

AS PRINCIPAIS CRÍTICAS À INTELIGÊNCIA ARTIFICIAL

THE MAIN CRITICISMS OF ARTIFICIAL INTELLIGENCE

<https://doi.org/10.26512/rfmc.v13i2.57831>

Philippe Lacour*
Universidade de Brasília

<http://lattes.cnpq.br/6424210911031934>
<https://orcid.org/0000-0003-3226-584X>
unb@philippelacour.net

* Professor Doutor do Departamento de Filosofia da Universidade de Brasília.

RESUMO

Este artigo pretende oferecer contribuições para o debate acerca da Inteligência Artificial e seus impactos sociais, buscando entender os aspectos fundamentais do seu funcionamento; identificar e explicar as características do seu modo de pensar e listar as principais críticas que recebe. A partir da análise do estado da arte das tecnologias de aprendizado de máquina aliada à reflexão sobre obras de pensadores e pensadoras que desenvolvem pesquisas sobre o fundamento e as consequências desse tipo de tecnologia, almeja-se desenvolver críticas inéditas que auxiliem no aprofundamento das questões. Sabe-se que hoje a Inteligência Artificial tem sido cada vez mais presente em diversos âmbitos da sociedade e essa presença tem ocasionado espanto, admiração, curiosidade ou medo. Por esse motivo fazem-se necessárias pesquisas que abordem a natureza dessa tecnologia, bem como seus efeitos na sociedade. O artigo estabelece uma lista de dezesseis críticas diferentes, separadas por razões metodológicas, apesar de ser frequentemente misturadas. Analisa primeiramente as limitações internas da IA antes de examinar as críticas externas. Por fim, destaca a originalidade de uma crítica em particular, em torno da noção de conjectura, indicando as perspectivas tecnológicas que ela abre.

Palavras-chave: IA. Críticas. Grandes modelos de linguagem. Tecnologia. Conjectura.

Abstract: This article aims to contribute to the debate on Artificial Intelligence and its social impacts, seeking to understand the fundamental aspects of how AI works, identify and explain the characteristics of its way of thinking and list the main criticisms it receives. By analyzing the state of the art of machine learning technologies and reflecting on the works of thinkers who have carried out research into the foundations and consequences of this type of technology, the goal is to develop new criticisms that will help to deepen the issues. It is well known that Artificial Intelligence has become increasingly present in various areas of society and this presence has caused astonishment, admiration, curiosity or fear. For this reason, there is a need for research into the nature of this technology, as well as its effects on society. The article lists sixteen different criticisms, separated for methodological reasons, although they are often mixed up. It first looks at the internal limitations of AI before examining the external criticisms. Finally, it highlights the originality of one criticism in particular, regarding the notion of conjecture, indicating the technological perspectives it opens up.

Keywords: AI. Criticisms. Large Language Models. Technology. Conjecture.

INTRODUÇÃO

A partir dos anos 1990, os estudos sobre inteligência artificial se transformaram e a estatística voltou a ser o fundamento dessa tecnologia (Cardon, 2018; 2019). A partir daí, com os cálculos probabilísticos, foi possível fazer com que a máquina aprenda de acordo com as experiências que é treinada. Esse método de aprendizado de máquina de baixo pra cima consiste em expor o *software* a uma grande quantidade de informações para que ele possa reconhecer padrões a partir do método de tentativa e erro (Sumpter, 2019, p. 144). Ao lidar com os dados, a IA procede por método indutivo para fazer generalizações ou previsões (Mitchell, 1997, p. 23).

O problema desse modelo de inteligência artificial é que ele é muito menos acurado do que se costuma acreditar. O matemático David Sumpter, em *Dominados pelos números* (2019), pôs à prova diversos algoritmos, entrevistou vários desenvolvedores e apresentou como conclusão a tese de que os algoritmos são muitas vezes ineficientes e básicos e, ainda que possa oferecer resultados satisfatórios em diversas ocasiões, também apresenta diversos erros que costumamos não reparar. As pessoas ficam espantadas quando recebem propagandas de coisas que queriam comprar, mas não percebem a quantidade de produtos que lhes é oferecida, mas que não desperta interesse.

Essa falta de precisão resulta na criação de vieses que fazem com que as decisões tomadas pelas IAs sejam frequentemente consideradas preconceituosas. A formação de viés ocorre, por sua vez, pelo próprio fundamento do aprendizado de máquina: o método indutivo. A indução é crença de que as coisas se repetirão como foram. Assim, a partir da observação de um certo número de objetos, infere-se uma repetição nas características. De acordo com a conclusão de que todos os cisnes analisados são brancos, por exemplo, afirma-se que *todo cisne é branco*. Essa formação de viés poderia ser atenuada pelo trabalho humano que poderia suavizar os vieses inerentes a tal tecnologia. Contudo, empresas

como a *OpenAI* costumam dar pouca atenção a esse tipo de trabalho. Como foi noticiado pela revista *Time* (<https://time.com/6247678/openai-chatgpt-kenya-workers/>), trabalhadores responsáveis por essa função possuem condições precárias de trabalho e recebem menos de 10 R\$ por hora de trabalho (Perrigo, 2023).

Nesse sentido, nota-se que a IA, ainda que apresente conclusões acerca da realidade, não pode ser considerada *stricto sensu* como uma ciência (que implica um método – Anderson, 2008). A forma como os *softwares* fazem correlações e sínteses estatísticas faz dessa tecnologia uma espécie de *heurística*. Isso quer dizer que ela é capaz de encontrar respostas no meio dos dados a que é exposta, mas não consegue explicá-las, nem sistematizar conclusões eficientes diante de questões complexas.

Outro ponto que faz com que se deva ter atenção às decisões tomadas pelas inteligências artificiais é o fato de que nas redes neurais (forma mais avançada de aprendizado de máquina) não é possível compreender o que levou o software à conclusão alcançada. Isso acontece porque a quantidade de números calculados para a tomada de decisão é tão grande que torna impossível a criação de explicações humanamente compreensíveis sobre a forma de como a máquina chegou à resposta. Isso quer dizer que as soluções apresentadas pela inteligência artificial não podem ser auditadas, não é possível conhecer quais foram os argumentos em que se basearam.

Nos primeiros meses de 2023, o *ChatGPT* colocou novamente a IA nas capas de jornais, sites de notícias, reportagens televisivas, dentre outros veículos de mídias (Amaral, 2023). A capacidade do *chatbot* de emular o discurso humano e a aparente precisão nas respostas surpreendeu o público e até mesmo profissionais da área (Gefen, 2023). De fato, a *OpenAi*, empresa que desenvolveu o *ChatGPT*, apresentou um *software* com capacidade de automatizar diversas tarefas que envolvem criação de textos e códigos através de conteúdos presentes na *Internet*. Na mesma linha, as aplicações, que criam imagens a partir do comando escrito, também se tornaram populares e levantaram o debate acerca dos direitos autorais. Além disso, o nível de realismo das produções também tem gerado preocupação, porque imagens criadas pelas inteligências

artificiais foram tomadas como imagens reais, gerando confusão e desinformação. Em outros âmbitos da sociedade, como na medicina e no direito, existem também aplicações que realizam trabalhos que outra dependeriam exclusivamente do trabalho humano. Na medicina já existem *softwares* capazes de tecer diagnósticos e prognósticos a partir da avaliação de exames. No direito é possível encontrar aplicações que analisam documentos e emitem pareceres que podem decidir o resultado de um processo.

Percebe-se, dessa forma, que a IA é uma tecnologia cada vez mais presente na sociedade e tomada como uma fonte segura e objetiva para se alcançar as melhores decisões. Contudo, uma análise sobre seus fundamentos lógicos e seu método revela que as conclusões alcançadas por essas são baseadas em formas de conhecimento não seguras.

Neste artigo contemplarei a questão das críticas à inteligência artificial, abordando a sua estrutura e seus impactos sociais. Trata-se, primeiramente, de estabelecer um mapa geral das várias críticas a serem feitas a essa tecnologia. Esse esforço é particularmente importante visto que muitas objeções se repetem, são misturadas, ou ao contrário, exclusivas umas das outras. Essa confusão e parcialidade provocam um estado de esquecimento de algumas críticas relevantes. Daí a necessidade de identificar as várias objeções, listá-las e ordená-las, de tal modo se possam assim entender melhor as inúmeras combinações entre elas, que aparecem regularmente nas formulações mais frequentes. Além disso, utilizarei o caso das tecnologias da tradução, sendo elas não apenas um caso particular de aplicação da IA, mas o lugar de nascimento da IA, para destacar uma crítica (à noção de conjectura) de particular acuidade para a avaliação dos desafios e limites da IA em geral.

DEFINIÇÕES PRELIMINARES

O termo Inteligência Artificial designa hoje em dia tantas coisas que é difícil fixar uma definição que seja ao mesmo tempo robusta e opera-

tória. A *inteligência* faz referência a uma capacidade de compreensão por via do estabelecimento de uma inter-relação entre várias coisas ou eventos; mas também designa um jeito de agir e até de sentir emoção: se opõe a ignorância e também a estupidez, simetricamente vista como incapacidade de juntar impressões. A noção de *artifício* tem uma história um pouco complexa: designava antigamente uma fabricação humana (de caráter regular e repetitivo), em oposição as produções da natureza. Mas Gilbert Simondon sublinha que o artificial designa o fato que o homem dever intervir para manter um objeto na existência, protegendo-o do mundo natural (Simondon, 2012); é então algo interior a ação artificializante do ser humano (uma flor obtida em estufas só pode se reproduzir por intervenção humana). Aparece aqui uma ambivalência interessante que vai ter um certo papel na IA contemporânea.

Uma observação deve ser adicionada em torno da última geração de IA. O famoso *ChatGPT* é uma aplicação (para o diálogo e a pesquisa de informação) de um Grande Modelo de Linguagem; outras aplicações incluem: criação artística, fabricação de imagens, assistência a composição de música, ao diagnóstico, às decisões e pesquisas jurídicas, ao ensino, etc. Agora o modelo *GPT* constitui a matriz de *ChatGPT*, sendo ele apenas mais um exemplo de tecnologia de (*Very*) *Large Language Model*: outros modelos do mesmo gênero estão sendo desenvolvidos nesse momento (*LightOn*, *Bloom*, etc). Esses Grandes Modelos de Linguagens constituem um desenvolvimento da IA da geração anterior (aprendizagem profunda), em termo de *quantidade* de informação mobilizada, de velocidade e poder de cálculo, e de adequação das pesquisas à informação considerada (*fine tuning*). No mais, são baseadas sobre os mesmos princípios que as IA anteriores: cálculo, probabilidade, presunção, indução, previsão. Por esses motivos, fazem-se necessárias políticas públicas para lidar com essa tecnologia de forma mais produtora, no momento em que vários países tentam definir uma política pública nesse setor, como por exemplo, os EUA, a China, a Rússia, a EU (AI Act de 2023), a França (Villani, 2018), etc.

Agora, pode-se dizer de maneira liminar que a IA designa ao mesmo tempo um conhecimento e uma tecnologia, e que procede por via de cálculo probabilístico, conjectural e indutivo. O pensamento por “cál-

culo” foi idealizado por Leibniz no século XVII, quando ele imaginou, através da noção de “característica universal”, o princípio da lógica contemporânea. A ambição desse “pensamento cego” consistia notadamente a substituir o cálculo a disputa, inclusive nas controvérsias diplomáticas e nas deliberações políticas. Três séculos depois, o programa de Leibniz recebeu uma confirmação extraordinária pela invenção da máquina (abstrata) de Turing (e da noção de “calculabilidade”), e a sua primeira encarnação concreta na forma de um “calculador” (*computer*). Deve ser, porém, notado que, na mesma época que Leibniz, o filósofo Pascal insistiu sobre a originalidade de um “espírito” de fineza, ao lado do espírito da geometria (dividindo assim a razão teórica), e que o contemporâneo de Turing, Gödel, traçou limites insuperáveis para a aplicação da noção de cálculo.

Desde o início, então, e na sua base mais fundamental, a IA parece sofrer de uma certa ambiguidade. E pode-se aprofundar esse mal-estar quando examinar as outras noções implicadas na sua definição (probabilidades, presunção, indução). Na verdade, se torna particularmente importante uma lista completa das críticas a serem feitas a IA, classificadas de maneira gradual, a partir das características próprias da IA, até a questão dos seus usos, e finalmente das consequências dessas utilizações. Obviamente, muitas dessas críticas estão interligadas, e a separação na lista seguinte é apenas para fins metodológicos. Classifiquei essas críticas em duas categorias, sendo as primeiras internas à IA (as limitações são aquelas dos conceitos diretores da tecnologia), enquanto as segundas são mais externas.

AS CRÍTICAS INTERNAS

Algumas críticas já foram formuladas e estão conseqüentemente bem estabelecidas. Mas como elas podem ser frequentemente esquecidas, é importante retomá-las e esclarece-las com novos instrumentos conceituais, hoje mais avançados. Outras críticas à IA estão um pouco esque-

cidas ou escondidas, e mereceriam, por essa razão, ser investigadas com mais precisão.

1. Limitação interna da noção de cálculo: A noção de cálculo é definida na lógica formal com certas características (não-contradição, completude e decidibilidade) que aparecem todas ao nível básico do cálculo das proposições, mas que já começam a sumir com a lógica da primeira ordem (que perde a decidibilidade) e a matemática (que perde também a completude). Esses aspectos fundamentais, profundamente sublinhados por Gödel nos seus famosos teoremas de incompletude, destacam o que um comentador inspirado chamou de “limitações internas dos formalismos” (Ladrière, 1992). Se essa noção lógica já não é suficiente para capturar toda a diversidade criadora dos objetos matemáticos, deve-se insistir que é *a fortiori* o caso para sistemas simbólicos dotados de propriedades ainda mais complexas que aquelas da matemática, como os aspectos pragmáticos das línguas naturais, por exemplo (Granger, 1994). Ora esse argumento é particularmente importante, visto que as IA da última geração se definem como Grandes Modelos de Linguagens ((*Very*) *Large Language Models*), como *BERT*, *GPT* ou *Bloom*, que analisam repartições de palavras em textos de línguas naturais.
2. Debate sobre a interpretação da natureza ontológica e epistêmica das probabilidades. Nem todo cálculo se reduz a um cálculo de probabilidade: em particular, podem existir cálculos formais de significações linguísticas que seguem mais as regras gramaticais (Desclés, 2020). Mas a IA é baseada sobre a ideia de um cálculo de probabilidade. Tradicionalmente, se diferenciam dois aspectos das ciências do aleatório: a estatística, que visa formas de eventos passados, efetivamente acontecidos, isto é: atuais; e a probabilidade, que visa formas de eventos quaisquer (não-atuais), inclusive do futuro (*cf.* Granger, 1995). Ora várias interpretações das probabilidades existem, em torno de seu valor ontológico (“é ‘real’ uma probabilidade de 50% de um evento – por exemplo um acidente – acontecer?). Uma vertente objetiva insiste sobre a dimensão absoluta dos resultados do cálculo probabilístico, que podem ser vistos como frequências. Outra, mais subjetiva (ou bayesiana), sublinha que se deve entender

mais como um grau de certeza da crença pessoal – aqui entram diversas variantes (Laplace, Savage, Ramsey, Carnap). Essa(s) alternativa(s) colocam em questão a dimensão supostamente definitiva, convincente, inquestionável e até constrangedora dos resultados das IAs: existe de fato um problema reflexivo, filosófico -propriamente crítico- de saber como interpretá-los. A tecnologia da aprendizagem profunda poderia parecer constituir uma objeção, no sentido que, com ele, a máquina aprende precisamente a “improvisar”. Na verdade, essa improvisação é também guiada por regras indutivas, de tal maneira que esteja assim duplicado e não resolvido o problema indutivo, sem garantia de conseguir capturar a contingência dos eventos, nem de prever as variações irregulares.

3. Modelagem e interpretação: Sempre é possível fazer uma modelagem formal de um material empírico, inclusive humano para representá-lo e explica-lo melhor. Isso vale também para a realidade empírica linguística (inclusive, aliás, para as produções da linguagem literária). A modelagem pode ser do tipo estatístico, como no caso de uma navegação sobre um corpus, por exemplo. E permite observar novos fenômenos, até agora despercebidos, através da identificação de correlações. Mas o modelo formal, qualquer que seja, deve ser fabricado a partir do material empírico e, uma vez constituído, deve ser aplicado a outro material (Granger, 1988). Ora, esses dois gestos de ida e volta são operações da interpretação, que não pode, ela mesma, ser formalizada (calculada) (Granger, 1991).
4. Crítica das ambições (racionais) exclusivas desse tipo de presunção feitas por cálculo probabilístico, do tipo do cálculo estocástico (muito utilizado para modelar processos complexos, ou estudar sistemas que variam através do tempo de maneira aleatória, isto é: sequencias probabilísticas convergentes). É particularmente útil para pensar a evolução de sistemas complexos, não deterministas, por exemplo o movimento browniano (fenômenos de difusão), com aplicações de predições optimais na meteorologia ou nas antecipações financeira. A pretensão correlativa desse tipo de afirmação é que não existe outra maneira *racional* de fazer presunção (ou conjectura); em particular não existe uma avaliação fina, comparativa (no

sentido contrastivo), interpretativa e aproximada (como se encontra nas ciências humanas). Irei aprofundar essa crítica no final do artigo.

5. Pretensão preditiva da indução: por abstração universal e necessária, então válida para o futuro (antecipação). A indução consiste em passar por inferências de vários casos particulares para um resultado mais abstrato, sobre o qual pode-se apoiar uma predição. Desde Leibniz, se sabe que só se pode conseguir um resultado geral, e não um que seja universal. Desde Hume, se sabe que a previsão não é necessária, mas apenas contingente. Consequentemente, pretender raciocinar indutivamente sobre fatos estatísticos do passado para antecipar um evento futuro de maneira probabilística consiste é sempre possível, mas nunca resulta em certezas absolutas, pois não consegue totalmente capturar as variações imprevisíveis e contingentes.
6. Crítica da pretensão da IA ao “pensamento”. Essa crítica é recorrente, mesmo que tome uma acuidade suplementar desde que um engenheiro de Google diante declarou sua perplexidade de uma IA que teria chegado a nível avançado de consciência, mostrando que ela teria medo da morte. A filosofia analítica sublinhou que o “pensamento” assim associado as máquinas vem do projeto inicial de Turing de dissociar o material (hardware) do software (a computação): através de uma analogia, a hipótese funcionalista idealizou uma isomorfia descritiva entre o pensamento humano e o funcionamento computacional (Putnam, 1975). Nasceu a IA simbólica, baseada sobre o tratamento das regras de sintaxe, sem consideração de semântica. Apesar das dificuldades, isso incentivou a pensar numa “linguagem do pensamento” (Fodor, 1975), que recebeu um certo apoio da linguística (Chomsky) ou das neurociências (Marr). Mas se essas idealizações passam o teste de Turing, a experiência de pensamento do “cérebro na cuba” (Putnam) mostra que o programa da máquina não pensa: ele simula a linguagem e a visão de maneira mecânica, sem possibilidade de atingir a sua existência ou aquela do mundo. No programa de computação, na se trata de consciência, mas de simulação da consciência segundo um conjunto de regras algorítmicas.

cas que permite manipular símbolos sem que nem estes nem a sua manipulação faça sentido para a máquina (Corteel, 2020b; 2023). Esses símbolos não tem significação, nem referência – por isso a máquina é “cega” diante do mundo. O paralogismo da analogia (da máquina que pensa) continua na época da aprendizagem profunda, que reduz o conhecimento a indução física neuronal: por ser dessa vez *semântica*, a abordagem não permite fazer a diferença entre o cérebro e uma máquina, pois se perdem a intenção e a significação própria do pensamento humano (Searle, 1994). Faltariam qualidades perceptivas (*qualia*, como aponta a experiência de *Bock*). Como o ilustram bem as dificuldades da tradução, a significação de um poema não se reduz as condições materiais da sua compreensão (cérebro, nervos óticos, transdução física da luz sobre os nervos óticos).

AS CRÍTICAS EXTERNAS

Em uma outra perspectiva:

1. Crítica aos vieses normativos. É a crítica mais conhecida, que aponta as preferências dos algoritmos, que se podem detectar na leitura dos resultados, a favor ou ao detrimento de tal ou tal categoria de gênero, de raça, social ou política (por exemplo, fazer uma pesquisa sobre “enfermagem” e obter resultados sobre “enfermeiras”; ou a associação preconceituosa de “imigrante” com “delinquência”). Essas tendências naturais dos algoritmos justificam a presença, dentro dos times de IA, de novos “justiceiros” que leem, avaliam e corrigem os resultados, dando às aprendizagens uma inflexão ligada a normas práticas (éticas, jurídicas, políticas) e não apenas teóricas (correções cognitivas do tipo: “isso é um gato, isso ao contrário não é”). Se pode relacionar a essa abordagem a crítica dos abusos do poder algorítmico (Berns e Rouvroy, 2013; Cardon 2015).
2. Crítica social da exploração dos trabalhadores digitais. De inspiração operária, esse crítica investiga um fato menos conhecido que

os vieses, mas que foi, todavia um pouco estudado por alguns autores, que mostram que a automatização da IA é ligada a toda uma série de operações humanas, tanto antes quanto depois do cálculo algorítmico, mas também ao longo do processo (treino, ajustes, retificações, etc.). Esse trabalho frequentemente invisibilizado, mal pago, repetitivo e desvalorizado (tal como o famoso Amazon Turk), e feito por um conjunto de “pequenas mãos”. As análises socialistas da exploração de uma classe operária por uma rica burguesia funcionam bem também nesse setor industrial. Se adicionam a elas as críticas “luddistas” da absorção dos saberes-fazer humanos pelas máquinas, então de uma dimensão descartável dos trabalhadores. As recentes mobilizações de certas categorias profissionais (atores nos EUA, atores de legendagem na França) mostram que o medo de substituição do trabalho humano pelas máquinas (que aspiram os saberes-fazer) tem uma base objetiva.

3. Crítica das condições ecológicas da IA. Se trata de uma reivindicação crescente a favor da conscientização da crise climática contemporânea. Boa parte das tecnologias da IA, baseadas sobre a automação, uma grande potência de cálculo, imensas bases de dados (que vão se enriquecendo de maneira constante) e um processo de aprendizagem permanente, funcionam com custos energéticos gigantescos. Aproximadamente 10% das emissões com efeito de estufa são ligadas ao setor digital em 2025. As dimensões que os economistas chamam pudicamente de “externalidades negativas” são frequentemente negligenciadas pelas análises tradicionais em termos de relação custo-benefício (resfriamento dos servidores, desvio dos rios, processos extremamente consumidores de energia, do tipo blockchain), como alguns autores apontam (Strubell et al., 2019). Porém, deveriam ser colocadas em relação com as necessidades ecológicas, numa perspectiva de prudência, dever ou até de sustentabilidade razoável (McNeill 2001, Bonneuil e Fressoz, 2013), inclusive no domínio das tecnologias da linguagem (Stibbe 2015). Assim sendo, novos critérios devem ser valorizados, como aquele da eficiência (Schwartz et al., 2019)? No caso particular das tecnologias da tradução, não se trata de sonhar de uma de-mecanização, mas talvez de passar da alta para a baixa tecnologia (Cronin, 2017, cap. 4;

Bihouix, 2014), ou até de promover, contra a competição de velocidade contra a máquina (Erik Brynjolfsson and Brian McAfee, 2011; 2014) uma tradução lenta, a partir do modelo da alimentação lenta (slow food – cf Cronin, 2017, cap. 2)?

4. Crítica da dialética entre ciência e engenharia. A Inteligência Artificial é, de maneira ambivalente, ao mesmo tempo um projeto interdisciplinar teórico (de análise da inteligência, como as ciências cognitivas estudam o conhecimento, ou as ciências médicas o ser humano são e doente) e um projeto prático de engenharia, pesquisa e produção de alguns resultados (Bachimont, 2000). Essa tensão é ilustrada no percurso do robótico Rodney Brooks que, a partir da IA clássica (anos 80-90, fundada sobre as várias etapas a serem realizadas sobre representações internas), orientou (nos anos 2000) suas pesquisas numa direção menos estritamente calculadora, e mais inspirada por modelos biológicos e behavioristas (adaptação). A IA parece assim hesitar entre uma ambição teórica e uma vocação prática, a ciência e a engenharia.
5. Crítica de uma certa visão da tecnologia (automatização). A Inteligência Artificial é um tópico abrangente que tende a referir-se a tudo e ao seu oposto ao mesmo tempo. No entanto, o termo tende a concentrar-se numa certa tendência tecnológica que, embora possa ser majoritária, não é a única. É a *automatização* (por exemplo, 99,99% das tecnologias da linguagem baseiam-se nos postulados do processamento *automático* da linguagem), que Simondon sublinhou claramente ser apenas a nível zero da interação homem-máquina – e, portanto, nem a única nem a mais interessante. Existem também relações que não são de delegação exaustiva, mas sim de sugestão, assistência, apoio (Chateauraynaud, 2019), por exemplo as ferramentas de ajuda à tradução (Computer Assisted Translation Tools).
6. Crítica pela história. Compreender, de maneira documentada, as origens da IA e sua evolução permite relativizar as supostas novidades, e assim evitar cair num certo fetichismo tecno-utopista. Essa crítica é mais erudita, mas extremamente sugestiva, pois coloca o estado atual da IA (com suas ambições mais recentes) numa pers-

pectiva longa, em relação com o projeto inicial. Esse desvio mostra que a IA foi imaginada para tentar modelizar e entender melhor a inteligência humana, e de jeito nenhum para se substituir a ela (como ela é frequentemente vista hoje), ou ser imitada pela inteligência humana (cf. Bénel, 2020).

7. Crítica pelo imaginário. E hoje fundamental analisar as representações sociais da IA, que carregam muitos fantasmas e angústias exageradas, mas reveladoras. Se trata de uma crítica promissora que se interessa nas representações associadas ao projeto mesmo da IA, no espírito dos vários projetistas, utilizadores, e também prospectivistas que, da literatura ao cinema de ciência-ficção, utilizam o tópico para investigar certos aspectos despercebidos, certas consequências inesperadas, e até certas lógicas latentes. Se trata de uma análise das imagens seduzentes ou ilusórias, das mitologias, e até das ideologias, tal que, por exemplo: a máquina pensa (tem uma consciência, uma liberdade, uma vida, etc.) – esse fantasma de inferioridade intelectual permitiu disseminar ferramentas, que são na verdade de inteligência fraca, mas perfeitamente capazes de juntar os nossos dados pessoais para manipular as nossas decisões, valores e desejos, na linha de uma nova servidão voluntária. Existe por exemplo um projeto, coordenado pelo professor Alexandre Gefen (CNRS, França) sobre a história “cultural” da IA, que vai nessa direção: (CulturaIA: <https://cis.cnrs.fr/culturaia/>), como também os trabalhos de antropologia das técnicas que colocam a IA no conjunto das relações sociais, e até das dimensões simbólicas que a constituem e a tornam possível.
8. Críticas dos efeitos éticos dessas tecnologias. O desenvolvimento da inteligência artificial fez com que muitas decisões pudessem, pelo menos supostamente, ser automatizadas em diversos âmbitos da sociedade. Contemporaneamente há um grande investimento em carros autônomos, por exemplo. Esses veículos devem decidir não só o melhor caminho, como também o que fazer em caso de acidente iminente. Na internet os algoritmos escolhem também quais conteúdos devem ter destaque ou não; nesse contexto há muito debate sobre os critérios e os impactos das escolhas feitas pela inteligência

artificial. Sobre essas questões, surgem vários problemas: quem beneficia dessa captura da atenção (Brennan, 2020), e também quem responsabilizar? (Coeckelbergh, 2020, p. 111) Visto que a máquina não possui senso moral, livre arbítrio, emoções etc., é preciso responsabilizar os humanos que a programaram. Contudo, o problema fica mais complicado quando se percebe que a cadeia de agências presente na programação é exaustivamente longa e complexa. Esses assuntos provocam reflexões até nos níveis nacionais, por exemplo na França (CNIL, 2017).

9. Crítica dos impactos jurídicos dessas tecnologias. Enquanto norma da esfera prática, junto com a ética ou a política, o direito está afetado pelo impacto da IA. De maneira direta, para começar, no sentido que os avanços da IA no domínio jurídico correspondem, como na medicina, as promessas da automatização do juízo, das antecipações ou previsões, e da personalização. É nesse sentido que o jurista Antoine Garapon chama a atenção da profissão sobre os riscos de uma tecnologia que deixaria de dar uma assistência a decisão para tomar decisões sozinha (Garapon e Lassègue, 2018). Além desse impacto sobre o judiciário, são todas *as regras jurídicas sobre a IA* que devem ser examinadas, pois correspondem as estratégias diferentes (entre os EUA, a EU, a China e o Brasil, por exemplo). A questão é particularmente sensível para países como o Brasil ou no continente africano, que não possuem uma tecnologia soberania para segurar o controle nacional de algumas utilizações sensíveis da IA.
10. Crítica dos efeitos políticos dessas tecnologias. É indiscutível que o mundo político se transformou após a popularização das redes sociais, de modo que a presença na internet e a preocupação de como a imagem é reproduzida no meio digital são inevitáveis para quem almeja iniciar ou manter uma carreira política (Empoli 2020). A partir de meados da década passada, a partir da publicização da influência de empresas como a Cambridge Analytica nas eleições dos Estados Unidos e no Brexit, tornou-se necessário debater o funcionamento das redes sociais (enquanto alimentadas por IA malignas) e como elas têm interferido na vida pública (Kaiser, 2020). Pensar o papel das *fake News*, das câmaras de eco e da leniência com dis-

curso de ódio tornou-se fundamental para a teoria social contemporânea. Daí a importância de medir o papel das IA nesse domínio também, como também naquele de luta contra as desigualdades políticas (O’Neil, 2016). Como justificar, aliás, a proibição da utilização de IA por certos países (China, Rússia, Itália para ChatGPT, por exemplo)?

A NOÇÃO DE CONJECTURA

A quarta crítica esboçada (cf *supra*) merece uma explicação mais detalhada, para destacar melhor a originalidade dela, através de um desvio histórico. A noção de *conjectural* (em grego antigo, *stokhastikos*, de *stokhos*: objetivo, alvo, conjectura) é etimologicamente dotada de uma rica ambiguidade, que a evolução do termo preservou durante a Idade Média e o Renascimento. Ela possui uma dimensão que é, ao mesmo tempo, objetiva e subjetiva, gnosiológica e ontológica, aproximada e vaga. No sentido mais amplo, conjectura refere-se à avaliação de uma situação de forma aproximada, sem qualquer pretensão de certeza, ainda que o senso comum a associe mais a uma hipótese ainda não confirmada – esse é também o sentido atribuído por Popper (2006).

Alguns significados técnicos podem ser instrutivos: em matemática, a conjectura refere-se a uma afirmação para a qual não há demonstração, mas que se acredita firmemente ser verdadeira, seja por não haver contraexemplo conhecido ou por generalizar resultados já demonstrados. No direito, refere-se a uma consequência que a lei (ou o juiz) infere de um fato conhecido para um fato desconhecido: a presunção, portanto, isenta a pessoa que ela favorece de qualquer prova. Na filologia, a conjectura é uma forma de correção realizada pelo estudioso ao se deparar com uma lacuna ou passagem alterada em um manuscrito: baseia-se em uma hipótese plausível, construída a partir do conhecimento do idioma, dos hábitos do autor e do contexto. Finalmente, na retórica, corresponde a um dos momentos de determinação do estado de uma questão no raciocínio tópico, antes da qualificação (Thouard, 2015).

De modo geral, a conjectura refere-se ao aspecto inacabado do conhecimento e do raciocínio humano, ambos envolvidos em um processo de retificação e aprimoramento perpétuos. Foi nesse sentido que Nicolas de Cues distinguiu o conjectural do confuso e do certo (*De conjecturiis*, 1444), e que Bacon, no *Novum Organum*, insistiu em seu papel na ampliação do domínio da experiência – uma ideia criticamente adotada por Popper. Uma definição semelhante pode ser encontrada na medicina, pelo menos até a virada formal do provável na era clássica (Corteel, 2020a).

Associado à conjectura, o termo *probabilis* significava “aquilo que é digno de aprovação”, em um sentido muito amplo, com uma nuance de autoridade – um uso do qual só restam vestígios negativos (descrever um evento como “improvável” é considerá-lo inacreditável). Na maioria das vezes, era uma questão de opinião ou conhecimento com base em signos (ainda mal distinguidos de indícios), como na medicina, astrologia, jurisprudência ou história. É provável que a exclusão tomista e escolástica da opinião (relacionada a fatos contingentes) e da ciência (visando a verdades necessárias) tenha retardado o desenvolvimento de uma reflexão sobre o conhecimento probabilístico (de eventos singulares), embora isso já fosse conhecido pela tradição, principalmente por meio de referências à obra de Carnéades. Foi somente na era clássica que a noção realmente tomou forma e se tornou mais precisa, por duas razões combinadas: o desenvolvimento de uma concepção formal de probabilidade e a criação do conceito de cálculo.

Inicialmente, com o trabalho de Pascal e Fermat sobre as partes (por volta de 1650), o termo assumiu um aspecto altamente matemático (Coqui 2015). Em seguida, o cálculo de chances se desenvolveu com Huyghens, Bernouilli, Leibniz (cujas reflexões sobre a “lógica da probabilidade” e os graus de prova anunciaram a lógica indutiva) e, no século XVIII, Bayes, Condorcet, Lagrange e Laplace. Foi só então que a noção, entendida como uma aproximação da certeza, perdeu sua nuance de autoridade e tornou-se distinta da noção de plausibilidade, associada apenas ao poder de persuasão (com Lambert e Kant). Quanto ao que *realmente* são probabilidades, interpretações rivais compartilham o en-

tendimento em termos de frequência (Granger, 1995), propensão (Popper, 1992) ou grau de crença subjetiva (Granger, 1988).

Em segundo lugar, a noção lógica de cálculo começou a ser inventada por Leibniz por volta da mesma época: na forma de pensamento “cego”, o filósofo de Hanover pretendia substituir as disputas estéreis por uma maneira rigorosa de pensar e decidir, por meio de cálculos. Esse programa de lógica formal só foi realmente revivido no final do século XIX, encontrando uma conclusão particularmente convincente no trabalho de Turing. Sua máquina abstrata é o modelo para “computação”, e sua encarnação concreta gradual marcou a história da computação, desde as primeiras conquistas até os supercomputadores e microcomputadores pessoais de hoje. Nesse ponto, então, Leibniz estava certo: a computação é um tipo de pensamento – se é o único, ou se deve receber algum tipo de preeminência, é mais duvidoso.

Os dois conceitos se uniram nos tempos modernos, começando no século XIX com a disseminação do pensamento estatístico – um ponto que foi particularmente bem estudado (Hacking, 1990, 2002; Corteel, 2020b) – e no século XX com o desenvolvimento da inteligência artificial. Na sua forma mais recente, a inteligência artificial pode ser definida pela combinação de cinco conceitos: o conceito lógico de cálculo; o conceito matemático de probabilidade; o conceito formal de conjectura (provável); o conceito (indutivo) de previsão; e uma visão automatizada da tecnologia. Os desenvolvimentos recentes em aprendizagem profunda, vinculados aos avanços em algoritmos e à massa de dados consultados, com as promessas associadas de geração inventiva, na realidade simplesmente estendem os mesmos mecanismos, pois se baseiam nos mesmos princípios. No campo da tradução, por exemplo, pode-se demonstrar que, independentemente do progresso recente, os limites *estruturais* dessa abordagem impossibilitam a compreensão de textos minimamente complexos, como a passagem bíblica sobre o assassinato de Abel por Caim (Bénel, Falip, Lacour, 2024). Portanto, há dimensões do pensamento e do raciocínio que estão definitivamente fora do alcance da IA.

Esses são exatamente os tipos de questões que se enquadram no domínio do *plausível* e que a tradição não abandonou completamente em favor do provável formalizado. A plausibilidade ainda está ligada a um esforço de aproximação interpretativa, ponderada, contrastiva e reflexiva de uma situação ou evento (que pode ser futuro no caso de uma tentativa de previsão). Com suas origens hermenêuticas, o termo manteve a noção de semelhança, rejeitou probabilidades formais externas e alimenta-se de seu caminho interpretativo com uma sensibilidade para detalhes significativos. Pascal, o próprio inventor do cálculo de probabilidades, teve o cuidado de distinguir entre o espírito da geometria – e o espírito da exatidão, útil na física –, um espírito de *fineza* adaptado ao conhecimento das maneiras e dos homens na sociedade. Longe de estar fora do escopo da racionalidade, este último constitui uma de suas dimensões porque, em ambos os casos, é uma questão de raciocínio a partir de princípios (os do espírito de fineza são familiares, mas difíceis de perceber e numerosos). O espírito de fineza não demonstra, mas, de maneira tácita, penetra subitamente na realidade como um todo, com um julgamento inventivo e flexível, até mesmo insolente, que o tira da confusão. Depois de Pascal, outros epistemólogos das ciências da cultura (Geertz, com a interpretação (Lacour, 2020); Ginzburg, com o conhecimento indexical (Ginzburg, 1989, Thouard, 2007); Passeron e Revel, com a casuística (Lacour, 2005)) insistiram nessa dimensão, que é ao mesmo tempo alternativa e racional.

O reconhecimento da *ambivalência insuperável da noção de conjectura* tem implicações tecnológicas fortes, pois limita imediatamente as pretensões da IA contemporânea. Nessa perspectiva, gostaria, em particular, de mostrar que uma abordagem diferente para as humanidades digitais é possível, desde que concordemos em ancorar seus princípios de forma imanente nos gestos intelectuais tradicionais dos usuários, em vez de pretender aplicar sem reflexão (de forma autoritária ou implícita), uma ciência da computação derivada da engenharia ortodoxa. Desse ponto de vista, as ciências da cultura *estão nos forçando* a inventar uma forma de computação centrada no ser humano, colaborativa e hermenêutica, e que prioriza o *design* do usuário em relação à engenharia tradicional.

CONCLUSÃO

A principal ambição desse artigo era de identificar os postulados que orientam as ferramentas contemporâneas de tradução por IA, especialmente aquelas de última geração, como as baseadas em redes neurais, aprendizagem profundo, grandes modelos de linguagem (ChatGPT). Trata-se de compreender os aspectos fundamentais do seu funcionamento, isto é, identificar e explicar as características do seu modo de “pensar” e listar as principais críticas que recebem. A IA designa simultaneamente um campo de conhecimento e uma certa tecnologia, que opera por meio de cálculo probabilístico, modelagem, conjectura e raciocínio indutivo, com a ambição de previsão e de automatização. Compreender a natureza dessas noções é fundamental, pois cada uma delas possui limitações estruturais, as quais são, em essência, insuperáveis. O conceito de cálculo enfrenta limitações estruturais demonstradas pelo *Teorema de Gödel* (perda da completude na matemática). O conceito de probabilidade, por sua vez, pode ser interpretado de várias maneiras (objetiva, subjetiva, ou variantes semi-subjetivas), o que implica opções diferentes na maneira de conceber a relação entre os resultados do cálculo e a realidade. A modelagem estatística, embora possível, deve ser completada por elementos de interpretação; enquanto o conceito de presunção (*stoichazestai*, em grego), é mais rico do que a sua mera redução a um cálculo probabilístico, visto que pretende também ser uma conjectura, uma avaliação refinada, interpretativa, comparativa (no sentido contrastivo do termo), contextual.

Além disso, a indução não atinge o universal, mas apenas o geral, que admite exceções (variações irregulares). Por fim, a ambição de “pensamento”, já idealizada por Leibniz para caracterizar o cálculo, deve ser considerada como uma metáfora a ser manipulada cuidadosamente, de forma a evitar paralogismos que confundam a significação com as suas condições materiais. Enquanto às críticas mais “externas” (vieses, exploração social, impacto ecológico, hesitação entre ciência e engenharia, preconceito da automatização, análise da história da IA e de suas representações imaginárias, impactos éticos, jurídicos e políticos), elas

correspondem as abordagens clássicas da reflexão filosófica sobre tecnologia.

Por fim, frisei a originalidade da crítica da ambivalência da noção de conjectura, para mostrar que esta não pode ser reduzida a um cálculo do provável, pois se deixa também pensar racionalmente na forma da interpretação do plausível. Para justificar esse esforço de renovação da dualidade da razão, destacada por Pascal, entre “espírito de fineza” e “espírito de geometria”, teria que adicionar uma comprovação técnica, por exemplo no domínio das tecnologias da linguagem (Desjardin *et al.* 2020). De fato, a construção de uma ferramenta colaborativa e multilíngue para a tradução de textos culturais se inspira desses princípios de avaliação fina do plausível (Bénel *et al.* 2014, 2022). Trata-se da invenção de uma informática baseada no conceito de *documento* (a ser interpretado) e não de *data* (a ser calculada), ainda mais comprovada na versão 3 da ferramenta *TraduXio* (<https://traduxio.org>), fundamentada no software *HyperGlosae* (<https://github.com/Hypertopic/HyperGlosae>).

Não se trata, pois, de negar a relevância das ferramentas de tradução automática baseadas em *cálculo probabilístico*, mas sim enfatizar que as operações de interpretação, voltadas para a avaliação do plausível e aplicadas a uma abordagem conjectural da tradução, possuem maior pertinência. Em vez de buscar a instanciação de tendências, a tradução humana de qualidade foca no que é singular e diferente – de uma língua a outra, de um texto a outro, de um trecho a outro; valoriza a singularidade do sentido, a originalidade do contexto, a unicidade de significação visada e a operação de interpretação que procura o “equivalente sem identidade” (Ricoeur, 2003). No caso de textos culturais singulares, cuja semântica é complexa, a tecnologia não deve visar à automatização, mas sim privilegiar a *interação rica* entre humano e máquina: ela se torna instrumento de *sugestão, comparação, assistência ao juízo e à decisão*. A noção de *documento* é mais antiga do que a de *dados*, e as ciências da documentação não esperaram o advento da noção de computação (Turing) ou a revelação de seus limites (Gödel) para classificar, inserir notas, adicionar observações e comentários ou relacionar documentos entre si (Zacklad 2003).

REFEPRÊNCIAS

AMARAL, Olavo. O ChatGPT e os limites da inteligência artificial. *Revista piauí*, abr. 2023.

ANDERSON, Chris. The end of theory: the data deluge makes the scientific method obsolete. *Wired*, jun. 2008. Disponível em: <https://www.wired.com/2008/06/pb-theory/>. Acesso em: 12 nov. 2024. Ver também: CALUDE, Cristian; LONGO, Giuseppe. The deluge of spurious correlations in Big Data. CDMTCS-488. Auckland: University of Auckland, 2015. Disponível em: <https://www.cs.auckland.ac.nz/research/groups/CDMTCS/researchreports/488cris.pdf>. Acesso em: 12 nov. 2024.

BACHIMONT, Bruno. L'intelligence artificielle comme écriture dynamique: de la raison graphique à la raison computationnelle. In: PETITOT, Jean; FABBRI, Paolo (org.). *Au nom du sens: autour de l'œuvre d'Umberto Eco*. Paris: Grasset, 2000. p. 290-319.

BÉNEL, Aurélien. Modéliser ce qui résiste à la modélisation. *Revue Ouverte d'Intelligence Artificielle*, v. 1, n. 1, p. 71-88, 2020. Disponível em: https://roia.centre-mersenne.org/issues/ROIA_2020__1_1/. Acesso em: 12 nov. 2024.

BÉNEL, Aurélien; FALIP, Joris; LACOUR, Philippe. “Quand Abel tue Caïn”: ce qui échappe à la traduction automatique. In: LEVIN, François; OLLION, Étienne (org.). *Qu'est-ce qui échappe à l'intelligence artificielle*. Paris: Hermann, 2024. p. xx-xx.

BÉNEL, Aurélien; et al. Towards a participative platform for cultural texts translators. In: *Cross-Cultural Interaction: Concepts, Methodologies, Tools and Applications*. Hershey, PA: IGI Global, 2014. p. 1695-1705. DOI: 10.4018/978-1-4666-4979-8.ch096. Disponível em: <https://hal.science/hal-02362473>. Acesso em: 12 nov. 2024.

BERNS, Thomas; ROUVROY, Antoinette. Gouvernamentalité algorithmique et perspectives d'émancipation. *Réseaux*, n. 177, p. 163-196, 2013. Disponível em: <https://www.cairn.info/revue-reseaux-2013-1-page-163.htm>. Acesso em: 12 nov. 2024.

BIHOUIX, Philippe. *L'âge des low tech: vers une civilisation techniquement soutenable*. Paris: Seuil, 2014.

BONNEUIL, Christophe; FRESSOZ, Jean-Baptiste. *L'événement Anthropocène: la Terre, l'histoire et nous*. Paris: Seuil, 2013.

BRENNAN, Matthew. *Attention factory: the story of TikTok and China's ByteDance*. Independently published, 2020.

BRYNJOLFSSON, Erik; McAFEE, Andrew. *The race against the machine*. Lexington, MA: Digital Frontier Press, 2011.

BRYNJOLFSSON, Erik; McAFEE, Andrew. *The second machine age: work, progress, and prosperity in a time of brilliant technologies*. New York: W. W. Norton, 2014.

CARDON, Dominique. *À quoi rêvent les algorithmes: nos vies à l'heure des big data*. Paris: Seuil, 2015.

CARDON, Dominique. Intelligence artificielle. In: *Culture numérique*. 2019. p. 385-398.

CARDON, Dominique; COINTET, Jean-Philippe; MAZIÈRES, Antoine. La revanche des neurones: l'invention des machines inductives et la controverse de l'intelligence artificielle. *Réseaux*, n. 211, p. 173-220, 2018.

CHATEAURAYNAUD, Francis. Petit traité de contre-intelligence artificielle: retour sociologique sur des expérimentations numériques. *Zilsel*, n. 5, 2019. Disponível em: <https://www.cairn.info/revue-zilsel-2019-1-page-174.htm>. Acesso em: 12 nov. 2024.

CNIL – Commission Nationale de l'Informatique et des Libertés (França). *Comment permettre à l'Homme de garder la main? Rapport sur les enjeux éthiques des algorithmes et de l'intelligence artificielle*. Dez. 2017. Disponível em: <https://www.cnil.fr/fr/comment-permettre-lhomme-de-garder-la-main-rapport-sur-les-enjeux-ethiques-des-algorithmes-et-de>. Acesso em: 12 nov. 2024.

COECKELBERGH, Mark. *AI ethics*. Cambridge, MA: MIT Press, 2020.

COQUI, Guillaume. Probabilité. In: BERNER, Christian; THOUARD, Denis (org.). *L'interprétation: un dictionnaire philosophique*. Paris: Vrin, 2015. p. 343-351.

CORTEEL, Mathieu. La médecine comme *ars conjectandi*. *Histoire, médecine et santé*, n. 15, p. 109-124, 5 nov. 2020. DOI: 10.4000/hms.2236. Disponível em: <https://doi.org/10.4000/hms.2236>. Acesso em: 12 nov. 2024.

CORTEEL, Mathieu. *Le hasard et le pathologique*. Paris: Presses de Sciences Po, 2020b.

CORTEEL, Mathieu; KYROU, Ariel; MOULIER-BOUTANG, Yann. Pour une culture critique de l'IA. *Multitudes*, 2020. Disponível em: <https://www.cairn.info/revue-multitudes-2020-1-page-51.htm>. Acesso em: 12 nov. 2024.

CORTEEL, Mathieu. Pourquoi les IA ne pensent-elles pas? *AOC Media*, jan. 2023. Disponível em: <https://aoc.media/>. Acesso em: 12 nov. 2024.

CRONIN, Michael. *Eco-translation: translation and ecology in the age of the Anthropocene*. London: Routledge, 2017. DOI: 10.4324/9781315689357. Acesso em: 12 nov. 2024.

DESCLÉS, Jean-Pierre. Vers un calcul des significations dans l'analyse des langues. *Revista de Filosofia Moderna e Contemporânea*, v. 8, n. 1, 2020. Disponível em: <https://periodicos.unb.br/index.php/fmc/article/view/31016>. Acesso em: 12 nov. 2024.

DESJARDINS, Renée; LARSONNEUR, Claire; LACOUR, Philippe. *When translation goes digital: case studies and critical reflections*. London; New York: Palgrave Macmillan, 2020.

EMPOLI, Giuliano da. *Os engenheiros do caos*. Trad. Arnaldo Bloch. São Paulo: Vestígio, 2020.

FODOR, Jerry. *The language of thought*. Cambridge, MA: Harvard University Press, 1975.

GARAPON, Antoine; LASSÈGUE, Jean. *Justice digitale: révolution graphique et rupture anthropologique*. Paris: PUF, 2018.

GEFEN, Alexandre. *Vivre avec ChatGPT*. Paris: Éditions de l'Observatoire, 2023.

GINZBURG, Carlo. Traces. Racines d'un paradigme indiciare. In: GINZBURG, Carlo. *Mythes, emblèmes, traces: morphologie et histoire*. Paris: Flammarion, 1989.

GRANGER, Gilles-Gaston. *Essai d'une philosophie du style*. Paris: Odile Jacob, 1988.

GRANGER, Gilles-Gaston. Must a science of artificial intelligence be necessarily reductionist? In: AGAZZI, Evandro (ed.). *The problem of reductionism in science*. Dordrecht: Kluwer Academic Publishers, 1991. p. 173-182.

GRANGER, Gilles-Gaston. *Les conditions protologiques des langues naturelles*. In: GRANGER, Gilles-Gaston. *Formes, opérations, objets*. Paris: Vrin, 1994.

GRANGER, Gilles-Gaston. *Le possible, le probable et le virtuel: essai sur le non-actuel dans la pensée scientifique*. Paris: Odile Jacob, 1995.

HACKING, Ian. *The taming of chance*. Cambridge: Cambridge University Press, 1990.

HACKING, Ian. *L'émergence des probabilités*. Paris: Seuil, 2002.

KAISER, Brittany. *Manipulados*. Trad. Roberta Clapp; Bruno Fiuzza. Rio de Janeiro: HarperCollins, 2020.

LACOUR, Philippe. Penser par cas, ou comment remettre les sciences sociales à l'endroit. *EspacesTemps.net – Revue électronique des sciences humaines et sociales*, 31 mai 2005. Disponível em: <https://www.espacestemp.net/articles/remettre-les-sciences-sociales-a-endroit/>. Acesso em: 12 nov. 2024.

LACOUR, Philippe. Interprétation, traduction et rationalité clinique dans l'anthropologie de Clifford Geertz. *Texto!* [en ligne], 2020. Disponível em: http://www.revue-texto.net/docannexe/file/3352/texto_lacour_geertz_modif.pdf. Acesso em: 12 nov. 2024.

LADRIÈRE, Jean. *Les limitations internes des formalismes: étude sur la signification du théorème de Gödel et des théorèmes apparentés dans la théorie des fondements des mathématiques*. Paris: J. Gabay, 1992. (Les grands classiques). Edição original: Louvain; Paris: Nauwelaerts-Gauthier-Villars, 1957.

MCNEILL, John R. *Something new under the sun: an environmental history of the twentieth-century world*. London: Norton, 2001.

MITCHELL, Tom. *Machine learning*. New York: McGraw-Hill, 1997.

O'NEIL, Cathy. *Weapons of math destruction: how big data increases inequality and threatens democracy*. New York: Penguin Books, 2016.

PERRIGO, Billy. OpenAI used Kenyan workers on less than \$2 per hour to make ChatGPT less toxic. *TIME*, 18 jan. 2023. Disponível em: <https://time.com/6247678/openai-chatgpt-kenya-workers/>. Acesso em: 12 nov. 2024.

POPPER, Karl. *Un univers de propensions*. Trad. Alain Boyer. Combas: L'Éclat, 1992.

POPPER, Karl. *Conjectures et réfutations: la croissance du savoir scientifique*. Trad. Michelle-Irène. Paris: Payot, 2006.

PUTNAM, Hilary. *Mind, language and reality: philosophical papers*. Cambridge: Cambridge University Press, 1975.

RICOEUR, Paul. *Sur la traduction*. Paris: Bayard, 2003.

SCHWARTZ, Roy; DODGE, Jesse; SMITH, Noah A.; ETZIONI, Oren. Green AI. *arXiv preprint arXiv:1907.10597*, 13 ago. 2019. Disponível em: <https://arxiv.org/abs/1907.10597>. Acesso em: 12 nov. 2024.

SEARLE, John. *The rediscovery of mind*. Cambridge, MA: MIT Press, 1994.

SIMONDON, Gilbert. *Du mode d'existence des objets techniques*. Paris: Aubier, 2012.

SUMPTER, David. *Dominados pelos números*. Trad. Anna Maria Sotero; Marcello Neto. Rio de Janeiro: Bertrand Brasil, 2019.

STIBBE, Arran. *Ecolinguistics: language, ecology and the stories we live by*. London; New York: Routledge, 2015.

STRUBELL, Emma; GANESH, Ananya; McCALLUM, Andrew. Energy and policy considerations for deep learning in NLP. *arXiv preprint arXiv:1906.02243*, jun. 2019. Disponível em: <https://arxiv.org/abs/1906.02243>. Acesso em: 12 nov. 2024.

THOUARD, Denis. Conjecture. In: BERNER, Christian; THOUARD, Denis (org.). *L'interprétation: un dictionnaire philosophique*. Paris: Vrin, 2015. p. 102-107.

THOUARD, Denis. Réflexions sur une hypothèse vingt-cinq ans après. In: THOUARD, Denis. *L'interprétation des indices: enquête sur le paradigme indiciaire avec Carlo Ginzburg*. Villeneuve d'Ascq: PUL, 2007.

VILLANI, Cédric (dir.). *Donner un sens à l'intelligence artificielle: pour une stratégie française et européenne*. Rapport de la mission parlementaire, mar. 2018. Disponível em: https://www.aiforhumanity.fr/pdfs/9782111457089_Rapport_Villani_accessible.pdf. Acesso em: 12 nov. 2024.

ZACKLAD, Manuel. Processus de documentarisation dans les Documents pour l'Action (DopA): statut des annotations et technologies de la coopération associées (nouvelle version corrigée). 2003. Disponível em: <https://www.enssib.fr/bibliotheque-numeri->

que/documents/1209-processus-de-documentation-dans-les-docu-
ments-pour-l-action-dopa.pdf. Acesso em: 12 nov. 2024.

Recebido em 07 de abril de 2025
Aprovado em 24 de julho de 2025
Publicado em 21 de outubro de 2025

RFMC