



POR QUE ACREDITAR NA RESPOSTA? CONFIANÇA E AUTORIDADE EPISTÊMICA NO ENSINO JURÍDICO DIANTE DA INTELIGÊNCIA ARTIFICIAL GENERATIVA

WHY BELIEVE THE ANSWER? TRUST AND EPISTEMIC AUTHORITY IN LEGAL EDUCATION IN THE AGE OF GENERATIVE ARTIFICIAL INTELLIGENCE

¿POR QUÉ CREER EN LA RESPUESTA? CONFIANZA Y AUTORIDAD EPISTÉMICA EN LA ENSEÑANZA DEL DERECHO ANTE LA INTELIGENCIA ARTIFICIAL GENERATIVA

Leonardo Peteno Magnusson¹; Olívia Inocência²; Caio Cezar Pontim Scholz³

¹Doutorando. Faculdade Autônoma de Direito de São Paulo. leonardopetenomagnusson@gmail.com

²Doutoranda. Faculdade Autônoma de Direito de São Paulo. olivaiinocencioo@gmail.com

³Doutor. Universidade Estadual do Oeste do Paraná. cezar.cs@hotmail.com

RESUMO

O artigo examina as condições sob as quais um estudante de Direito deve conceder crédito a respostas produzidas por sistemas de inteligência artificial generativa, a partir do problema mais geral da confiança epistêmica na formação jurídica. Parte-se da constatação de que a autonomia intelectual não pressupõe autossuficiência, pois o conhecimento jurídico depende estruturalmente de fontes, autoridades e testemunhos que o próprio sujeito não pode integralmente verificar. O artigo delimita deliberadamente seu referencial teórico em torno de duas linhas, a epistemologia do testemunho e da autoridade epistêmica, e a literatura sobre confiança calibrada em sistemas automatizados e sobre confiança apropriada, evitando cobertura panorâmica em favor de aprofundamento vertical. Sustenta-se que a inteligência artificial generativa desloca a confiança epistêmica ao converter propriedades da apresentação discursiva, como fluência e segurança linguística, em substitutos práticos dos critérios pelos quais a autoridade deveria ser reconhecida. O objetivo consiste em identificar esse deslocamento e propor critérios para uma confiança epistemicamente calibrada no uso educacional da inteligência artificial. Adota-se abordagem qualitativa, de natureza bibliográfica, com finalidade exploratória e analítica e raciocínio predominantemente dedutivo. Os resultados indicam que a plausibilidade discursiva de uma resposta generativa não constitui, por si mesma, razão suficiente para lhe atribuir autoridade epistêmica. Conclui-se que a formação jurídica deve desenvolver, além da capacidade de avaliar informações, a capacidade de calibrar a confiança segundo quatro dimensões articuladas: a diferenciação da tarefa e da proposição, a rastreabilidade das razões e da autoridade, a proporcionalidade da verificação e a revisabilidade da confiança concedida.

Palavras-chave: Ensino jurídico. Inteligência artificial generativa. Confiança epistêmica. Autoridade epistêmica. Epistemologia do testemunho.

ABSTRACT

This article examines the conditions under which a law student should grant credence to answers produced by generative artificial intelligence systems, starting from the broader problem of epistemic trust in legal education. It begins from the observation that intellectual autonomy does not presuppose self-sufficiency, since legal knowledge structurally depends on sources, authorities, and testimony that the subject cannot fully verify independently. The article deliberately delimits its theoretical framework around two lines, the epistemology of testimony and epistemic authority, and the literature on calibrated trust in automated systems and on appropriate trust, favoring vertical depth over panoramic coverage. It argues that generative artificial intelligence displaces epistemic trust by converting properties of discursive presentation, such as fluency and linguistic confidence, into practical substitutes for the criteria by which authority should be recognized. The objective is to identify this displacement and to propose criteria for epistemically calibrated trust in the educational use of artificial intelligence. The study adopts a qualitative and bibliographic approach, with exploratory and analytical purposes and predominantly deductive reasoning. The results



indicate that the discursive plausibility of a generative response does not, by itself, constitute sufficient reason to attribute epistemic authority to it. It concludes that legal education must develop, beyond the capacity to evaluate information, the capacity to calibrate trust along four articulated dimensions: the differentiation of task and proposition, the traceability of reasons and authority, the proportionality of verification, and the revisability of the trust granted.

Keywords: Legal education. Generative artificial intelligence. Epistemic trust. Epistemic authority. Testimony epistemology.

RESUMEN

El artículo examina las condiciones bajo las cuales un estudiante de Derecho debe otorgar credibilidad a las respuestas producidas por sistemas de inteligencia artificial generativa, a partir del problema más general de la confianza epistémica en la formación jurídica. Se parte de la constatación de que la autonomía intelectual no presupone autosuficiencia, puesto que el conocimiento jurídico depende estructuralmente de fuentes, autoridades y testimonios que el propio sujeto no puede verificar íntegramente. El artículo delimita deliberadamente su marco teórico en torno a dos líneas: la epistemología del testimonio y de la autoridad epistémica, y la literatura sobre la calibración de la confianza en sistemas automatizados y sobre la confianza apropiada, privilegiando la profundidad analítica frente a una revisión panorámica. Se sostiene que la inteligencia artificial generativa desplaza la confianza epistémica al convertir propiedades de la presentación discursiva, como la fluidez y la seguridad lingüística, en sustitutos prácticos de los criterios conforme a los cuales debería reconocerse la autoridad. El objetivo consiste en identificar este desplazamiento y proponer criterios para una confianza epistémicamente calibrada en el uso educativo de la inteligencia artificial. Se adopta un enfoque cualitativo, basado en la investigación bibliográfica, con una finalidad exploratoria y analítica y un razonamiento predominantemente deductivo. Los resultados indican que la plausibilidad discursiva de una respuesta generativa no constituye, por sí sola, una razón suficiente para atribuirle autoridad epistémica. Se concluye que la formación jurídica debe desarrollar, además de la capacidad de evaluar información, la capacidad de calibrar la confianza conforme a cuatro dimensiones interrelacionadas: la diferenciación de la tarea y de la proposición, la trazabilidad de las razones y de la autoridad, la proporcionalidad de la verificación y la revisabilidad de la confianza otorgada.

Palabras clave: Enseñanza del Derecho. Inteligencia artificial generativa. Confianza epistémica. Autoridad epistémica. Epistemología del testimonio.

1 INTRODUÇÃO

A inteligência artificial generativa passou a ocupar, na rotina de estudo do Direito, uma posição que antes pertencia a professores, manuais, tratados e colegas mais experientes. Diante de uma dúvida sobre a vigência de uma norma, o alcance de um precedente ou a posição consolidada da doutrina, o estudante pode obter, em segundos, uma resposta articulada e discursivamente segura. Essa disponibilidade não resolve, apenas desloca, uma pergunta que a formação jurídica sempre precisou enfrentar. Não interessa apenas se a resposta está correta, mas por que razões, e em que medida, ela merece ser incorporada ao raciocínio de quem a recebeu.

Essa pergunta não é nova. Toda formação intelectual pressupõe alguma dependência de fontes, testemunhos e autoridades que o próprio sujeito não pode integralmente reconstruir ou checar. O que a inteligência artificial generativa acrescenta a esse problema já conhecido é uma nova espécie de interlocutor, capaz de produzir, em linguagem natural e com aparência de

fundamentação, respostas que se assemelham a explicações, mas cuja relação com a verdade das proposições que contêm não pode ser inferida da fluência com que são apresentadas. O estudante precisa, assim, decidir quanto confiar em algo que fala como uma autoridade sem necessariamente possuir as propriedades que tornariam essa autoridade justificada em uma perspectiva epistemológica.

Entrementes, não se pretende, neste artigo, analisar os formatos pedagógicos por meio dos quais o professor conduz a aula ou avalia o desempenho do estudante, nem avaliar, como objeto principal, a exatidão técnica dos sistemas de inteligência artificial generativa em tarefas jurídicas, questão sobre a qual já existe literatura empírica específica. O recorte concentra-se na relação epistêmica que se estabelece entre o estudante e a resposta que recebe, e na pergunta sobre as condições que tornam racional, ou não, conceder-lhe crédito.

Essa opção de recorte impõe delimitar o referencial teórico mobilizado. Em vez de percorrer panoramicamente a bibliografia sobre inteligência artificial na educação, aprofunda-se duas linhas escolhidas por sua aderência direta ao problema formulado. A primeira é a epistemologia do testemunho e da autoridade epistêmica, lida em contraste com a teoria da autoridade jurídica em sentido institucional. A segunda é a literatura sobre confiança em sistemas automatizados, fundada no diagnóstico sobre uso, abuso e subutilização da automação e na distinção entre confiança e dependência, mobilizada como critério analítico para explorar o que significa confiar de maneira epistemicamente calibrada em respostas de inteligência artificial generativa.

A relevância do tema decorre de uma lacuna perceptível sobre inteligência artificial e ensino jurídico. As análises se concentram, preponderantemente, em medir a exatidão das respostas geradas, perquirir taxas de alucinação, desempenho comparado ao de estudantes ou confiabilidade de ferramentas especializadas de pesquisa jurídica, inclusive em estudos recentes voltados ao contexto brasileiro, prevalentemente centrados na integração curricular e nas competências técnicas envolvidas. Discute-se menos como o estudante deveria se relacionar, em ótica epistêmica, com uma resposta cuja exatidão ele próprio não tem, no momento em que a recebe, condições de avaliar integralmente.

Diante desse cenário, pergunta-se *em que medida a confiança depositada por estudantes de Direito em respostas de inteligência artificial generativa pode ser justificada de forma epistêmica, e quais critérios permitem distinguir uma confiança formativamente responsável de uma deferência indevida à plausibilidade discursiva dessas respostas. Essa é a problemática que orienta a investigação.*

A proposta inicial é a de que essa confiança não pode ser resolvida pela alternativa entre aceitação indiscriminada e desconfiança generalizada, mas depende de uma calibração que relacione seu grau e modalidade à natureza da tarefa, à possibilidade de rastrear as razões e autoridades subjacentes, ao risco associado ao eventual erro e à disposição para revisar o crédito concedido diante de novos elementos. Essa calibração não é espontânea. Depende de critérios que a formação jurídica precisa desenvolver propositadamente, e não presumir como subproduto automático do contato do estudante com a tecnologia.

Nesse sentido, o objetivo do artigo consiste em explorar como a inteligência artificial generativa desloca a confiança epistêmica na formação jurídica e propor critérios para uma confiança calibrada, em perspectiva epistêmica, em seu uso educacional. Especificamente, busca-se compreender os fundamentos da dependência epistêmica, da confiança e da autoridade na formação do jurista; examinar como a inteligência artificial generativa tensiona esses fundamentos ao aproximar indevidamente plausibilidade discursiva de autoridade epistêmica; e estabelecer critérios concretos para seu uso responsável no ensino jurídico.

A pesquisa adota abordagem qualitativa, de natureza bibliográfica, com finalidade exploratória e analítica e raciocínio predominantemente dedutivo, sem pretensão de mensurar empiricamente os padrões de confiança de estudantes de Direito em sistemas generativos.

Dessa forma, o artigo está organizado em três partes. A primeira investiga os fundamentos da dependência epistêmica, da confiança e da autoridade na formação jurídica, distinguindo-os da autoridade jurídica em sentido institucional. A segunda analisa como a inteligência artificial generativa desloca essa estrutura ao converter propriedades da apresentação discursiva em substitutos práticos dos critérios de autoridade. A terceira propõe critérios para uma confiança calibrada, organizados em quatro dimensões articuladas, capazes de orientar o uso educacional responsável dessa tecnologia no ensino jurídico.

2 CONFIAR PARA CONHECER: DEPENDÊNCIA, CONFIANÇA E AUTORIDADE EPISTÊMICA NA FORMAÇÃO JURÍDICA

A formação intelectual é frequentemente associada à autonomia, entendida como a capacidade de examinar informações, avaliar razões e formar juízos sem se limitar a reproduzir conclusões recebidas de terceiros. Essa autonomia, contudo, não pressupõe autossuficiência. Uma parcela expressiva daquilo que se conhece depende de informações e razões às quais o sujeito não teve acesso direto, mas que foram produzidas ou transmitidas por outras pessoas.



Coady (1992) identifica no testemunho uma fonte de conhecimento historicamente subestimada pela epistemologia, e Hardwig (1985) chega ao mesmo problema a partir da especialização, mostrando que mesmo quem domina determinado campo precisa apoiar parte de suas crenças no conhecimento alheio. A dependência epistêmica não constitui, portanto, deficiência excepcional, mas característica ordinária da produção social do conhecimento. Essa constatação modifica a própria noção de autonomia. Se o sujeito não pode reproduzir pessoalmente todas as investigações nem dominar todos os campos dos quais depende aquilo que sabe, a racionalidade não pode ser medida exclusivamente pela independência com que forma suas crenças.

Hardwig (1985) observa que alguém pode ter boas razões para acreditar em uma proposição porque tem boas razões para considerar que outra pessoa dispõe de razões adequadas para sustentá-la, ainda que não tenha acesso direto às evidências correspondentes, pois a transmissão de conhecimento por outros integra o funcionamento ordinário de práticas epistêmicas socialmente organizadas (Coady, 1992; Faulkner, 2011). A autonomia intelectual deve, por isso, ser distinguida da independência epistêmica. Formar um sujeito autônomo não significa torná-lo independente de todo conhecimento alheio, mas desenvolver condições para que reconheça como, quando e em que medida essa dependência pode ser racionalmente incorporada à formação de seus próprios juízos.

A passagem da dependência para a confiança, entretanto, não é automática. Hardwig (1991) acentua que a confiança integra a própria estrutura de empreendimentos cognitivos complexos, nos quais conhecer depende da integridade e da competência de outros participantes. Faulkner (2011) formula o problema de maneira útil ao distinguir o que torna razoável aceitar aquilo que alguém afirma daquilo que confere justificação à crença formada com base nesse testemunho. A confiança pode tornar racional a dependência em relação à palavra de outro, sem que isso signifique que a verdade da proposição decorra do próprio ato de confiar.

Ketzer (2016), por sua vez, aprofunda essa distinção ao apresentar que confiar epistemicamente em um testemunho não se reduz a simplesmente fiar-se em alguém, exigindo antes uma disposição racionalmente sensível às razões que tornam esse testemunho crível. Confiar não elimina, assim, a necessidade de razões. Modifica a estrutura pela qual essas razões se tornam acessíveis ao sujeito.

A alternativa à autossuficiência não pode ser, porém, a credulidade. A mesma abertura comunicativa que amplia aquilo que um indivíduo pode conhecer o expõe a informações equivocadas, incompletas ou enganosas. Sperber *et al.* (2010) descrevem esse problema pela



ideia de vigilância epistêmica. A dependência comunicativa exige mecanismos capazes de avaliar tanto a fonte quanto o conteúdo da informação recebida, pois os benefícios do conhecimento social coexistem com o risco da desinformação. A competência epistêmica não se encontra, portanto, em nenhum dos extremos entre independência absoluta e deferência indiscriminada, mas na capacidade de administrar relações de dependência sem convertê-las em aceitação irrefletida.

O problema se torna mais complexo quando quem depende de conhecimento alheio não dispõe da mesma competência técnica de quem o produz. Goldman (2001) formula essa dificuldade a partir da posição do não especialista, que precisa escolher entre autoridades epistêmicas possíveis sem condições de reproduzir integralmente o juízo técnico que pretende avaliar, mas mostra que essa escolha não precisa ser cega. O sujeito pode considerar a qualidade dos argumentos acessíveis, o reconhecimento por outros especialistas, possíveis vieses e o histórico de desempenho de quem reivindica expertise, apoiando a confiança em indicadores de credibilidade que não exigem reprodução completa do conhecimento especializado.

Fricker (2007) oferece contraponto relevante ao mostrar que essa assimetria não é apenas cognitiva, mas pode ser atravessada por preconceitos que reduzem indevidamente a credibilidade atribuída a determinados falantes, fenômeno que a autora denomina injustiça testemunhal. Essa advertência não invalida os critérios de Goldman (2001), mas recorda que atribuir credibilidade nunca é operação puramente técnica.

Essa possibilidade permite distinguir dependência de autoridade. Nem toda fonte da qual se obtém informação ocupa, apenas por isso, posição de autoridade epistêmica. Zagzebski (2012) situa a autoridade epistêmica em uma teoria mais ampla da confiança e sustenta que reconhecer a autoridade de outros pode ser compatível com a autonomia de quem acredita, desde que existam razões para considerar que acompanhar o juízo alheio oferece melhores condições de alcançar uma crença verdadeira do que insistir na formação inteiramente independente do próprio juízo.

Fricker (2006), em estudo específico sobre autonomia epistêmica e testemunho, converge com esse diagnóstico ao defender que aceitar racionalmente o testemunho de outrem, quando existem boas razões para tanto, não compromete a autonomia epistêmica, distinguindo essa aceitação mediada por razões da deferência simplesmente cega. Essa relação reivindica, porém, cautela adicional. Se a deferência a uma autoridade exigisse verificação integral prévia

de cada razão, a autoridade perderia sua função epistemológica; se dispensasse qualquer juízo sobre as razões para deferir, a autonomia seria substituída por submissão intelectual.

Zagzebski (2012) supera essa oposição ao sustentar que aceitar autoridade pode resultar do próprio exercício da reflexão conscienciosa, e a preocupação de Goldman (2001) com a identificação de especialistas evidencia que essa deferência pode ser racionalmente mediada por razões disponíveis ao não especialista. A autonomia assume, então, configuração relacional, manifesta também na capacidade de reconhecer dependências, selecionar autoridades, estabelecer o alcance da deferência e conservar condições para revisá-la.

Na formação jurídica, essa estrutura adquire características particulares. O conhecimento do Direito é construído mediante pluralidade de fontes e interlocutores aos quais se atribuem funções distintas. O estudante aprende por meio de professores, manuais e artigos científicos, mas também precisa operar com textos legislativos, decisões judiciais e atos produzidos por instituições dotadas de competências constitucional e legalmente estabelecidas, que apresentam diferentes tipos de autoridade, não apenas diferentes graus de confiabilidade.

Com efeito, a opinião de um jurista reconhecido pode constituir razão relevante para uma interpretação; a decisão de um tribunal pode possuir efeitos jurídicos que não decorrem de sua superioridade epistêmica; e, por conseguinte, o texto legislativo ocupa posição institucional que não se reduz à credibilidade de quem o produziu. Ferraz Junior (2003) situa essa pluralidade ao distinguir, na dogmática jurídica brasileira, a autoridade que decorre do reconhecimento institucional das fontes daquela que decorre da persuasão dos argumentos, expondo que a decidibilidade dos conflitos jurídicos depende de critérios de autoridade que não se confundem com a plausibilidade intrínseca do enunciado.

Lopes (2023), ao reconstruir historicamente a formação do conhecimento jurídico, descreve que a autoridade das fontes do Direito nunca decorreu de sua evidência intrínseca, mas de práticas institucionais específicas de cada tradição jurídica que autorizam certos enunciados e não outros, o que reforça a distinção entre autoridade epistêmica e autoridade institucional. Schauer (2009) apresenta, em sentido convergente, que aprender a distinguir autoridade vinculante de autoridade meramente persuasiva constitui parte central do raciocínio jurídico (Thamay; Magnusson; Pagani, 2025), distinção que precisa ser explicitamente ensinada, e não presumida como intuitiva.

Sem embargo, autoridade epistêmica e autoridade jurídica não são categorias equivalentes. A primeira diz respeito às razões pelas quais se reconhece em outro uma posição diferenciada para formar ou orientar crenças; a segunda decorre de estruturas institucionais que

atribuem a certos atos e agentes capacidade de produzir razões juridicamente relevantes. Raz (1979) leciona que a autoridade do Direito não se reduz à maior probabilidade de acerto de quem emite uma diretiva, mas relaciona-se à maneira pela qual razões são mediadas por fontes institucionalmente reconhecidas.

Um pesquisador pode possuir elevada autoridade epistêmica sobre determinado problema jurídico sem poder produzir uma decisão vinculante; um tribunal pode decidir uma controvérsia sem que cada afirmação de seu pronunciamento deva ser tratada como epistemicamente superior. A epistemologia jurídica brasileira já explorou distinção próxima. Badaró (2023), ao se debruçar sobre o convencimento judicial acerca dos fatos, esclarece que a confiabilidade de uma fonte de prova não pode ser presumida a partir de sua forma de apresentação, exigindo critérios específicos de valoração conforme sua natureza e origem, disciplina metodológica que, transposta ao problema posto, deveria orientar igualmente o estudante diante de respostas de sistemas generativos cuja forma discursiva não corresponde necessariamente à solidez das razões que a sustentam.

Essa particularidade impede que a confiança no campo jurídico seja reduzida à probabilidade de uma informação estar correta. Uma proposição pode reproduzir adequadamente o texto de uma decisão e ainda atribuir-lhe alcance que ela não possui, ou enunciar corretamente uma norma sem considerar elementos que condicionam sua aplicação. Saber em que confiar envolve, portanto, reconhecer não apenas a confiabilidade de quem fornece a informação, mas a procedência da afirmação, a natureza da fonte e a autoridade que lhe corresponde.

A formação jurídica pressupõe, assim, desenvolver a capacidade de distribuir confiança, o que significa reconhecer quando a dependência é inevitável, distinguir fontes dignas de crédito, identificar a espécie de autoridade que exercem e conservar condições para revê-la diante de razões contrárias. A autonomia intelectual não desaparece quando o estudante confia. Ela se manifesta na forma pela qual essa confiança é concedida, limitada e revisada.

3 QUANDO A PLAUSIBILIDADE SE TORNA AUTORIDADE: INTELIGÊNCIA ARTIFICIAL GENERATIVA E DESLOCAMENTO DA CONFIANÇA EPISTÊMICA

A incorporação da inteligência artificial generativa às práticas de aprendizagem acrescenta nova característica ao problema da confiança, mas não inaugura o problema da dependência tecnológica. Parasuraman e Riley (1997) distinguem situações de uso, uso

excessivo e subutilização da automação, mostrando que tanto a dependência inadequada quanto a rejeição de recursos úteis podem produzir resultados indesejáveis.

Lee e See (2004) desenvolvem a questão a partir da relação entre confiança e *reliance*, que orienta a disposição do sujeito para apoiar-se em sistemas automatizados quando sua complexidade torna impraticável compreender integralmente seu funcionamento, de modo que o problema não é uma escolha binária entre confiar e desconfiar, mas uma questão sobre as condições em que determinada dependência se mostra adequada à tarefa.

Skitka, Mosier e Burdick (1999) documentam empiricamente o viés da automação, a tendência de usuários a aceitar recomendações automatizadas mesmo diante de evidências contraditórias, deixando de processá-las por confiarem excessivamente no sistema, achado que antecede em décadas a inteligência artificial generativa, mas descreve o mecanismo que respostas fluentes tendem a intensificar.

A inteligência artificial generativa introduz, contudo, diferença relevante em relação a formas de automação tradicionalmente examinadas pela literatura. Sistemas generativos não apenas executam operações ou apresentam recomendações. Produzem, em linguagem natural, respostas que podem assumir a forma de explicações, sínteses e argumentos, aproximando a interação de práticas comunicativas por meio das quais ordinariamente se adquire conhecimento, embora isso não autorize equiparar a relação com a inteligência artificial às relações de testemunho ou autoridade. Aquilo que se apresenta discursivamente como explicação pode influenciar a disposição do usuário para aceitar uma proposição sem que a forma da resposta forneça, por si mesma, razões suficientes para reconhecer no sistema uma autoridade epistêmica.

Essa organização torna relevante distinguir propriedades da apresentação de propriedades epistemicamente pertinentes à proposição apresentada. Coerência textual e segurança linguística podem tornar uma resposta inteligível e persuasiva, mas não validam que suas informações sejam verdadeiras. A literatura sobre fluência de processamento já demonstrava que a facilidade com que uma informação é processada pode funcionar como pista em julgamentos sobre propriedades não diretamente observáveis, inclusive sua veracidade (Reber; Schwarz, 1999).

Alter e Oppenheimer (2009) mostram que a fluência opera como sinal metacognitivo utilizado em julgamentos de verdade, familiaridade e confiança, porque o sujeito tende a interpretar processamento fácil como indício de correção. Essa perspectiva não permite concluir que estudantes tomem a fluência de uma resposta generativa como prova de sua correção, mas

revela que a apresentação pode participar da formação de juízos de credibilidade sem corresponder às propriedades que se pretende inferir, o que se relaciona à vigilância epistêmica.

Se, como sustentam Sperber *et al.* (2010), a comunicação amplia o que pode ser conhecido ao mesmo tempo em que expõe o destinatário a informações enganosas, a atribuição de crédito depende da avaliação da fonte e do conteúdo, o que a inteligência artificial generativa torna particularmente complexo ao produzir respostas com elevada uniformidade discursiva mesmo quando o estatuto epistemológico das proposições que as compõem é muito diferente.

Essa dificuldade aparece de maneira mais direta nas análises sobre inteligência artificial e aprendizagem. Liu, Jou e Hao (2026) identificaram, em pesquisa com estudantes do ensino superior, processo que classificam como *epistemic trust transfer*, associado ao *cognitive offloading* e à redução da responsabilidade pela aprendizagem, com maior utilização da inteligência artificial associada a maior *cognitive offloading* e este à transferência de confiança epistêmica e a menor *learning accountability*. Essa reorganização não precisa ser decisão consciente de reconhecer a inteligência artificial como autoridade. Pode ocorrer quando a resposta do sistema passa a funcionar como ponto de partida não questionado para a investigação subsequente, de modo que a confiança pode ser deslocada sem atribuição explícita de autoridade ao sistema.

No campo jurídico, essa diferença assume especial relevância, pois modelos de linguagem produzem respostas estruturadas segundo formas reconhecíveis do discurso jurídico, o que pode torná-las plausíveis sem assegurar correspondência com o material jurídico pertinente. Dahl *et al.* (2024), ao estudarem sistematicamente respostas de modelos de linguagem a questões de jurisprudência federal norte-americana, encontraram elevada incidência de alucinações, dificuldade dos próprios modelos para identificar respostas incorretas e tendência a aceitar premissas jurídicas contrafactuais apresentadas pelos usuários, indicando que uma resposta compatível com o discurso jurídico pode não preservar correspondência adequada com os fatos que pretende representar.

A dificuldade tampouco desaparece quando a tecnologia é desenvolvida especificamente para pesquisa jurídica. Magesh *et al.* (2025) avaliaram ferramentas comerciais que combinam modelos generativos com mecanismos de recuperação de informações jurídicas e, embora tenham reduzido a frequência de alucinações em comparação com modelos generalistas, encontraram respostas alucinatórias em proporções entre 17% e 33%.

Alimardani (2024), em avaliação de sistemas generativos em provas de Direito Penal, chega a diagnóstico complementar. Os sistemas superaram a média dos estudantes em tarefas

dissertativas abertas, mas tiveram desempenho inferior nas questões que exigiam análise jurídica detalhada, o mesmo registro genérico em que a resposta tende a soar mais plausível sem corresponder a maior confiabilidade. Ou seja, a vinculação da geração a bases específicas pode alterar significativamente o desempenho, mas não elimina a necessidade de examinar a relação entre aquilo que a resposta afirma e as fontes que efetivamente lhe dão suporte.

Abal e Pilati (2025), ao relatarem experiências de utilização de sistemas generativos personalizados em disciplinas jurídicas, observam que a integração pedagógica dessas ferramentas exige acompanhamento ponderado do professor, sob pena de que a facilidade de obtenção de respostas substitua o exercício que a atividade pretendia provocar, confirmando a partir da sala de aula o mesmo risco que se documenta a partir do desempenho técnico dos sistemas.

Impende que, em registro mais teórico, anunciado por Peixoto e Silva (2019) e Maranhão, Florêncio e Almada (2021), a inteligência artificial deixou de ser mera ferramenta auxiliar externa ao raciocínio jurídico para integrar a própria estrutura da argumentação e da decisão, o que amplia a relevância de aprofundar criticamente a autoridade que se atribui às suas respostas.

O problema, por isso, ultrapassa a existência de referências inventadas ou informações falsas. Uma resposta pode indicar fonte existente e atribuir-lhe conteúdo que ela não contém, ou reproduzir adequadamente uma norma sem considerar sua vigência ou âmbito de aplicação. A confiabilidade jurídica de uma resposta não constitui, assim, propriedade unitária, podendo exigir verificar a existência da fonte, a correspondência entre a fonte e aquilo que lhe é atribuído, a adequação da interpretação, a autoridade da fonte e a força jurídica da razão dela extraída, decomposição que constitui construção analítica proposta neste artigo, não atribuída aos autores citados anteriormente.

Cada uma dessas operações responde a uma pergunta diferente. Confirmar que uma decisão existe não demonstra que sustenta a proposição para a qual foi citada, e a resposta generativa tende a reunir essas operações em um produto discursivo contínuo, enquanto a aprendizagem jurídica exige que elas possam ser distinguidas.

É nesse ponto que a plausibilidade se aproxima indevidamente da autoridade, não porque a inteligência artificial adquira autoridade epistêmica pela capacidade de produzir linguagem persuasiva, mas porque os sinais que tornam sua resposta plausível podem funcionar como substitutos dos critérios pelos quais a autoridade deveria ser reconhecida. Lustig e Nardi (2015) descrevem esse fenômeno, em sistemas cuja legitimidade decorre da forma como processam informação, e não de deliberação institucional reconhecível, como autoridade algorítmica,



sustentada na aparência de objetividade do processamento, e não nas razões que efetivamente incorpora.

No debate contemporâneo, Streck (2025) converge com esse diagnóstico ao advertir que a fluência discursiva dos sistemas generativos não deve ser confundida com compreensão hermenêutica do sentido jurídico, crítica sintetizada na imagem dos aprendizes de feiticeiro tecnológicos, incapazes de substituir a compreensão que apenas o intérprete pode produzir. Grau (2009), em teoria da interpretação já consolidada, converge com esse diagnóstico ao sustentar que interpretar e aplicar o Direito é sempre ato normativo situado, e não reprodução mecânica de um sentido preexistente, o que explica por que a fluência discursiva de um sistema generativo não equivale à compreensão hermenêutica que a interpretação jurídica exige. Uma resposta organizada pode aumentar a disposição para aceitá-la sem que nenhuma dessas propriedades estabeleça a competência epistêmica da fonte.

Essa distinção impede que o problema seja reduzido à prevenção do erro. Se apenas respostas falsas fossem epistemicamente problemáticas, bastaria aperfeiçoar a precisão dos sistemas, mas respostas verdadeiras também podem ser aceitas por razões inadequadas, quando o estudante as incorpora não porque reconheceu a autoridade e avaliou as razões, mas porque a resposta lhe pareceu convincente. Zhou, Jurafsky e Hashimoto (2026) confirmam a gravidade dessa distinção ao mostrar que modelos de linguagem tendem a expressar suas respostas com grau de segurança discursiva que não acompanha, de maneira confiável, a probabilidade real de sua correção, de modo que a confiança comunicada pelo sistema não constitui indicador seguro de sua própria confiabilidade.

A inteligência artificial generativa coloca, desse modo, um problema de confiança que não pode ser resolvido nem pela aceitação automática nem pela desconfiança generalizada. A primeira converte propriedades da apresentação em substitutos da justificação; a segunda ignora que sistemas tecnológicos podem fornecer informações úteis à própria investigação jurídica. Torna-se necessário formar critérios capazes de relacionar o grau e a modalidade da confiança às características da tarefa, à natureza da proposição, à possibilidade de verificação e à autoridade das fontes envolvidas. Se a formação jurídica pressupõe aprender a distribuir confiança, a presença da inteligência artificial torna necessário avançar uma etapa. Aprender a calibrá-la.

4 DA CONFIANÇA AUTOMÁTICA À CONFIANÇA CALIBRADA: CRITÉRIOS PARA O USO EPISTEMICAMENTE RESPONSÁVEL DA IA NO ENSINO JURÍDICO



A identificação dos riscos associados à confiança em respostas generativas não conduz à conclusão de que o uso responsável da inteligência artificial dependa de desconfiança geral. Se a aceitação indiscriminada converte a plausibilidade da resposta em substituto da justificação, a rejeição indiscriminada desconsidera situações nas quais o sistema pode contribuir para localizar informações, explorar argumentos ou identificar aspectos relevantes de um problema, razão pela qual a literatura sobre interação humano-automação formula o objetivo em termos de confiança adequadamente calibrada, ou *appropriate reliance* (Parasuraman; Riley, 1997; Lee; See, 2004).

Aplicada à formação jurídica, contudo, essa formulação exige desenvolvimento adicional. Não basta perguntar quanto confiar na inteligência artificial, mas em que resultado, para qual finalidade e mediante quais razões. Nesse contexto, Mehrotra *et al.* (2024) identifica considerável heterogeneidade entre conceitos como *appropriate trust*, *calibrated trust* e *appropriate reliance*, empregados a partir de diferentes tradições.

Em formulação recorrente, *calibrated trust* corresponde à relação entre a confiança do usuário e a capacidade efetivamente apresentada pelo sistema, enquanto *appropriate reliance* tende a enfatizar se o comportamento do usuário responde adequadamente à qualidade da recomendação recebida (Lee; See, 2004). Impende que, não há definição inteiramente uniforme desses conceitos; a distinção adotada neste texto serve à organização analítica do argumento, sem pretender estabelecer terminologia consensual para o campo. Para fins formativos, interessa menos produzir determinado nível subjetivo de confiança do que desenvolver a capacidade de decidir adequadamente quando um resultado pode ser incorporado ao próprio raciocínio.

Essa constatação permite distinguir duas formas de calibração que não devem ser confundidas. A primeira incide sobre o sistema, quando o usuário ajusta suas expectativas às capacidades e limitações que a tecnologia demonstra em determinada classe de tarefas. A segunda incide sobre a resposta, quando o usuário precisa determinar se existem razões suficientes para apoiar-se naquele resultado particular. Um sistema com desempenho elevado não torna verdadeira cada resposta que produz, assim como erros anteriores não tornam falsas suas respostas posteriores.

Calibrar a confiança no sistema não equivale, portanto, a justificar a confiança na resposta, distinção que dialoga com a leitura de *appropriate reliance*, atenta a decisões situadas sobre aceitar ou rejeitar recomendações, sem corresponder a uma taxonomia já consolidada

nesse campo. Essa distinção é particularmente importante diante de modelos generativos porque uma mesma interação pode reunir proposições com estatutos epistemológicos distintos, como uma informação normativa facilmente verificável ao lado de uma conclusão interpretativa aberta à controvérsia, sem que todas mereçam o mesmo grau de confiança apenas por integrarem o mesmo texto.

Essa sensibilidade afasta uma regra frequentemente apresentada como solução suficiente para os problemas da inteligência artificial, a recomendação genérica de verificar as respostas, que não constitui operação uniforme. Du e Yuan (2026) observam que a dependência em relação à inteligência artificial se torna educacionalmente problemática quando o resultado tecnológico assume função epistemicamente diretiva e desloca operações de julgamento que deveriam permanecer com o estudante, entre elas verificar, interpretar e justificar.

A questão formativa não está em exigir que todo conteúdo produzido pela inteligência artificial seja reconstruído independentemente, exigência que tornaria a ferramenta praticamente inútil, mas em desenvolver a capacidade de reconhecer quais operações não podem ser transferidas sem comprometer aquilo que se pretende aprender.

Vu, Cummings e Park (2026), em experimento com tarefas educacionais envolvendo *ChatGPT*, constataram que treinamento destinado a alertar estudantes para limitações do sistema aumentou comportamentos observáveis de verificação, embora não tenha produzido efeito significativo sobre a intenção declarada de verificar, com redução da checagem sob pressão temporal. É preciso, portanto, transformar a consciência abstrata do risco de erro em práticas de investigação capazes de identificar quando a confirmação externa é necessária.

A verificação deve, por isso, ser compreendida de maneira proporcional ao risco epistemológico envolvido, que depende não apenas da probabilidade abstrata de erro do sistema, mas da função que a proposição desempenha no raciocínio e das consequências de aceitá-la equivocadamente. Utilizar a inteligência artificial para organizar um parágrafo não demanda o mesmo controle exigido para afirmar a vigência de uma norma ou sustentar que específica interpretação corresponde à posição consolidada de um tribunal, formulação empregada neste artigo como construção analítica de proporcionalidade epistêmica.

No Direito, essa proporcionalidade pode ser concretizada mediante decomposição. Se uma resposta atribui determinada tese a um precedente, a verificação necessária dependerá da utilização pretendida, bastando em alguns casos confirmar a existência da decisão e, em outros, analisar se a tese possui caráter vinculante e permanece aplicável ao caso em questão.

Essa operação também permite deslocar o objeto da confiança. Em vez de permanecer concentrada na capacidade geral da inteligência artificial, a confiança pode ser progressivamente transferida para razões cuja autoridade possa ser independentemente reconhecida. Se o sistema indica determinada regra jurídica, o ponto de chegada da investigação não pode ser a confiança na afirmação de que a regra existe, mas o acesso à fonte normativa correspondente; se atribui uma tese a determinado tribunal, a confiança pode ser reconstruída mediante a consulta à decisão. A rastreabilidade não depende que todo resultado seja descartado até sua reprodução integral por meios independentes, mas que, quando a proposição desempenhar função epistemicamente relevante, existam condições para deslocar o fundamento de sua aceitação da autoridade aparente da resposta para as razões que efetivamente podem sustentá-la.

Esse movimento é especialmente importante porque a verificação não se confunde com a simples localização de uma fonte compatível. Procurar uma referência depois de receber a resposta da inteligência artificial pode converter-se em exercício confirmatório se o estudante buscar apenas materiais capazes de corroborar a conclusão previamente fornecida. A verificação epistemicamente relevante precisa preservar a possibilidade de que a resposta esteja errada ou incompleta, admitindo fontes independentes e avaliando se a evidência encontrada efetivamente sustenta a proposição, em vez de apenas compartilhar seu vocabulário. Verificar não é confirmar aquilo que a inteligência artificial afirmou. É submeter a afirmação a condições sob as quais ela possa ser confirmada, modificada ou rejeitada, distinção que constitui desenvolvimento próprio deste artigo.

A confiança epistemicamente responsável precisa, por isso, permanecer revisável. Uma expectativa inicial de confiabilidade pode justificar que uma resposta seja considerada hipótese ou auxílio, mas não imuniza seu conteúdo diante de razões posteriores, da mesma forma que identificar um erro não torna racional rejeitar indiscriminadamente toda contribuição futura do sistema. Os estudos sobre *appropriate trust*, a exemplo de Mehrotra *et al.* (2024) enfatiza o caráter dinâmico da calibração, que deve responder às informações disponíveis sobre as capacidades e limitações do sistema e às condições concretas da interação. No contexto educacional, essa revisabilidade adquire significado adicional, pois o estudante precisa conservar a autoridade para aumentar, reduzir ou retirar o crédito inicialmente concedido conforme avança na investigação.

Pode-se, a partir dessas considerações, organizar a confiança epistemicamente calibrada no uso educacional da inteligência artificial em quatro dimensões relacionadas. A primeira é a

diferenciação da tarefa e da proposição, pela qual se reconhece que diferentes utilizações e afirmações demandam diferentes modalidades de confiança. A segunda é a rastreabilidade das razões e da autoridade, que permite deslocar o fundamento da aceitação da plausibilidade da resposta para fontes e razões capazes de sustentá-la. A terceira é a proporcionalidade da verificação, mediante a qual a intensidade do controle acompanha a relevância epistemológica e jurídica do possível erro. A quarta é a revisabilidade da confiança, que preserva a possibilidade de alterar a dependência diante de novas evidências.

Essas dimensões não reproduzem taxonomia encontrada na literatura de *calibrate*, mas constituem proposta analítica para transportar o problema da confiança ao contexto da formação jurídica, generalizando para esse contexto a distinção entre confiança no sistema e confiança na resposta e a exigência de proporcionalidade já examinadas nesta seção, e combinando a literatura sobre *appropriate reliance* com a estrutura de autoridade epistêmica examinada.

A formulação permite atribuir à calibração um significado mais exigente do que a simples correspondência entre confiança subjetiva e desempenho médio da tecnologia. Confiança epistemicamente calibrada, no ensino jurídico, é a disposição para apoiar-se em resultados de inteligência artificial em medida ajustada à natureza da tarefa e da proposição, à possibilidade de rastrear as razões que as sustentam, ao risco associado à aceitação equivocada e às evidências disponíveis para manter ou revisar o crédito concedido.

O objetivo não é maximizar, minimizar ou exigir verificação universal, mas formar condições para que a confiança seja discriminada, fundamentada e corrigível. O estudante epistemicamente competente não é aquele que consegue verificar independentemente tudo o que recebe da inteligência artificial, mas aquele que reconhece o que precisa ser verificado, por que e com qual intensidade, diante da finalidade da tarefa. A confiança deixa, então, de ser atitude passiva diante da tecnologia e passa a integrar o próprio exercício do juízo jurídico.

5 CONCLUSÃO

A investigação partiu do problema de saber em que medida a confiança depositada por estudantes de Direito em respostas de inteligência artificial generativa pode ser epistemicamente justificada, e quais critérios permitem distinguir uma confiança formativamente responsável de uma deferência indevida à sua plausibilidade discursiva. O exame da dependência, da confiança e da autoridade epistêmica mostrou que a autonomia intelectual nunca pressupõe autossuficiência, dependendo sempre da capacidade de reconhecer dependências, distinguir

fontes dignas de crédito, identificar a espécie de autoridade que exercem e conservar condições para revisar a confiança concedida.

A investigação apresentou, em seguida, que a inteligência artificial generativa desloca essa estrutura ao produzir respostas cuja coerência textual e segurança discursiva podem funcionar como substitutos práticos dos critérios pelos quais a autoridade epistêmica deveria ser reconhecida, deslocamento que não depende de decisão consciente de atribuir autoridade ao sistema, podendo ocorrer de maneira mais elementar, quando a resposta generativa ocupa a posição que deveria ser preenchida por fontes e razões independentemente reconhecíveis.

A problemática formulada pode, assim, ser respondida em sentido qualificado. A confiança depositada em respostas de inteligência artificial generativa não é, em si mesma, epistemicamente ilegítima, mas apenas quando calibrada segundo a natureza da tarefa, a rastreabilidade das razões, a proporcionalidade da verificação e a revisabilidade do crédito concedido. Não se trata de alternativa entre confiar e não confiar, nem de escala uniforme de maior ou menor confiança, mas de disposição diferenciada, sensível à proposição específica sobre a qual se pretende apoiar e às consequências de um eventual erro.

Os critérios propostos oferecem parâmetros para essa diferenciação, articulados em quatro dimensões, a diferenciação da tarefa e da proposição, a rastreabilidade das razões e da autoridade, a proporcionalidade da verificação e a revisabilidade da confiança concedida. Não pretendem esgotar as possibilidades de uso responsável da inteligência artificial no ensino jurídico, mas oferecer vocabulário comum para decisões que, de outra forma, tendem a ser tomadas de maneira improvisada, disciplina a disciplina, resposta a resposta. A contribuição proposta consiste, portanto, em deslocar a discussão sobre confiabilidade da inteligência artificial do plano da precisão técnica dos sistemas para o plano da relação epistêmica que o próprio estudante estabelece com suas respostas.

O presente estudo possui natureza teórico-analítica e propositiva, sem pretensão de generalização empírica sobre os padrões efetivos de confiança de estudantes de Direito em sistemas generativos. A taxonomia proposta foi formulada a partir do referencial teórico delimitado na introdução, e não a partir de observação sistemática de sua aplicação em disciplinas reais, de modo que sua utilidade prática permanece, neste momento, hipótese de trabalho a ser testada. A opção por verticalizar a discussão em torno de um número delimitado de referenciais, em vez de buscar cobertura panorâmica da literatura disponível, constitui escolha metodológica deliberada, que permitiu maior profundidade sobre a estrutura epistêmica

da confiança na formação jurídica, ao custo de não explorar exaustivamente outras tradições filosóficas eventualmente relevantes, limite que também deve ser reconhecido.

Estudos futuros poderão examinar empiricamente os padrões de confiança de estudantes de Direito em respostas de inteligência artificial generativa, avaliando se as quatro dimensões aqui propostas efetivamente orientam decisões mais calibradas do que a alternativa entre aceitação indiscriminada e desconfiança generalizada. Cabe também investigar como essa calibração se articula com a formação da identidade e da responsabilidade profissional do jurista, tema já anunciado em outro texto desta série, e com os próprios formatos pedagógicos por meio dos quais poderia ser deliberadamente ensinada e exercitada em sala de aula.

REFERÊNCIAS

ABAL, Felipe Cittolin; PILATI, Adriana Fasolo. Inteligência artificial no ensino jurídico: experiências com GPTs personalizados. **International Journal of Digital Law**, Belo Horizonte, v. 6, e609, p. 3-21, 2025.

ALIMARDANI, Armin. Generative artificial intelligence vs. law students: an empirical study on criminal law exam performance. **Law, Innovation and Technology**, v. 16, n. 2, p. 777-819, 2024.

ALTER, Adam L.; OPPENHEIMER, Daniel M. Uniting the tribes of fluency to form a metacognitive nation. **Personality and Social Psychology Review**, v. 13, n. 3, p. 219-235, ago. 2009.

BADARÓ, Gustavo Henrique. **Epistemologia judiciária e prova penal**. 2. ed. São Paulo: Revista dos Tribunais, 2023.

COADY, Cecil Anthony John. **Testimony**: a philosophical study. Oxford: Clarendon Press, 1992.

DAHL, Matthew *et al.* Large legal fictions: profiling legal hallucinations in large language models. **Journal of Legal Analysis**, v. 16, p. 64-93, 2024.

DU, Yiran; YUAN, Yijia. Epistemic dependence in AI-mediated learning. **AI & Society**, p. 1-11, 2026. DOI: <https://doi.org/10.1007/s00146-026-03294-1>.



FAULKNER, Paul. **Knowledge on trust**. Oxford: Oxford University Press, 2011.

FERRAZ JUNIOR, Tercio Sampaio. **Introdução ao estudo do direito**: técnica, decisão, dominação. 4. ed. São Paulo: Atlas, 2003.

FRICKER, Elizabeth. Testimony and epistemic autonomy. In: LACKEY, Jennifer; SOSA, Ernest (org.). **The epistemology of testimony**. Oxford: Clarendon Press, 2006.

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

GOLDMAN, Alvin I. experts: which ones should you trust? **Philosophy and Phenomenological Research**, v. 63, n. 1, p. 85-110, jul. 2001.

GRAU, Eros Roberto. **Ensaio e discurso sobre a interpretação/aplicação do direito**. 5. ed. São Paulo: Malheiros, 2009.

HARDWIG, John. Epistemic dependence. **The Journal of Philosophy**, v. 82, n. 7, p. 335-349, jul. 1985.

HARDWIG, John. The role of trust in knowledge. **The Journal of Philosophy**, v. 88, n. 12, p. 693-708, dez. 1991.

KETZER, Patricia. A irredutibilidade do conceito de confiança na epistemologia do testemunho. **Conjectura: Filosofia e Educação**, Caxias do Sul, v. 21, n. 3, p. 496-513, set./dez. 2016.

LEE, John D.; SEE, Katrina A. Trust in automation: designing for appropriate reliance. **Human Factors**, v. 46, n. 1, p. 50-80, spring 2004.

LIU, Yu; JOU, Min; HAO, Yungwei. From cognitive offloading to learning accountability erosion: how generative AI reshapes students' epistemic trust and educational engagement. **Interactive Learning Environments**, p. 1-17, 2026. DOI: <https://doi.org/10.1080/10494820.2026.2680276>.



LOPES, José Reinaldo de Lima. **O direito na história**: lições introdutórias. 7. ed. Barueri: Atlas, 2023.

LUSTIG, Caitlin; NARDI, Bonnie. Algorithmic authority: the case of bitcoin. In: **Proceedings of the 2015 48th Hawaii International Conference on System Sciences**, p. 743-752, 2015. DOI: 10.1109/HICSS.2015.95.

MAGESH, Varun *et al.* Hallucination-free? Assessing the reliability of leading AI legal research tools. **Journal of Empirical Legal Studies**, v. 22, n. 2, p. 216-242, 2025.

MARANHÃO, Juliano Souza de Albuquerque; FLORÊNCIO, Juliana Abrusio; ALMADA, Marco. Inteligência artificial aplicada ao direito e o direito da inteligência artificial. **Suprema: Revista de Estudos Constitucionais**, Brasília, v. 1, n. 1, p. 154-180, jan./jun. 2021.

MEHROTRA, Siddharth *et al.* A systematic review on fostering appropriate trust in human-AI interaction: trends, opportunities and challenges. **ACM Journal on Responsible Computing**, v. 1, n. 4, art. 26, p. 1-45, nov. 2024.

PARASURAMAN, Raja; RILEY, Victor. Humans and automation: use, misuse, disuse, abuse. **Human Factors**, v. 39, n. 2, p. 230-253, 1997.

PEIXOTO, Fabiano Hartmann; SILVA, Roberta Zumblick Martins da. **Inteligência artificial e direito**. v. 1. Curitiba: Alteridade, 2019.

RAZ, Joseph. **The authority of law**: essays on law and morality. Oxford: Clarendon Press, 1979.

REBER, Rolf; SCHWARZ, Norbert. Effects of perceptual fluency on judgments of truth. **Consciousness and Cognition**, v. 8, n. 3, p. 338-342, 1999.

SCHAUER, Frederick. **Thinking like a lawyer**: a new introduction to legal reasoning. Cambridge: Harvard University Press, 2009.



SKITKA, Linda J.; MOSIER, Kathleen L.; BURDICK, Mark. Does automation bias decision-making? **International Journal of Human-Computer Studies**, v. 51, n. 5, p. 991-1006, 1999.

SPERBER, Dan *et al.* Epistemic vigilance. **Minds & Language**, v. 25, n. 4, p. 359-393, set. 2010.

STRECK, Lenio Luiz. **Robô não desce escada e trapezista na voa**: os limites dos aprendizes de feiticeiro. Avaré: Contracorrente, 2025.

THAMAY, Rennan; MAGNUSSON, Leonardo Peteno; PAGANI, Lucas Augusto Gaioski. Sistema de precedentes e democracia participativa: *amicus curiae* enquanto elemento de correção do déficit democrático. **Revista dos Tribunais**, São Paulo, v. 1077, p. 151-181, jul. 2025.

VU, Chi B.; CUMMINGS, James J.; PARK, Daniel Y. Student engagement with ChatGPT for educational tasks: Effects of inoculation training on verification intentions and behavior. **Computers and Education Open**, v. 10, art. 100335, p. 1-21, 2026.

ZAGZEBSKI, Linda Trinkaus. **Epistemic authority**: a theory of trust, authority, and autonomy in belief. Oxford: Oxford University Press, 2012.

ZHOU, Kaitlyn; JURAFSKY, Dan; HASHIMOTO, Tatsunori. Navigating the grey area: how expressions of uncertainty and overconfidence affect language models. In: **Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing**, Singapura, p. 5506-5524, dez. 2023. Disponível em: <https://aclanthology.org/2023.emnlp-main.335/>. Acesso em: 26 set. 2026.