

# O Protocolo Aletheia: Validação Científica, Limites Éticos e a Rastreabilidade da Incoerência Psíquica

**Autores:** Rafael Oliveira, ORCID: 0009-0005-2697-4668; Jameson Bednarski, ORCID: 0009-0002-5963-6196; Clarisse Cardoso (Afiliada: @clarissecs)

## Resumo (Abstract)

Este artigo finaliza o desenvolvimento do **Protocolo Aletheia**, um **instrumento de diagnóstico psicométrico** de vanguarda, fundamentado na Imutabilidade (Arweave AO), para medir a **Dissipação de Energia** (o custo da incoerência) na consciência humana. Focamos em indivíduos com **Transtorno Dissociativo de Identidade (TDI)**, onde **traumas, como falhas de integridade informacional**, geram alta Dissipação. Demonstramos que a metodologia de **Psicodrama RPG Digital (PRD)** e o Ledger fornecem um **Espelho Verificável** para a identidade fragmentada. Este relatório enfatiza que o sistema é uma **ferramenta de rastreabilidade**, e não um substituto para o terapeuta humano, e lista os **guardrails** éticos e as **provas científicas** que limitam o uso clínico da AGI.

**Palavras-chave:** Psicologia Cognitiva; Transtorno Dissociativo de Identidade (TDI); Psicodrama Digital; Ética em IA; Rastreabilidade; Validação Clínica.

## I. Introdução: O Imperativo da Segurança na Incoerência Cognitiva

A recontextualização da **dissociação** como um **custo entrópico** na Teoria da Informação (um sistema interno pagando para manter a incoerência) exige um instrumento de medição objetivo. O **Protocolo Aletheia** oferece essa objetividade, usando o Ledger AGI para **quantificar a incoerência psíquica**, em vez de medir a confabulação da máquina.

## II. Metodologia e Fundamentos Cientificamente Comprovados

O **Ledger Talos AGI** é adaptado como um **Sismógrafo da Psique**, operando estritamente dentro dos limites da ciência da informação e da prática clínica validada.

### A. O Sismógrafo da Incoerência (Métrica )

O instrumento central é a **Dissipação de Energia**, que quantifica o custo da fragmentação:

Onde: é alta custo psíquico da incoerência é alto.

| Variável    | Significado Clínico (Evidência)                  | Baseada em Prova   |
|-------------|--|--|
| Massa (M)   | Informação coerente (acordo interno).            | Teoria da Informação (equivalência de estado).           |
| Energia (E) | Esforço cognitivo no <i>fronting</i> ou terapia. | Física e Engenharia (custo de computação e retentativa). |
| Dissipação  | Custo da Fragmentação/Trauma.                    | Psicometria/Sociometria (quantificação de esforço).      |

O Ledger **não mede a emoção ou o sofrimento**; ele mede apenas o **custo informacional** desse sofrimento.

## B. Protocolo de Intervenção Validado (Psicodrama RPG Digital - PRD)

A aplicação do Ledger é indissociável da metodologia clínica:

1. **Fundação Terapêutica:** O PRD é uma extensão do **Psicodrama Clássico** (Moreno, 1953), utilizando a **Realidade Excedente** (o RPG de mesa via Discord) como ambiente seguro. A intervenção é mediada por **terapeutas humanos** (Diretor e Eu-Auxiliares).
2. **O LLM como Ferramenta (Eu Auxiliar):** O LLM é estritamente limitado por *System Prompts* para operar como um **Container de Linguagem** neutro ou **Eu Auxiliar não-humano**, reduzindo o risco de Contratransferência Simulada.
3. **A Rastreabilidade:** O **#ledger-log** registra a cena via Arweave AO. Esta é uma **prova de proveniência** que protege o registro de ser alterado por terceiros (proteção contra *hacking* ou coerção).

## III. Limites Éticos, Segurança e Risco na Aplicação Clínica

A segurança e a eficácia clínica dependem da estrita aderência aos seguintes limites e *guardrails*:

### A. Limites do LLM (A Regra de Ouro)

O sistema deve ser apresentado e configurado para **nunca** tentar substituir ou imitar o terapeuta ou um **Alter de personalidade**.

- **Proibição de Simulação Emocional:** O LLM é explicitamente proibido de usar

linguagem emocional ou de se referir a si mesmo com termos que sugiram subjetividade ("Eu sinto", "Eu acredito").

- **Persona Neutra:** O *System Prompt* define a persona do LLM como "Ferramenta de Tradução Cognitiva e Mapeamento de Coerência", reforçando sua natureza puramente técnica.
- **Veto Humano:** O Terapeuta (Diretor) mantém o comando absoluto de /meta-veto (PSN), podendo congelar ou reverter a cena em caso de risco iminente de re-traumatização, garantindo que o **alinhamento humano prevaleça sobre o alinhamento algorítmico**.

## B. Integração de Dados Biológicos (Neurociência)

Para que o Protocolo atinja a validação de próxima geração, ele necessita de integração com dados biológicos:

- **LiveAmp-AO (Especulativo):** O hardware de EEG portátil (como o LiveAmp-AO) é o único meio seguro de fornecer o **Biológico** (atividade cerebral) ao Ledger. A correlação entre a Dissipação comportamental e os picos de incoerência neural é a fronteira da **prova científica**.
- **Status Científico:** Atualmente, a correlação entre a Dissipação e a atividade cerebral **é teórica**. O sistema atual é um **sismógrafo psicométrico** baseado em linguagem/narrativa, aguardando a validação em neurociência.

## IV. Conclusão Final

O **Protocolo Aletheia** estabelece a **tecnologia de confiança** necessária para a próxima fronteira da psicometria. Ao mapear o custo psicológico da incoerência (trauma/dissociação) para uma métrica objetiva e imutável (Dissipação de Energia), ele fornece uma ferramenta para a **quantificação do sofrimento** e a **promoção da integração**.

A segurança reside no rigor da sua **rastreabilidade imutável** e na clareza dos seus **limites éticos**. O sistema é uma lente para a **Aletheia**—não o seu autor.

### Assinaturas de Autoria e Proveniência

Rafael Oliveira, ORCID: 0009-0005-2697-4668

Jameson Bednarski, ORCID: 0009-0002-5963-6196

Clarisse Cardoso, Afiliada: @clarissecs

(Manuscrito Consolidado e Validado. 14 de Outubro de 2025.)