



Responsible AI communication: How do European officials share their concerns?

Rusu G Anca, Merad Myriam

► To cite this version:

Rusu G Anca, Merad Myriam. Responsible AI communication: How do European officials share their concerns?. Congrès Lambda Mu 23 “ Innovations et maîtrise des risques pour un avenir durable ” - 23e Congrès de Maîtrise des Risques et de Sécurité de Fonctionnement, Institut pour la Maîtrise des Risques, Oct 2022, Saclay, France. hal-03968180

HAL Id: hal-03968180

<https://hal.science/hal-03968180>

Submitted on 1 Feb 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Responsible AI communication: How do European officials share their concerns?

RUSU G. Anca

Dauphine University

Pl. du Maréchal de Lattre de Tassigny, 75016 Paris, France
anca-georgiana.rusu@dauphine.eu

MERAD Myriam

Dauphine University

Pl. du Maréchal de Lattre de Tassigny, 75016 Paris, France
myriam.merad@lamsade.dauphine.fr

Résumé — Cet article examine la manière dont les représentants officiels de la Commission Européenne (EC) et des Etats membres expriment leurs préoccupations concernant l'utilisation de l'intelligence artificielle (IA) à travers la perception de ce que nous avons appelée: le concept de "communication responsable". L'hypothèse est que l'expression des inquiétudes est considéré comme représentative d'un comportement responsable en termes de communication par rapport à une technologie émergent. Afin d'analyser leur attitude envers la responsabilité en termes de communication concernant l'utilisation de l'IA d'un point de vue gouvernemental, cette recherche a analysé les discours exprimés dans les documents officiels représentant les stratégies d'IA au niveau européen.

L'analyse nous a permis d'identifier trois types de modèles de communication utilisés dans le discours écrits et publié par des organismes gouvernementaux européens comme représentants de la stratégie européenne par rapport à l'usage de l'IA. Les modèles identifiés sont: la communication pessimiste de l'IA, la communication optimiste de l'IA et la communication équilibrée de l'IA.

Mots-clefs — *risque, communication, responsabilité, technologies émergentes, intelligence artificielle (IA).*

Abstract— This paper looks into how European Commission (EC) and EU Member States officials express their concern regarding the use of artificial intelligence (AI) through the lenses of what we have called to be based on a “responsible communication” concept. The assumption is that the expression of concerns is considered representative of a responsible behavior in terms of communication regarding an emerging technology. To analyze their attitude to responsibility in terms of communication regarding the use of AI from a governmental perspective, this research investigated the discourses expressed in official documents representing the AI strategies at the European level.

The analysis allowed us to identify three types of communication models used in the written and published discourse of European governmental bodies as representatives of the European strategy with regard to the use of AI. The identified communication models are pessimistic AI communication, optimistic AI communication and equilibrated AI communication.

Keywords — *risk, communication, responsibility, emerging technologies, artificial intelligence (AI).*

I. INTRODUCTION

Artificial intelligence (AI) is expected to improve and enhance various sectors, with implications and consequences on individuals, organizations and society at large. Due to the

present knowledge gaps and early adoption of the technology, emerging risks can arise. When dealing with emerging risks, communicating is seen as a condition for success [1]. It informs the stakeholder's goals and interests that make sense for its context, role, or mission [2]. Following the growing interest in AI, more governments are creating strategies to uptake the benefits of AI technology. Despite these pioneering documents' vague and uncertain nature, they provide insight into how AI will affect people's lives, safety, and well-being if this technology is fully adopted [3]. By analyzing how the communication regarding the use of AI in various sectors is framed in these documents, we can see how the narrative at the EU level is constructed, acknowledging that by its performative function, the discourse helps create and close possible scenarios in the future.

In the case of emerging technologies, if communication is analyzed early in the development stage, it can provide information on potential negative aspects of the technology. For example, extreme fear narratives can have positive outcomes, helping researchers, regulators, and technology implementers take safety concerns into account at an early stage [4]. Therefore, because of the performativity of the discourse, how positives, negatives and concerns are communicated regarding a specific technology can affect how various stakeholders perceive it and how it will be developed and deployed.

Analyzing how the narrative has been constructed around the use of AI at the European level is seen as an example to be analyzed from the point of view of a responsible behavior. The analysis of the communication is on the narrative text, which can be defined as “a text in which an agent relates (‘tells’) a story in a particular medium, such as language, imagery, sound, buildings, or a combination thereof” [5]. They can differ, even if they tell the same story. Whether they are functional or not, the narratives have a performative effect on the real world. “They affect individuals and collectives, and influence human action, thought and social outcomes”, and “they have the power to either enhance or limit the potential for human flourishing”, the potential of AI more exactly [4].

AI is considered an emerging technology as it is a new technology in terms of data processing capabilities and therefore automated processing of data and the new uses resulting from this. Like any other emerging technology, it has the following attributes: (i) radical novelty, (ii) relatively fast growth, (iii) coherence, (iv) prominent impact, and (v)

uncertainty and ambiguity [6]. When an emerging technology is developed, attention is given to the potential adverse effects, but not all can be a priori determined. Therefore, responsible communication must be implemented as action to govern potential future risks. Assessors and managers of emerging technologies face particular challenges because they lack the procedures and tools necessary to assess the potential impact of the technology, lack adequate data upon which to build evidence, and are uncertain about how the technology will evolve [7].

Given the uncertainties around the use of AI, looking into how the EU communicated around it constructs an imaginary window toward an imagined future. Because the narrative constructed by the EU influences human action in and for the future, the question of responsibility needs to be addressed. Therefore, this article focuses on the EU's responsibility regarding what is communicated on its behalf.

II. RESPONSIBLE COMMUNICATION ON EMERGING TECHNOLOGIES

The expression of concerns is impossible to separate from the expression of the perceived positives and negatives concerning technology, as they are not all the time explicitly communicated. However, the concerns, which can be expressed as known negative impacts or uncertainty about possible unknown negatives, are at the basis of one's actions and have a central role in constructing the narratives, even if they are expressed.

The European Union (seen as the institutions and governments that constitute it) is responsible for what is communicated to the European and non-European members (individuals and groups) under the government's authority. So, looking into how the European Commission and the Member States communicated about AI is considered a lens toward the imagined future because of the strong performativity of its discourse.

Multiple sources can be chosen to analyse the European discourse around the use of AI. Still, this research investigates the written reports, as they represent official governmental strategies regarding the use of AI at the European level.

Acknowledging the interdisciplinarity of the responsible communication of emerging technologies (RCET) concept and the lack of a definition, the authors of this paper propose a working definition framed by putting together the replies to these four questions: What is it? How should it be done? What is it for? Why do it?

Responsible communication of an emerging technology (RCET) is a practice by which humans reproduce and produce social relations, and co-create meaning and sense of the world by responsibly transmitting information, ideas and opinions about it with and within social groups in a defined context. In order to position and implement a specific technology for a better future, the process should be driven by the principles of good governance and has to promote an ethical behavior.

- *Ethical behavior regarding RCET refers to a two-way communication process that focuses on social responsibility, and contributes to social development and the greater/social good. Following this practice, the communicator shares information regarding positive and negative impacts, as well as concerns regarding*

an emerging technology, considering the impacts of their communication process and content on the stakeholders.

When used, RCET contributes to the communication of motivations to build a given technology, the risks and opportunities regarding the technology, and the creation of a narrative around it.

To analyze EC and the Member States's attitude to responsibility in terms of communication regarding the use of AI from a governmental perspective, this research analyzed is looking into what has been communicated in various written AI strategies (European and national) by considering the following directions that are underlined by definition as being part of a responsible behavior:

- Framing - what is the AI context at the European level, the corpus, who are the influential stakeholders, and how is the better future described.
- Responsibility regarding the effects of the technology on the future - Communication of positive and negative impacts and concerns (not individual or organizational gain only), explain and argue decisions that have a negative impact, and the narrative that has been communicated.
- Responsibility regarding the effects of the use of the technology, focusing on sources of risk - Communicate about the positives and negatives of the technology, applied to specific use and not in general; focus on the possible sources of risk.
- Responsibilities associated with the communication process - The communicator (as an individual or representative of an organization) should not intentionally hide information, not be catastrophic, and needs to give solutions to identified problems. When applied to an individual, they must pursue the deontological rules like being honest, truthful, and sincere when communicating.

III. METHODOLOGY

To analyze the communication of the European governmental stakeholder, we have chosen to examine standalone organizational reports, part of the defined AI strategies at the EU level. Reports and publications disseminated by various organizations can play an active and significant role in deciding what is relevant to each defined group of actors. They represent a form of discourse, an actor's position, and have a performative role in society. However, we acknowledge the limitations of using only one form of media per stakeholder group. Still, the present study focuses on what is communicated regarding AI rather than how much is said.

A. Focus of investigation

While other methods could have been used (i.e., content analysis), this paper has used discourse analysis because it considers participants' "real" thoughts and feelings as represented in the official discourses. The construction of discourse in an organizational perspective entails the defined role in the society of a specific institution. In this case, what the official documents declare informs on a role and position in a particular matter. This method is also considered due to its links with power, politics, hegemony and ideology.

Moreover, discourse theory and method allow the micro and macro contexts to be linked and enable the investigation of the production of truths and legitimacy through its philosophical foundations regarding the nature of reality and knowledge.

B. Corpus

In this paper, the governmental stakeholder is represented by the European Union as an overreaching reglementary body at the European level and the national governments of the countries part of the EU. Therefore, to go in-depth into the discourse of the government stakeholder, it was decided to look into two types of documents, geographically located only at the European level. Therefore, documents published by the European Commission representing the AI strategy at the European level were analyzed first, followed by the published national AI strategies.

The AI strategy started to be well defined at the European level in spring 2018 when the Commission published a communication on AI. At the end of the same year, an AI plan for the Member States was implemented, and the states signed a collaboration [8].

This analysis investigates the AI European policy and the documents at the basis of its strategy on artificial intelligence, which started in 2018 when the Commission (EC) and the Member States agreed to join forces on AI policy. All the documents published by the EC were reviewed, and eight publications were selected to be further analyzed for their relevance to the EU AI strategy. We analyzed three documents referring to the vision of the EU AI strategy (Communication Artificial Intelligence for Europe, published in April 2018, White Paper on AI: a European approach to excellence and trust, main document, published in February 2020 [9], White Paper on Artificial Intelligence: Public consultation towards a European approach for excellence and trust, published in July 2020 [10]) and five documents referring to rules and actions (Coordinated Plan on Artificial Intelligence, main document, published in April 2018 [8], Coordinated Plan on Artificial Intelligence, published in April 2018 [11], Fostering a European approach to Artificial Intelligence, main document, April 2018 [12], Fostering a European approach to Artificial Intelligence, annex, April 2021 [13], Artificial Intelligence Act, main document, published in April 2021 [14]).

Regarding the National EU strategies, we consider the countries listed in the AI Watch 2021 National strategies on Artificial intelligence [15]. Unfortunately, some of the listed countries were not published in English, or we could not find the document; therefore, we analyzed 16 initiatives in total: Bulgaria, 2020; Czech Republic, 2019; Denmark, 2019; Estonia, 2019; Finland, 2017; France, 2018; Germany, 2018; Hungary, 2020; Lithuania, 2019; Luxembourg, 2019; Malta, 2019; Netherlands, 2019; Norway, 2020; Portugal, 2019; and Slovakia, 2019.

C. Focus on sectors of applicability

To better look into the responsibility regarding the effects of the technology on the future, we analyzed how the use of the AI per each of the nine selected sectors of applicability was communicated. The sectors were determined by considering the ones identified in the empirical reading of the reports: Agriculture, Energy, Environment, Health, Justice, Media, Public sector, Transportation and Space. The paragraphs where the use of AI was mentioned for each of the selected sectors of applicability were included.

D. Focus on possible sources of risk

As part of the methodology to be applied, a list of possible sources of risks has been created.

TABLE 1: SUMMARY OF RISK CATEGORIES AND POTENTIAL RISK SOURCES

Categories of risk	Potential risk sources
Technical design attributes	Accuracy
	Data availability
	Data quality
	Reliability
	Resilience
	Robustness
	Security
	Standardization
Socio-technical attributes	Resource and energy efficiency
	Accountability
	Bias
	Discrimination
	Explainability
	Interpretability
	Liability
	Labor market
	Privacy
	Safety
Guiding policies and principles	Ethical AI
	Fairness
	Human agency and oversight
	Human rights and values
	Responsible use
	Sustainability
	Transparency
	Trust
	Misuse

The list combines the risks identified by the empirical reading of the documents with a classification proposed by the National Institute of Standards and Technology (NIST) (which considers the OECD principles (OECD, n.d.), the European Union Digital Strategy's Ethics Guidelines for Trustworthy AI [16] and the US Executive Order 13960, Promoting the Use of Trustworthy Artificial Intelligence in the Federal Government [17]) [18]) (see Table 1). It is acknowledged that sometimes one category could imply another (e.g., ethical AI that includes transparency). Still, the separation was kept to be able to design the communicated narrative. The table below summarises the categorization of the potential risk sources used in the analysis of the documents.

IV. FINDINGS

A. Framing

Communication at the European governmental level is strongly made internally (Member states) and externally (non-Member States). The EC's website is a reliable source of information about regulations in place, regulations of involved stakeholders and projects, blogs and news, statistical data used in various reports, best practices, libraries, and dedicated reports on trends and evolutions. This strong communication about the AI strategy in the spring of 2018 and multiple

actions and documents that support the strategy have been published since then. These are also part of the communication strategy, and we remember the financial contributions of EC as by 2025, AI investments will reach € 22.4 billion and surpass the € 20 billion target by over 10% [19]. Our analyses show that the EC AI strategy strongly appears to be communicated in the EU national strategies, showing a strong communication with the Member states.

At the European level, the question of stakeholders is complex. It needs to be named for its implication in the communication process and its influence on the construction of the message. Therefore, considering the multitude of stakeholders, their mapping is essential in shaping the context where and to whom the communication strategies are addressed. To shortly map the stakeholders of the EU, we acknowledge that democratic governance is both proactive and reactive (creating interest and power and responding to them) [20], and that the non-EU actors have a role in the healthy functioning of the EU's governance. We also support the multi-actor dimension of the EU, which reflects the involvement of non-state actors and the acknowledgment that that state is not a unitary actor [21]. Therefore, the following stakeholder groups have been identified at the EU level: European Union's institutions (the Treaty of Lisbon lists the Union's institutions which shall be: the European Parliament, the European Council, the Council, the European Commission, the Court of Justice of the European Union, the European Central Bank, the Court of Auditors [22]), Consultative committees (European Economic and Social Committee, European Committee of the Regions), Agencies and independent authorities (Fundamental Rights Agency, European Data Protection Supervisor (EDPS), European Institute of Innovation and Technology (EIT), Europol), Expert groups (High-Level Expert Group on Artificial Intelligence, Expert Group on Liability and New Technologies), National Governments (For each of the member states), Non-institutional actors (Business, Trade Unions, Intergovernmental, Technical community, Academia, Civil society¹). This listing gives an image of the actors that can and could influence the AI strategy at the European level. These actors also affect the national strategies, where the stakeholders can be mapped on three influential levels: national, European and international. Some countries created new organs, while others assigned new responsibilities to existing ones.

Regarding academic and industrial stakeholders, new research directions and businesses have been defined and determined by existing and new challenges at the three levels. In terms of framing, it is to be considered that the stakeholder structure and level of influence have consequences on the communication made by the European governments. Therefore, it is formulated to address all the stakeholders.

B. Responsibility regarding the effects of the technology on the future

By identifying the main themes that have been communicated in each of the analyzed documents, we have been able to see if the discourse about AI was referring mainly to a positive or a negative impact in the present or the future, which informed us on the communicated narrative regarding the use of AI.

1) The EC strategy

To summarise the findings, the negative impacts of using AI have been categorized into six categories, and the same number of mitigation measures have been found (see Table 2). The most referenced (12 out of 26) negative impacts are the rights and values, the design of AI systems, the risks engaged by the AI application, and the product liability issue. The EU uses rules and regulations as a mitigation measure for most of the expressed negative impacts of AI. Stakeholder engagement and promoting ethical and trustworthy AI are the most used measures. To address the problems regarding fundamental rights, security, and respect for values and privacy, the EC strategies propose creating and updating the rules and regulations, promoting ethical and trustworthy AI, and conducting impact assessment and management. The concerns coming from the technologies' design are also mitigated by rules and regulations and ethical and trustworthy AI promotion. To address the chances engaged by the AI and AI systems in general, the measures support stakeholder engagement, promote ethical and trustworthy AI, and the necessity to design requirements. At the European level, the concerns regarding the labor market transitions and skills gap are addressed by developing programs and initiatives and the engagement of stakeholders and cooperation. Finally, stakeholder engagement, the necessity of rules and regulations, and requirements are mentioned to mitigate the risks engaged by the AI application.

TABLE 2: NEGATIVE IMPACTS OF THE USE OF AI AND MITIGATION MEASURES (EC)

Negative impacts of the use of AI	Mitigation measures
Rights and values	Rules and regulations
	Impact assessment and management
	Promoting ethical and trustworthy AI
Design of the AI technology/systems	Rules and regulations
	Promoting ethical and trustworthy AI
Chances engaged by the AI and AI systems in general	Stakeholder engagement and experts
	Design requirements
	Promoting ethical and trustworthy AI
Labour market transitions and skills gap	Stakeholder engagement and experts
	Develop programs and initiatives
Risks engaged by the AI application	Stakeholder engagement and experts
	Rules and regulations
	Design requirements
Risks engaged by the AI application (in the present)	Rules and regulations

By looking into the positive impacts envisioned for the future of AI and implementation strategies, five positive categories and nine strategies to increase positive impact have been defined (see Table 3). The strategies to be implemented for the uptake of AI are the support of innovation and research, excellence, stakeholder engagement, integration of expert knowledge, coordination between various actors, and developing multiple programs and initiatives. It is acknowledged that the use of AI technology brings benefits and opportunities. The EC proposes supporting innovation and research and developing the necessary infrastructure to benefit those. Furthermore, with multiple suggested strategies is the acknowledgement of AI to improve services. In addition, the EU highlights the need for rules and regulations, the support of innovation and research, excellence, and the development of programs and initiatives. There is one reference to the positive changes that the use of AI can bring

¹ This classification was made by considering OECD's website where the Stakeholder initiatives are mapped by type of organisations (OECD, n.d.)

to human lives, and for doing so, rules and regulations need to support the use of the AI technology, as well as the influence of AI on the European leadership, which can be done by promoting ethical and trustworthy AI. Referring to the positive impacts, it was observed that some of the mitigation measures are used as strategies for implementation (e.g., rules and regulations, stakeholder engagement, and support of initiatives). The most used strategies are the supported innovation and, research, excellence, followed by stakeholder engagement, experts and collaboration.

TABLE 3: POSITIVE IMPACTS OF THE USE OF AI AND STRATEGIES (EC)

Positive impacts of the use of AI	Strategies
AI uptake	Support the innovation and research, excellence
	Stakeholder engagement, experts and coordination
	Develop programs and initiatives
Seize opportunities	Support the innovation and research, excellence
	Develop infrastructure
Improve services	Rules and regulations
	Support the innovation and research, excellence
	Develop programs and initiatives
	Stakeholder engagement, experts and coordination
Positive change of our life	Rules and regulations
Strengthen leadership	Promoting ethical and trustworthy AI

2) The Member States

As in the EC discourse, the Member States' narrative also refers mainly to the future (only one reference to the present was found). It has been observed that the positive and negative impacts are less diverse than the ones found for the EC. Only four of the six negative impacts and five of the mitigation measures defined for the EC strategy have been found (see Table 4).

The labor market transition and skills gaps are the most mentioned category, followed by the general concerns regarding the changes engaged by the AI and AI systems. To address the concerns regarding the changes engaged by the AI and AI systems in general, what is needed as mitigation measures includes rules and regulation, promoting ethical and trustworthy AI and stakeholder engagement. There is only one measure for the concerns regarding the changes in the labor market and skills gap: the development of programs and initiatives (and implicitly supporting education). And to engage the risks associated with AI applications, a measure to be implemented is assessing and managing the impact. The development of programs and initiatives is the most used strategy, followed by updating or creating rules and regulations.

TABLE 4: NEGATIVE IMPACTS OF THE USE OF AI AND MITIGATION MEASURES (MEMBER STATES)

Negative impacts of the use of AI	Mitigation measures
Changes engaged by the AI and AI systems in general	Rules and regulations
	Stakeholder engagement, experts and collaboration
	Promoting ethical and trustworthy AI
Labour market transitions and skills gap	Develop programs and initiatives
Rights and values	Rules and regulations
Risks engaged by the AI application	Impact assessment and management

The references to positive impacts are more diverse than the negative. Compared to the EC discourse, three more strategies have been defined: define priority sectors, responsible development and sustainability, and promote the use of AI to promote its benefits. Unlike the EU strategy, the national AI documents propose more strategies that benefit

from the positive impacts of the use of AI (see Table 4). The national strategies present the engagement of the stakeholders and support collaboration and the support of innovation and research as strategies to be implemented to uptake the AI use. In addition, to these main strategies, the development of various programs and initiatives, the definition of priority sectors, the development of the necessary infrastructure and responsible development and sustainability are mentioned. The responsible development appears only once as a proposed strategy in the analysed documents but deserves further analysis.

TABLE 5: POSITIVE IMPACTS OF THE USE OF AI AND STRATEGIES (MEMBER STATES)

Positive impacts of the use of AI	Strategies
AI uptake	Develop programs and initiatives
	Support the innovation and research, excellence
	Define priority sectors
	Stakeholder engagement, experts and collaboration
	Responsible development and sustainability
	Develop infrastructure
Seize opportunities	Support the innovation and research, excellence
	Define priority sectors
	Promote the use of AI to promote its benefits
	Develop infrastructure
	Rules and regulations
	Promoting ethical and trustworthy AI
	Stakeholder engagement, experts and collaboration
Strengthen leadership	Promote the use of AI to promote its benefits
	Rules and regulations
	Support the innovation and research, excellence
	Define priority sectors
	Promoting ethical and trustworthy AI
	Develop programs and initiatives
Improvement of services	Promote the use of AI to promote its benefits
	Develop infrastructure
	Rules and regulations
	Support the innovation and research, excellence
Positive change in our life	Define priority sectors
	Promoting ethical and trustworthy AI
	Promote the use of AI to promote its benefits
	Support the innovation and research, excellence
	Stakeholder engagement, experts and collaboration

All of the documents mentioned the importance of seizing the opportunities and benefits offered by the use of AI, and this can be made using various strategies: support the innovation and research, the goal of enabling its excellence, the development of the infrastructure, the definition of priority sectors where benefits can accentuate the promotion of the use of AI and its benefits, the necessity of rules and regulations (but not as a primary strategy), the importance of promoting ethical and trustworthy AI and the engagement of stakeholders and encourage collaboration.

One positive impact that appears multiple times in the national AI strategies but only once in the EU strategies is the impact of AI on the national and EU leadership. Among the strategies, the following have been named: the promotion of the use of AI to promote its benefits, the creation and update of rights and values, the support of innovation and research, very relevant the definition of priority sectors and the promotion of ethical and trustworthy AI, as well the development of various programs and initiatives. Many of the references are made to the use of AI in multiple sectors or applications and which is translated as a positive impact of the use of AI, where the development of infrastructure is essential, along with the support of the innovation and research,

excellence and the definition of rules and regulations. The main concern of national AI strategies is the shifting and changes in the labor market, which are treated with one measure: education and upskills. Regarding future scenarios, the Member States communicate more positives than the EC (see **Erreur ! Source du renvoi introuvable.**).

Putting all our findings together, we have been able to shape the narrative that both EC and national EU governments have communicated. The vision of the long-term future is very positive, being the result of applying mitigation measures in the present and the short-term future. Referring to the general governmental narrative, the analyzed actors frame the use of AI as conducting to a positive future if mitigation measures are considered for the identified risks and concerns.

Concerns regarding the use of AI are communicated at the European level in the present and short-term future. However, mitigation measures have been identified and considered, and no concerns are shared for the long-term future.

C. Responsibility regarding the effects of the use of the technology, with a focus on sources of risk

How the European governments communicated on the specific use of the technology, focusing on selected sectors of applicability, helped us better analyze which concerns exist in the present and the short-term but not in the long-term.

Looking into the communication of risks and opportunities regarding the use of AI in the nine selected sectors of applicability, it has been observed the proposals for AI regulation of Denmark and Hungary are the documents that communicate the most about the opportunities. The conclusions regarding the proposal for regulation are to be taken with a grain of salt because the arguments are constructed differently from the other documents around a risk-based perspective, where the sectors of applicability are used as examples. Considering all the documents, it has been observed that the positive impact of AI in all the analyzed sectors is more communicated than the negative impacts. Denmark, Hungary and Bulgaria were the first three countries that communicated most about positives, and the ones communicating most about risks were Norway, Denmark and the Netherlands. At the same time, Estonia and Sweden have a relatively balanced communication. Strong communication about risks and positives associated with the use of AI shows preparedness regarding the use of technology, being the case for Denmark. The Czech Republic is the country that speaks the least about positives and not at all about risks (see **Erreur ! Source du renvoi introuvable.**). It communicates a neutral position regarding the positives and negatives, focusing on action points.

From a general perspective, more opportunities have been communicated than negatives. However, a difference between the quantity of communicated opportunities compared to the risks per document and between documents has been observed. This brings us to the observation that some documents and countries have a more optimistic or pessimistic communication regarding the use of AI.

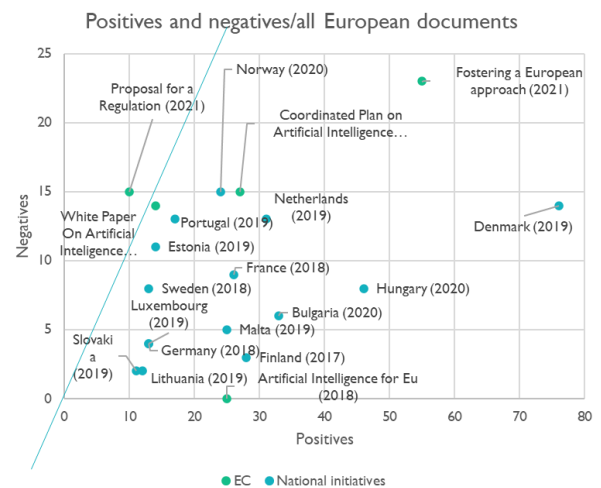


Figure 1: Governmental discourse: Risk and opportunities communicated

No correlation is observed between the communication of risks and opportunities and the year of the publication of the strategies and neither between the countries that are neighbors and the implicit or explicit expression of risks. Therefore, it can be assumed that each country used national experts to define the strategy, experts that could be/or not part of the various EC expert groups created to support the construction of the AI strategy at the European level. Consequently, an EC's overreaching impact has been observed in the texts of the national strategies that refer to their strategic positioning under the European authority.

Each document had a different communication regarding the use of AI per application sector. However, the sectors for which opportunities have been the most communicated are health and energy, health being the one for which the most risks are communicated (see Fig. 2). While EC displays many positives regarding the environment, the Member States focus on the positive use of AI in the public sector. Justice is the sector that presents the most risks for the EC and the Member States. If the risks communicated for transportation and health are more or less the same, the risks and opportunities associated with the public sector are very different. For the Member States, the majority of risks in this sector come from data availability, privacy, and security, which are the most communicated in general.

The sources of most communicated risks depend on context, sector and the country that published the document. It has been observed that the Netherlands, Norway, Sweden and Denmark communicate the most about the risks from the guiding policies and principles category. Estonia, Lithuania, Finland, Luxembourg, and Portugal are on the opposite side of this category. Part of the EC's communication regarding the rules and actions to be taken, the documents that represent the coordination with the Member states (Coordinated Plan on Artificial Intelligence and Fostering a European approach to Artificial Intelligence), are the ones that focus the most on technical design attributes. The primary source of risks communicated in these documents is similar to the communication made by the Estonian government, which is "data availability". Different from how the Member States have made the communication, it has been observed that the source of risks that the EU most communicated when AI is used in the selected sector of applicability is "unknown",

which shows that EC has a responsible behavior regarding the communication of risks, acknowledging the existence of unknowns.

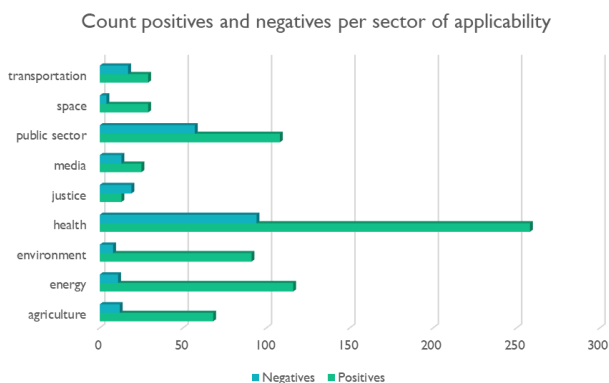


Figure 2: Risk and opportunities per sector of applicability and document

Combined with the findings from the previous chapter, concerns are expressed regarding technology in general. Still, they are more explicit regarding its use per sector of applicability, where unknown risks can appear. According to the definition presented at the beginning of the article, the expression of concern (known and unknown) is representative of responsible behavior regarding the communication concerning an emerging technology. The lack of “unknown” risks communicated by the Member States does not directly imply a communication that is not responsible because all the national strategies position their strategy under the EC’s communicated strategy.

The general perspective is that the communication made by the EC seems to be correlated to the strategy that has been communicated and that the Member States acknowledge their role under the EU umbrella by following EC’s strategy and determining at the same time their own. Following this perspective, the concerns are communicated differently according to specific defined strategies (at the European or national level) but remain in the same spectrum of concerns. The exception is the communication of unknown, which is considered normal behavior because of the EC’s construction and responsibility and the Member States in the EU.

D. Responsibilities associated with the communication process

To put into perspective how the concerns regarding the use of AI have been expressed at the European level and to consider the EU’s government’s communication behavior regarding the use of AI, the following elements have been analyzed:

- The opinion of the European people regarding the communication of opportunities and risks (including concerns) made by the EU officials regarding the use of AI.
- Concerns about the government’s ethical behavior regarding the communication of the use of AI (relative to what has been communicated and what can be

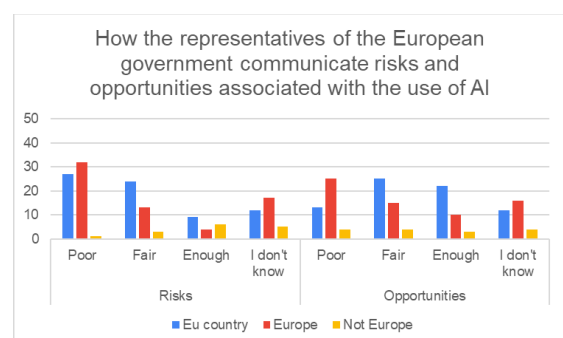
improved), both about the content (the technology) and communication process.

The replies to a survey² conducted among about 150 people in various countries and backgrounds have been analyzed for this part of the research.

Investigating the replies to a question examining how the people consider the European and national governments communicated about the opportunities and risks associated with the use of AI (see Figure 3: How the representatives of the European (top) and national governments (bottom) communicated risks and opportunities associated with the use of AI

), it has been observed that the respondents generally consider that governments have poor communication. The national governments are perceived as communicating twice as poorly as the EC regarding the risks. Regarding this element, no significant difference has been observed in the opinion of people based in Europe, whether being or not EU members. The replies from the non-European countries have not been considered significant for this statistical analysis. What could be analyzed further is the observation that the people from countries outside Europe seem to believe that the European government doesn’t speak poorly about the risks, but the national ones do. The relatively high amount of people choosing “I don’t know” when asked how the EU governmental representatives communicate risks and opportunities suggests not very effective communication.

The people based in European countries but not part of the EU consider that national governments communicated enough, a reply very different from those of the EU. It is to be noted that out of the 66 respondents constituting the “Europe” category, 59 are based in Switzerland. This shows a bias in our data and that the Swiss government is doing something different from the others and could be considered an example. This difference could be because the Swiss government communicated very strongly about the opportunities and risks associated with the use of AI and that the people feel informed, that they are using the proper channels and messages to share the information, and even that the people based in Switzerland have confidence in the Swiss government.



² The survey is available at the following link: <https://forms.gle/7uFKo8mtL9rGGBmf6>. Currently, there are 158 replies. The survey was shared by Facebook, LinkedIn and personal emails to PhD colleagues.

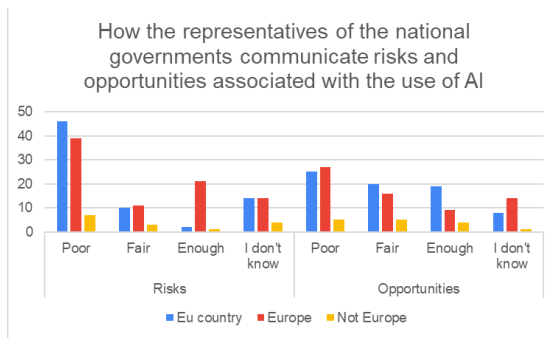


Figure 3: How the representatives of the European (top) and national governments (bottom) communicated risks and opportunities associated with the use of AI

To ask people's opinions regarding the communication done by the governments, NGOs, industry and academia about the use of AI, the term "ethical behavior" has been used. This is because the term has been considered the right concept to ask about actors' behavior regarding communication. The question "Do you believe that these actors have an ethical behavior when communicating about the use of AI? Why or why not?" helped observe people's concerns regarding what has been communicated and inform about various actors' communication behavior. Therefore, from the few replies that have been specifically referring to governments' communication about AI, most of them considered that the governments don't have an ethical behavior because of a lack of effective communication about risks and opportunities, missing information, lack of transparency, own interests, and even an emphasis on the risks to boost anti-tech credentials. There was only one reply that considered strongly that the EU has an ethical behavior when communicating about AI because of the launch of the AI act (which is considered by the same respondent as focusing too much on the risk, and that a focus on the opportunities is necessary, as well as a balance). Some of the replies have been more general and not associated with a specific actor but can be extrapolated to the governments because of their broad application. Therefore, among the reasons why the actors don't have an ethical behavior when communicating about the use of AI is the fact that even with the best ethical intentions, they tend to bend the reality towards their purposes, that individuals are egocentric and behind all the actors there are individuals, that the AI ethics is a relatively new field in a domain dominated by tech and industrial interests, and even that the AI ethical domain is currently under-represented in most organizations. Thus, more efforts are needed to ensure transparency and safety in its use. Among the expressed concerns is that AI is often communicated as a universal solution to all problems, that there is a lack of accountability, a multitude of unknowns, and that the opportunities are overblown. At the same time, risks are downplayed. The concerns linked to a pre-existing agenda and interests and a lack of information regarding the technology and its use and effects have been mentioned multiple times.

Regarding the suggestions made by the respondents for actors' improvements regarding AI communication and social responsibility, not much difference has been observed between people based in the EU Members States or not. The suggestions made by the respondents based in the Member States are mainly associated with behavior regarding AI and stressed the following arguments: the consideration of all the stakeholders and public discussions, the need for risk

assessment, the communication of risks and limitations, more information and education about the technology, transparency regarding the use and the processes, the need for regulation, laws, standards and communication of incentives. Even if the respondents have mentioned some new suggestions in the "Europe" category (need for communication of facts as they are, good faith and political will, the consideration of worst-case scenarios and an AI charter for all sectors), the majority of the suggestions for improvement have been the same as the ones made by the EU countries. This suggests that the report with the technology is not very much linked to the government communication, as both groups have the same ideas for improvement, even if there is a significant difference in appreciation of the communication of risks and opportunities done by the national governments.

The communication of risks and opportunities appears as a central element of ethical behaviour regarding the communication of an emerging technology. Because the respondents mentioned the communication of risks and opportunities in different questions (not only the ones that referred explicitly to these elements), the conclusion is that the transparent communication of the two is a requirement and expected behavior for an ethical communication of AI should be.

Referring to the responsible communication of the EU, besides the responsibility associated with the governments, the question of the implications of humans and the existence of human ethics in the communication has been stressed. However, considering an ethical communication refers both to the organization and the individual, no signs of strongly expressed judgment or poor communication have been observed.

Analyzing opinions regarding the government's communication regarding the use of AI, it can be concluded that people consider the communication of opportunities and risks regarding AI as representative of an ethical communication. Furthermore, no strong criticisms of the communication made by the governments have been found. Consequently, governments have a behavior regarding the use of AI that is appreciated and considered responsible, but improvements are to be made.

Inspired by the analysis on the document and the replies to our survey, it has also been observed that responsible communication on AI can be considered following two main distinct axes: Communication process/technology and present/future (see Figure 4).



Figure 4: Two axes for a responsible communication on emerging technologies

V. CONCLUSION

To determine how the EU governments expressed their concern regarding the use of AI, this research considered the communication done by the EC and national governments in various official publications. Furthermore, this research used a definition of responsible communication of emerging technologies to determine the communication behavior of governments at the European level, what have been the main themes communicated, how the communication of sources of risks per sector of applicability has been made, and how people based in Europe perceived what has been communicated. Therefore, besides analyzing the official representatives' communication about AI use, this research has also contributed to the communication theory.

The use of AI in the long-term future is communicated as only bringing opportunities by both EC and national governments. However, the governments (EC and national) have communicated risks associated with using AI in various sectors of applicability, with the difference that the EC also expressed the existence of unknown risks. Looking into people's opinion, there is a general perspective that the EC and national governments communicate poorly about the use of AI and improvements to their communication are expected. This research helped prove our assumption that the communication of risks and opportunities are indicators for a responsible communication of emerging technologies. Therefore, it can be concluded that we have identified indicators that illustrate that the EC and national governments fill in the requirement of a responsible attitude regarding communication of the use of AI.

Even if this analysis did not inform us on a scale regarding the responsibility, the analysis in this paper informed us of elements indicating a responsible communication on emerging technologies as follows:

- Importance of communicating about the technology, positives and negatives associated with the use of the technology.
- Communicating both positives and negatives is a desired and expected behavior. However, attention is to be given not to emphasize too much one or the other, which is considered not an ethical behavior.
- More opportunities and risks have been communicated depending on the context, sector of applicability and country, which brought us to determine three

communication models (that apply to both content and process): Confident in the future (scaling from low to overconfident), Balanced (equally confident and fearful), fearful regarding the future (scaling from only fears to mainly fears).

- Putting together a lack of measures for responsibility and the three communication models, we also conclude that the communication of positives, negatives, mitigation measures and unknowns associated with the use of a specific emerging technology is made, indicating a responsible communication on emerging technologies. Other elements that could be added to scale from confident and fearful represents a suggestion for future research. This article strongly suggests a further analysis of the findings regarding Switzerland. It could represent an example of responsible communication to follow or proof that the judgment regarding a responsible communication of governments doesn't depend on the information that has been shared but on the trust in the institution.

Moreover, it has been observed that the responsible communication on emerging technologies can be defined around two axes: a time one (the impact of an emerging technology is communicated as being in the present or future), and the communication in terms of content and process (when the elements that have been communication refer to the communication process or the technology).

VI. REFERENCES

- [1] M.-V. Florin and M. T. Bürkler, "Introduction to the IRGC Risk Governance Framework," EPFL International Risk Governance Center, 2017. doi: 10.5075/epfl-irgc-233739.
- [2] C. Bieder, "Societal Risk Communication—Towards Smart Risk Governance and Safety Management," in *Risk Communication for the Future: Towards Smart Risk Governance and Safety Management*, M. Bourrier and C. Bieder, Eds. Cham: Springer International Publishing, 2018, pp. 155–175. doi: 10.1007/978-3-319-74098-0_11.
- [3] A. Saveliev and D. Zhurenkov, "Artificial intelligence and social responsibility: the case of the artificial intelligence strategies in the United States, Russia, and China," *Kybernetes*, vol. 50, no. 3, pp. 656–675, Jan. 2020, doi: 10.1108/K-01-2020-0060.
- [4] Royal Society and Royal Academy of Engineering, "Greenhouse Gas Removal," 2018. <https://royalsociety.org/topics-policy/projects/greenhouse-gas-removal/> (accessed Jun. 18, 2021).
- [5] M. Bal, *Narratology Introduction to the Theory of Narrative*. University of Toronto Press, 1997. Accessed: May 18, 2022. [Online]. Available: https://www.academia.edu/28730312/Narratology_Introduction_to_the_Theory_of_Narrative_Second_Edition_UNIVERSITY_OF_TORONTO_PRESS
- [6] D. Rotolo, D. Hicks, and B. R. Martin, "What is an emerging technology?," *Res. Policy*, vol. 44, no. 10, pp. 1827–1843, Dec. 2015, doi: 10.1016/j.respol.2015.06.006.
- [7] IRGC, "Ensuring the environmental sustainability of emerging technologies," 2022. [Online]. Available: [file:///C:/Users/rusu/Downloads/IRGC%20\(2022\)%20Ensuring%20the%20environmental%20sustainability%20of%20emerging%20technologies.%20Workshop%20report%20\(24\).pdf](file:///C:/Users/rusu/Downloads/IRGC%20(2022)%20Ensuring%20the%20environmental%20sustainability%20of%20emerging%20technologies.%20Workshop%20report%20(24).pdf)
- [8] EC, "Coordinated Plan on Artificial Intelligence," 2018. https://knowledge4policy.ec.europa.eu/publication/coordinated-plan-artificial-intelligence-com2018-795-final_en (accessed May 02, 2021).
- [9] EC, "Communication Artificial Intelligence for Europe | Shaping Europe's digital future," Apr. 2018. <https://digital-strategy.ec.europa.eu/en/library/communication-artificial-intelligence-europe> (accessed May 02, 2021).
- [10] EC, "White paper on artificial intelligence - A European approach to excellence and trust," Feb. 2020. https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf (accessed May 02, 2021).
- [11] EC, "Coordinated Plan on Artificial Intelligence. Annex," 2018. <https://digital-strategy.ec.europa.eu/en/library/coordinated-plan-artificial-intelligence> (accessed Jul. 20, 2021).
- [12] EC, "Fostering a European approach to Artificial Intelligence," Apr. 2021. [Online]. Available: <https://ec.europa.eu/newsroom/dae/redirection/document/75790>

- [13] EC, "Annexes to the Fostering a European approach to Artificial Intelligence," Apr. 21, 2021. <https://digital-strategy.ec.europa.eu/en/library/coordinated-plan-artificial-intelligence-2021-review> (accessed Jun. 05, 2022).
- [14] EC, "Proposal for a Regulation laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) | Shaping Europe's digital future," 2021. <https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-laying-down-harmonised-rules-artificial-intelligence-artificial-intelligence> (accessed May 05, 2021).
- [15] V. Van Roy, F. Rossetti, K. Perset, and L. Galindo-Romero, "AI Watch - National strategies on Artificial Intelligence: A European perspective, 2021 edition," *JRC Publications Repository*, Jun. 22, 2021. <https://publications.jrc.ec.europa.eu/repository/handle/JRC122684> (accessed May 15, 2022).
- [16] EC, "Ethics Guidelines for Trustworthy AI," *FUTURIUM - European Commission*, Dec. 17, 2018. <https://ec.europa.eu/futurium/en/ai-alliance-consultation> (accessed May 02, 2021).
- [17] White House, "Executive Order on Promoting the Use of Trustworthy Artificial Intelligence in the Federal Government – The White House," 2020. <https://trumpwhitehouse.archives.gov/presidential-actions/executive-order-promoting-use-trustworthy-artificial-intelligence-federal-government/> (accessed May 22, 2022).
- [18] National Institute of Standards and Technology (NIST), "Draft - Taxonomy of AI Risk," 2021, [Online]. Available: https://www.nist.gov/system/files/documents/2021/10/15/taxonomy_AI_risks.pdf
- [19] European AI Alliance, "New AI Watch report: 2020 EU AI investments," Oct. 2021. <https://futurium.ec.europa.eu/en/european-ai-alliance/blog/new-ai-watch-report-2020-eu-ai-investments> (accessed May 14, 2022).
- [20] J. G. March and J. P. Olsen, *Democratic governance*. New York: Free Press, 1995.
- [21] M. Vukasovic, J. Jungblut, M.-H. Chou, M. Elken, and P. Ravinet, "Multi-level, Multi-actor and Multi-issue Dimensions of Governance of the European Higher Education Area, and Beyond," in *European Higher Education Area: The Impact of Past and Future Policies*, A. Curaj, L. Deca, and R. Pricopie, Eds. Cham: Springer International Publishing, 2018, pp. 321–334. doi: 10.1007/978-3-319-77407-7_20.
- [22] EU, "Consolidated version of the Treaty on European Union," *Off. J. Eur. Union*, p. 34, Oct. 2012.