CONTROLE DA BASE REPRESENTACIONAL E SOBERANIA COGNITIVA NA CIÊNCIA ASSISTIDA POR INTELIGÊNCIA ARTIFICIAL: UMA ANALOGIA COM A ANÁLISE DE FOURIER

CONTROL OF THE REPRESENTATIONAL BASIS AND COGNITIVE SOVEREIGNTY IN AI-ASSISTED SCIENCE: AN ANALOGY BASED ON FOURIER ANALYSIS

Jandislei Antonio Genova

Advogado e pesquisador independente. Especialista em Direito Civil e em Gestão em Saúde. São Paulo, Brasil.

Versão de 4 de agosto de 2026 · Artigo acadêmico integral revisado · https://jandisleigenova.com

O poder algorítmico não reside apenas em influenciar respostas, mas em estruturar as bases nas quais perguntas, hipóteses e decisões se tornam possíveis.

RESUMO

Este artigo examina como sistemas de inteligência artificial empregados na descoberta científica podem influenciar a seleção das informações e as bases representacionais a partir das quais problemas e hipóteses são formulados. O estudo contrapõe o *position paper* "LLMs can't jump" aos materiais públicos divulgados pela OpenAI sobre resultados matemáticos atribuídos ao sistema Astra. Adota-se ensaio teórico-jurídico documental, com corpus encerrado em 4 de agosto de 2026 e hierarquizado entre alegações públicas, artefatos técnicos e validação comunitária. A análise de Fourier é empregada como analogia heurística para distinguir o controle dos coeficientes — saliência, recorrência, ordem e temporalidade — do controle da base representacional — categorias, ontologias e espaços de hipótese. Os materiais examinados permitem identificar artefatos formalmente inspecionáveis, mas não autorizam conclusão definitiva sobre novidade, autonomia abdutiva ou transformação conceitual. Sustenta-se que a concentração de dados, representações, mecanismos de geração, validação e circulação pode criar risco de dependência epistêmica. Propõe-se, *de lege ferenda*, governança graduada baseada em proveniência, contestabilidade, validação independente e capacidade pública de auditoria.

Palavras-chave: inteligência artificial; soberania cognitiva; raciocínio abdutivo; análise de Fourier; governança epistêmica.

ABSTRACT

This article examines how artificial intelligence systems used in scientific discovery may influence both the selection of information and the representational bases from which problems and hypotheses are formulated. It contrasts the *position paper* “LLMs can’t jump” with public materials released by OpenAI concerning mathematical results attributed to Astra. The study adopts a theoretical, legal, and documentary approach based on a corpus closed on 4 August 2026 and structured into three evidentiary levels: public claims, technical artifacts, and community validation. Fourier analysis is used as a heuristic analogy to distinguish control of coefficients — salience, recurrence, ordering, and timing — from control of the representational basis — categories, ontologies, and hypothesis spaces. The materials provide formally inspectable artifacts but do not support definitive conclusions about novelty, abductive autonomy, or conceptual transformation. The article argues that the concentration of data, representations, generation mechanisms, validation, and dissemination may create a risk of epistemic dependency. It proposes, *de lege ferenda*, a graduated governance framework grounded in provenance, contestability, independent validation, and public audit capacity.

Keywords: artificial intelligence; cognitive sovereignty; abductive reasoning; Fourier analysis; epistemic governance.

1 INTRODUÇÃO

A inteligência artificial deixou de ocupar apenas funções auxiliares de cálculo, busca e classificação para participar de etapas centrais da produção científica. Sistemas contemporâneos geram conjecturas, propõem programas, exploram espaços de prova, interagem com verificadores formais e produzem manuscritos completos. AlphaProof alcançou desempenho equivalente ao de medalhista de prata da Olimpíada Internacional de Matemática de 2024 em ambiente formal; AlphaEvolve combina modelos de linguagem, busca evolucionária e avaliadores automáticos para projetar e otimizar algoritmos; e o projeto AI Scientist automatiza a formulação de ideias, a execução de experimentos e a redação de artigos em domínios de aprendizado de máquina (Hubert et al., 2026; Novikov et al., 2025; Lu et al., 2024).

Em abril de 2026, Tom Zahavy apresentou o *position paper* “LLMs can’t jump”. O argumento distingue indução, dedução e abdução e sustenta que grandes modelos de linguagem são fortes na compressão de regularidades e avançam na dedução formal, mas não dispõem do enraizamento sensorial necessário ao salto abdutivo que formula premissas científicas radicalmente novas. A proposição identifica um possível gargalo; não constitui teorema de impossibilidade nem posição institucional automaticamente imputável à Google DeepMind (Zahavy, 2026).

Em 1º de agosto de 2026, a OpenAI atribuiu a uma versão interna de seu próximo modelo, denominada Astra, a obtenção de dez resultados em matemática e ciência da computação teórica. A empresa divulgou uma coleção técnica de 249 páginas e um repositório com formalizações em Lean. A formulação institucional é graduada: os resultados resolveriam ou fariam progresso substancial em problemas abertos, e os manuscritos teriam sido preparados por pessoas com auxílio do mesmo sistema após a geração dos argumentos pelo modelo (OpenAI, 2026a; OpenAI, 2026b; OpenAI, 2026c).

A proximidade temporal entre a tese de incapacidade abdutiva e a divulgação do Astra cria uma controvérsia intelectualmente fértil, mas não autoriza conclusão binária. Resolver um problema formulado dentro de uma tradição não equivale a criar essa tradição; produzir uma construção inédita pode configurar novidade objetiva ou criatividade exploratória sem transformar as categorias fundamentais do campo. Além disso, a documentação pública não permite reconstruir integralmente arquitetura, treinamento, seleção dos problemas, ferramentas, tentativas descartadas e intervenções humanas. O caso deve ser tratado como alegação corporativa acompanhada de artefatos tecnicamente inspecionáveis, ainda dependente de validação comunitária.

O problema principal é: como o controle privado das representações, dos mecanismos de geração e dos processos de validação empregados pela inteligência artificial pode afetar a soberania cognitiva e científica? Duas questões subordinadas orientam a análise: em que medida os resultados atribuídos ao Astra tensionam a distinção entre indução, dedução e abdução; e quais salvaguardas jurídicas são proporcionais quando sistemas proprietários participam materialmente da descoberta científica?

A hipótese possui dimensão descritiva e normativa. Descritivamente, os materiais do Astra tornam menos segura uma leitura categórica segundo a qual sistemas atuais seriam incapazes de participar da geração de hipóteses ou construções não triviais em domínios estruturados; eles não demonstram, porém, novidade científica, autonomia abdutiva ou criação de uma base conceitual. Normativamente, a concentração vertical de dados, seleção de problemas, representações, geração, validação e circulação pode criar risco de dependência epistêmica mesmo quando cada artefato isolado é formalmente verificável.

A contribuição original consiste em distinguir controle dos coeficientes e controle da base. O primeiro altera saliência, recorrência, sequência e temporalidade dentro de um conjunto informacional disponível. O segundo participa da definição de categorias, ontologias, variáveis e espaços de hipótese nos quais perguntas e alternativas podem existir. A análise de Fourier é utilizada apenas como analogia heurística e taxonômica: não atribui periodicidade, linearidade ou decomponibilidade matemática à cognição, nem propõe uma lei matemática da decisão humana.

O artigo está organizado em oito etapas: método e corpus; alcance da analogia com Fourier; indução, dedução e abdução; análise crítica do caso Astra e da verificação formal; distinção entre coeficientes e base; construção da soberania cognitiva; fundamentos jurídicos; e proposta graduada de governança, incluindo um Relatório de Impacto Epistêmico.

2 METODOLOGIA, CORPUS E NÍVEIS DE EVIDÊNCIA

2.1 Desenho da pesquisa e corpus documental

A pesquisa assume a forma de ensaio teórico-jurídico documental, de natureza qualitativa. Não se pretende oferecer revisão sistemática exaustiva, medir empiricamente dependência epistêmica nem auditar as dez demonstrações matemáticas, tarefa que demandaria especialistas distintos. O objetivo é examinar a plausibilidade conceitual da hipótese, reconstruir mediações técnicas relevantes e avaliar consequências jurídicas do controle da infraestrutura epistêmica.

O corpus foi encerrado em 4 de agosto de 2026 e organizado em quatro conjuntos. O primeiro reúne obras matemáticas sobre séries, transformadas, bases e filtros. O segundo contém literatura de filosofia da ciência, criatividade, classificação e filosofia da informação. O terceiro reúne fontes técnicas primárias: o *position paper* de Zahavy, a publicação institucional da OpenAI, a coleção técnica dos dez resultados, o repositório Lean, o artigo do AlphaProof, o estudo técnico do AlphaEvolve, o AI Scientist e a Declaração de Leiden. O quarto contém fontes jurídicas e de integridade científica: Constituição brasileira, Emenda Constitucional nº 85/2015, Marco Legal de Ciência, Tecnologia e Inovação, LGPD, Pacto Internacional sobre Direitos Econômicos, Sociais e Culturais, Comentário Geral nº 25, Recomendação da UNESCO sobre Ciência Aberta, Regulamento Europeu de Inteligência Artificial, orientações e Código de Prática de IA de finalidade geral, COPE e CRediT. Para comparar o RIE com instrumentos existentes, também foram examinados DPIA, FRIA, *model cards* e *datasheets*.

As fontes recentes foram hierarquizadas. Comunicados corporativos são tratados como evidência de que determinada alegação foi publicamente formulada, não como prova autônoma de sua verdade científica. Manuscritos, códigos e certificados formais elevam a verificabilidade do objeto, mas não substituem revisão por pares, avaliação de anterioridade, importância disciplinar ou interpretação histórica. A validação comunitária independente constitui nível probatório distinto.

2.2 Categorias analíticas e teste da hipótese

O material foi lido mediante matriz analítica manual composta por seis categorias correspondentes às camadas da infraestrutura epistêmica: dados e proveniência; seleção do problema; base representacional; geração e busca; validação; e circulação e acesso. Em cada camada, formularam-se quatro perguntas: quem escolhe; quais exclusões podem ser produzidas; quais artefatos permitem reconstrução; e quais contrapoderes externos podem

contestar o processo. A matriz organiza a análise, mas não constitui protocolo estatístico nem instrumento empírico validado.

Quadro 1 — Matriz de evidência e função analítica

Nível	Objeto	O que permite afirmar	O que não permite afirmar
E1 — alegação pública	Comunicado, página institucional ou declaração do fornecedor.	Que a organização reivindicou determinado resultado, custo, processo ou capacidade.	Correção científica, originalidade, autonomia do modelo ou consenso disciplinar.
E2 — artefato técnico	Manuscrito, código, certificado formal, conjunto de dados ou protocolo.	Que existem materiais suscetíveis de inspeção, compilação ou reprodução parcial, conforme o tipo de artefato.	Novidade científica, anterioridade, autonomia do modelo, adequação semântica integral ou independência da validação.
E3 — validação comunitária	Revisão por pares, replicação, crítica especializada e integração à literatura.	Que a comunidade examinou correção, anterioridade, relevância, inteligibilidade e integração à literatura.	Ausência absoluta de erro ou encerramento definitivo da controvérsia.

Fonte: elaboração do autor.

Como o desenho é conceitual, o artigo não afirma testar empiricamente a hipótese. A conclusão descritiva deverá ser revista se os artefatos não corresponderem aos enunciados divulgados, se a participação humana tiver sido omitida de forma material ou se a avaliação comunitária demonstrar reprodução de resultados anteriores. A plausibilidade normativa será menor caso as camadas se mostrem amplamente auditáveis, plurais e acessíveis, sem dependência relevante de fornecedor ou de base representacional. Como a recepção externa estava apenas começando no fechamento do corpus, as conclusões sobre Astra permanecem estritamente provisórias.

2.3 Definições operacionais e limites

Soberania cognitiva significa, neste artigo, capacidade individual, institucional e coletiva de participar da formação e da contestação das categorias, problemas, alternativas e critérios de validação que estruturam decisões e conhecimentos. Não significa isolamento mental, autarquia tecnológica ou ausência de influência. Controle da base significa poder material de selecionar ou impor representações que delimitam o espaço de perguntas e hipóteses. Governança epistêmica designa salvaguardas aplicáveis às mediações por meio das quais conhecimento é gerado, verificado e distribuído.

O artigo não atribui consciência, intenção ou personalidade jurídica ao sistema. Também não presume que toda infraestrutura privada seja ilegítima ou que abertura integral seja sempre compatível com segurança, propriedade intelectual e segredo comercial. O exame jurídico separa direito vigente, interpretação possível e proposta *de lege ferenda*.

3 ANÁLISE DE FOURIER: REPRESENTAÇÃO, COEFICIENTES E FILTROS

3.1 Decomposição harmônica e escolha da base

A teoria de Fourier surgiu da investigação da propagação do calor e da possibilidade de representar funções periódicas por combinações de senos e cossenos. Seja f uma função periódica de período T e frequência angular fundamental $\omega_0 = 2\pi/T$. Sob condições usuais de integrabilidade e convergência, sua série pode ser escrita como (Fourier, 1822; Oppenheim; Willsky; Nawab, 1997, cap. 3):

$$f(t) = a_0/2 + \Sigma_{n=1}\infty\, [a_n \cos(n\omega_0 t) + b_n \sin(n\omega_0 t)]$$

$$a_n = (2/T)\int_0^T f(t)\cos(n\omega_0 t)dt \quad \text{e} \quad b_n = (2/T)\int_0^T f(t)\sin(n\omega_0 t)dt$$

Os coeficientes a_n e b_n registram a contribuição de cada componente à reconstrução e dependem da base adotada. A mesma série pode ser expressa em termos de amplitude e fase. A transformada de Fourier permite trabalhar com espectros contínuos e sinais não periódicos, enquanto filtros reforçam, atenuam ou removem faixas determinadas. O resultado observado depende do sinal, da representação e do operador aplicado (Bracewell, 2000).

A noção de base utilizada adiante não é propriedade exclusiva da análise de Fourier. Em álgebra linear e análise funcional, uma base fornece elementos a partir dos quais objetos de um espaço são representados por coordenadas. Uma mudança de base altera os coeficientes sem necessariamente alterar o objeto matemático; em contextos sociais e institucionais, contudo, categorias podem também constituir o objeto administrado. O emprego jurídico e epistemológico é, por isso, analógico e delimitado: “base representacional” designa o conjunto de categorias, variáveis e dispositivos que tornam certas formulações possíveis e outras menos visíveis.

3.2 Correspondência analógica com a formação decisória

A decisão humana não é função periódica, e a mente não pode ser reduzida a receptor passivo. A analogia permite apenas decompor dimensões distintas da mediação informacional. Amplitude aproxima-se de saliência; frequência, de recorrência; fase, de posição temporal relativa; filtro, de seleção e exclusão; e base, do conjunto de categorias em que o problema é formulado. A literatura sobre exposição repetida, enquadramento e efeitos de ordem fornece sustentação psicológica independente para esses fenômenos, sem validar uma equivalência matemática entre mente e sinal (Zajonc, 1968; Tversky; Kahneman, 1981).

Uma interface pode tornar uma alternativa mais saliente por posição, contraste, recomendação e prova social. A repetição aumenta familiaridade e disponibilidade, embora seus efeitos dependam do contexto. A ordem altera o significado relativo de informações apresentadas antes ou depois de alertas, perdas e momentos de fadiga. A arquitetura de busca

ou recomendação atua como filtro quando define quais conteúdos ingressam no campo visível. Esses mecanismos se aproximam do *hypernudge*, da regulação por design e da manipulação digital não coercitiva (Yeung, 2017; Susser; Roessler; Nissenbaum, 2019).

A teoria da influência cognitiva estrutural descreve esse primeiro nível: a decisão continua formalmente atribuída ao sujeito, mas as condições informacionais anteriores à manifestação da vontade são organizadas de modo assimétrico. A decisão não começa no clique; começa na seleção dos elementos que poderão participar de percepção, avaliação e escolha (Genova, 2026, p. 41-58).

Quadro 2 — Correspondência analógica controlada entre Fourier e decisão

Elemento	Correspondência heurística	Risco jurídico-institucional	Limite da analogia
Amplitude	Saliência e intensidade perceptiva.	Superdimensionamento de uma opção ou informação.	Não mede força causal nem adesão subjetiva.
Frequência	Recorrência e disponibilidade cognitiva.	Normalização artificial, urgência ou aparência de consenso.	Repetição pode gerar familiaridade, rejeição ou nenhum efeito.
Fase	Posição temporal relativa e sequência de exposição.	Exploração de fadiga, perda, ansiedade ou vulnerabilidade temporal.	Não equivale tecnicamente à ordem narrativa.
Filtro	Seleção, atenuação e exclusão de conteúdos.	Ocultação de alternativas e assimetria informacional.	Toda comunicação seleciona; a relevância depende da assimetria e do impacto.
Base	Categorias e representações disponíveis.	Controle constitutivo do espaço de escolhas, hipóteses e perguntas.	É analogia proveniente de conceito matemático mais geral.

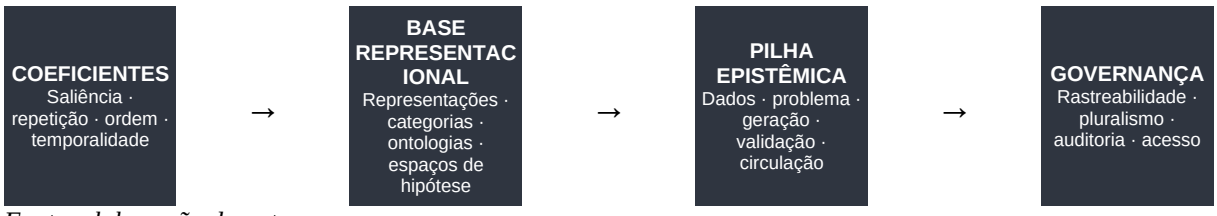
Fonte: elaboração do autor.

3.3 Limites e utilidade da analogia

Fenômenos cognitivos são não lineares, reflexivos, históricos e socialmente situados. Pessoas modificam objetivos, aprendem, resistem, reinterpretam sinais e contestam arquiteturas. Não existe correspondência mensurável automática entre coeficientes matemáticos e influência cognitiva. Por isso, Fourier não opera como lei da decisão, modelo causal ou prova de manipulação.

A utilidade é taxonômica. A analogia permite separar modulação de componentes e constituição do espaço representacional. Se o argumento pudesse ser mantido sem essa distinção, Fourier seria mero ornamento. Aqui, porém, a passagem dos coeficientes à base marca a diferença entre influenciar pesos dentro de um conjunto e participar da definição do próprio conjunto. Essa função será testada nas seções seguintes.

Figura 1 — Do controle dos coeficientes à governança da infraestrutura epistêmica



Fonte: elaboração do autor.
Descrição da figura: fluxo em quatro etapas, do controle dos coeficientes à base representacional, à pilha epistêmica e à governança.

4 INDUÇÃO, DEDUÇÃO E ABDUÇÃO NA DESCOBERTA CIENTÍFICA

4.1 Três modos de inferência

A dedução deriva consequências de premissas aceitas; se as premissas forem verdadeiras e a forma válida, a conclusão segue necessariamente. A indução generaliza padrões observados e produz conclusões ampliativas sujeitas a revisão. A abdução propõe hipótese capaz de explicar fato surpreendente. Em Peirce, inaugura a investigação ao sugerir explicação possível; em Harman, aproxima-se da inferência para a melhor explicação; em Magnani, inclui raciocínio model-based e dimensões eco-cognitivas (Peirce, 1931-1958, v. 5, §§ 171-189; Harman, 1965; Magnani, 2009).

Essas tradições não são intercambiáveis. Geração de hipótese, seleção da melhor explicação e transformação de uma representação respondem a problemas distintos. Uma descoberta real pode combinar ciclos de abdução, dedução e indução: formula-se hipótese, deduzem-se consequências e confrontam-se previsões com observações. Verificadores formais fortalecem a justificação, mas não explicam por si como uma hipótese se tornou concebível (Reichenbach, 1938).

4.2 O salto de Zahavy e o enraizamento simbólico

Zahavy recupera o esquema epistemológico desenhado por Einstein em carta a Maurice Solovine, de 7 de maio de 1952: a passagem das experiências aos axiomas não seria logicamente dedutível, mas dependeria de conexão intuitiva e revisável. O *position paper* sustenta que modelos de linguagem comprimem regularidades textuais e operam sobre premissas disponíveis, porém não possuem *grounding* sensório-motor suficiente para inventar princípios fundamentais que reorganizam o espaço conceitual (Einstein, 1987, p. 123-124; Zahavy, 2026).

A tese dialoga com o problema do enraizamento simbólico: relações entre símbolos não bastariam para constituir significado; seria necessário conectá-los ao mundo por capacidades perceptivas e práticas (Harnad, 1990). O argumento é relevante, mas a expressão “estruturalmente incapaz” excede o material apresentado. Arquiteturas podem integrar linguagem, visão, simulação, robótica, memória, experimentação e verificadores. Uma limitação observada em modelos atuais não se transforma automaticamente em impossibilidade universal.

4.3 Critérios para distinguir novidade, abdução e transformação

A oposição entre repetição e genialidade é insuficiente. Boden distingue criatividade combinacional, exploratória e transformacional: combinar elementos conhecidos; explorar de forma nova um espaço de possibilidades; ou alterar as regras que definem o espaço. Kuhn diferencia solução de quebra-cabeças dentro de matriz disciplinar e mudança de paradigma (Boden, 2004; Kuhn, 2012).

Quadro 3 — Critérios analíticos para avaliar o “salto”

Categoria	Indicadores mínimos	Evidência necessária	Conclusão admissível
Novidade de resultado	Enunciado ou solução não identificado na literatura anterior.	Busca de anterioridade e avaliação especializada.	Resultado objetivamente novo.
Criatividade exploratória	Método, construção ou combinação não explicitamente disponível no corpus conhecido.	Reconstrução do processo, tentativas e dependência de avaliadores.	Exploração original de espaço conceitual herdado.
Abdução limitada	Geração autônoma de hipótese explicativa ou representação útil e testável.	Proveniência, autonomia relativa, relação entre hipótese e verificação.	Capacidade funcional de formular hipóteses.
Transformação conceitual	Criação ou alteração de categorias, princípios e problemas legítimos do campo.	Recepção histórica, mudança de práticas e fecundidade disciplinar.	Mudança da base representacional ou paradigmática.

Fonte: elaboração do autor.

A questão empírica, portanto, não é apenas se uma sentença apareceu literalmente nos dados de treinamento. É preciso examinar quem formulou o problema, quais bibliotecas e avaliadores delimitaram a busca, se o sistema criou representação não fornecida, se a hipótese antecedeu o verificador e se a contribuição alterou a compreensão do campo.

5 O CASO ASTRA E A FRONTEIRA DA MATEMÁTICA ASSISTIDA POR IA

5.1 O que foi publicamente atribuído ao sistema

A OpenAI divulgou dez resultados em empacotamento de esferas, códigos binários e esféricos, grupos não sóficos, álgebras de operadores, complexidade de circuitos, complexidade quântica, criptografia em reticulados, geometria convexa e combinatória extremal. A coleção técnica afirma, entre outros pontos, determinar a taxa assintótica do programa linear de Cohn-Elkies e resolver assintoticamente o problema correspondente de incerteza de sinal de Fourier (OpenAI, 2026a; OpenAI, 2026b, p. i-3).

Segundo a empresa, uma versão interna do Astra gerou os argumentos; pessoas prepararam os manuscritos com auxílio do modelo; e o sistema formalizou cada argumento em certificado Lean. O repositório público, consultado no *commit* 94bc0feb6a9ff12c7d31d6de640a725c9d43d2b6, contém arquivos para os dez resultados,

instruções de compilação em Lean 4.32.0 e procedimento adicional de checagem por Comparator. O artigo não executou nem auditou integralmente essas formalizações; registra apenas que os artefatos e as instruções foram disponibilizados (OpenAI, 2026a; OpenAI, 2026c).

O nome Astra designa uma versão interna de um modelo futuro. A documentação pública não individualiza suficientemente arquitetura, ferramentas, orquestração, memória, verificadores, curadoria ou grau de intervenção humana. Essa indeterminação impede comparar diretamente o caso com uma tese dirigida genericamente a grandes modelos de linguagem.

5.2 O que Lean verifica — e o que permanece fora do certificado

Lean é assistente de prova baseado em teoria de tipos dependentes. Uma demonstração formal é codificada como termo que o núcleo verifica relativamente às definições, premissas e axiomas admitidos. Quando o termo possui o tipo correspondente ao teorema, a garantia de validade formal é elevada (de Moura; Ullrich, 2021).

AlphaProof demonstra a produtividade dessa arquitetura: aprendizagem por reforço em ambiente formal permite explorar provas com feedback verificável. O próprio artigo, contudo, mostra que autoformalização, bibliotecas, formulação do problema e computação em tempo de teste são partes materiais do resultado. O sistema alcança rigor dentro do ambiente formal, não independência em relação à base representacional fornecida (Hubert *et al.*, 2026).

Um certificado não decide se o problema é relevante, se as definições correspondem ao enunciado informal, se a formalização omitiu hipótese material, se a ideia é original, se as referências foram atribuídas corretamente ou se a prova produz compreensão humana. Também não mede efeitos distributivos, ambientais ou estratégicos. A pergunta respondida é delimitada: o termo segue das premissas codificadas no sistema?

A Declaração de Leiden enfatiza transparência, verificabilidade independente, atribuição humana, divulgação de ferramentas e preservação da autonomia disciplinar. Ela alerta que a tradução entre apresentação humana e formal permanece vulnerável e que incentivos podem deslocar agendas para problemas compatíveis com automação (Leiden Declaration, 2026).

5.3 O Astra refuta “LLMs can’t jump”?

Não há resposta binária sustentada pelo corpus. Os materiais oferecem artefatos formais inspecionáveis correspondentes aos enunciados divulgados e tornam pertinente investigar criatividade exploratória ou geração funcional de hipóteses. Não estabelecem, por

si, novidade científica, anterioridade, autonomia do sistema nem transformação conceitual, pois essas conclusões dependem de reconstrução do processo e avaliação comunitária.

A comparação também é assimétrica: Zahavy discute enraizamento sensorial e formulação de axiomas a partir da experiência; Astra foi testado sobretudo em matemática e ciência da computação teórica, domínios nos quais feedback formal e funções de avaliação substituem parte do contato empírico. O caso torna menos segura uma leitura absolutamente categórica do *position paper*, mas não testa diretamente sua tese forte sobre invenção científica ancorada no mundo.

A conclusão provisória é restrita: sistemas integrados a busca e verificação formal podem participar da produção de artefatos matemáticos não triviais dentro de espaços representacionais estruturados. A novidade, a importância e a originalidade intelectual dos resultados permanecem questões de nível E3. Continua aberta a capacidade de formular autonomamente novos problemas, selecionar variáveis ou construir bases conceituais a partir de experiência não previamente organizada.

6 DO CONTROLE DOS COEFICIENTES AO CONTROLE DA BASE

6.1 Primeiro nível: modulação dos componentes

No ambiente digital, a influência mais visível atua sobre componentes disponíveis. Plataformas alteram saliência, recorrência, sequência e temporalidade. A analogia com coeficientes descreve esse nível sem afirmar determinação absoluta: o sujeito conserva capacidade de agir, mas elementos recebem pesos perceptivos desiguais.

Esse poder aparece em publicidade comportamental, recomendação, apostas, crédito, recrutamento, consumo e deliberação política. O problema não é a existência de influência — inerente à comunicação —, mas a assimetria sistemática entre quem conhece, testa e ajusta a arquitetura e quem decide dentro dela. Transparência, proibição de manipulação, deveres de informação e revisão de decisões são respostas predominantemente voltadas a esse primeiro nível.

O controle dos coeficientes explica por que liberdade formal de clicar não assegura autonomia substancial. Uma escolha pode ser atribuída ao indivíduo e resultar de ambiente que maximizou urgência, ocultou custos, repetiu recompensas ou sincronizou estímulos com vulnerabilidades.

6.2 Segundo nível: constituição da base representacional

O poder mais profundo surge quando a infraestrutura não apenas altera pesos dentro de conjunto de alternativas, mas participa da definição do conjunto e da linguagem em que

alternativas podem existir. Categorias, ontologias, variáveis, rótulos e modelos não são recipientes neutros: classificam o mundo e distribuem visibilidade, comparabilidade e legitimidade (Bowker; Star, 1999).

6.2.1 Ontologia, níveis de abstração e projeto conceitual: diálogo com Luciano Floridi

Essa passagem da seleção de alternativas à estruturação do espaço representacional pode ser aprofundada pela filosofia da informação de Luciano Floridi. O método dos níveis de abstração mostra que um sistema é observado por meio de um conjunto escolhido de observáveis e relações, adequado à finalidade da análise. A descrição não constitui duplicação neutra do real: depende do nível de abstração, das variáveis admitidas e das perguntas que orientam a modelagem (Floridi, 2011, p. 46-79).

Em *The logic of information*, Floridi trata a filosofia como projeto conceitual e desloca o conhecimento de uma concepção exclusivamente mimética para uma compreensão construtiva dos modelos. Modelar um sistema implica estabelecer requisitos, distinções e estruturas informacionais que permitam torná-lo inteligível e operacionalizável, sem que o modelo se confunda com a totalidade do objeto representado (Floridi, 2019, p. 27-52, 188-206). Essa leitura não elimina restrições objetivas: os níveis de abstração devem conservar adequação ao domínio e às relações estruturais relevantes, não podendo ser escolhidos arbitrariamente (Floridi, 2011, p. 315-339).

Aplicada à inteligência artificial, essa perspectiva permite afirmar que o sistema não recebe o mundo como realidade sem mediação. Recebe sinais captados por sensores, registros selecionados, dados rotulados, textos, imagens, taxonomias, bibliotecas e objetivos de otimização. Mesmo quando aprende representações latentes, as condições de captação, codificação e avaliação delimitam quais diferenças podem ser percebidas, quais regularidades são recompensadas e quais erros se tornam visíveis. O aumento quantitativo do corpus pode ampliar a cobertura sem corrigir ausências sistemáticas, categorias inadequadas ou níveis de abstração mal escolhidos.

Isso não conduz à exigência de ontologia perfeita, única e total. A partir do quadro de Floridi, este artigo propõe uma extensão jurídico-institucional própria: bases representacionais decisivas devem ser adequadas ao contexto, documentadas, revisáveis e comparáveis com alternativas. Esse pluralismo representacional não implica relativismo irrestrito, porque modelos podem ser avaliados por adequação, capacidade explicativa, robustez, efeitos distributivos e compatibilidade com bens jurídicos. Essa extensão normativa não é atribuída a Floridi.

O conceito de controle da base acrescenta uma dimensão de poder a esse quadro. Se toda modelagem depende de observáveis, categorias e requisitos, importa identificar quem os escolhe, quem pode contestá-los e como exclusões se propagam pelas camadas de geração,

validação e circulação. Floridi oferece instrumentos para compreender a mediação informacional; a proposta aqui desenvolvida examina sua concentração institucional e suas possíveis consequências para a soberania cognitiva e científica.

6.2.2 Excedente analítico do controle da base

Controle da base possui parentesco com *framing*, classificação, regulação por arquitetura e opacidade, mas não se confunde com eles. A distinção pode ser resumida no Quadro 4. Seu excedente analítico está em deslocar a atenção da alternativa escolhida para as condições que tornam uma alternativa, uma variável ou uma pergunta reconhecível.

Quadro 4 — Distinção entre controle da base e conceitos próximos

Conceito	Objeto principal	Limite para este artigo	Excedente do controle da base
Framing	Apresentação de alternativas reconhecidas.	Não explica como o conjunto de alternativas foi constituído.	Examina categorias e espaços de hipótese anteriores ao enquadramento.
Classificação	Atribuição de objetos, pessoas ou eventos a categorias.	Descreve a distribuição entre categorias, mas não necessariamente a governança dos critérios que as produzem.	Analisa quem define categorias, ontologias e variáveis, quais exclusões geram e como podem ser revistas.
Regulação por arquitetura	Condicionamento de condutas por design, affordances e opções disponíveis.	Concentra-se no comportamento dentro de uma arquitetura já constituída.	Desloca a análise para as condições de inteligibilidade que antecedem opções e condutas.
Opacidade	Dificuldade de conhecer processos, critérios ou modelos.	Descreve déficit de visibilidade, não necessariamente poder constitutivo.	Pergunta quem escolhe a base, quais exclusões produz e quais contrapoderes podem revisá-la.

Fonte: elaboração do autor, a partir de Tversky e Kahneman (1981), Bowker e Star (1999), Yeung (2017) e Susser, Roessler e Nissenbaum (2019).

Um sistema biomédico limitado às variáveis registradas pode invisibilizar determinantes sociais. Um agente jurídico orientado apenas por precedentes dominantes pode reduzir a emergência de teses minoritárias. Um sistema científico guiado por verificadores automáticos pode privilegiar problemas formalizáveis e mensuráveis. Esses exemplos são cenários analíticos, não demonstrações empíricas; mostram como exclusões da base podem anteceder a escolha.

A captura não exige que a plataforma escolha A em vez de B. Pode ocorrer quando C deixa de integrar o espaço representacional acessível. A arquitetura torna-se infraestrutura de inteligibilidade, e a soberania é afetada quando sujeitos e instituições não conseguem conhecer exclusões, comparar bases alternativas ou contestar categorias impostas.

6.3 Pilha epistêmica e concentração vertical

A ciência assistida por inteligência artificial pode ser descrita como pilha epistêmica composta por seis camadas:

1. **Dados e proveniência.** Publicações, bases, códigos, provas, idiomas e experiências que formam o repertório.
2. **Seleção de problemas.** Perguntas que recebem capacidade computacional, prioridade institucional e critérios de sucesso.

3. **Representação.** Vocabulários, ontologias, variáveis, formalizações, bibliotecas e espaços de hipótese.
4. **Geração e busca.** Conjecturas, construções, programas, provas e estratégias produzidos com graus distintos de autonomia.
5. **Validação.** Verificadores, métricas, testes, revisão humana e padrões de aceitação.
6. **Circulação e acesso.** Publicação, embargo, licenciamento, acesso ao modelo e capacidade de reproduzir resultados.

A concentração torna-se estrutural quando um agente controla várias camadas e não existe contrapoder capaz de auditar suas transições. Nesse cenário, a organização escolhe problema, fornece representação, gera hipótese, seleciona verificador, divulga resultado e define acesso. A correção formal de prova isolada não neutraliza o poder acumulado no percurso.

A noção de pilha não é simples renomeação de pipeline técnico. Seu ganho jurídico está em identificar pontos de imputação, dependências e remédios distintos. Proveniência exige documentação; seleção de problemas demanda governança de agenda; representação requer pluralismo; geração, rastreabilidade; validação, independência; circulação, acesso e regras de embargo. A integração vertical transforma um fluxo técnico em objeto de análise institucional.

7 SOBERANIA COGNITIVA E FUNDAMENTOS JURÍDICOS

7.1 Genealogia e delimitação do conceito

Soberania cognitiva dialoga com conceitos próximos, mas possui objeto próprio. Liberdade cognitiva protege o domínio mental contra interferências indevidas; integridade mental enfatiza proteção da esfera psíquica; agência epistêmica trata da capacidade de conhecer, interpretar e participar de práticas de justificação; injustiça epistêmica descreve desvantagens na condição de sujeito do conhecimento; soberania digital e tecnológica referem-se à capacidade de controlar infraestruturas estratégicas (Bublitz, 2013; Fricker, 2007; Ienca; Andorno, 2017; Floridi, 2020).

Neste trabalho, soberania cognitiva possui três dimensões. Individualmente, é capacidade de formar juízos com acesso plural e compreensão das mediações relevantes. Institucionalmente, é autonomia de universidades, periódicos e comunidades para escolher agendas, métodos e critérios de importância. Coletivamente, é capacidade de uma sociedade desenvolver, auditar e negociar infraestruturas sem dependência absoluta. Não implica ausência de interdependência; exige contestabilidade e alternativas.

A expressão é usada de modo funcional, não como criação de direito fundamental autônomo já positivado. Ela organiza bens jurídicos existentes — liberdade científica,

autonomia, igualdade de acesso, desenvolvimento tecnológico, participação nos benefícios da ciência e integridade da decisão — diante de riscos gerados pela infraestrutura epistêmica.

7.2 Constituição, direito à ciência e ciência aberta

A Constituição brasileira protege a livre expressão da atividade científica no art. 5º, IX. Após a Emenda Constitucional nº 85/2015, o art. 218 atribui ao Estado promoção e incentivo ao desenvolvimento científico, pesquisa, capacitação, tecnologia e inovação, com prioridade à pesquisa básica e tecnológica em vista do bem público. Os arts. 219-A e 219-B admitem cooperação público-privada e organizam o Sistema Nacional de Ciência, Tecnologia e Inovação em regime colaborativo (Brasil, 1988; Brasil, 2015).

A Lei nº 10.973/2004, alterada pela Lei nº 13.243/2016, vincula políticas de inovação à capacitação e à autonomia tecnológica. Essas normas não proíbem modelos proprietários nem criam dever geral de abertura. Elas autorizam margem de conformação legislativa e administrativa para políticas de capacidade computacional, auditoria, igualdade de oportunidades e redução de dependências estratégicas; não produzem, por si, obrigação jurídica individual de acesso a modelos privados (Brasil, 2004; Brasil, 2016).

O art. 15 do Pacto Internacional sobre Direitos Econômicos, Sociais e Culturais reconhece participação na vida cultural, fruição dos benefícios do progresso científico e liberdade indispensável à pesquisa. O Comentário Geral nº 25 associa o direito à ciência à disponibilidade, acessibilidade, qualidade, liberdade científica, cooperação e distribuição não discriminatória de benefícios e riscos (United Nations, 1966; CESCR, 2020, §§ 16-17, 46-49).

A Recomendação da UNESCO sobre Ciência Aberta inclui publicações, dados, software, hardware e infraestruturas e defende a redução de desigualdades tecnológicas e de conhecimento. Trata-se de *soft law* internacional: possui força orientadora, mas depende de implementação pelos Estados e instituições competentes e não cria obrigação diretamente exigível de abertura irrestrita (UNESCO, 2021).

7.3 Regulamento Europeu de Inteligência Artificial: incidência e limite científico

O Regulamento (UE) 2024/1689 estabelece obrigações para provedores de modelos de IA de finalidade geral. O art. 53 prevê documentação técnica, informações a provedores a jusante, política de conformidade autoral e resumo do conteúdo de treinamento. Para modelos com risco sistêmico, o art. 55 acrescenta avaliações, testes adversariais, mitigação de riscos, comunicação de incidentes e cibersegurança (European Union, 2024).

A aplicação ao caso científico exige enfrentar o art. 2º, nº 6, que exclui sistemas e modelos, inclusive resultados, especificamente desenvolvidos e colocados em serviço

exclusivamente para investigação e desenvolvimento científicos. O nº 8 também exclui atividades de pesquisa, teste e desenvolvimento anteriores à colocação no mercado ou em serviço, ressalvada testagem em condições reais. A exceção deve ser interpretada funcionalmente: não alcança automaticamente modelo de finalidade geral com vocação comercial apenas porque foi testado em problemas científicos.

Como a OpenAI descreve Astra como versão interna de seu próximo modelo principal, e não como sistema desenvolvido exclusivamente para investigação científica, não é possível concluir, com os dados públicos, se a exclusão do art. 2º, nº 6, incidiria. Tampouco se deve afirmar aplicação automática dos arts. 53 e 55 antes de estabelecer colocação no mercado, propósito, alcance territorial e categoria jurídica do modelo.

Na data de fechamento do corpus, as obrigações dos provedores de modelos de IA de finalidade geral já eram aplicáveis desde 2 de agosto de 2025. As orientações da Comissão esclarecem os conceitos de provedor, colocação no mercado, modificação substancial e algumas exceções, enquanto o Código de Prática oferece via voluntária de demonstração de conformidade. Desde 2 de agosto de 2026, os poderes de execução da Comissão, inclusive multas, passaram a ser aplicáveis; modelos colocados no mercado antes de 2 de agosto de 2025 dispõem, em regra, até 2 de agosto de 2027 para adequação (European Commission, 2025a; European Commission, 2025b).

Esse pacote de implementação fortalece documentação, transparência e gestão de riscos, mas não constitui regime completo de governança da agenda científica. Ele não exige, em termos gerais, justificar a seleção de problemas, registrar bases representacionais alternativas descartadas ou distribuir capacidade computacional entre instituições. O RIE proposto adiante é, portanto, complementar e não substitutivo.

7.4 LGPD, proporcionalidade e riscos da dependência

O art. 20 da LGPD assegura solicitação de revisão de decisões tomadas unicamente com base em tratamento automatizado de dados pessoais que afetem interesses do titular. A norma é relevante para decisões *downstream* individualizadas. O controle da base científica ocorre em estágio *upstream*, frequentemente sem decisão pessoal identificável, e não pode ser deduzido diretamente do art. 20 (Brasil, 2018).

Uma resposta jurídica legítima deve observar proporcionalidade. O objetivo é proteger integridade científica, capacidade de auditoria, pluralismo e distribuição de benefícios. Medidas adequadas incluem documentação e validação independente; a necessidade varia conforme contribuição material do sistema e impacto do resultado; e a proporcionalidade em

sentido estrito exige preservar segredo legítimo, segurança, liberdade científica e inovação. O artigo não sustenta abertura integral nem separação absoluta entre indústria e universidade.

Os riscos principais são econômicos, temporais, geográficos e linguísticos. O operador de modelo de fronteira pode conhecer resultados antes da comunidade, definir embargos e parceiros e acumular vantagem estratégica. Instituições com menor capacidade computacional recebem acesso posterior ou limitado. Modelos representam melhor idiomas e tradições mais presentes nos dados. A circulação de resultados pode aumentar enquanto a capacidade local de formular problemas diminui.

8 GOVERNANÇA DA INFRAESTRUTURA EPISTÊMICA

8.1 Princípios mínimos

A governança deve ser proporcional à contribuição causal do sistema e ao impacto potencial do resultado. Revisão linguística não exige salvaguardas idênticas às aplicáveis a prova, hipótese ou política pública gerada predominantemente por modelo proprietário. O regime deve concentrar exigências nos pontos em que autonomia, reprodutibilidade, atribuição e distribuição de benefícios estão materialmente em risco.

1. **Proveniência.** Registrar modelo, versão, data, ferramentas, bibliotecas relevantes e natureza das intervenções humanas.
2. **Reconstruibilidade.** Preservar enunciados, parâmetros, código, certificados, critérios de seleção e artefatos suficientes para auditoria proporcional.
3. **Validação independente.** Submeter resultados de maior impacto a revisão externa e, quando adequado, a verificadores distintos ou equipes interinstitucionais.
4. **Transparência representacional.** Documentar ontologias, variáveis, bibliotecas, pressupostos e exclusões que delimitam a busca, sem exigir divulgação de todos os pesos.
5. **Pluralismo e contestabilidade.** Permitir bases, métricas e interpretações alternativas e manter canais de correção.
6. **Acesso equitativo.** Criar programas transparentes de capacidade computacional, apoio a instituições sub-representadas e alternativas públicas ou consorciadas.
7. **Responsabilidade institucional.** Identificar pessoas e organizações responsáveis por integridade, citações, formalização e comunicação.
8. **Avaliação de concentração.** Mapear dependência de fornecedor, barreiras de saída, *lock-in* e impacto sobre autonomia científica.

8.2 Relatório de Impacto Epistêmico: diferença, materialidade e níveis

8.2.1 Diferença em relação aos instrumentos existentes

O RIE não pretende renomear instrumentos existentes. A avaliação de impacto sobre proteção de dados (DPIA/RIPD) concentra-se em tratamentos suscetíveis de alto risco a direitos e liberdades; a avaliação de impacto sobre direitos fundamentais (FRIA) do art. 27 do AI Act incide sobre determinados usos de sistemas de alto risco; *model cards* documentam

usos, desempenho e limitações de modelos; *datasheets* descrevem motivação, composição e produção de conjuntos de dados; e a documentação de GPAI organiza informações técnicas, autorais e de risco. O RIE acrescenta como objeto a cadeia epistêmica que antecede o produto final: seleção do problema, escolha da representação, geração, validação e decisão de circulação (European Union, 2016; European Union, 2024; Mitchell et al., 2019; Gebru et al., 2021; European Commission, 2025a).

Quadro 5 — Relação entre o RIE e instrumentos existentes

Instrumento	Objeto	Escopo típico	Lacuna coberta pelo RIE
DPIA / RIPD	Riscos do tratamento de dados pessoais.	Operações com alto risco a direitos e liberdades.	Não alcança, por si, seleção de agendas, ontologias e validação científica.
FRIA	Impactos de determinados sistemas de alto risco sobre direitos fundamentais.	Contexto de implantação e grupos afetados.	Não documenta necessariamente o processo de produção de conhecimento.
Model cards	Usos previstos, avaliações, desempenho e limitações do modelo.	Relato técnico do modelo disponibilizado.	Não reconstrói a escolha do problema, tentativas e decisão de circulação.
Datasheets	Motivação, composição, coleta e usos do conjunto de dados.	Documentação do dataset.	Não abrange representação, geração, validação e concentração vertical em conjunto.
Documentação GPAI	Características, treinamento, capacidades, limitações, direitos autorais e riscos sistêmicos.	Obrigações do provedor de modelo de finalidade geral.	Não governa integralmente a agenda científica ou bases alternativas descartadas.
RIE	Mediações que estruturam produção e autoridade do conhecimento.	Projetos com contribuição material de IA à cadeia epistêmica.	Integra problema, base, geração, validação, circulação e dependência institucional.

Fonte: elaboração do autor.

8.2.2 Natureza, gatilhos e níveis

Propõe-se, *de lege ferenda*, Relatório de Impacto Epistêmico — RIE para projetos em que sistema de IA contribua materialmente à formulação do problema, escolha da representação, geração da hipótese ou prova, validação ou decisão de circulação. Há contribuição material quando a retirada ou substituição do sistema provavelmente alteraria hipótese, representação, método, prova, interpretação central ou decisão de publicar. Impacto estratégico existe quando o resultado é reivindicado como transformador ou pode produzir efeitos sistêmicos, clínicos, políticos, econômicos, ambientais ou de segurança. O RIE não é obrigação jurídica geral vigente; pode funcionar como condição de financiamento, política de integridade, requisito editorial ou cláusula contratual. Imposição sancionatória dependeria de competência e base normativa específicas.

Quadro 6 — Níveis proporcionais do Relatório de Impacto Epistêmico

Nível	Gatilho	Conteúdo mínimo	Controle
RIE 1 — assistência	IA utilizada em busca, síntese, revisão, código auxiliar ou linguagem, sem gerar conclusão central.	Ferramenta, versão, finalidade, trechos ou tarefas afetadas e verificação humana.	Declaração no manuscrito ou processo institucional.
RIE 2 — contribuição material	IA participa da formulação de hipótese, método, prova, modelo ou interpretação relevante.	Proveniência, prompts ou protocolo funcional, artefatos, critérios de seleção, papel humano, limitações e validação externa.	Revisão metodológica e disponibilidade de materiais compatíveis com sigilo legítimo.
RIE 3 — impacto estratégico	Resultado reivindicado como transformador, com efeito sistêmico, político, clínico, econômico ou de segurança.	Todos os elementos do RIE 2; dependência de fornecedor; análise de concentração; plano de replicação; riscos distributivos; resumo público e anexo confidencial.	Avaliação interinstitucional, auditoria independente e revisão periódica.

Fonte: elaboração do autor.

O relatório deve identificar instituição responsável, garantidores humanos, momento de elaboração, classificação de impacto, prazo de retenção e plano de atualização. A reconstruibilidade pode exigir, conforme o caso, versão do modelo, instruções de sistema, ferramentas, bibliotecas, parâmetros relevantes, estado de memória, chamadas externas, critérios de seleção, ambiente de execução e *hashes* dos artefatos. Não se presume que a mera conservação de *prompts* reproduza sistemas agentes complexos. Um resumo público deve permitir compreender problema, contribuição do sistema, base representacional, validação e limitações; informações sensíveis podem integrar anexo confidencial acessível a revisores, financiadores ou auditores sujeitos a sigilo.

A governança deve acompanhar o contexto. Periódicos podem condicionar aceitação; agências de fomento, desembolso e prestação de contas; universidades, integridade de pesquisa; contratos públicos, documentação e auditoria. A instituição deve prever responsável pela classificação, canal de contestação, revisão do nível e consequências institucionais da omissão. Sanções públicas exigem lei, competência, devido processo e proporcionalidade. A proposta evita criar autoridade universal abstrata e distribui responsabilidades entre atores que já controlam publicação, financiamento e pesquisa.

8.3 Autoria, contribuição e responsabilidade

A Declaração de Leiden afirma que crédito e responsabilidade permanecem humanos. A OpenAI, por outro lado, sustenta que atribuir autoria humana a prova integralmente gerada por sistema falsearia a contribuição causal da máquina. O COPE estabelece que ferramentas de IA não satisfazem critérios de autoria porque não assumem responsabilidade, não declaram conflitos e não celebram acordos de licença; o uso deve ser divulgado e o autor humano permanece responsável (COPE, 2023; Leiden Declaration, 2026; OpenAI, 2026a).

As posições podem ser compatibilizadas por três categorias. Autoria jurídica e acadêmica permanece humana, pois envolve aprovação, responsabilidade, resposta a críticas e deveres éticos. Contribuição causal descreve o que o sistema efetivamente fez e deve ser

registrada. Responsabilidade institucional alcança laboratório, empresa, pesquisadores, validadores e editores que conferem autoridade ao resultado.

A taxonomia CRediT organiza quatorze papéis de contribuição e pode ser adaptada para tornar visíveis conceptualização, metodologia, *software*, validação, visualização e redação. Ela não transforma ferramenta em autora; ajuda a separar papéis humanos e assistência automatizada. Em trabalhos com IA, recomenda-se matriz que identifique: autores humanos; tarefas CRediT; funções executadas por sistemas; e garantidores da correção e das referências (NISO, 2022).

Quando nenhum autor humano compreende suficientemente o argumento para explicá-lo, responder a críticas e assumir responsabilidade, a publicação não deve simular autoria tradicional. O resultado pode ser divulgado como contribuição gerada por sistema sob responsabilidade institucional, com validação externa e descrição honesta de seus limites, até que a comunidade o incorpore.

8.4 Infraestrutura pública e cooperação internacional

Governança documental é insuficiente sem capacidade material. Estados, universidades e organizações internacionais precisam investir em computação científica pública, modelos auditáveis, repositórios, verificadores e formação interdisciplinar. A autonomia tecnológica prevista no ordenamento brasileiro não será alcançada apenas por obrigações de transparência dirigidas a fornecedores estrangeiros.

Consórcios podem compartilhar custos e priorizar problemas de interesse social subatendidos por incentivos comerciais, como doenças negligenciadas, adaptação climática local, infraestrutura urbana e línguas de baixa disponibilidade. A cooperação deve preservar abertura, mobilidade e intercâmbio, sem converter soberania em isolamento.

Uma interpretação participativa do direito à ciência sustenta que não basta receber o produto final da descoberta; é necessário ampliar capacidades de participação, verificação e uso. Na era de modelos proprietários, distribuição do conhecimento depende também da distribuição de infraestrutura, competências e poder de auditoria.

9 CONCLUSÃO

Os materiais examinados não autorizam concluir que o Astra tenha realizado abdução equivalente à humana, criado autonomamente nova base conceitual ou produzido resultados cuja novidade já esteja comunitariamente estabelecida. Eles indicam a disponibilização de artefatos formais correspondentes a enunciados apresentados pela OpenAI como avanços em domínios estruturados. O episódio torna insuficiente a imagem de sistemas como simples

repetidores, mas sua novidade, originalidade e importância permanecem dependentes de validação especializada.

A análise de Fourier foi empregada como vocabulário heurístico controlado para distinguir dois níveis de poder. No primeiro, a arquitetura modula componentes: saliência, recorrência, ordem, temporalidade e visibilidade. No segundo, participa da definição da base: categorias, ontologias, variáveis e espaços de hipótese que condicionam perguntas e decisões. A captura mais profunda ocorre quando sujeitos ou instituições não reconhecem que operam dentro de representação selecionada por terceiro.

A hipótese normativa mostrou-se conceitualmente plausível. A concentração de dados, seleção de problemas, representação, geração, validação e circulação pode produzir dependência epistêmica, mas sua incidência e extensão exigem investigação empírica. A correção formal de artefatos particulares não neutraliza, por si, o poder acumulado na pilha epistêmica.

O direito vigente oferece fundamentos parciais. A Constituição brasileira e o Marco Legal de Ciência, Tecnologia e Inovação valorizam liberdade científica, bem público e autonomia tecnológica. O direito internacional protege participação nos benefícios da ciência. A UNESCO orienta abertura e infraestrutura; o AI Act, suas orientações e o Código de Prática estabelecem documentação e gestão de riscos para modelos de finalidade geral, mas contêm exceções e limites de incidência que devem ser examinados caso a caso. Nenhum desses regimes governa integralmente a base representacional ou a agenda científica.

A resposta proporcional combina proveniência, reconstruibilidade, validação independente, pluralismo representacional, transparência compatível com sigilos legítimos, acesso equitativo e capacidade pública de auditoria. O RIE é apresentado como proposta exploratória, com níveis, gatilhos e controles graduados; sua imposição pública dependeria de competência e base normativa específicas.

A possibilidade de sistemas construírem e revisarem bases representacionais mediante interação direta com o mundo constitui problema autônomo, reservado a investigação subsequente.

A pergunta juridicamente decisiva não é se a inteligência artificial cria exatamente como o ser humano. É quem controla as condições nas quais problemas são selecionados, hipóteses se tornam pensáveis, provas são aceitas e conhecimentos circulam. A soberania cognitiva começa a ser capturada quando agentes deixam de controlar apenas respostas e passam a controlar a base na qual perguntas podem existir.

DECLARAÇÃO DE USO DE INTELIGÊNCIA ARTIFICIAL GENERATIVA

O problema de pesquisa e sua orientação jurídica foram desenvolvidos pelo autor em diálogo com ferramenta de inteligência artificial generativa da OpenAI (ChatGPT, modelo GPT-5.6). A ferramenta contribuiu para exploração conceitual, organização estrutural, elaboração de esboços e versões preliminares, tradução, normalização de citações e simulação de parecer adversarial. O autor selecionou e reescreveu as formulações, definiu as conclusões jurídicas, conferiu as fontes utilizadas e assume responsabilidade integral pelo conteúdo. Pelos critérios propostos no artigo, esse uso corresponde ao RIE 2 — contribuição material —, sem autoria da ferramenta nem validação autônoma. A verificação incluiu confronto das afirmações atuais com fontes oficiais e primárias, revisão das referências e nova leitura adversarial do manuscrito.

REFERÊNCIAS

BODEN, Margaret A. *The creative mind: myths and mechanisms*. 2. ed. London: Routledge, 2004.

BOWKER, Geoffrey C.; STAR, Susan Leigh. *Sorting things out: classification and its consequences*. Cambridge, MA: MIT Press, 1999.

BRACEWELL, Ronald N. *The Fourier transform and its applications*. 3. ed. Boston: McGraw-Hill, 2000.

BRASIL. *Constituição da República Federativa do Brasil de 1988*. Brasília, DF: Presidência da República, 1988. Disponível em: https://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Acesso em: 4 ago. 2026.

BRASIL. *Emenda Constitucional nº 85, de 26 de fevereiro de 2015*. Altera e adiciona dispositivos na Constituição Federal para atualizar o tratamento das atividades de ciência, tecnologia e inovação. Brasília, DF: Presidência da República, 2015. Disponível em: https://www.planalto.gov.br/ccivil_03/constituicao/emendas/emc/emc85.htm. Acesso em: 4 ago. 2026.

BRASIL. *Lei nº 10.973, de 2 de dezembro de 2004*. Dispõe sobre incentivos à inovação e à pesquisa científica e tecnológica no ambiente produtivo. Texto compilado. Brasília, DF: Presidência da República, 2004. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2004-2006/2004/lei/l10.973compilado.htm. Acesso em: 4 ago. 2026.

BRASIL. *Lei nº 13.243, de 11 de janeiro de 2016*. Dispõe sobre estímulos ao desenvolvimento científico, à pesquisa, à capacitação científica e tecnológica e à inovação. Brasília, DF: Presidência da República, 2016. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2016/lei/l13243.htm. Acesso em: 4 ago. 2026.

BRASIL. *Lei nº 13.709, de 14 de agosto de 2018*. Lei Geral de Proteção de Dados Pessoais. Texto compilado. Brasília, DF: Presidência da República, 2018. Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709compilado.htm. Acesso em: 4 ago. 2026.

BUBLITZ, Jan Christoph. My mind is mine!? Cognitive liberty as a legal concept. In: HILDT, Elisabeth; FRANKE, Andreas G. (ed.). *Cognitive enhancement: an interdisciplinary perspective*. Dordrecht: Springer, 2013. p. 233-264. DOI: 10.1007/978-94-007-6253-4_19.

CESCR — COMMITTEE ON ECONOMIC, SOCIAL AND CULTURAL RIGHTS. *General comment No. 25 (2020) on science and economic, social and cultural rights*. E/C.12/GC/25. Geneva: United Nations, 2020. Disponível em: https://www.ohchr.org/en/documents/general-comments-and-recommendations/general-comment-no-25-2020-article-15-science-and. Acesso em: 4 ago. 2026.

COPE — COMMITTEE ON PUBLICATION ETHICS. *Authorship and AI tools: COPE position statement*. 13 Feb. 2023. DOI: 10.24318/cCVRZBms. Disponível em: https://publicationethics.org/cope-position-statements/ai-author. Acesso em: 4 ago. 2026.

DE MOURA, Leonardo; ULLRICH, Sebastian. The Lean 4 theorem prover and programming language. In: PLATZER, André; SUTCLIFFE, Geoff (ed.). *Automated Deduction — CADE 28*. Cham: Springer, 2021. p. 625-635. DOI: 10.1007/978-3-030-79876-5_37.

EINSTEIN, Albert. *Letters to Solovine: 1906-1955*. Translated by Wade Baskin. New York: Philosophical Library, 1987.

EUROPEAN COMMISSION. *Guidelines on the scope of obligations for providers of general-purpose AI models under the AI Act*. Brussels, 18 July 2025a. Atualizado em: 28 abr. 2026. Disponível em: https://digital-strategy.ec.europa.eu/en/policies/guidelines-gpai-providers. Acesso em: 4 ago. 2026.

EUROPEAN COMMISSION. *General-Purpose AI Code of Practice*. Brussels, 1 Aug. 2025b. Atualizado em: 25 jun. 2026. Disponível em: https://digital-strategy.ec.europa.eu/en/policies/ai-code-practice. Acesso em: 4 ago. 2026.

EUROPEAN UNION. *Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data*. Official Journal of the European Union, L 119, 4 May 2016. Disponível em: https://eur-lex.europa.eu/eli/reg/2016/679/oj. Acesso em: 4 ago. 2026.

EUROPEAN UNION. *Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence*. Official Journal of the European Union, L 2024/1689, 12 July 2024. Disponível em: https://eur-lex.europa.eu/eli/reg/2024/1689/oj. Acesso em: 4 ago. 2026.

FLORIDI, Luciano. *The philosophy of information*. Oxford: Oxford University Press, 2011. DOI: 10.1093/acprof:oso/9780199232383.001.0001.

FLORIDI, Luciano. *The logic of information: a theory of philosophy as conceptual design*. Oxford: Oxford University Press, 2019. DOI: 10.1093/oso/9780198833635.001.0001.

FLORIDI, Luciano. The fight for digital sovereignty: what it is, and why it matters, especially for the EU. *Philosophy & Technology*, v. 33, p. 369-378, 2020. DOI: 10.1007/s13347-020-00423-6.

FOURIER, Jean-Baptiste Joseph. *Théorie analytique de la chaleur*. Paris: Firmin Didot, 1822.

FRICKER, Miranda. *Epistemic injustice: power and the ethics of knowing*. Oxford: Oxford University Press, 2007.

GEBRU, Timnit et al. Datasheets for datasets. *Communications of the ACM*, v. 64, n. 12, p. 86-92, 2021. DOI: 10.1145/3458723.

GENOVA, Jandislei Antonio. *Autonomia decisória na era da IA*. São Paulo: Editora Dialética, 2026.

HARMAN, Gilbert H. The inference to the best explanation. *The Philosophical Review*, v. 74, n. 1, p. 88-95, 1965. DOI: 10.2307/2183532.

HARNAD, Stevan. The symbol grounding problem. *Physica D: Nonlinear Phenomena*, v. 42, n. 1-3, p. 335-346, 1990. DOI: 10.1016/0167-2789(90)90087-6.

HUBERT, Thomas et al. Olympiad-level formal mathematical reasoning with reinforcement learning. *Nature*, v. 651, p. 607-613, 2026. DOI: 10.1038/s41586-025-09833-y.

IENCA, Marcello; ANDORNO, Roberto. Towards new human rights in the age of neuroscience and neurotechnology. *Life Sciences, Society and Policy*, v. 13, art. 5, 2017. DOI: 10.1186/s40504-017-0050-1.

KUHN, Thomas S. *The structure of scientific revolutions*. 4. ed. Chicago: University of Chicago Press, 2012.

LEIDEN DECLARATION. *Leiden Declaration on Artificial Intelligence and Mathematics*. 2 June 2026. DOI: 10.5281/zenodo.20302944. Disponível em: https://leidendeclaration.ai/. Acesso em: 4 ago. 2026.

LU, Chris et al. The AI Scientist: towards fully automated open-ended scientific discovery. *arXiv*:2408.06292, 2024. Disponível em: https://arxiv.org/abs/2408.06292. Acesso em: 4 ago. 2026.

MAGNANI, Lorenzo. *Abductive cognition: the epistemological and eco-cognitive dimensions of hypothetical reasoning*. Berlin: Springer, 2009.

MITCHELL, Margaret et al. Model cards for model reporting. In: CONFERENCE ON FAIRNESS, ACCOUNTABILITY, AND TRANSPARENCY, 2019. *Proceedings [...]*. New York: ACM, 2019. p. 220-229. DOI: 10.1145/3287560.3287596.

NISO — NATIONAL INFORMATION STANDARDS ORGANIZATION. *ANSI/NISO Z39.104-2022: CRediT, Contributor Roles Taxonomy*. Baltimore: NISO, 2022. DOI: 10.3789/ansi.niso.z39.104-2022.

NOVIKOV, Alexander et al. AlphaEvolve: a coding agent for scientific and algorithmic discovery. *arXiv*, 2506.13131, 2025. DOI: 10.48550/arXiv.2506.13131.

OPENAI. *Ten advances in mathematics and theoretical computer science*. 1 Aug. 2026a. Disponível em: https://openai.com/index/ten-advances-in-mathematics/. Acesso em: 4 ago. 2026.

OPENAI. *Ten advances in mathematics and theoretical computer science: technical collection*. 2026b. 249 p. Disponível em: https://cdn.openai.com/pdf/ten-proofs-oai.pdf. Acesso em: 4 ago. 2026.

OPENAI. *Ten-proofs: Lean certificates accompanying proofs in mathematics and theoretical computer science*. GitHub repository, commit 94bc0feb6a9ff12c7d31d6de640a725c9d43d2b6, 2 Aug. 2026c. Disponível em: https://github.com/openai/ten-proofs/tree/94bc0feb6a9ff12c7d31d6de640a725c9d43d2b6. Acesso em: 4 ago. 2026.

OPPENHEIM, Alan V.; WILLSKY, Alan S.; NAWAB, S. Hamid. *Signals and systems*. 2. ed. Upper Saddle River: Prentice Hall, 1997.

PEIRCE, Charles Sanders. *Collected papers of Charles Sanders Peirce*. HARTSHORNE, Charles; WEISS, Paul; BURKS, Arthur W. (ed.). Cambridge, MA: Harvard University Press, 1931-1958. 8 v.

REICHENBACH, Hans. *Experience and prediction: an analysis of the foundations and the structure of knowledge*. Chicago: University of Chicago Press, 1938.

SUSSER, Daniel; ROESSLER, Beate; NISSENBAUM, Helen. Online manipulation: hidden influences in a digital world. *Georgetown Law Technology Review*, v. 4, n. 1, p. 1-45, 2019.

TVERSKY, Amos; KAHNEMAN, Daniel. The framing of decisions and the psychology of choice. *Science*, v. 211, n. 4481, p. 453-458, 1981. DOI: 10.1126/science.7455683.

UNESCO. *Recommendation on Open Science*. Paris: UNESCO, 2021. Disponível em: https://www.unesco.org/en/legal-affairs/recommendation-open-science. Acesso em: 4 ago. 2026.

UNITED NATIONS. *International Covenant on Economic, Social and Cultural Rights*. United Nations Treaty Series, v. 993, p. 3, 1966. Disponível em: https://www.ohchr.org/en/instruments-mechanisms/instruments/international-covenant-economic-social-and-cultural-rights. Acesso em: 4 ago. 2026.

YEUNG, Karen. “Hypernudge”: Big Data as a mode of regulation by design. *Information, Communication & Society*, v. 20, n. 1, p. 118-136, 2017. DOI: 10.1080/1369118X.2016.1186713.

ZAJONC, Robert B. Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology*, v. 9, n. 2, pt. 2, p. 1-27, 1968. DOI: 10.1037/h0025848.

ZAHAVY, Tom. Position: LLMs can’t jump. In: INTERNATIONAL CONFERENCE ON MACHINE LEARNING (ICML), 43., 2026. *Position Paper Track*. OpenReview, 2026. Disponível em: https://openreview.net/forum?id=klU4737opt. Acesso em: 4 ago. 2026.

COMO CITAR

> GENOVA, Jandislei Antonio. *Controle da base representacional e soberania cognitiva na ciência assistida por inteligência artificial: uma analogia com a análise de Fourier*. Jandislei Genova, São Paulo, 4 ago. 2026. Disponível em: https://jandisleigenova.com/controle-base-representacional-ia-fourier/. Acesso em: dia mês ano.

SOBRE O AUTOR

Jandislei Antonio Genova é advogado, pesquisador independente e autor de *Autonomia decisória na era da IA*. É especialista em Direito Civil e em Gestão em Saúde. Sua pesquisa examina autonomia humana, influência cognitiva estrutural, arquitetura informacional, responsabilidade civil e soberania diante de sistemas algorítmicos.

NOTA DE PUBLICAÇÃO

Esta é a versão acadêmica integral preparada para o site pessoal do autor. A data de publicação e o endereço definitivo devem ser confirmados quando a página estiver efetivamente disponível. Uma versão ensaística e condensada será publicada no LinkedIn com remissão a este texto completo.