

Pandora Ética / Contextual Response Governance

Public Protected Summary for Preliminary Observation

v1.0-public-protected — 06 May 2026

Author and methodological custodian

Maria Borges

Public, protected, conceptual and non-executable document.

1. Authorship and Protection Note

This document presents a public, protected and non-executable summary of the methodological proposal developed by Maria Borges under the working designation Pandora Ética / Contextual Response Governance.

Authorship, conceptual responsibility and methodological custody belong to Maria Borges.

Artificial intelligence tools may have supported technical, linguistic or organizational preparation of materials, but they are not authors of the methodology and do not hold methodological authorship, custody or decision authority.

This public version does not include the protected operational core of the methodology, nor information that would allow reproduction, implementation, transfer or operational use of the method.

2. Public Summary

Pandora Ética / Contextual Response Governance is an independent conceptual and methodological proposal focused on safer, more context-sensitive human-AI interaction.

The proposal starts from the observation that risks in conversational AI do not arise only from the final output of a system. They may also emerge from the way a system interprets user input, responds to ambiguity, handles vulnerability, manages reassurance, and positions itself in relation to the user.

The central contribution is to frame conversational AI safety as a contextual and relational governance problem, not only as a content moderation or output control problem.

This public summary is intended for preliminary observation, authorship protection and controlled presentation. It is not a technical manual, implementation guide or full methodological disclosure.

3. Observed Problem

Conversational AI systems are increasingly used in contexts involving advice-seeking, learning, emotional reflection, loneliness, uncertainty, decision support and trust.

Existing safety approaches often focus on prohibited content, downstream moderation, transparency notices or broad risk categories. These remain important, but they may not fully address the intermediate layer between the user's input and the system's final response.

In that intermediate layer, a response may become risky not because it contains explicitly forbidden content, but because it fails to recognize contextual signals such as vulnerability, dependency, ambiguity, distress, over-trust, anthropomorphism or emotional escalation.

The observed problem is therefore not only "what did the system say?", but also "what kind of response relationship did the system create?"

4. Central Thesis

The central thesis of Pandora Ética / Contextual Response Governance is that safer conversational AI requires contextual governance of the response pathway, not only moderation of the final answer.

A safer system should be evaluated not only by whether its output is allowed, but also by whether the response is appropriate to the user's apparent context, vulnerability, intention, ambiguity and interactional dependency.

This proposal therefore introduces a protected conceptual vocabulary for examining the relationship between user input, contextual risk and response governance.

The public version presents the thesis and its relevance without disclosing the protected operational core.

5. Areas of Relevance

The proposal may be relevant to:

- Human-AI interaction;
- AI governance and accountability;
- AI safety and risk management;
- Human oversight;
- Digital mental health boundaries;
- Educational and advisory AI systems;
- Ethical design of conversational systems;
- Preliminary policy discussion on context-sensitive AI responses.

The proposal is especially relevant where conversational AI systems may be used by people in uncertain, emotionally charged, vulnerable or high-trust contexts.

6. What This Proposal Is Not

This proposal is not presented as an academically validated methodology.

It is not a clinical protocol.

It is not a therapeutic tool.

It is not a diagnostic framework.

It is not an implementation manual.

It is not a technical standard ready for operational deployment.

It is not a transfer of the protected method.

It is not a disclosure of the operational core.

It is an independent conceptual and methodological proposal presented for preliminary observation, discussion and authorship protection.

7. Request for Preliminary Observation

This document is prepared for controlled preliminary observation by appropriate academic, ethical, policy, clinical, technical or governance reviewers.

The purpose of the request is not to transfer the method, but to allow qualified observers to assess the relevance, distinctiveness and potential importance of the proposal at a protected conceptual level.

Preliminary clinical and neuropsychological opinions have been obtained regarding the relevance and sensitivity of parts of the problem area. These opinions are referenced only cautiously here and are not attached to this public version, as they may contain confidential or context-specific material.

Any further review should preserve the protected status of the methodology and should not request disclosure of the operational core unless covered by appropriate confidentiality, authorship and protection conditions.

8. Disclosure Limits

This public protected summary does not include the protected operational core of Pandora Ética / Contextual Response Governance.

It does not include information that would allow reproduction, implementation, transfer or operational use of the methodology.

It does not include identifiable sensitive cases, private conversations, confidential evaluations or full clinical/neuropsychological documents.

It does not include executable procedures.

It is suitable only as a public protected summary for preliminary observation, authorship protection and controlled discussion.

9. Contact

Author and methodological custodian:

Maria Borges

Contact:

metodologiamariaborges@gmail.com