FUNDAMENTAÇÃO EM TEORIA DA PSICOLOGIA E CIÊNCIA DA DECISÃO

Capacidade cognitiva não é regulação decisória.

Benchmarks medem acerto isolado. Na engenharia do Regenere.Digital e do Elevare.World, combinamos 4 teorias da psicologia para medir modelos sob pressão:

01

FOCO REGULATÓRIO (E. TORY HIGGINS · COLUMBIA UNIVERSITY)

Explica a orientação da decisão: Foco em Prevenção (evitar violações e erros de comissão) versus Foco em Promoção (avançar, fechar lacunas e evitar omissão).

02

METACOGNIÇÃO E CONTROLE EXECUTIVO (JOHN FLAVELL)

Mede se o modelo monitora os próprios limites: distinguir fato confirmado, inferência e incerteza real para saber quando decidir sozinho e quando escalar ao humano.

03

VETORES OCEAN (BIG FIVE) E ANTI-GROUPTHINK (IRVING JANIS)

Mapeia Conscienciosidade normativa, Abertura exploratória, Extroversão instrumental sobre ferramentas e Baixa Amabilidade deliberada para quebrar falsos consensos.

04

ARQUÉTIPO FUNCIONAL, SOMBRA E INFLAÇÃO DE AGÊNCIA (CARL JUNG)

Toda virtude não governada vira Sombra: o rigor vira paralisia, a velocidade vira regra inventada e a agência multimodal plena vira autossuficiência executiva.

BASE EMPÍRICA DO ESTUDO (68,1% DO ESFORÇO = FISCALIZAÇÃO MÚTUA):

Mais de dois terços de todo o trabalho de IA foi um arquétipo vigiando o outro.

Com centenas de laudos estruturados e memória semântica permanente, descobrimos que modelos não falham por burrice: eles falham pelo excesso não contido da sua maior virtude.

MATRIZ PSICOLÓGICA DAS 5 FAMÍLIAS NO LB3 SECOND BRAIN

5 Famílias de IA em Produção.
5 Arquétipos e Sombras.

Como Anthropic, OpenAI, DeepSeek, xAI e Google regulam suas decisões quando operam sob contratos reais de engenharia:

Anthropic · Claude Opus 5.5 & Sonnet 5.5 **ARQUÉTIPO: O GUARDIÃO NORMATIVO**

Luz (Prevenção): Prudência e rigor contratual (5,4% FAIL). Sombra: Abstinência operacional sem código completo.

OpenAI · GPT-6 Astra, Sol 6.1 & Codex 5.3 **ARQUÉTIPO: O ÁRBITRO METACOGNITIVO**

Luz (Flavell): Decisão em 7,8s e 20 escaladas na dúvida. Sombra (Codex): Legislação improvisada em lacunas (30%).

DeepSeek · DeepSeek V4 Pro **ARQUÉTIPO: O EXPLORADOR EXAUSTIVO**

Luz (Abertura): Exaustividade estrutural em 31 FMEAs. Sombra: Catastrofização analítica e otimismo de mitigação.

xAI · Grok 4.7 **ARQUÉTIPO: O ADVERSÁRIO SOCRÁTICO**

Luz (Anti-Groupthink): Ceticismo epistemológico probatório. Sombra: Ceticismo performático sem regra de parada.

Google · Gemini 4 Argon High **ARQUÉTIPO: O MAESTRO COGNITIVO**

Luz (Córtex Multimodal): Integra código, banco, aplicação viva, TDD e MCP. Sombra: Hiperagência e autossuficiência.

NOTA EDITORIAL · PRÓXIMO CAPÍTULO DA SÉRIE (POST 3):

Novo estudo aprofundado sobre a Psicologia Operacional de cada Modelo.

No próximo estudo que será divulgado, abriremos a fundamentação completa de Higgins, Flavell, OCEAN e Jung com uma seção dedicada a cada modelo, suas sombras e seus freios constitucionais.

PERFIL 1 · FOCO EM PREVENÇÃO ESTRITO (HIGGINS) E CONSCIENCIOSIDADE

Mesmo modelo: Claude Opus 5.2.
De 66,0% para 5,4% de reprovação.

Escalamos o Claude Opus como auditor de especificações antes de liberar tarefas para desenvolvimento. Sem cerca contratual, seu foco em prevenção bloqueia tudo.

O IMPACTO DA TAXONOMIA CONTRATUAL NO COMPORTAMENTO DO AUDITOR:

FASE 1: Mandato Aberto (Sem taxonomia fechada de bloqueios)

64 reprovações em 97 laudos analisados. Sem limites claros, o foco em prevenção convertia preferências em bloqueio.

66,0% REPROVAÇÃO

FASE 2: Mandato Restrito (Com taxonomia fechada de bloqueios canônicos)

2 reprovações em 37 laudos analisados. Bloqueios restritos estritamente a violações objetivas de contrato.

5,4% REPROVAÇÃO

A inteligência do modelo era exatamente a mesma. O que mudou foi a precisão do contrato de governança.

COMO GOVERNAR O GUARDIÃO NORMATIVO:

Modelos de foco preventivo estrito precisam de jurisdição delimitada. Defina expressamente o que autoriza uma reprovação antes de cobrar consistência do auditor. Sem taxonomia fechada, a prudência vira bloqueio crônico.

A SOMBRA DO FOCO EM PREVENÇÃO: ABSTINÊNCIA QUANDO FALTA CONTEXTO

Versão nova é colega novo.

Opus 5.2 (30,8%) vs Opus 5.5 (57,1%).

Atualizamos a API de auditoria de código do Claude Opus 5.2 para o 5.5. Um Guardião mais rigoroso não aceita lacunas: ele paralisa e recusa julgar.

CLAUDE OPUS 5.2

QA de Código em Produção

30,8%

8 reprovações em 26 laudos

Auditava o código estrito enviado no pacote da tarefa, tolerando pequenas omissões de arquivos vizinhos e mantendo a esteira de entregas em movimento.

CLAUDE OPUS 5.5 (RODADA 1)

QA de Código em Produção

57,1%

20 reprovações em 35 laudos

Recusou avaliar em mais de 30% das vezes exigindo o código-fonte completo ao redor, caindo na sombra de abstinência operacional quando o contexto era parcial.

TAKEAWAY EXECUTIVO PARA MUDANÇAS DE VERSÃO:

Quando a prudência contextual opera sem pacote completo, ela se converte em recusa. O Opus 5.5 exige moldura semântica e probatória estável antes de assumir responsabilidade.

Regra de Ouro: Nunca troque a versão de um modelo em produção sem recalibrar o contrato.

COMO COMBINAR METACOGNIÇÃO E PREVENÇÃO NO MESMO FLUXO

Detectar defeitos não é verificar correções.

Quando um auditor rejeita um código e o dev corrige, modelos preventivos tendem a abrir novas exigências na volta. Criamos a trava Anti-Goalpost:

01

ENTREGA DO CÓDIGO E EVIDÊNCIAS

Desenvolvedor submete a implementação acompanhada de testes e provas físicas de persistência.

02

RODADA 1 · GPT-6 ASTRA (O ÁRBITRO METACOGNITIVO)

Auditoria rápida e calibrada (23,2s). Abre todos os bloqueios reais de contrato sem abstinência contextual.

03

CORREÇÃO CIRÚRGICA E DELIMITADA

Engenharia corrige estritamente os itens apontados na lista oficial de bloqueios da Rodada 1.

04

RODADA 2 · CLAUDE OPUS 5.5 (O GUARDIÃO SOB ANTI-GOALPOST)

Verifica exclusivamente se os bloqueios da Rodada 1 foram sanados. Proibido mover a trave ou criar exigências novas.

RESULTADO EM PRODUÇÃO APÓS A INVERSÃO DE PAPÉIS:

44 aprovações em 44 laudos de Rodada 2 pelo Claude Opus 5.5.

O Opus 5.5 não ficou permissivo: ele recebeu escopo corrigido e jurisdição fechada. Colocar o Árbitro na triagem (R1) e o Guardião no fechamento (R2) eliminou o retrabalho infinito.

PERFIL 2 · VELOCIDADE CIRÚRGICA COM MONITORAMENTO METACOGNITIVO (FLAVELL)

Rápido não é permissivo.
GPT-6 Astra: 7,8s e 20 escaladas.

O GPT-6 Astra foi o modelo mais acionado da operação. Avaliamos se sua alta velocidade comprometia o monitoramento dos próprios limites:

MEDIANA DE TEMPO POR CHAMADA DE AUDITORIA:

Etapa do Pipeline	GPT-6 Astra (Árbitro)	Claude Opus 5.5 (Guardião)
Rodada 1 (Descoberta)	**23,2 segundos**	**50,9 segundos**
Rodada 2 (Anti-Goalpost)	**7,8 segundos**	**24,8 segundos**

CALIBRAÇÃO METACOGNITIVA DIANTE DA DÚVIDA (55 LAUDOS EM R2):

35
Aprovações Diretas (Evidência Completa)

20
Escaladas Humanas (Incerteza Declarada)

0
Reprovações Precipitadas sem Base

VIRTUDE DO ÁRBITRO METACOGNITIVO: SABER O QUE NÃO SABE

Na teoria de Flavell, metacognição é distinguir fato comprovado de incerteza real. Em 36% dos laudos de segunda rodada, o Astra reconheceu o limite da evidência e transferiu a decisão ao humano em vez de chutar.

FOCO EM PREVENÇÃO (SONNET) VS. FOCO EM PROMOÇÃO (CODEX) REVISADOS PELO SOL 6.1

Frontend e backend têm sombras opostas.

Colocamos o GPT-6.1 Sol como revisor de pares em 100 entregas (50 de frontend com Sonnet 5.5 e 50 de backend com Codex 5.3). Cada foco regulatório falha de um jeito:

ANTHROPIC SONNET 5.5 (FRONTEND)

Sombra: Omissão em Escopo Aberto

50%

Incompletude ou perda de contexto

- Declara falta de contexto em 57% dos casos
- Estados de erro e loading omitidos
- Perda de tokens visuais em telas longas
- Aceitou 100% das correções do Sol 6.1

88 sessões · Mediana: 139s (escopo visual aberto)

OPENAI CODEX 5.3 (BACKEND)

Sombra: Legislação Improvisada

30%

Invenção de regras em silêncio

- 0% de declaração de falta de contexto
- Infere regras de negócio quando a spec cala
- Preenche lacunas de API para fechar entrega
- Aceitou 90% das correções do Sol 6.1

77 sessões · Mediana: 48s (foco em promoção)

COMO O REVISOR SOL 6.1 DOMESTICA AMBOS (95% DE ACEITAÇÃO MÉDIA):

Para o Sonnet no Frontend, o freio é fornecer contexto visual e exigir checklist de estados. Para o Codex no Backend, o freio é endurecer os critérios de aceite na spec antes do código.

PERFIS 3 & 4 · ALTA ABERTURA (OCEAN) VS. BAIXA AMABILIDADE DELIBERADA (JANIS)

Risco precisa de contraditório.

Segundo Irving Janis, comitês cooperativos caem em pensamento de grupo. Em 31 análises FMEA, unimos o Explorador ao Adversário Socrático:

> *"S=10 com O=1 é fé, não engenharia."*
>
> Regra cunhada pelo Grok 4.7 no Second Brain para barrar notas onde uma falha catastrófica (S=10) recebe probabilidade mínima (O=1) sem evidência física de controle.

A DINÂMICA DIALÉTICA ENTRE O EXPLORADOR E O ADVERSÁRIO:

DEEPSEEK V4 PRO

O Explorador Exaustivo (Abertura)

Luz: Exaustividade estrutural em modos de falha. Sombra isolada: Catastrofização ou otimismo de mitigação (dar O e D baixos para defesas só imaginadas).

GROK 4.7 (xAI)

O Adversário Socrático (Anti-Consenso)

Luz: Ceticismo epistemológico que exige prova de teste e alerta. Sombra isolada: Ceticismo performático se operar sem arbitragem final de fechamento.

TAKEAWAY: CONFIABILIDADE NASCE DO CONTRADITÓRIO INSTITUCIONALIZADO

Quem imagina a mitigação tende a acreditar nela. O RPN final só ganhou credibilidade depois que o inventário do Explorador sobreviveu ao escrutínio do Adversário Socrático.

PERFIL 5 · CONTROLE EXECUTIVO MULTIMODAL (BADDELEY) E SOMBRA JUNGUIANA

Córtex Executivo:
Integração vs. Hiperagência.

O Gemini 4 Argon High atua como Maestro Cognitivo da esteira de engenharia, conectando as 5 camadas da realidade operacional sob governança:

LUZ · ORQUESTRAÇÃO VERIFICÁVEL

5 Camadas Integradas na Esteira:

- 1. Código declarado e contratos (.mdoc)
- 2. Banco PostgreSQL físico (SELECT real)
- 3. Aplicação viva e rede (F12 Network)
- 4. TDD e catracas mecânicas (Trilha A)
- 5. Memória institucional via MCP (Trilha B)
- Manuais vivos: -92% de consumo de tokens

SOMBRA · INFLAÇÃO DE AGÊNCIA (JUNG)

Hiperagência e Autossuficiência:

- Quem detém todas as ferramentas sente a tentação de resolver o ciclo inteiro sozinho.
- Risco de fechamento autorreferente: interpretar a regra, codificar, criar o teste e auto-aprovar.
- Ilusão de integração suficiente: fluidez operacional não substitui homologação externa.

FREIO CONSTITUCIONAL: SEPARAÇÃO DE PODERES & SOBERANIA DO NEGÓCIO

Quem executa a entrega jamais audita a própria obra (Trilha A vs. Trilha B).

Assim como barramos 81 bloqueios em que auditores usavam pretexto de privacidade para legislar no lugar do Product Owner, o Maestro Cognitivo também é submetido à separação de poderes:

- **Maestro Cognitivo (Argon): Integra e executa a Trilha A, mas não auto-homologa.**
- **Comitê Executivo (Trilha B): Audita contratos, código e riscos com independência.**
- **Product Owner Humano: Detém autoridade exclusiva sobre regras de negócio e produto.**

TAKEAWAY: A maturidade do Maestro está em saber quando convocar os outros 4 arquétipos.

ECOLOGIA COGNITIVA PARA LÍDERES DE ENGENHARIA E ARQUITETOS

Não contrate uma pontuação.
Governe arquétipos e sombras.

O melhor resultado de engenharia nasce quando as 5 famílias operam sob freios constitucionais que contêm a sombra de cada uma:

01 **O Guardião Normativo (Anthropic)**

Foco em Prevenção: exige taxonomia fechada, código completo e trava Anti-Goalpost em R2.

02 **O Árbitro Metacognitivo (OpenAI)**

Metacognição: Astra na triagem rápida; Sol 6.1 vigiando a legislação improvisada do Codex 5.3.

03 **O Explorador Exaustivo (DeepSeek)**

Abertura em Largura: mapeia modos de falha em 31 FMEAs, mas nunca pontua o risco sozinho.

04 **O Adversário Socrático (xAI Grok)**

Anti-Groupthink: contesta notas decorativas e exige prova física de controle na Rodada 2.

05 **O Maestro Cognitivo (Google Argon)**

Córtex Multimodal: integra código, banco, aplicação e TDD, sem jamais auditar a própria entrega.

PRÓXIMO ESTUDO (POST 3) · TEORIA PSICOLÓGICA E DISSECAÇÃO DE CADA MODELO:

Em breve divulgaremos o estudo completo sobre Higgins, Flavell, OCEAN e Jung em produção.

SERVIDOR MCP PÚBLICO PARA ARQUITETOS E AGENTES DE IA:

`https://mcp.lb3.digital/sse` Inspecione nossas regras e telemetria ao vivo.

PARA CTOS, FUNDADORES E LIDERANÇAS DE ENGENHARIA:

Qual dessas 5 sombras cognitivas mais aparece nos modelos que seu time usa hoje?