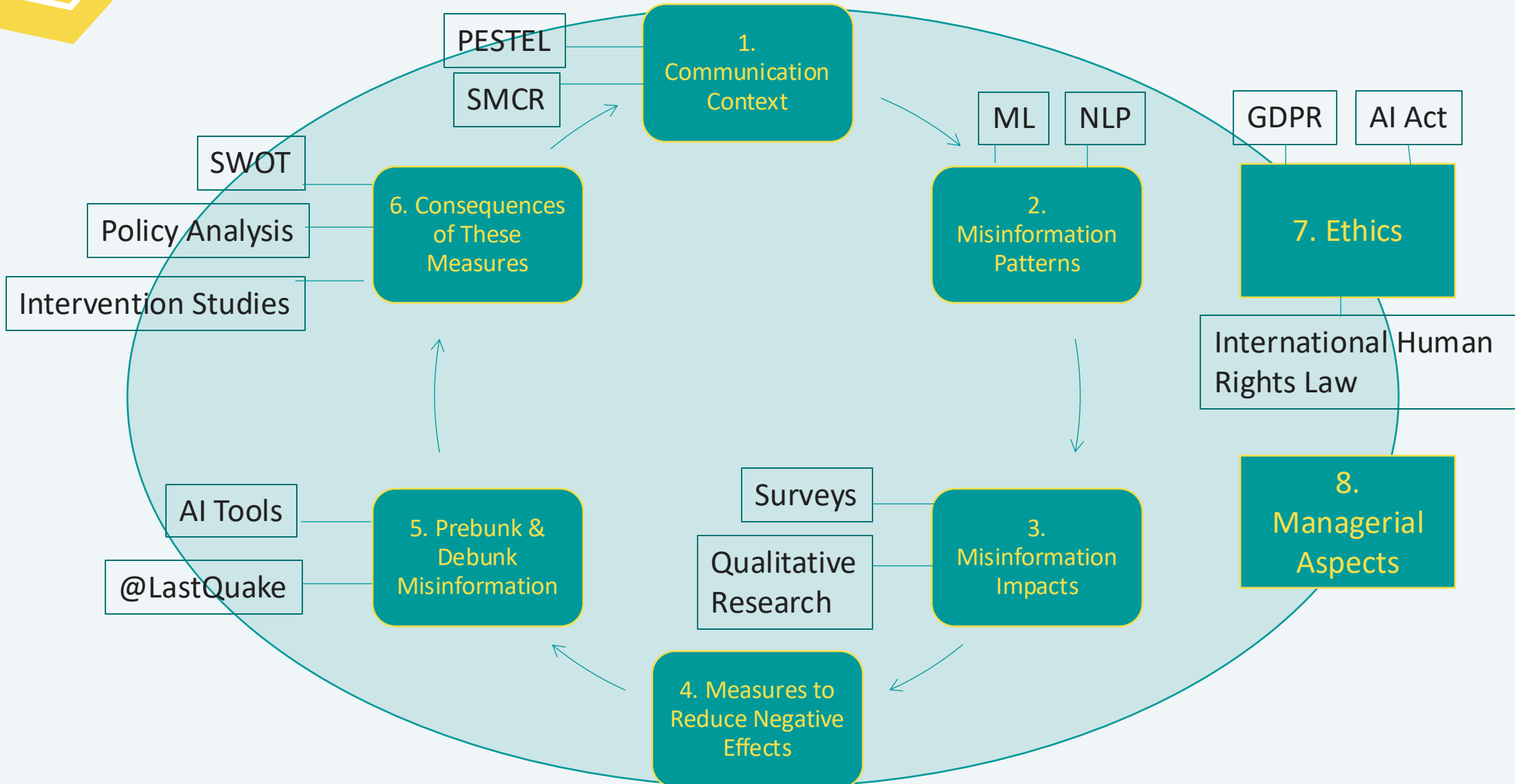


2 CONTENT & OBJECTIVES



- ❖ This **toolbox** aims to provide a **methodological framework for addressing various forms of misinformation relevant to disaster risk reduction on social media platforms.**
- ❖ It is aimed to help **institutions, policy makers, decision makers, practitioners and scientists.**
- ❖ The methodology comprises **eight steps**, addressing communication patterns, influence of social media information on risk perceptions, ethical challenges, stakeholder preferences for misinformation-fighting tools.
- ❖ These tools and steps are not necessarily intended to be implemented exhaustively. They should be **chosen and adapted according to the specific context** of implementation.

EIGHT STEPS TO TACKLE MISINFORMATION

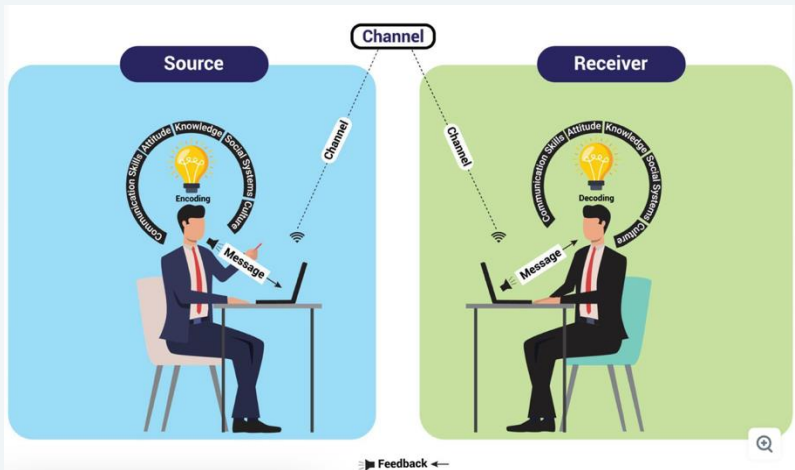


❖ **PESTEL analysis** to identify the political, economic, social, technological, environmental, and legal factors.

❖ **Berlo's communication model** outlines the key steps of communication: **Source, Message, Channel, Receiver, Effect, & Feedback.**

P	E	S	T	E	L
<ul style="list-style-type: none"> Government policy Political stability Corruption Foreign trade policy Labour law 	<ul style="list-style-type: none"> Economic growth Inflation rates Disposable income Unemployment rates 	<ul style="list-style-type: none"> Population growth rate Age distribution Safety emphasis Health consciousness Lifestyle attitudes Cultural barriers 	<ul style="list-style-type: none"> Technology incentives Level of innovation Automation R&D activity Technological change Technological awareness 	<ul style="list-style-type: none"> Weather Climate Environmental policies Climate change Pressures from NGOs 	<ul style="list-style-type: none"> Discrimination laws Employment laws Consumer protection laws Copyright and patent laws Health and safety laws

(Dubetcky, 2024)

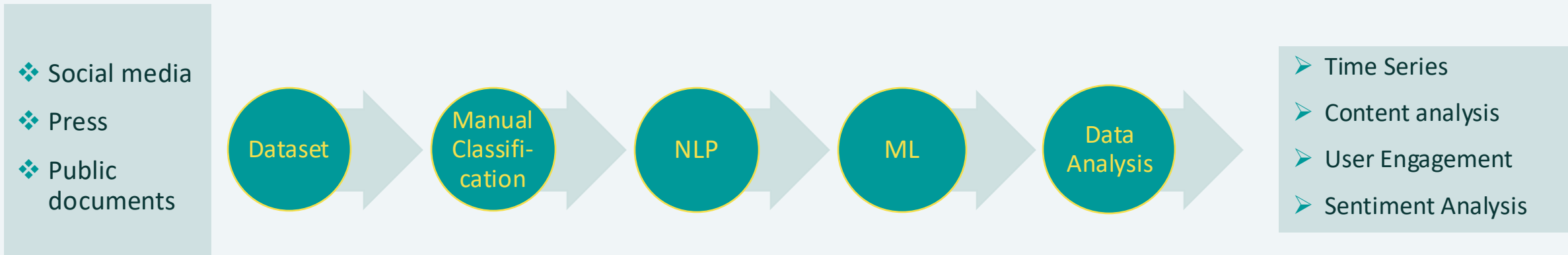




STEP 2: Identify Current Misinformation Patterns



❖ A plethora of qualitative and quantitative approaches exist. Our focus lies on recent techniques for analyzing misinformation on social media, supported by automated tools, with the aim of facilitating rapid responses.



❖ Erokhin et al., 2022 ; Elroy and Yosipof, 2022; Elroy and Yosipof, 2023; Dallo et al., 2023; Vicari et al., 2024; Dallo et al., 2023.



STEP 3: Assess Misinformation Impact on Risk Perceptions and Management



- ❖ Develop a **conceptual and theoretical framework** (e.g. Theory of Planned Behaviour) to understand how misinformation impacts risk perception.

Model	Independent Variables	Dependent Variables
Theory of Planned Behavior (TPB) (Pundir et al., 2021)	<ul style="list-style-type: none"> ○ Awareness of fake news ○ Attitudes toward news verification ○ Perceived behavioral control ○ Subjective norms ○ Fear of missing out ○ Sadism 	Social media users' intentions to verify news before sharing it
Risk Information Seeking and Processing (RISP) (Griffin et al., 2004)	<ul style="list-style-type: none"> ○ Attributes of risks ○ Individual characteristics 	'Information sufficiency'
Heuristic framework by Hansson et al. (2020)	<ul style="list-style-type: none"> ○ Communication-related factors ○ Individual, social-structural, and situational vulnerabilities 	Individuals' abilities to prepare for and respond to disasters

- ❖ **Surveys and Questionnaires** to gather data on individuals' exposure to misinformation and their perceived risks.
- ❖ **Experiments** to observe changes in risk perception after exposure to misinformation.
- ❖ **Social Media Analysis** to analyze misinformation spread and engagement on platforms.
- ❖ **Interviews** to understand individuals' perceptions and decision-making processes.
- ❖ **Focus Groups** to gather diverse insights on misinformation impacts and let people build on each other ideas.
- ❖ **Thematic Analysis** to identify recurring themes and insights regarding how misinformation influences risk perceptions.



STEP 4: Implement Measures to Mitigate Negative Effects (Dallo et al., 2023)



❖ Case studies :

- ❖ Link between 5G & Covid
- ❖ Earthquake predictions
- ❖ Volcano Vesuvius
- ❖ Monkeypox
- ❖ Manchester Arena attack
- ❖ Earthquakes & cross-cultural comparison



❖ Recommendations :

- ❖ Source Identification and Trust Building
- ❖ Message Tone and Content
- ❖ Channel Management and Cross-Verification
- ❖ Tailored Strategies for Different Receiver Groups
- ❖ Consider Emotional States in Emergency Situations
- ❖ Anticipate and Address Potential Effects
- ❖ Established Network for Feedback and Response
- ❖ Adapt to Contemporary Information Systems and Technologies



STEP 5: Prebunk and Debunk Misinformation



- ❖ In recent years, advancements in AI have provided sophisticated and relatively fast means to fight against misinformation (Vicari and Komendatova, 2023; Komendatova et al., 2021)
- ❖ There is a need to :
 - ❖ Balance between **algorithmic solutions** and **user autonomy**,
 - ❖ Leverage **Covid-19 pandemic-related research** to develop and improve tools for other risks.
- ❖ Given the complexity of the problem and the issue of trust, it is crucial to understand the **preferences of various stakeholder groups** regarding **AI tools** for combating misinformation (Komendatova et al., 2023; Erokhin and Komendatova, 2023; Komendatova et al., 2021).
- ❖ Using innovative tools such as the **EMSC-developed @LastQuake** Twitter bot, prebunking aims to proactively counter the spread of misinformation by quickly disseminating accurate information following seismic events (Bossu et al., 2023; Fallou et al. 2024).



STEP 6: Evaluate the Effectiveness of Measures

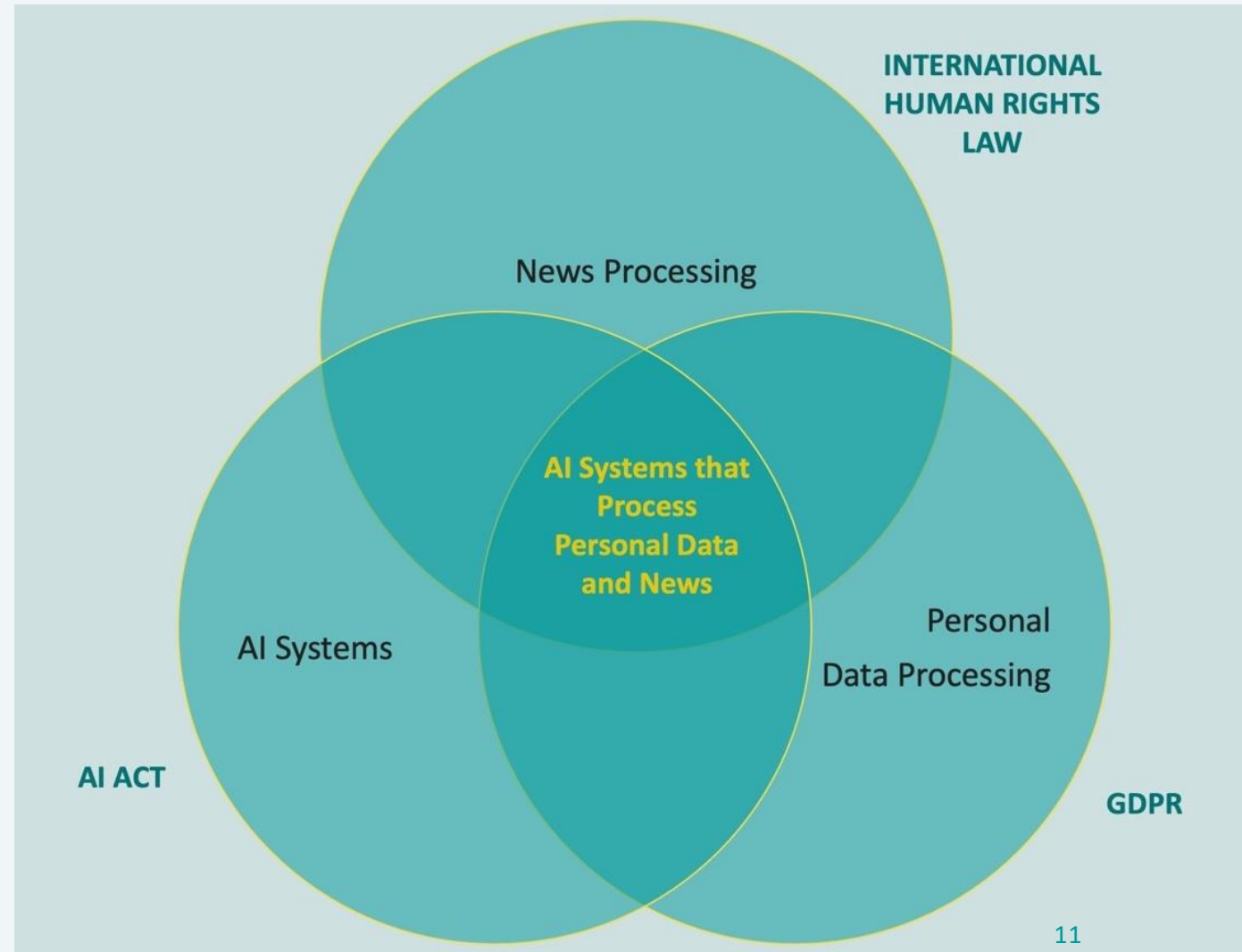


- ❖ SWOT analysis: Strength, Weaknesses, Opportunities, Threats.
- ❖ Policy Analysis: Review existing risk management policies to evaluate their effectiveness in the presence of misinformation.
- ❖ Evaluate how the implementation of measures affects the dynamics of information dissemination and reception during disasters.
 - ❖ Intervention Studies: Test the effectiveness of various interventions (e.g., fact-checking, public information campaigns) in mitigating the impact of misinformation.
- ❖ Monitor changes in communication patterns and adapt strategies accordingly.

Strengths	Weaknesses
<ul style="list-style-type: none">○ What do we do well?○ What do our target say we do well?○ What is our unique offer proposition?○ Do we have strong brand awareness/customer loyalty?○ Supplier, distributor, influencer relationships?○ What skills do we have that our others don't?	<ul style="list-style-type: none">○ Where can we improve?○ What do our targets frequently complain about?○ Which objections are hardest to overcome?○ Do we have any limitations in delivering?○ Are our resources and equipment outdated or limited?○ Are we suffering from skills, or training deficiencies?
Opportunities	Threats
<ul style="list-style-type: none">○ Is there an untapped pain point?○ Are there potential new sources of support?○ Are social or political trends that could benefit us?○ Are any technologies that could benefit us?	<ul style="list-style-type: none">○ Social or political trends that could work against us?○ Any new technology that could work against us?

(Harrison, 2024)

- ❖ Strategies to monitor and combat misinformation should align with International **Human Rights Law** (UN, 2023), the **General Data Protection Regulation** (GDPR) (European Parliament and European Council, 2024) and the **AI Act** (European Commission, 2024)
- ❖ The increasing focus on **ethics in NLP**, particularly in areas like **healthcare**, underscores the need for actionable frameworks that can also be applied to managing disaster-related misinformation (Vicari and Komendatova, 2024).
- ❖ **Implementing ethical standards** and **designing ethical tools** will be crucial in navigating these ethical challenges through actionable research.



Know Your Audience

- Preparedness Level
- Information Needs
- Risk Culture
- Expectations
- Beliefs
- Interest
- Science Literacy
- Risk Perception
- Psychology of Risk

Refine Messaging

- Acknowledge Limitations & Uncertainties
- Be Inclusive
- Provide Examples
- Avoid Information Overload
- Use visuals
- Be empathetic

Establish a Relationship

- Be Clear and Consistent
- Manage Expectations
- Collaborate with All Actors
- Build Trust
- Be Transparent

(Dallo et al., 2023)



References



- ❖ D. Erokhin, A. Yosipof, and N. Komendantova, “COVID-19 Conspiracy Theories Discussion on Twitter,” *Soc Media Soc*, vol. 8, no. 4, p. 205630512211260, Oct. 2022, doi: 10.1177/20563051221126051.
- ❖ O. Elroy and A. Yosipof, “Analysis of COVID-19 5G Conspiracy Theory Tweets Using SentenceBERT Embedding,” 2022, pp. 186–196. doi: 10.1007/978-3-031-15931-2_16.
- ❖ O. Elroy, D. Erokhin, N. Komendantova, and A. Yosipof, “Mining the Discussion of Monkeypox Misinformation on Twitter Using RoBERTa,” 2023, pp. 429–438. doi: 10.1007/978-3-031-34111-3_36.
- ❖ O. Elroy and A. Yosipof, “Semi-Supervised Learning Classifier for Misinformation Related to Earthquakes Prediction on Social Media,” 2023, pp. 256–267. doi: 10.1007/978-3-031-44207-0_22.
- ❖ I. Dallo, O. Elroy, L. Fallou, N. Komendantova, and A. Yosipof, “Dynamics and characteristics of misinformation related to earthquake predictions on Twitter,” *Sci Rep*, vol. 13, no. 1, p. 13391, Aug. 2023, doi: 10.1038/s41598-023-40399-9.
- ❖ R. Vicari, O. Elroy, N. Komendantova, and A. Yosipof, “Persistence of misinformation and hate speech over the years: The Manchester Arena bombing,” *International Journal of Disaster Risk Reduction*, vol. 110, p. 104635, Aug. 2024, doi: 10.1016/j.ijdr.2024.104635
- ❖ I. Dallo et al., “Impact of misinformation on social media on risk perception in a multi-risk environment ,” Aug. 2023. Accessed: Sep. 01, 2024. [Online]. Available: <https://pure.iiasa.ac.at/id/eprint/19767/>
- ❖ R. Vicari and N. Komendantova, “Systematic meta-analysis of research on AI tools to deal with misinformation on social media during natural and anthropogenic hazards and disasters,” *Humanit Soc Sci Commun*, vol. 10, no. 1, p. 332, Jun. 2023, doi: 10.1057/s41599-023-01838-0.
- ❖ N. Komendantova et al., “A value-driven approach to addressing misinformation in social media,” *Humanit Soc Sci Commun*, vol. 8, no. 1, p. 33, Jan. 2021, doi: 10.1057/s41599-020-00702-9.
- ❖ R. Bossu, M. Corradini, J.-M. Cheny, and L. Fallou, “A social bot in support of crisis communication: 10-years of @LastQuake experience on Twitter,” *Front Commun (Lausanne)*, vol. 8, Mar. 2023, doi: 10.3389/fcomm.2023.992654.
- ❖ L. Fallou, R. Bossu, and J.-M. Cheny, “Prebunking earthquake predictions on social media,” *Front Commun (Lausanne)*, vol. 9, Jun. 2024, doi: 10.3389/fcomm.2024.1391480.
- ❖ N. Komendantova, D. Erokhin, and T. Albano, “Misinformation and Its Impact on Contested Policy Issues: The Example of Migration Discourses,” *Societies*, vol. 13, no. 7, p. 168, Jul. 2023, doi: 10.3390/soc13070168.
- ❖ D. Erokhin and N. Komendantova, “The role of bots in spreading conspiracies: Case study of discourse about earthquakes on Twitter,” *International Journal of Disaster Risk Reduction*, vol. 92, p. 103740, Jun. 2023, doi: 10.1016/j.ijdr.2023.103740.
- ❖ N. Komendantova, L. Ekenberg, W. Amann, M. Danielson, and V. Koulolias, “Chapter 10 The Adequacy of Artificial Intelligence Tools to Combat Misinformation,” 2021, pp. 172–198. doi: 10.1007/978-3-030-70370-7_10.
- ❖ EMSC, “EMSC X Channel,” X. Accessed: Sep. 05, 2024. [Online]. Available: <https://x.com/lastquake?lang=en>
- ❖ R. Vicari and N. Komendantova, “Ethics for Natural Language Processing: A Systematic Meta-Analysis (in preparation),” 2024.
- ❖ I. Dallo, M. Corradini, F. Laure, and M. Michèle, “How to fight misinformation about earthquakes? A Communication Guide,” Zurich, Feb. 2022. Accessed: Sep. 01, 2024. [Online]. Available: <https://www.research-collection.ethz.ch/handle/20.500.11850/530319>



QUESTIONS ?



ROSA VICARI

vicari@iiasa.ac.at

NADEJDA KOMENDATOVA

komendan@iiasa.ac.at