

AS MÁQUINAS PODEM JULGAR? do jogo da imitação ao jogo do julgamento

Kaio Vinícius da Silva Vasconcelos¹
Azel Lins de Oliveira¹
Guilherme Antônio Lopes de Oliveira²

RESUMO

Em 1950, Alan Turing substituiu a pergunta “as máquinas podem pensar?” por um problema operacional: seria possível distinguir, pelo comportamento, uma máquina de um ser humano? Este artigo transpõe essa estratégia para o Direito e concentra a análise no denominado jogo do julgamento. Como experimento mental, propõe-se que um mesmo caso seja submetido separadamente a um magistrado humano e a um sistema de inteligência artificial, com posterior anonimização das decisões e avaliação cega por juristas. O objetivo não é provar que a máquina possui consciência ou que exerce jurisdição, mas verificar o que realmente se demonstra quando uma decisão artificial se torna exteriormente comparável à humana. A análise sustenta que passar no jogo do julgamento comprovaria, no máximo, equivalência funcional em determinados critérios observáveis. Isso não elimina a diferença entre produzir uma decisão, reproduzir comportamentos associados ao julgamento e possuir legitimidade para exercer poder jurisdicional. Conclui-se que a questão mais útil não é simplesmente saber se a inteligência artificial “julga”, mas identificar até onde a imitação do produto do julgamento pode avançar sem ser confundida com o próprio exercício legítimo da jurisdição.

Palavras-chave: Inteligência artificial. Alan Turing. Jogo da imitação. Jogo do julgamento. Poder Judiciário.

1 INTRODUÇÃO

Em 1950, Alan Mathison Turing publicou *Computing Machinery and Intelligence* e iniciou sua investigação com a pergunta “As máquinas podem pensar?”. A força do texto, contudo, não está apenas na pergunta. Turing percebeu que definir previamente os termos “máquina” e “pensar” poderia transformar o debate em uma disputa semântica. Em vez de tentar resolver diretamente a essência do pensamento, substituiu a questão por outra suscetível de observação: o jogo da imitação (TURING, 1950, p. 433).

Na formulação turingiana, o interrogador permanece separado dos demais participantes e deve identificá-los a partir das respostas recebidas. Quando uma máquina ocupa uma das

posições, o foco deixa de ser aquilo que ela é internamente e passa a ser aquilo que consegue fazer durante a interação. A mudança é metodológica: não se prova consciência, sentimento ou identidade entre mente humana e máquina; testa-se a capacidade de produzir um comportamento que, sob certas condições, não seja distinguido de um comportamento humano (TURING, 1950, p. 433-434).

O desenvolvimento contemporâneo da inteligência artificial torna essa operação especialmente provocativa para o Direito. Sistemas computacionais já interpretam linguagem natural, organizam informações, identificam relações entre dados e produzem textos argumentativos. Como a atividade jurídica também depende de linguagem, interpretação, comparação, fundamentação e decisão, surge uma adaptação quase inevitável da pergunta de Turing: se antes se perguntava se máquinas poderiam pensar, pode-se perguntar agora se máquinas podem julgar.

A formulação, porém, repete o problema conceitual original. Se “julgar” for definido como uma atividade necessariamente humana, dependente de consciência, experiência e sensibilidade, a resposta negativa estará embutida na própria definição. No extremo oposto, se julgar significar apenas receber informações e produzir uma conclusão, inúmeros sistemas automatizados já poderiam ser chamados de julgadores. Os dois caminhos simplificam demais a atividade jurisdicional.

Por isso, este artigo não tenta decidir de início o que constitui a essência metafísica do julgamento. O foco é mais restrito e, ao mesmo tempo, mais verificável: quais manifestações exteriormente associadas ao julgamento podem ser reproduzidas por uma inteligência artificial e o que seria legítimo concluir caso essa reprodução se tornasse indistinguível da produção humana?

A ideia central é o “jogo do julgamento”, uma releitura jurídica do jogo da imitação. Um mesmo caso seria entregue a um magistrado humano e a uma inteligência artificial. Cada um produziria, de forma independente, uma decisão fundamentada. Depois, as decisões seriam anonimizadas e submetidas a avaliadores juridicamente qualificados, que tentariam identificar qual texto foi produzido pelo humano e qual foi produzido pela máquina. A hipótese desloca o debate da pergunta abstrata “a IA julga?” para uma comparação concreta de desempenho.



O ponto decisivo aparece depois do teste. Se os avaliadores não conseguirem distinguir consistentemente as decisões, isso significa que a máquina julgou ou apenas que conseguiu imitar o produto de um julgamento? A distinção entre essas duas conclusões organiza todo o trabalho. O artigo sustenta que a equivalência observável pode demonstrar capacidade funcional, mas não resolve, sozinha, questões de consciência, responsabilidade, competência e legitimidade jurisdicional.

2 OBJETIVO GERAL

Analisar, a partir da releitura do jogo da imitação de Alan Turing, o denominado jogo do julgamento, investigando o significado e os limites de uma eventual indistinguibilidade entre decisões judiciais produzidas por magistrados humanos e por sistemas de inteligência artificial, especialmente quanto à diferença entre imitação funcional do julgamento e exercício legítimo da jurisdição.

3 MÉTODO

A pesquisa possui natureza qualitativa, exploratória e teórico-bibliográfica, desenvolvida sob a forma de ensaio teórico. O eixo principal é Computing Machinery and Intelligence, de Alan Turing, especialmente a substituição da pergunta conceitual pelo jogo da imitação e as objeções relacionadas à consciência, à programação, à informalidade do comportamento e à aprendizagem das máquinas (TURING, 1950). A Teoria Pura do Direito, de Hans Kelsen, é utilizada apenas como apoio para evitar a redução caricatural do julgamento à aplicação mecânica da literalidade normativa (KELSEN, 2009). Não é realizado experimento empírico. O jogo do julgamento funciona como experimento mental: primeiro reconstrói-se a lógica do teste turingiano; em seguida, transpõe-se essa estrutura para a decisão judicial; por fim, examinam-se as conclusões que poderiam ou não ser extraídas caso uma inteligência artificial apresentasse desempenho exteriormente equivalente ao de um magistrado.

4 RESULTADOS E DISCUSSÃO

A releitura do texto de Turing indica que a principal contribuição do jogo da imitação não é oferecer uma definição definitiva de pensamento, mas obrigar o debate a abandonar respostas circulares. Aplicado ao Judiciário, o mesmo procedimento impede duas conclusões fáceis: negar qualquer possibilidade de julgamento artificial apenas porque a máquina não é humana, ou afirmar que ela julga simplesmente porque consegue produzir uma sentença. Entre



esses extremos existe uma questão mais precisa: qual é o alcance probatório de uma decisão artificial que se mostra funcionalmente comparável à decisão humana?

4.1 DO JOGO DA IMITAÇÃO AO JOGO DO JULGAMENTO

O jogo do julgamento conserva a operação epistemológica central de Turing, mas altera o objeto comparado. Em vez de respostas de uma conversa, comparam-se decisões jurídicas. Um mesmo conflito, com os mesmos fatos, argumentos e elementos probatórios relevantes, é apresentado separadamente a um magistrado e a uma inteligência artificial. Autor e réu não competem no experimento; fornecem o conflito. O terceiro papel pertence ao avaliador, que recebe os textos sem identificação de autoria.

O teste não exige que humano e máquina cheguem à mesma conclusão. Essa exigência seria artificial, pois dois magistrados humanos podem interpretar o mesmo caso de maneira diferente sem que um deles deixe de exercer atividade jurisdicional. O que se observa é se a decisão artificial apresenta estrutura argumentativa, enfrentamento dos argumentos, utilização dos elementos do caso, interpretação normativa e fundamentação suficientemente plausíveis para ser reconhecida como produto possível de um julgador humano.

Essa diferença também impede que o jogo seja reduzido a uma prova de “acerto”. Em conflitos jurídicos complexos, nem sempre existe uma única solução mecanicamente extraível da norma. O jogo examina a qualidade funcional do desempenho decisório e a capacidade de a máquina participar do espaço argumentativo no qual decisões juridicamente defensáveis são construídas.

Se avaliadores qualificados não conseguirem identificar de modo consistente a origem artificial, a conclusão mínima será relevante: segundo os critérios observáveis adotados, a máquina reproduziu manifestações associadas ao julgamento humano. É exatamente aqui, porém, que o teste deve parar. A incapacidade de distinguir dois produtos não demonstra automaticamente identidade entre os processos, agentes ou fundamentos que os produziram.

4.2 O QUE SIGNIFICARIA “PASSAR” NO JOGO DO JULGAMENTO?

Passar no jogo do julgamento significaria parecer julgar segundo critérios previamente definidos. Não significaria provar que a máquina possui consciência, experiência subjetiva, responsabilidade moral ou competência jurídica. A distinção é semelhante ao cuidado exigido pelo próprio jogo da imitação: um critério comportamental é útil para comparar desempenho,

mas se torna enganoso quando a conclusão é ampliada para além daquilo que foi efetivamente observado.

A objeção da consciência ilustra esse limite. Turing discute a tese de que uma máquina só poderia ser considerada pensante se tivesse consciência daquilo que faz e experimentasse internamente pensamentos ou emoções correspondentes. Ele observa que o acesso direto à experiência interna também é impossível entre seres humanos, de modo que a vida social depende amplamente de inferências feitas a partir do comportamento (TURING, 1950, p. 445-446).

No julgamento, a objeção reaparece sob a forma da sensibilidade. Um magistrado possui biografia, experiências, emoções e referências culturais que uma máquina não possui. Mas a pergunta relevante para o jogo é outra: o Direito exige a experiência interna da sensibilidade ou exige que seus efeitos juridicamente relevantes apareçam na decisão? Uma IA pode não sentir compaixão e, ainda assim, reconhecer vulnerabilidade, consequências graves e circunstâncias que justificam tratamento individualizado.

Isso não torna a experiência subjetiva irrelevante. Apenas impede que sua relevância seja presumida como resposta suficiente. Se uma decisão artificial considera adequadamente as particularidades do caso e produz fundamentação comparável à de um julgador sensível, será necessário explicar por que a ausência da experiência interna invalida juridicamente o desempenho observado. O jogo, portanto, não resolve a objeção; obriga-a a se tornar mais específica.

4.3 IMITAR O JULGAMENTO É JULGAR?

A pergunta central do artigo surge justamente quando a máquina obtém sucesso. Uma decisão indistinguível da humana pode ser descrita, com segurança, como imitação funcional do julgamento. A passagem dessa afirmação para “a máquina julgou” exige uma definição adicional do próprio verbo julgar. O teste não elimina essa discussão; apenas impede que ela seja resolvida antes da análise do desempenho.

A objeção de Lady Lovelace reforça o problema. Segundo a formulação examinada por Turing, máquinas fariam apenas aquilo que seus criadores soubessem ordenar. Turing rejeita a conclusão de que programação implique ausência absoluta de novidade, lembrando que sistemas podem produzir resultados não antecipados individualmente por seus criadores

(TURING, 1950, p. 450-451). No contexto jurídico, isso significa que dizer “a máquina foi programada” não basta para provar que cada decisão é uma resposta previamente escrita.

A aprendizagem amplia a dificuldade. Turing propõe a ideia de uma “máquina-criança”, formada por estrutura inicial, educação e experiências, cujo comportamento se modifica durante o processo de aprendizagem (TURING, 1950, p. 456-459). Um sistema jurídico artificial poderia, em tese, aprender padrões a partir de normas, decisões, argumentos e correções sem possuir uma instrução individual para cada caso. Quanto maior essa capacidade, menos convincente se torna a imagem de um mecanismo que apenas consulta um catálogo fixo de respostas.

Por outro lado, a imprevisibilidade do resultado não cria autonomia jurídica. Uma máquina produzir algo não antecipado por seu programador demonstra apenas que programação não é sinônimo de previsibilidade absoluta. Não demonstra que o sistema se tornou sujeito de responsabilidade, titular de competência ou agente legitimado para exercer poder estatal. É nesse ponto que a imitação funcional encontra uma fronteira que o próprio desempenho não consegue atravessar.

4.4 REGRAS, CASOS CONCRETOS E LEGITIMIDADE

Turing também enfrenta a objeção da informalidade do comportamento e reconhece a dificuldade de construir um conjunto de regras capaz de determinar antecipadamente o que deve ser feito diante de toda circunstância concebível (TURING, 1950, p. 452). A aproximação com o Direito é evidente: normas gerais incidem sobre combinações fáticas que não podem ser integralmente previstas, e o julgamento frequentemente exige determinar alcance, peso e sentido diante do caso concreto.

Isso enfraquece a ideia de um juiz artificial concebido apenas como aplicador literal de regras. A referência a Kelsen não autoriza a caricatura de que a atividade jurídica seja simples reprodução automática do texto normativo. Mesmo em uma concepção fortemente normativista, interpretação e aplicação não se confundem com localizar uma frase e copiá-la para a decisão (KELSEN, 2009). Um sistema que apenas associasse fatos a respostas prontas poderia automatizar operações jurídicas sem, por isso, reproduzir adequadamente a complexidade do julgamento.

O jogo do julgamento, contudo, admite uma hipótese mais exigente: uma máquina capaz de considerar contexto, argumentos concorrentes, precedentes, particularidades e consequências juridicamente relevantes. Se ela alcançasse desempenho indistinguível do humano, a objeção puramente técnica perderia força. A discussão passaria então da capacidade para a legitimidade.

Essa mudança é decisiva. Produzir uma decisão, reproduzir funcionalmente comportamentos de julgamento e exercer jurisdição são três categorias diferentes. A primeira diz respeito à geração de um produto decisório. A segunda envolve desempenho suficientemente complexo para satisfazer critérios associados ao julgamento. A terceira envolve autoridade para impor consequências jurídicas, responsabilidade pelo exercício do poder, possibilidade de controle e inserção na estrutura institucional do Estado.

Assim, mesmo uma inteligência artificial que “passasse” no jogo não receberia, por consequência lógica, autorização para substituir autonomamente o magistrado. A competência técnica pode justificar uso como ferramenta de apoio, análise ou elaboração. A atribuição de poder decisório exige fundamentos jurídicos e institucionais que o teste comportamental não fornece. Confundir esses níveis seria transformar eficiência em legitimidade por um salto argumentativo.

O resultado mais importante do experimento mental é, portanto, deslocar a fronteira. Talvez o limite da inteligência artificial no Judiciário não esteja necessariamente no ponto em que ela deixa de conseguir produzir uma boa decisão. Pode estar no ponto em que a qualidade da decisão deixa de ser a única pergunta relevante e passa a importar quem possui autoridade para decidir, responder pelo ato e exercer legitimamente o poder jurisdicional.

4.5 O PARADOXO CENTRAL: QUANTO MELHOR A IMITAÇÃO, MAIOR O PROBLEMA JURÍDICO

O jogo do julgamento produz um paradoxo. Enquanto a decisão artificial é claramente inferior, a discussão jurídica permanece relativamente simples: a máquina pode ser tratada como ferramenta auxiliar cuja produção depende de revisão humana. O problema se torna mais difícil justamente quando a tecnologia melhora. Se o texto artificial passa a enfrentar argumentos, interpretar normas, individualizar circunstâncias e justificar conclusões em nível



comparável ao de um magistrado, a crítica baseada apenas na incapacidade técnica perde parte de sua força.

Nesse cenário, a pergunta deixa de ser “a máquina consegue fazer?” e passa a ser “o que muda juridicamente quando ela consegue?”. O jogo da imitação é útil porque mostra que aparência funcional e natureza do agente são planos diferentes. Um sistema pode reproduzir com enorme precisão o comportamento esperado sem que disso decorra identidade entre os agentes comparados. No jogo do julgamento, essa cautela é ainda mais importante, pois a decisão judicial não é apenas um texto tecnicamente bem construído: é também um ato institucional que produz efeitos sobre direitos, patrimônio, liberdade e deveres.

Imagine-se, por exemplo, que os avaliadores recebam cem pares de decisões e não consigam identificar a autoria acima do nível esperado pelo acaso. Esse resultado seria forte evidência de que os sinais textuais usados para reconhecer uma decisão humana deixaram de ser suficientes. Ainda assim, o experimento continuaria sem responder quem deve assumir responsabilidade por erro, viés, omissão ou consequência indevida. O teste mede aquilo que aparece na decisão; responsabilidade e legitimidade dependem também da posição jurídica ocupada por quem decide.

Essa separação evita dois exageros opostos. O primeiro é o excepcionalismo humano: presumir que todo resultado produzido por um magistrado contém, apenas por sua origem, uma qualidade inacessível à máquina. O segundo é o reducionismo tecnológico: concluir que, se a máquina reproduz o resultado observável, todas as diferenças restantes se tornam irrelevantes. O jogo do julgamento não autoriza nenhum dos dois. Ele exige que cada diferença alegada seja examinada conforme sua relevância para a atividade jurisdicional.

A sensibilidade oferece um exemplo claro. Se ela for defendida como elemento indispensável, será preciso indicar de que modo sua presença deve influenciar uma decisão e por que uma reprodução funcional desses efeitos seria insuficiente. O mesmo vale para autonomia: se a crítica consiste em afirmar que a máquina apenas segue condicionamentos anteriores, deve-se explicar por que isso a distingue, em termos juridicamente relevantes, de julgadores humanos também formados por educação, linguagem, precedentes, instituições e

limites normativos. A comparação não elimina diferenças, mas impede que elas sejam simplesmente presumidas.

No fim, o jogo funciona como uma espécie de filtro argumentativo. Antes dele, é fácil afirmar genericamente que máquinas nunca julgarão porque não são humanas. Depois dele, caso a imitação seja bem-sucedida, a objeção precisa indicar precisamente qual elemento ausente permanece indispensável e por quê. É nesse ponto que o debate deixa de ser apenas tecnológico ou filosófico e se torna propriamente jurídico: mesmo que a performance seja equivalente, a jurisdição continua sendo uma forma institucionalizada de poder cuja atribuição não pode ser deduzida de um teste de desempenho.

5 CONSIDERAÇÕES FINAIS

A releitura do jogo da imitação permite formular o problema das máquinas julgadoras de maneira mais precisa. Perguntar simplesmente se uma inteligência artificial “julga” tende a produzir respostas dependentes da definição adotada. O jogo do julgamento substitui essa disputa inicial por um teste conceitual de desempenho: diante do mesmo caso, seria possível distinguir de modo consistente uma decisão humana de uma decisão artificial?

Se a resposta fosse negativa, o resultado demonstraria algo importante, mas limitado. A máquina teria reproduzido manifestações exteriormente associadas ao julgamento em grau suficiente para confundir avaliadores qualificados. Não estaria demonstrado, apenas por isso, que possui consciência, sensibilidade humana, autonomia jurídica ou legitimidade para exercer jurisdição.

A oposição simples entre “máquinas julgam” e “máquinas não julgam” esconde etapas distintas. Sistemas artificiais podem produzir decisões; podem, em tese, aproximar-se funcionalmente do desempenho de julgadores; e ainda assim permanecer sem autoridade para exercer autonomamente a função jurisdicional. Capacidade técnica, equivalência funcional e legitimidade são problemas relacionados, mas não idênticos.

Por isso, o jogo do julgamento funciona menos como uma prova destinada a coroar um “juiz artificial” e mais como um instrumento para localizar a verdadeira controvérsia. Quanto melhor a máquina imitar o produto do julgamento, menos suficiente será recusá-la apenas com base na afirmação de que “não pensa” ou “não sente”. Ao mesmo tempo, quanto mais



**XVII Semana
Científica do
UniChrisfapi**

RESILÊNCIA E SAÚDE MENTAL NA FORMAÇÃO ACADÊMICA E PROFISSIONAL



Unichrisfapi
CENTRO UNIVERSITÁRIO CHRISFAPI

competente se tornar, mais necessário será impedir que desempenho seja confundido automaticamente com autorização para exercer poder.

A questão final, portanto, não é apenas se a máquina consegue elaborar uma sentença que pareça humana. É saber o que estamos autorizados a concluir quando ela consegue. O método turingiano ajuda a testar a fronteira da capacidade; o Direito continua responsável por definir a fronteira da legitimidade.

REFERÊNCIAS

KELSEN, Hans. Teoria pura do direito. Tradução de João Baptista Machado. 8. ed. São Paulo: WMF Martins Fontes, 2009.

TURING, Alan M. Computing machinery and intelligence. Mind, Oxford, v. 59, n. 236, p. 433-460, out. 1950.