ARTIFICIAL INTELLIGENCE AND THE FUTURE OF COMMUNICATION OPPORTUNITIES, RISKS AND ETHICAL RESPONSIBILITY

Ionuț VLĂDESCU[1]

[1]Assoc. Prof., PhD, "Apolonia" University of Iaşi, Romania
Corresponding author: Ionuț Vlădescu; e-mail: ionut772002@yahoo.com

Abstract

Artificial intelligence is reshaping communication by automating content production, enabling large-scale personalization, supporting multilingual interaction, and assisting people with disabilities. At the same time, generative systems can produce plausible falsehoods, reproduce social bias, weaken privacy, facilitate deepfakes, and blur the attribution of responsibility. This article examines these changes through an interdisciplinary framework that combines communication studies, applied ethics, and emerging AI governance. Its central argument is that the ethical quality of AI-mediated communication cannot be reduced to technical accuracy. It also depends on preserving human agency, traceability, contextual truth, fairness, and the possibility of contesting consequential outputs. Drawing on UNESCO principles, the OECD framework, the NIST AI Risk Management Framework, and the European Union Artificial Intelligence Act, the study proposes a model of responsible communication organized around five requirements: disclosure of AI involvement, human editorial control, verification proportional to risk, protection of personal data, and AI literacy. The article concludes that AI should extend human communicative capacity without replacing moral judgment or dissolving accountability. Institutions that integrate these safeguards can benefit from speed and accessibility while protecting trust, dignity, and democratic dialogue.

Keywords*: artificial intelligence; communication; generative AI; media ethics; deepfakes; transparency; accountability; AI literacy.*

1. INTRODUCTION

Artificial intelligence has entered communication not as a simple new channel, but as a technical participant capable of formulating, selecting, translating and personalizing messages. Recommendation engines order the visibility of public discourse, generative models write texts and produce images, and conversational agents simulate the continuity of a dialogue. Consequently, the central question is no longer only who conveys the message and through what medium, but also to what extent its form, content and circulation have been determined by an automated system.

There is a real ambivalence to this change. AI-based tools can reduce translation costs, make information accessible to people with disabilities, support writing and tailor explanations to the interlocutor. The same tools can produce false claims in a convincing form, exploit personal data, amplify stereotypes and fabricate audiovisual evidence. The power to communicate on a large scale is thus separated from the competence to verify and take responsibility for the consequences.

The article argues that the future of communication depends on the articulation of three levels: technical performance, relational quality, and moral responsibility. A message can be fluent, personalized, and fast, but it remains inappropriate if it misleads, violates dignity, or makes it impossible to identify the person responsible for it. UNESCO's Recommendation on the Ethics of Artificial Intelligence places human dignity, transparency, equity, and human oversight at the heart of technological governance. These principles provide a benchmark for the analysis of communication, as trust is not only built through efficiency, but through truth, explainability, and accountability. (UNESCO, 2021).

The method used is conceptual and interdisciplinary. The analysis correlates the communication theory with the literature on generative linguistic models and the main international normative frameworks. Four questions are pursued: what new capabilities does artificial intelligence bring; what risks affect the truth and the relationship between people; how responsibility is distributed between the developer, the institution and the user; what rules

can transform ethical principles into verifiable practices.

2. AI-MEDIATED COMMUNICATION

From tool to technical co-author

Any communication technology alters the relationship between author, message and audience. The press has expanded reproduction, radio and television have amplified simultaneity, and digital platforms have transformed the receiver into producer and distributor. Artificial intelligence introduces an additional difference: the system not only carries pre-formulated content, but generates probabilistic variants of it. The linguistic model anticipates plausible sequences based on training data and instructions received. The resulting fluency can create the impression of understanding, even when the system does not have secure access to the facts or the moral intention of the interlocutor.

Bender and collaborators cautioned that large-scale linguistic models produce language through statistical regularities, and the appearance of coherence should not be confused with understanding or an autonomous epistemic source. For communication studies, observation is decisive: the credible form of the utterance can be gleaned from referential verification. The user tends to attribute competence and intentionality to a well-articulated response, and this anthropomorphization can reduce critical vigilance. (Bender et al., 2021)

In professional communication, artificial intelligence already functions as a technical co-author. It proposes titles, summarizes documents, adapts the tone, generates responses and selects examples. The human author can save time, but remains obliged to verify the information and assume the published version. The use of the system does not automatically transfer moral responsibility to the technology provider. The more important the effect of the message is for health, education, reputation or rights, the more rigorous human control must be.

Personalization and audience change

Artificial intelligence allows the transition from communication addressed to large segments to almost individually adjusted messages. A system can modify the vocabulary, the length of the explanation, the examples and the time of transmission based on the profile of the user. In education, this ability can facilitate understanding; in administration, it can simplify institutional language; in relation to the public, it can quickly identify the expressed need.

Personalization becomes problematic when it is based on opaque inferences or pursues emotional vulnerabilities. An adapted message is not necessarily a respectful one. If the system infers fear, precariousness, illness, or political orientation and adjusts persuasion in order to achieve behaviour, efficiency turns into manipulation. The principle of autonomy requires that the interlocutor retains a real possibility of understanding the purpose of the message and refuse profiling.

Opportunities for contemporary communication

Accessibility and inclusion

The most compelling promise of artificial intelligence is expanding access. Machine transcription can support people with hearing impairments, text-to-speech can make texts accessible to the visually impaired, and rephrasing can reduce barriers produced by administrative or academic language. Computer-assisted translation facilitates cooperation between people who do not share the same language and can open up educational resources to wider communities.

However, inclusion should not be evaluated solely by the existence of the function. Quality varies across languages, dialects, and social groups. A repeated speech recognition error for certain accents can exclude the very audience that the tool claims to assist. Testing must include diverse users, and the institution must provide an alternative path when the system fails (McLuhan, 1994).

Editorial efficiency and assisted creativity

In newsrooms, universities and organizations, generative systems can take over repetitive tasks: initial drafts, summaries, classification of messages or comparison of variants. The time saving allows focusing on documentation, interpretation and relationship with the audience.

Assisted creativity works well when the model offers options, and the professional selects and justifies the end solution.

The value of the system increases when it is used as an exploration partner, not as a substitute for discernment. Asking for counterarguments, checking coherence or adapting a text for different audiences can improve the process. On the other hand, the full delegation of writing can standardize the style and introduce errors that are difficult to notice. Institutional communication risks becoming formally correct, but lacking concrete experience, voice and responsibility.

Support in crisis situations

In emergencies, artificial intelligence can quickly analyse large volumes of information, identify recurring questions, and generate multilingual versions of alerts. The benefit depends on integration with official sources and human approval. A quick but unverified response can increase panic or direct people towards dangerous behaviour.

The proportionality of control is essential. An internal message with little effect can be verified by survey, while a medical recommendation, a public alert or an election information must be fully verified before dissemination. NIST proposes risk management throughout the lifecycle and organizes practice around the functions of governance, mapping, measurement and management. Applied to communication, this logic requires establishing the possible consequences before speed becomes the dominant criterion.(National Institute of Standards and Technology, 2023).

Epistemic and social risks

Hallucination and the authority of form

Generative models can formulate non-existent sources, mix up facts, and answer questions with certainty for which they do not have sufficient information. The term hallucination describes a fluent but unsubstantiated result. The risk comes not only from error, but from the relationship between error and convincing form. The reader encounters reliable grammar, orderly explanations, and sometimes fabricated references; these stylistic clues mimic academic authority.

In public communication, the verification must follow the statement, the source and the timeliness. A text generated regarding legislation, health or public policy cannot be validated by a second reformulation of the same model. It is necessary to consult the primary document and record the date of access. In the academic environment, citations and references must be compared with the actual editions. The practice protects both the correctness and the credibility of the author.

Synthetic disinformation and deepfake

Cheap text, voice, image, and video generation reduces the cost of persuasive forgery. Deepfakes can attribute words to a person who hasn't said them, and networks of automated accounts can create the impression of consensus. The problem goes beyond isolated fake content. The abundance of fakes produces widespread uncertainty in which even authentic evidence can be denied. Public trust erodes when the receiver no longer has accessible criteria to distinguish the origin of the material.

The European Artificial Intelligence Regulation introduces transparency obligations for certain systems and content generated or manipulated, including reporting on deepfakes. Legal transparency is necessary, but not sufficient. Tagging must be intelligible, visible and preserved when a certain material is redistributed. Digital mechanisms are also useful, but they are not a substitute for media literacy and editorial responsibility (European Parliament and Council of the European Union, 2024).

Discrimination and communication inequality

Models learn from corpora that contain historical asymmetries, stereotypes, and unequal distributions of voices. The result can marginalize certain communities through offensive representations, poor translations, or unfavourable recommendations. Discrimination does not always appear as an explicit insult; it can take the form of invisibility, misclassification, or a cultural standard presented as universal.

Equity involves evaluating performance on relevant groups and including affected people

in design. Floridi and Cowls show that an ethic of artificial intelligence must combine beneficence, non-harm, autonomy, justice, and explainability. In communication, justice means both avoiding harm and distributing benefits equitably. An institution cannot invoke average efficiency if errors are systematically borne by the same groups. (Floridi & Cowls, 2019)

Privacy and data exploitation

Communication with a conversational system may contain intimate information, professional data or confidential documents. The user may treat the interface as a private interlocutor, without understanding the route of the data, the retention period or its use to improve the service. Institutions must establish clear rules on information that cannot be entered into public systems and choose solutions compatible with data protection obligations.

The principle of minimization requires the collection of only the data necessary for the declared purpose. Personalization does not justify the unlimited accumulation of profiles. Formal consent, hidden in a long document, does not resolve informational asymmetry. The interlocutor must know when communicating with a system, what data is processed and who he can contact for correction or deletion.

Dependency and degradation of skills

Permanent assistance can reduce the exercise of writing, argumentation, and listening. If the user delegates to the system the formulation of any difficult message, he misses the opportunity to negotiate the meaning and assume the vulnerability of a real relationship. In education, the problem is not only fraud, but the outsourcing of the process by which the student learns to think and express a position.

The appropriate response is not the general prohibition, but the design of tasks that make the process visible: successive versions, justification of choices, reflection on errors and oral discussion. The system can support learning if the user compares, corrects and explains. It becomes harmful when the fluent product hides the absence of understanding.

Ethical responsibility

Dignity, autonomy and human presence

Communication is more than the transmission of information. A system can simulate empathy through appropriate formulas, but it does not bear the consequences of the relationship and does not experience the vulnerability of the speaker. Therefore, in pastoral, psychological, medical or educational contexts, automatic assistance must be delimited from the professional and human relationship.

Respecting autonomy involves informing the interlocutor about the role of the system and maintaining a human choice for important decisions. It is unethical for a person to receive a decision with a major effect through an automated message that is impossible to dispute. UNESCO insists that artificial intelligence must not replace ultimate human responsibility. The institution must identify the person or structure competent to review the result and repair the damage. (UNESCO, 2021)

Transparency and explainability

Transparency has at least three forms. The first is identity transparency: the receiver knows that he or she is interacting with a system or that the material has been generated. The second is process transparency: the institution describes the relevant data, the limits and the criteria for use. The third is transparency of decision: the affected person receives sufficient explanation in order to understand and dispute the result.

Not all models can provide a complete explanation of the internal mechanism, but this does not justify organisational opacity. The institution can explain why it chose the system, what tests it has carried out, what are the known limits and who approves the messages. The OECD includes transparency, explainability and accountability among the principles of trusted artificial intelligence. The explanation must be tailored to the risk and the public, not reduced to a technical mention that is difficult to understand (OECD, 2019; World Health Organization, 2021).

Distributed liability without dilution of liability

The communication chain involves the model developer, the application provider, the institution that implements it, the professional who validates and the user who distributes. Responsibility is distributed, but it must not be diluted to the point of extinction. Each actor is responsible for the decisions he or she controls: the developer for evaluation and documentation, the institution for purpose and procedures, the professional for validation, the user for culpable or misleading use.

A serious policy names the owner of each risk. The formula according to which "the algorithm got it wrong" hides the human choice to use the result. For high-impact messages, the system version, the relevant instruction, the sources consulted, the editorial approval and the subsequent corrections must be preserved. Traceability allows for auditing and institutional learning.

A practical model for responsible communication

Ethical principles become useful when they are transformed into procedures. The model proposed below starts from five obligations that can be applied in media, education, administration and organizational communication.

Principle	Control question	Institutional practice	Proof
Transparency	Does the public know that AI has been used?	Clear labelling of interaction and synthetic content	Label and register
Human control	Who approves the final message?	Editorial review proportional to the impact	Name and date
Verification	What claims can cause damage?	Confirmation from independent primary sources	Sources sheet
Data protection	Are necessary and permitted data entered?	Minimize anonymization and access rules	Impact assessment
AI literacy	Do users understand the limits of the system?	Regular training and error detection exercises	Evaluation and update

Transparency of artificial intelligence engagement

The public must be informed when the system is carrying a conversation, when an image or voice is synthetic, and when a text has been substantially produced automatically. Disclosure is not necessary for any spell checking, but becomes mandatory when automation influences the content, the apparent identity of the author, or the receiver's decision. Institutions need to define thresholds and examples, so that the rule is applied consistently.

Editorial control proportional to risk

Full verification is necessary in high-impact areas. For current uses, control can be based on sampling and error indicators. A simple matrix can combine the probability, severity and reversibility of harm. The more personal, authoritative and difficult to correct the message, the stricter the supervision must be.

Provenance and verification of sources

Any important factual statement must be linked to an accessible source. In journalism and research, the model should not be treated as a source, but as a working tool. Audiovisual content requires verification of origin, metadata and context. Corrections must be published prominently, and the organization must keep the version history in order to understand how the error occurred.

Data protection and security

Before using a system, the institution must classify the data and prohibit the entry of sensitive information into unapproved tools. Contracts must clarify data retention, subsequent

entrainment and third-party access. The assessment is not limited to confidentiality: hidden instructions, manipulation of sources and the retrieval of unverified data can compromise the integrity of the message.

Artificial Intelligence Literacy

Traditional digital competence is not enough. The user needs to understand the probabilistic nature of the outcome, the possibility of error, the effect of instructions and the limits of automatic detection. Literacy also includes the ability to decide when the tool should not be used. UNESCO links responsible governance to education, civic participation, and data and ethics skills (UNESCO, 2021).

In universities, training should include comparing a generated response with primary sources, identifying false references, declaring use, and reflecting on personal contribution. Student assessment should aim at judgment, argumentation, and accountability, not just the end product. In organizations, training should be repeated because patterns and risks change.

Implications for education and communication professions

Communication professors have the responsibility of integrating artificial intelligence into the curriculum without turning the tool into authority. Students must learn both productive use and epistemic limits. A project may require documenting instructions, analysing the differences between versions, and justifying the final choice. Therefore, use becomes an object of reflection and does not remain a hidden operation.

For journalists and public relations specialists, editorial policy must specify the permitted uses, prohibited dates, the necessary verification and the manner of declaration. Generating an illustration, transcribing an interview and writing a news story do not have the same risk. The rules must be differentiated and the public informed when authenticity or attribution can be misinterpreted.

In pastoral communication and counselling, technology can organize materials, translate, and provide general information. It cannot replace the encounter with the person, the discernment, and the responsibility of the one who accompanies suffering. An automatic response that mimics compassion can be perceived as closeness, but it lacks awareness and does not assume the consequences. The limit must be communicated explicitly, especially to vulnerable people.

At the institutional level, responsible adoption begins with an inventory of the systems used and the affected audiences. This is followed by risk classification, designation of those responsible, testing, training, and incident reporting. NIST's profile for generative AI recommends treating specific risks based on the overall management framework and tailoring measures to suit organizational goals. This approach avoids both uncritical enthusiasm and general rejection of technology (National Institute of Standards and Technology, 2024).

3. DISCUSSION

The future of communication will not be decided solely by the performance of models. It will depend on the rules by which society determines what must be declared, what cannot be delegated and who repairs the damage. The essential difference is not between communication with and without artificial intelligence, but between traceable, contestable and person-oriented use and opaque, extractive and manipulative use.

European regulation provides a risk-based framework, but formal compliance with the law represents the minimum threshold. Ethics intervenes before the legal sanction and between the situations expressly provided. An organization can comply with an obligation to inform and yet formulate a label so vague that the public does not understand anything. It can obtain formal consent and yet exploit vulnerability. Moral evaluation concerns the real effect on autonomy and dignity.

International principles converge on values: human rights, equity, transparency, safety and accountability. The difficulty lies in operationalization. The model proposed in this article responds to this difficulty with simple evidence: etiquette, human approval, source sheet, data evaluation and verifiable training.

These elements do not eliminate the risk, but make the decision auditable and correctable.

An anthropological question also remains. If more and more messages are produced by optimizing probability, communication can become effective without being personal. The human voice is not reduced to lexical originality; it expresses experience, intention and willingness to respond. Protecting this dimension requires spaces where people speak on their own behalf, accept imperfection and assume the relationship (Habermas, 1984).

4. CONCLUSIONS

Artificial intelligence expands the ability to produce, adapt and distribute messages. The benefits are important in accessibility, translation, information organization and editorial support. The risks are just as concrete: convincing errors, audiovisual falsification, discrimination, invasion of privacy, personalized manipulation and weakening of human skills.

The ethical response does not consist in attributing a moral intention to the system, but in organizing human responsibility around it. The author, the professor, the journalist and the institution remain responsible for the messages they validate and distribute. They must declare the involvement of artificial intelligence when it is relevant, verify claims in proportion to the risk, protect the data and provide avenues for challenge.

The communication of the future will be credible if technology remains subordinate to the truth, dignity and freedom of the person. Artificial intelligence can amplify the human voice, but it cannot take over its consciousness. Institutions that maintain this distinction will therefore be able to use technological efficiency without sacrificing the trust on which any communication community depends.

References

Bender, E.M., Gebru, T., McMillan-Major, A. & Shmitchell, M. (2021) On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? In Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, pp. 610–623. New York: ACM.

European Parliament and Council of the European Union (2024) 2024/1689 (EU) Regulation laying down harmonised rules on artificial intelligence. Official Journal of the European Union.Available from: https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng [July 22, 2026].

Floridi, L. & Cowls, J. (2019) A Unified Framework of Five Principles for AI in Society. Harvard Data Science Review, 1, doi:10.1162/99608f92.8cd550d1.

Habermas, J. (1984) The Theory of Communicative Action. Vol. 1. Boston: Beacon Press.

McLuhan, M. (1994) Understanding Media: The Extensions of Man. Cambridge, MA: MIT Press.

National Institute of Standards and Technology (2023) Artificial Intelligence Risk Management Framework. Available from: https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf [July 22, 2026].

National Institute of Standards and Technology (2024) Artificial Intelligence Risk Management Framework Generative Artificial Intelligence Profile. Available from: https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.600-1.pdf [July 18, 2026].

OECD (2019) Recommendation of the Council on Artificial Intelligence. Available from: https://legalinstruments.oecd.org/en/instruments/oecd-legal-0449 [July 15, 2026].

UNESCO (2021) Recommendation on the Ethics of Artificial Intelligence. Paris: UNESCO.

World Health Organization (2021) Ethics and Governance of Artificial Intelligence for Health. Geneva: WHO.