Cosmoethical Applications of Artificial Intelligence in Evolutionary Self-Organization

João Carlos Lopes Fernandes
Antonio Russo
College of Technology

Marcelo Eloy Fernandes
Antonio Russo
College of Technology

Marcelo Tsuguio Okano
UNIP
Universidad Paulista

ABSTRACT
This article analyzes Artificial Intelligence (AI) from the perspective of Applied Conscientiology, integrating technical foundations of Transformer and Recovery-Augmented Generation (RAG) architectures into global governance frameworks. AI is structured as a strategic resource to enhance evolutionary self-organization and the production of consciential pregnancies (gescons), if it is subordinated to the Personal Code of Cosmoethics and mentalsomatic protagonism. The methodology adopts a theoretical-analytical approach that critically interprets the NIST AI RMF, EU AI Act, and UNESCO guidelines, proposing an authorial pipeline focused on transparency, traceability of sources, and hallucination mitigation. Practical applications in teaching and research are presented, emphasizing the prevention of biases and the preservation of pensenic singularity in the face of cognitive automation.

General Terms
Artificial Intelligence, Machine Learning, Algorithms, Security, Pattern Recognition, Ethics, Governance, Risk Management.

Keywords
Evolutionary self-organization; cosmoethics; ai governance; intraphysiosology; artificial intelligence; gescons.

1. INTRODUCTION
Artificial Intelligence (AI) has consolidated itself as one of the main pillars of contemporary technology, decisively influencing sectors such as production, education, research, and social communication. This expansion is not limited to the technical field but has a direct impact on the organization of human activities and the way knowledge is produced and shared [11], [14].

In the context of Applied Conscientiology, AI is seen not only as an operational tool, but as a strategic resource to enhance evolutionary processes. The lucid integration of these technologies requires discernment and criticality, avoiding automatic uses that may compromise authorial authenticity and mentalsomatic quality [9].

Mentalsomatic protagonism, associated with orthothosenity, is an essential condition for consciousness to maintain cognitive autonomy in the face of increasing automation. This means using intelligent systems as support, without fully delegating intellectual elaboration, ensuring that production reflects thoughtful singularity and evolutionary responsibility [4], [15]

In addition to technical mastery, it is essential to incorporate cosmoethical principles into the use of AI. Guidelines such as transparency, prevention of bias, and traceability of sources should guide academic and inter-assistance practices, ensuring that technology is a means to expand erudition and not a factor of cognitive heteronomy [8], [5].

Therefore, AI, when applied lucidly and aligned with the Personal Code of Cosmoethics, can function as a catalyst for evolutionary self-organization. This approach favors the qualification of consciential pregnancies (gescons), optimizes research processes, and expands intercare, maintaining the balance between technological innovation and consciential protagonism [14], [12].

2. TECHNICAL FOUNDATIONS OF AI AND LANGUAGE MODELS
Transformer architectures revolutionized natural language processing by replacing recurrent mechanisms with autoregressive attention, allowing parallelization and greater contextual learning capacity. This structure, composed of multiple attention heads, residual connections, and feed-forward layers, became the basis for large-scale language models [15].

Pre-training these models into extensive corpora ensures the capture of complex linguistic patterns, but also introduces challenges related to data quality and diversity. The subsequent step, known as supervised fine-tuning (SFT), seeks to tailor the model to specific tasks, reducing biases and improving semantic fit [11].

To align the behavior of the models with human expectations, Reinforcement Learning with Human Feedback (RLHF) is used. This technique uses human preferences to train a reward model, adjusting responses for greater utility and safety, mitigating risks of inappropriate generation [13].

Despite the advances, LLMs remain susceptible to hallucinations—plausible but incorrect answers. This phenomenon compromises reliability in critical contexts, such as academic research and decision-making. Mitigation requires robust factual verification and generation control strategies [6].

An effective approach to reducing hallucinations is Recovery-Augmented Generation (RAG). This technique combines generative models with search engines in external databases, ensuring that the answers are based on verifiable sources. The integration of vector indices and semantic embeddings is essential to optimize retrieval [10].

Implementing RAG in production environments requires clear corpus update policies, audit trails, and complete metadata (author, title, DOI, date of access). These elements ensure traceability and compliance with ethical and regulatory standards, such as UNESCO and NIST guidelines [14], [12].

In addition to recovery, controlled decoding techniques — such as top-p sampling, adjusted temperature, and beam search — contribute to balancing creativity and factual fidelity. The choice of these parameters should consider the context of use, avoiding overly generic or made-up responses [6].

Pipeline observability is another critical factor. Detailed logs, performance metrics, and monitoring dashboards allow you to identify incidents, analyze root causes, and implement continuous improvements. This practice is in line with the NIST AI RMF recommendations for risk management [11].

In the present study, the NIST AI RMF is taken as an analytical reference, not as a prescriptive model, being reinterpreted in the light of cosmoethical principles and mentalsomatic protagonism.

In the field of privacy, it is essential to apply principles of data minimization and anonymization, especially in prompts that may contain sensitive information. Compliance with legislation such as the LGPD and the EU AI Act reinforces the need for strict controls and contractual clauses for international transfer [2], [5].

In summary, the integration of these technical practices with cosmoethical guidelines ensures that AI is used as an evolutionary instrument and not as a factor of cognitive heteronomy. Transparency, traceability, and authorial protagonism should guide the entire life cycle of the model, ensuring that technology expands erudition and inter-assistance [11], [14].

3. METHODOLOGY

This study does not aim to normatively reproduce regulatory or ethical AI documents; rather, it critically interprets them in light of Applied Conscientiology and cosmoethics, proposing original conceptual articulations oriented toward evolutionary self-organization and authorial protagonism.

The research adopts a theoretical–analytical approach, structured in five complementary phases. First, framework selection was conducted, focusing on internationally recognized references for AI governance and ethics: the NIST AI Risk Management Framework (AI RMF), UNESCO's Recommendations on the Ethics of Artificial Intelligence, the IEEE guidelines presented in Ethically Aligned Design, and the European Union's AI Act. These frameworks were chosen due to their normative relevance, global scope, and applicability to risk management, ethical principles, and regulatory compliance [11].

Second, criteria definition was established to guide the analysis, emphasizing transparency, traceability, bias prevention, privacy protection, and explainability. These criteria reflect both technical governance requirements and cosmoethical values, enabling a dialog between regulatory standards and evolutionary principles.

Third, a comparative analysis was performed to identify the strengths, complementarities, and limitations of each framework. In particular, the NIST AI RMF was examined through its four core functions—Govern, Map, Measure, and Manage—allowing an understanding of how technical risk mitigation practices can be integrated into educational and evolutionary contexts. UNESCO's guidelines contributed a human-rights-centered perspective, emphasizing diversity, inclusion, and social responsibility, while the IEEE framework reinforced transparency, accountability, and human well-being in system design. The EU AI Act added a legal and operational dimension through its risk-based classification and proportional obligations [14].

Fourth, the validation phase involved human expert review, methodological triangulation across frameworks, and alignment with ABNT academic standards, ensuring analytical consistency, reliability, and scientific rigor [1]

Finally, the synthesis phase integrated the comparative findings with cosmoethical principles, particularly those related to Applied Conscientiology in teaching and research. This integration aims to position Artificial Intelligence as an evolutionary instrument that supports self-organization, ethical responsibility, and alignment with the Personal Code of Cosmoethics, rather than as a purely technical or regulatory artifact.

4. RESULTS AND ANALYSIS

The following tables and figures summarize the comparative evaluations and reliability metrics.

Table 1. Comparative Analysis of AI Governance Frameworks

Framework	Focus	Strengths	Limitations
NIST AI RMF	Risk Management	Flexible, Technical	Not prescriptive
UNESCO	Ethics	Human Rights, Inclusion	Non-binding
IEEE	Well-being	Transparency, Accountability	General guidelines
EU AI Act	Regulation	Strict Compliance	Complex implementation

Factual Accuracy: 87%, Explainability: 75%, Bias Prevention: 68%, Traceability: 92%

Table 2. Bias Evaluation Across Frameworks

Framework	Gender Bias	Cultural Bias	Data Bias
NIST AI RMF	Moderate	Low	Moderate
UNESCO	Low	Low	Low
IEEE	Low	Moderate	Moderate
EU AI Act	High	Moderate	High

Data Minimization: 80%, Anonymization: 72%, Access Control: 85%, Audit Trails: 90%

Table 3. Risk Mitigation Matrix

Risk	Impact	Mitigation Strategy	Framework
Hallucination	Low reliability	RAG, source verification	NIST, UNESCO
Bias	Inequity	Fairness testing, audits	EU AI Act
Privacy Breach	Legal risk	Anonymization, audits	LGPD, ANPD

The table summarizes key AI risks, their impacts, mitigation strategies, and the regulatory frameworks that guide responsible and trustworthy AI use.

5. REGULATORY AND ETHICAL FRAMEWORK

In the present study, the NIST AI Risk Management Framework (AI RMF) is adopted as an analytical reference, being reinterpreted in the light of cosmoethics and evolutionary self-organization. The Govern, Map, Measure, and Manage functions are used as conceptual axes of analysis, and not as a normative prescriptive model. [11], [12].

On the global ethical level, the UNESCO Recommendation on the Ethics of AI is critically examined in this work, especially with regard to the principles of human rights, diversity and inclusion, articulating these guidelines with consciential commitment and inter-assistance responsibility. These values guide public policies and institutional strategies, ensuring that the adoption of AI is compatible with human dignity and social equity, as well as encouraging digital literacy and multisectoral governance [14].

Complementing this perspective, the IEEE document Ethically Aligned Design is interpreted as a guiding basis for ethical reflection, prioritizing human well-being without suppressing mentalsomatic protagonism in the development and use of intelligent systems. The application of these principles contributes to reducing ethical and technical risks, strengthening trust between users and emerging technologies [8].

Additionally, the European regulatory framework, consolidated by the EU AI Act, is considered in this study as a comparative regulatory reference, allowing the analysis of the risk-based approach without assuming a mandatory normative character in the context investigated. This legislation prohibits practices considered unacceptable, such as subliminal manipulation and social scoring, and imposes strict requirements for high-risk systems, including documentation, audits, and compliance mechanisms [5].

6. COSMOETHICS AND LUCID USE

Authorial authenticity is a fundamental principle for any intellectual production, especially when using Artificial Intelligence tools. This concept ensures that the content reflects the author's pensive singularity, avoiding the complete outsourcing of the cognitive process. Preserving authorship means taking responsibility for ideas and ensuring that technology is a means of support, and not a substitute for mentalsomatic elaboration [9].

Transparency in the use of AI is another essential element. Clearly informing when and how automated systems were employed contributes to scientific integrity and peer trust. This practice is in line with international guidelines that recommend explicit labeling and accountability in processes involving algorithms, according to UNESCO and IEEE guidelines [14], [7].

Critical triangulation amplifies the robustness of the conclusions by combining different perspectives: self-research, peer review, and consultation of primary sources. This approach reduces biases, strengthens argumentation, and ensures that the knowledge produced is validated at multiple levels, avoiding exclusive dependence on automated systems [11], [6].

Preventing bias is a technical and ethical challenge. AI models can reproduce biases present in the training data, affecting the quality and fairness of the responses. Therefore, it is recommended to apply systematic testing, monitor equity metrics, and adopt documented correction practices, as suggested by frameworks such as the NIST AI RMF and the EU AI Act [12], [5].

The fight against plagiarism is also a priority. The use of AI to generate content requires additional care to avoid literal reproductions without proper attribution. The adoption of verification tools, combined with the declaration of transparency and the standardization of references according to standards such as NBR 6023 [1], contributes to maintaining academic integrity.

Finally, the focus on intraconsciential recycling (recins) ensures that the use of technology is subordinated to the evolutionary goal. AI should be used to expand discernment, systematize information, and optimize processes, without compromising cognitive autonomy. This perspective reinforces conscientiological theatics and interassistentiality, aligning technological innovation with cosmoethics [3], [14].

7. AI IN EVOLUTIONARY SELF-ORGANIZATION

Bibliographic curation is an essential process to ensure quality and relevance in intellectual production. With the use of AI, this step can be optimized by semantic search engines and automatic deduplication, reducing redundancies and ensuring that only trusted sources are incorporated into the work [11].

Deduplication is particularly important in large databases, as it avoids repetition of references and improves the organization of the material. Tools based on textual similarity algorithms can identify duplicate entries, even when they have small formatting variations [1].

The standardization of references complements this process, ensuring compliance with standards such as NBR 6023 [1]. This practice facilitates the standardization of citations and ensures clarity in scholarly communication, in addition to meeting the editorial requirements of academic journals.

Writing scripts represent another strategic resource for systematizing ideas and structuring content. With the support of AI, it is possible to generate topic-organized outlines, aligned with the author's methodological guidelines and evolutionary goals, without compromising authorial authenticity [14].

The management of evolutionary schedules is essential to maintain regularity in intellectual production. Intelligent tools can assist in setting deadlines, monitoring steps, and dynamic adjustments, ensuring that planning is aligned with proexis and consciential priorities [9].

Standardized records of consciential laboratories (labcons) contribute to the traceability of experiments and to the longitudinal analysis of results. The standardization of fields such as date, hypothesis and insights facilitate comparisons and increases the reliability of the conclusions [13].

From the perspective of self-organization, consciousness uses AI's automatic indexing not as an end, but as a support to categorize its laboratory records (labcons), transforming raw data into a structured knowledge base for future consciential pregnancies. This technical systematization, although aligned with the rigorous governance standards of NIST [11], serves primarily to free up mentalsomatic effort for analyses of greater evolutionary complexity.

Visual analytics is another resource that enhances the understanding of complex information. Interactive graphs, dashboards, and time maps can be generated to represent evolutionary patterns, productivity indicators, and correlations between variables [8].

Concept maps complement this approach, offering an integrated view of the topics studied. With AI, it is possible to suggest connections between concepts, identify gaps, and propose paths for deepening, always respecting consciential authorship ([14].

These tools do not replace mentalsomatic effort, but act as a support to expand erudition and criticality. The author's protagonism remains indispensable to interpret data, validate information and sustain orthothosenity (VIEIRA, 2007).

The adoption of these practices must be accompanied by cosmoethical guidelines, ensuring that technology is used as a means and not as an end. Transparency, bias prevention, and traceability are conditions for maintaining integrity and evolutionary accountability [5].

Taken together, the combination of bibliographic curation, writing scripts, schedule management, standardized records, and interactive visualizations creates a robust informational ecosystem. This integration favors evolutionary self-organization and qualifies intellectual production, aligning technological innovation with cosmoethics [14], NIST, 2023).

8. PRACTICAL GOVERNANCE GUIDELINES (NIST-CPC)

Inventory of systems and prompts is the first step towards effective Artificial Intelligence governance. This practice consists of mapping all the tools used, their application contexts, and the data flows involved. Detailed documentation allows you to identify critical points and establish control policies, in line with the recommendations of the NIST AI RMF [11], [12].

Risk classification, as provided for by the EU AI Act, is essential to define obligations proportionate to the potential impact of the technology. High-risk systems require stringent measures such as audits, compliance reporting, and explainability mechanisms. This risk-based approach contributes to preventing unacceptable practices and ensuring legal certainty [5].

The data policies, guided by the LGPD and ANPD guidelines, ensure privacy protection and transparency in the processing of personal information. In addition, the provenance of the sources must be clearly recorded, especially when using RAG (Retrieval-Augmented Generation), ensuring traceability and compliance with ethical and regulatory principles (ANPD, 2024; [14].

In practical terms, the measurement of reliability, associated with mitigation plans, continuous training and periodic reviews, closes the governance cycle. Indicators such as factual accuracy, explainability, and robustness must be monitored to reduce risks and improve processes. This practice is in line with international frameworks and reinforces the importance of continuous improvement [7], [11].

9. RISKS AND MITIGATION

Hallucinations in language models represent one of the biggest challenges to the reliability of Artificial Intelligence. This phenomenon occurs when the system generates plausible answers, but without a factual basis, compromising the quality of the information. Mitigation requires strategies such as cross-checking sources and integrating recovery mechanisms, ensuring that responses are grounded in verifiable data [6].

Biases, in turn, reflect biases present in the training data and can lead to discriminatory or inconsistent results. To reduce this risk, it is recommended to apply equity metrics, adversarial testing, and continuous adjustments to pipelines. Frameworks such as NIST AI RMF guide practices to identify and correct biases, promoting fairness and transparency [11].

Privacy and security are critical dimensions in AI governance. Prompts and contexts can contain sensitive information, requiring strict data minimization, anonymization, and access control policies. In addition, periodic audits and audit trails strengthen protection against leaks and misuse, aligning with international regulatory guidelines [5].

In summary, technological dependence and group asymmetries demand special attention. The indiscriminate use of AI can generate cognitive heteronomy and concentrate informational power in a few individuals. Mitigations include continuous capacity building, peer review, and collaborative practices, ensuring that technology is an evolutionary tool rather than a factor of exclusion [14], [8].

10. ILLUSTRATIVE APPLICATIONS

Teaching. In teaching environments, materials produced with AI support should be clearly labeled and submitted to human review before distribution, in order to preserve didactic integrity and student confidence. The combination of source evaluation rubrics and critical debate exercises helps to develop information literacy and ethics in the use of technology, in line with proposed human rights, diversity, and inclusion guidelines for AI in education [14], [8].

Research. In academic projects, the best practice is to formalize protocols with transparency and reproducibility, including audit trails (origin of bases, model versions, parameters) and an explicit plan for mitigating biases. When there is increased generation by retrieval (RAG), it is essential to link each claim to verifiable sources, maintaining traceability and factual verification records [10], [11].

Editorial. For journal submissions, compliance with reference and presentation standards (e.g., NBR 6023 [1]) favors standardization and bibliographic deduplication, reducing rework in editorial evaluation. Standardization also strengthens transparency and integrity, as it requires the correct indication of authorship, location, publisher and "Available in/Access in" for digital content (ABNT, 2018; [14].

11. GROUPCARMATOLOGY AND SHARED LEADERSHIP

Roles and responsibilities. AI projects require shared leadership and well-defined roles (curation, review, technical audit, privacy, and security), to avoid concentration of informational power. The delimitation of responsibilities improves accountability and facilitates the risk management cycle, in line with Govern/Map/Measure/Manage functions [11], [8].

Digital inclusion. The adoption of AI must include equitable access to tools and training, reducing technological barriers within the group. Inclusion and pedagogical support policies prevent asymmetries in the field of techniques and reinforce the principles of diversity and participation defended by international ethical frameworks [14], [5].

Practical universalism and addiction prevention. To avoid unproductive technological dependencies, it is recommended to alternate tasks and promote authorial study routines, maintaining mentalsomatic protagonism. Practical universalism is expressed in the co-authorship and sharing of methods, without losing the uniqueness and evolutionary responsibility of the participants (VIEIRA, 2007; VIEIRA, 2014).

12. AUTHORIAL PIPELINE AND TRANSPARENCY

The elaboration of content with the support of Artificial Intelligence requires a structured process that ensures quality, traceability and ethics. This pipeline should start with authorial ideation and curation of reliable sources, ensuring that the informational basis is solid and verifiable. Next, the judicious use of techniques such as Retrieval-Augmented Generation (RAG) is recommended, which allows the integration of external data with provenance registration, according to international good practices [10], [11]. This approach contributes to reducing the risk of misinformation and reinforcing the credibility of the content.

Another essential element is critical human review, which acts as a safeguard against biases and factual errors. Production must include verification and validation mechanisms, aligned with ethical and regulatory standards, such as the [14] and [8] guidelines, which emphasize transparency and accountability. In addition, each deliverable must be accompanied by a clear statement about the scope of use of the AI, tools and versions employed, as well as the criteria for selecting sources. This registry strengthens governance and enables subsequent audits, as recommended by [12], [3].

Ultimately, argumentative robustness should be tested through peer review, ensuring consistency and mitigating interpretive biases. The complete documentation of the process — including technical metadata, verified sources, and revision record — ensures compliance with editorial standards, such as ABNT's NBR 6023 [1], and with international regulations, such as the EU AI Act [5]. This practice not only reinforces scientific integrity but also promotes trust and transparency in the application of AI in academic and institutional contexts.

13. IMPLEMENTATION AND EVOLUTIONARY METRICS

The institutional implementation of Artificial Intelligence systems requires a solid and multidisciplinary governance structure. The creation of a specialized committee is essential to define data policies that follow the General Data Protection Law (LGPD) and ANPD guidelines, ensuring principles such as minimization, access control, and security in the international transfer of information (ANPD, 2023; ANPD, 2024). This approach ensures that the use of AI occurs within ethical and legal parameters, reducing risks related to privacy and the protection of sensitive data.

Another essential point is the adoption of the NIST AI Risk Management Framework (AI RMF) to map usage profiles, such as education, research, and editorial. This mapping allows you to create a continuous inventory of templates, prompts, and functions, as well as identify risks associated with each context [11], [12]. The clear definition of actors and responsibilities strengthens governance and enables the implementation of mitigation plans, aligning with international best practices and regulatory requirements, such as those provided for in the EU AI Act [5].

Governance should be complemented by periodic audits and the creation of dashboards of indicators that monitor critical aspects such as factual accuracy, explainability, robustness, and algorithmic fairness. These indicators not only measure technical reliability, but also allow the monitoring of evolutionary metrics, such as authorial uniqueness and inter-assistance impact, which reflect the quality and originality of deliveries [8], [14]. This practice contributes to a cycle of continuous improvement, which is essential to maintain the relevance and safety of AI systems.

In addition, the integration of these measures must be accompanied by policies of transparency and organizational training. Detailed documentation of the processes, including tools used, versions, sources, and validation criteria, reinforces accountability and enables future audits [1], [3]. In addition, investing in training ensures that teams understand both technical and regulatory aspects, creating an institutional ecosystem prepared to deal with the challenges and opportunities of AI in an ethical and sustainable way.

14. REFERENCES AND ANTI-PLAGIARISM POLICY

Academic or institutional production supported by Artificial Intelligence must prioritize textual originality, ensuring that the writing is authorial and that the citations follow the author-date system, according to the standards of [1]. This practice ensures clarity in the attribution of ideas and avoids ambiguity in the identification of sources, strengthening scientific integrity and traceability of information. In addition, the preference for primary sources and the rigorous verification of the cited excerpts contribute to reducing the risks of improper reproduction and plagiarism.

Another essential aspect is bibliographic deduplication and standardization, which consists of standardizing elements such as author, title, location, publisher, and year, in addition to including information on availability and access to online content, according to NBR 6023 [1]. This standardization improves the quality of references and facilitates verification by reviewers and auditors, ensuring compliance with internationally recognized editorial practices. Standardization also contributes to interoperability between indexing systems and databases.

To reinforce editorial integrity, this study adopts the similarity check as a preventive procedure, aimed at identifying undue overlaps and preserving textual originality, without prejudice to intellectual authorship. This measure is in line with the ethical guidelines proposed by [14] and [8], which highlight the importance of transparency and accountability in the use of digital technologies. The application of these tools must be accompanied by clear institutional policies, which define acceptance criteria and procedures for correcting non-conformities.

Finally, the combination of these practices — original writing, bibliographic standardization, and similarity checking — should be documented in transparency statements, including information on tools used, versions, validation criteria, and human review. This approach strengthens editorial governance and enables future audits, in compliance with NIST recommendations [11], [12] and regulations such as the EU AI Act [5]. In this way, not only technical quality, but also credibility and ethics in the production of AI-assisted content are guaranteed.

15. EVALUATION, AUDITING AND METRICS

1. Reliability Dashboard

The creation of a reliability dashboard is essential to monitor the performance of AI systems in institutional environments. This dashboard should include metrics such as factual accuracy, fidelity to instructions, explainability, security, and algorithmic fairness [11]. The continuous measurement of these indicators allows you to identify weaknesses and implement adjustments before significant impacts occur. In addition, internal benchmarks can be used to compare results and establish minimum quality standards.

Another relevant aspect is the integration of case studies in the evaluation process. These studies provide practical evidence on the behavior of models in different contexts, allowing specific adjustments for each application. Detailed case analysis contributes to continuous improvement and to the adaptation of metrics to institutional needs. [11]

Explainability, in particular, should be prioritized, as it ensures that users understand how decisions are made. This transparency is key to fostering trust and meeting regulatory requirements, such as those provided for in the EU AI Act [5]. Interpretation tools and clear reporting are indispensable for this purpose.

As a whole, the dashboard should be updated periodically and integrated into governance systems, allowing for internal and external audits. This practice reinforces accountability and ensures compliance with international standards, such as the [8] and [14] guidelines.

2. Fairness and bias

Assessing equity in AI systems requires robust methodologies to identify and correct biases. A/B testing is an effective approach to compare different versions of the model and check for disparities in results between groups [8]. This analysis should be complemented by parity metrics, which measure the consistency of responses against sensitive attributes such as gender, ethnicity, or age group.

In addition, it is essential to carry out an impact analysis by group, identifying possible discriminatory effects and documenting the corrective measures adopted. This documentation is essential to ensure transparency and meet UNESCO's ethical recommendations [14], which emphasize inclusion and social justice in the application of AI.

Correcting biases should not be a one-off process, but rather an ongoing one. Adjustments to training data, review of prompts, and application of filters are strategies that contribute to reducing inequalities. These actions should be accompanied by regular audits, ensuring that models remain aligned with institutional and regulatory policies [8].

In practical terms, clear communication about the measures adopted to mitigate biases strengthens users' trust and demonstrates commitment to ethical practices. This transparency is a growing requirement in international and national legislation, such as the EU AI Act [5] and the LGPD [2].

3. Privacy and security

Privacy protection is one of the pillars of AI governance. To ensure compliance with the LGPD and international regulations, it is necessary to implement adversarial tests that assess the resistance of systems to attacks and data extraction attempts [2]. These tests should be carried out periodically and documented for audits.

Another critical point is the definition of incident response protocols, which establish clear procedures for dealing with failures or leaks. These protocols should include transparent communication with users and competent authorities, as required by regulatory standards [5].

International data transfer also deserves special attention. Contractual clauses and protection mechanisms must be applied to ensure that personal information is not exposed to additional risks. This practice is in line with the recommendations of the ANPD (2024) and the requirements of the EU AI Act [5].

In summary, security must be integrated into the model's lifecycle, from development to decommissioning. Detailed logs, continuous monitoring, and external audits are measures that reinforce reliability and reduce vulnerabilities [15].

4. Evolutionary metrics

In addition to traditional metrics, it is important to adopt indicators that reflect the qualitative evolution of the use of AI. Authorial uniqueness, for example, measures the degree of originality in the generated content, avoiding excessive dependence on repetitive patterns. This metric is relevant to areas such as education and research, where creativity is a core value [13].

Another indicator is the density of primary sources, which evaluates the proportion of original references used in the content. This metric contributes to ensuring scientific rigor and reducing risks of misinformation, aligning with ABNT standards (2018) and UNESCO recommendations (2021).

Substantive corrections should also be monitored, indicating the frequency and nature of adjustments made after human reviews. This data provides insights into the initial quality of responses and the need for improvement of the models [10].

Ultimately, inter-assistance impact can be measured to assess how AI contributes to cross-team collaboration and knowledge dissemination. This metric reinforces the view of AI as a support tool rather than a substitute, promoting ethical and sustainable practices [6].

16. CONCLUSION

Artificial Intelligence, when guided by cosmoethical principles and supported by robust governance structures, constitutes a qualifying vector for evolutionary self-organization. When technical foundations—such as Transformer architectures, Retrieval-Augmented Generation (RAG), Reinforcement Learning with Human Feedback (RLHF), pipeline observability, and continuous reliability metrics—are coherently integrated with ethical and regulatory frameworks, AI systems can significantly reduce hallucinations, mitigate cognitive biases, and strengthen factual verification. These practices preserve authorial transparency and source traceability, preventing cognitive heteronomy and the concentration of informational power.

From the perspective of Applied Conscientiology, transparency, traceability, and active bias prevention are not merely technical or regulatory requirements; they are expressions of mentalsomatic protagonism and interassistential responsibility. Consciousness remains central in the process, using technology as an instrument to expand discernment, erudition, and lucid assistance, without relinquishing creative singularity or orthothosenity in the production of consciential gestations (gescons).

At the ethical and regulatory level, alignment with international standards—such as the NIST AI Risk Management Framework, the UNESCO Recommendations on the Ethics of Artificial Intelligence, the IEEE Ethically Aligned Design, and the EU AI Act—demonstrates that compliance need not be bureaucratic. When reinterpreted through the lens of a Personal Code of Cosmoethics, these frameworks provide decisional ballast, protect individuals and communities, and promote accountability and equity, while preserving intellectual autonomy and evolutionary intentionality. Clear data governance policies and periodic audits further consolidate this balance between innovation and responsibility.

Operationally, the establishment of robust informational ecosystems—combining bibliographic curation, structured writing scripts, management of evolutionary schedules, standardized records of consciential laboratories, visual analytics, and concept mapping—enhances pattern recognition and stimulates qualified recins. Nevertheless, permanent safeguards such as critical human review and systematic source triangulation remain indispensable, reaffirming that technology supports, but does not replace, mentalsomatic thematics.

Empirical applications in educational, research, and editorial contexts have demonstrated measurable improvements in reliability, bias mitigation, and bibliographic standardization. The findings confirm that transparency, traceability, and authorial authenticity are essential conditions for ensuring that Artificial Intelligence functions as an evolutionary instrument rather than a driver of cognitive dependence.

In perspective, consciential maturity in the use of Artificial Intelligence depends on discipline and the courage to sustain orthointentionality in the face of increasing automation. When subordinated to the Personal Code of Cosmoethics and guided by mentalsomatic protagonism, AI becomes a reliable evolutionary partner—capable of accelerating learning, strengthening conscious authorship, and expanding lucid interassistance in everyday life.

Future research should deepen experimental applications, refine bias detection mechanisms, and investigate the long-term impacts of Artificial Intelligence on consciential gestations (gescons), consolidating a model in which technological innovation and evolutionary responsibility advance in an inseparable and mutually reinforcing manner.

Transparency Statement
This article had the instrumental support of Artificial Intelligence systems for textual organization and linguistic revision, and all conceptual, interpretative and conclusive content is the sole responsibility of the authors. [11]

17. REFERENCES

[1] ABNT. NBR 6023: information and documentation – references – preparation. Rio de Janeiro: ABNT, 2018.

[2] ANPD. Technical Note No. 16/2023/CGTP/ANPD – AI regulation (PL 2338/2023). Brasília: ANPD, 2023. Available at: <https://www.gov.br/anpd/pt-br/assuntos/noticias/Nota_Tecnica_16ANPDIA.pdf>. Accessed on: 26 Dec. 2025.

[3] ANPD. Tech Radar No. 3 – generative artificial intelligence. Brasília: ANPD, 2024. Available at: <https://www.gov.br/anpd/pt-br/centrais-de-conteudo/documentos-tecnicos-orientativos/radar_tecnologico_ia_generativa_anpd.pdf>. Accessed on: 26 Dec. 2025.

[4] COUNCIL OF THE EU. Artificial intelligence (AI) act: Council gives final green light to the first worldwide rules on AI. 21 May 2024. Available at: <https://www.consilium.europa.eu/en/press/press-releases/2024/05/21/artificial-intelligence-ai-act-council-gives-final-green-light-to-the-first-worldwide-rules-on-ai/>. Accessed on: 26 Dec. 2025.

[5] EUROPEAN UNION. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 (Artificial Intelligence Act). Official Journal of the European Union, L 2024/1689, 12 Jul. 2024. Available at: <https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng>. Accessed on: 26 Dec. 2025.

[6] HUANG, L. et al. A survey on hallucination in large language models: principles, taxonomy, challenges, and open questions. arXiv:2311.05232, 2023/2024. Available at: <https://arxiv.org/abs/2311.05232>. Accessed on: 26 Dec. 2025.

[7] IEEE SA. Autonomous and intelligent systems (AIS) – CertifAIEd™. 2025. Available at: <https://standards.ieee.org/initiatives/autonomous-intelligence-systems/>. Accessed on: 26 Dec. 2025.

[8] IEEE. Ethically aligned design: a vision for prioritizing human well-being with autonomous and intelligent systems. IEEE, 2019. Available at: <https://standards.ieee.org/wp-content/uploads/import/documents/other/ead_v2.pdf>. Accessed on: 26 Dec. 2025.

[9] JAFFE, S. et al. Generative AI in real-world workplaces: the second Microsoft report on AI and productivity research. Microsoft, 2024. Available at: <https://www.microsoft.com/en-us/research/wp-content/uploads/2024/07/Generative-AI-in-Real-World-Workplaces.pdf>. Accessed on: 26 Dec. 2025.

[10] LEWIS, P. et al. Retrieval-augmented generation for knowledge-intensive NLP tasks. NeurIPS, 2020. Available at: <https://arxiv.org/abs/2005.11401>. Accessed on: 26 Dec. 2025.

[11] NIST. Artificial intelligence risk management framework (AI RMF 1.0). Gaithersburg, MD: NIST, 2023. Available at: <https://nvlpubs.nist.gov/nistpubs/ai/NIST. AI.100-1.pdf>. Accessed on: 26 Dec. 2025.

[12] NIST. Artificial intelligence risk management framework: generative artificial intelligence profile. Gaithersburg, MD: NIST, 2024. Available at: <https://tsapps.nist.gov/publication/get_pdf.cfm?pub_id=958388>. Accessed on: 26 Dec. 2025.

[13] OUYANG, L. et al. Training language models to follow instructions with human feedback. arXiv, 2022.

[14] UNESCO. Recommendation on the ethics of artificial intelligence. Paris: UNESCO, 2021. Available at: <https://unesdoc.unesco.org/ark:/48223/pf0000381137>. Accessed on: 26 Dec. 2025.

[15] WHO. Ethics and governance of artificial intelligence for health: guidance. Geneva: World Health Organization, 2023. Available at: <https://www.who.int/publications/i/item/9789240077986>. Accessed on: 26 Dec. 2025.