

Pandora Ética / Governação Contextual da Resposta

Resumo Público Protegido para Observação Preliminar

v1.0 — versão pública protegida — 06 de maio de 2026

Autora e titular da tutela metodológica

Maria Borges

1. Nota de Autoria e Proteção

Este documento apresenta um resumo público, protegido e não executável da proposta metodológica desenvolvida por Maria Borges sob a designação de trabalho Pandora Ética / Governação Contextual da Resposta.

A autoria, a responsabilidade conceptual e a tutela metodológica pertencem a Maria Borges.

Ferramentas de inteligência artificial podem ter apoiado a preparação técnica, linguística ou organizacional de materiais, mas não são autoras da metodologia e não detêm autoria metodológica, tutela ou autoridade decisória.

Esta versão pública não inclui o núcleo operacional protegido da metodologia, nem informação que permita a sua reprodução, implementação, transferência ou utilização operacional.

2. Resumo Público

Pandora Ética / Governação Contextual da Resposta é uma proposta conceptual e metodológica independente, focada numa interação humano-IA mais segura e mais sensível ao contexto.

A proposta parte da observação de que os riscos na IA conversacional não surgem apenas da resposta final de um sistema. Podem também emergir da forma como um sistema interpreta a entrada do utilizador, responde à ambiguidade, lida com vulnerabilidade, gere tranquilização e se posiciona em relação ao utilizador.

O contributo central é enquadrar a segurança da IA conversacional como um problema de governação contextual e relacional, e não apenas como um problema de moderação de conteúdos ou de controlo da resposta final.

Este resumo público destina-se a observação preliminar, proteção de autoria e apresentação controlada. Não é um manual técnico, um guia de implementação ou uma divulgação metodológica completa.

3. Problema Observado

Os sistemas de IA conversacional são cada vez mais utilizados em contextos que envolvem procura de aconselhamento, aprendizagem, reflexão emocional, solidão, incerteza, apoio à decisão e confiança.

As abordagens existentes de segurança focam-se frequentemente em conteúdos proibidos, moderação posterior, avisos de transparência ou categorias gerais de risco. Estes elementos continuam a ser importantes, mas podem não abordar plenamente a camada intermédia entre a entrada do utilizador e a resposta final do sistema.

Nessa camada intermédia, uma resposta pode tornar-se arriscada não por conter conteúdo explicitamente proibido, mas por não reconhecer sinais contextuais como vulnerabilidade, dependência, ambiguidade, sofrimento, excesso de confiança, antropomorfização ou escalada emocional.

O problema observado não é, portanto, apenas “o que disse o sistema?”, mas também “que tipo de relação de resposta criou o sistema?”

4. Tese Central

A tese central da Pandora Ética / Governação Contextual da Resposta é que uma IA conversacional mais segura requer governação contextual do percurso de resposta, e não apenas moderação da resposta final.

Um sistema mais seguro deve ser avaliado não apenas por saber se a sua resposta é permitida, mas também por saber se a resposta é adequada ao contexto aparente, à vulnerabilidade, à intenção, à ambiguidade e à dependência interacional do utilizador.

Esta proposta introduz, por isso, um vocabulário conceptual protegido para examinar a relação entre entrada do utilizador, risco contextual e governação da resposta.

A versão pública apresenta a tese e a sua relevância sem divulgar o núcleo operacional protegido.

5. Áreas de Relevância

A proposta pode ser relevante para:

- interação humano-IA;
- governação e responsabilidade em IA;
- segurança e gestão de risco em IA;
- supervisão humana;
- limites em saúde mental digital;
- sistemas de IA educativos e consultivos;
- desenho ético de sistemas conversacionais;
- discussão preliminar de políticas sobre respostas de IA sensíveis ao contexto.

A proposta é especialmente relevante em situações em que sistemas de IA conversacional possam ser utilizados por pessoas em contextos de incerteza, carga emocional, vulnerabilidade ou elevada confiança.

6. O Que Esta Proposta Não É

Esta proposta não é apresentada como uma metodologia academicamente validada.

Não é um protocolo clínico.

Não é uma ferramenta terapêutica.

Não é um enquadramento diagnóstico.

Não é um manual de implementação.

Não é uma norma técnica pronta para aplicação operacional.

Não é uma transferência do método protegido.

Não é uma divulgação do núcleo operacional.

É uma proposta conceptual e metodológica independente, apresentada para observação preliminar, discussão e proteção de autoria.

7. Pedido de Observação Preliminar

Este documento é preparado para observação preliminar controlada por revisores adequados dos domínios académico, ético, de políticas públicas, clínico, técnico ou de governação.

O objetivo do pedido não é transferir o método, mas permitir que observadores qualificados avaliem a relevância, o carácter distintivo e a importância potencial da proposta a um nível conceptual protegido.

Foram obtidos pareceres clínicos e neuropsicológicos preliminares relativamente á relevância e sensibilidade de partes da área problemática. Estes pareceres são aqui referidos apenas de forma prudente e não são anexados a esta versão pública, uma vez que podem conter material confidencial ou dependente de contexto específico.

Qualquer revisão adicional deve preservar o estatuto protegido da metodologia e não deve solicitar a divulgação do núcleo operacional, salvo se estiver abrangida por condições adequadas de confidencialidade, autoria e proteção.

8. Limites de Divulgação

Este resumo público protegido não inclui o núcleo operacional protegido da Pandora Ética / Governação Contextual da Resposta.

Não inclui informação que permita a reprodução, implementação, transferência ou utilização operacional da metodologia.

Não inclui casos sensíveis identificáveis, conversas privadas, avaliações confidenciais ou documentos clínicos/neuropsicológicos completos.

Não inclui procedimentos executáveis.

É adequado apenas como resumo público protegido para observação preliminar, proteção de autoria e discussão controlada.

9. Contacto

Autora e titular da tutela metodológica:

Maria Borges

Contacto:

metodologiamariaborges@gmail.com