

ESCOLA DE HUMANIDADES  
PROGRAMA DE PÓS-GRADUAÇÃO EM FILOSOFIA  
MESTRADO EM FILOSOFIA

GUILHERME AFFONSO FERREIRA DE CAMARGO

**JUSTIÇA E INTELIGÊNCIA ARTIFICIAL:  
ALGORITMOS E RESOLUÇÃO DE CONFLITOS**

Porto Alegre

2023

PÓS-GRADUAÇÃO - *STRICTO SENSU*



Pontifícia Universidade Católica  
do Rio Grande do Sul

GUILHERME AFFONSO FERREIRA DE CAMARGO

**JUSTIÇA E INTELIGÊNCIA ARTIFICIAL:  
ALGORITMOS E RESOLUÇÃO DE CONFLITOS**

Dissertação de mestrado, requisito parcial para aprovação no curso de Mestrado em Filosofia da Escola de Humanidades da Pontifícia Universidade Católica do Rio Grande do Sul, dentro da linha de pesquisa “Ética e Filosofia Política”.

Orientador: Prof. Dr. Thadeu Weber

Porto Alegre

2023

GUILHERME AFFONSO FERREIRA DE CAMARGO

**JUSTIÇA E INTELIGÊNCIA ARTIFICIAL:  
ALGORITMOS E RESOLUÇÃO DE CONFLITOS**

Dissertação de mestrado, requisito parcial para aprovação no curso de Mestrado em Filosofia da Escola de Humanidades da Pontifícia Universidade Católica do Rio Grande do Sul, dentro da linha de pesquisa “Ética e Filosofia Política”.

Orientador: Prof. Dr. Thadeu Weber

Aprovada em: 07 de Agosto de 2023

BANCA EXAMINADORA:

---

Prof. Dr. Nythamar Hilario Fernandes de Oliveira Jr.

---

Prof. Dr. Paulo Antonio Caliendo Velloso da Silveira

Porto Alegre

2023

## Ficha Catalográfica

C172j Camargo, Guilherme Affonso Ferreira de

Justiça e inteligência artificial : algoritmos e resolução de conflitos / Guilherme Affonso Ferreira de Camargo. – 2023.  
86.

Dissertação (Mestrado) – Programa de Pós-Graduação em Filosofia, PUCRS.

Orientador: Prof. Dr. Thadeu Weber.

1. Direito. 2. Tecnologia. 3. Justiça. 4. Algoritmos. 5.  
Inteligência artificial. I. Weber, Thadeu. II. Título.

“O real não é senão uma modalidade do possível”  
(François Ost)

“Ó rodas, ó engrenagens, r-r-r-r-r eterno!  
Forte espasmo retido dos maquinismos em fúria!  
Em fúria fora e dentro de mim  
(...)  
Ah, poder exprimir-me todo como um motor se exprime!  
Ser completo como uma máquina!”

(Fernando Pessoa [Álvaro de Campos]: “Ode Triunfal”)

Para Andiará, Olivia e Catarina.

Obrigado pela paciência.

## **Agradecimentos**

Ao adentrar os corredores da PUCRS em 2015, economista formado em 1999, não esperava que ao longo desses anos me tornaria também advogado, historiador e, agora, filósofo. Foram três graduações e duas jornadas de mestrado (PUCRS e UFRGS) que me permitiram explorar diferentes ângulos do conhecimento e expandir as fronteiras do pensamento. No entanto, além das conquistas acadêmicas, o verdadeiro tesouro que levo comigo são as amizades que sei que irão perdurar pela vida e que continuarão a me desafiar intelectualmente.

Estes anos também foram testemunhas de mudanças significativas em minha vida pessoal. Enquanto eu me dedicava às páginas dos livros e às discussões jurídico-histórico-filosóficas, Olivia cresceu, Catarina nasceu. Hoje, olho para minhas filhas com esperança, desejando que elas herdem minha curiosidade inesgotável pelo mundo.

Nenhum percurso acadêmico é trilhado sozinho, e é com imensa gratidão que reconheço aqueles que me guiaram e me inspiraram ao longo dessa jornada. Ao professor Thadeu Weber, meu orientador, expresso minha profunda apreciação por sua orientação, sabedoria e apoio. Aos membros da banca, Nythamar Oliveira e Paulo Caliendo, agradeço por serem interlocutores não apenas no momento crucial da defesa, mas por compartilharem sua expertise ao longo de toda minha trajetória acadêmica.

Seria injusto mencionar alguns nomes e omitir outros, pois a lista de interlocutores que me ajudaram a moldar minha visão de mundo é extensa. A todos vocês, colegas, amigos, professores e companheiros de jornada, minha gratidão mais sincera.

Por fim, uma menção especial a Andiará, minha companheira incansável ao longo desses anos. Seu apoio inabalável, paciência e amor foram fundamentais para que eu conseguisse enfrentar os desafios que se apresentaram. Você esteve ao meu lado em cada etapa, compartilhando alegrias e superando obstáculos. Minha jornada não teria sido a mesma sem você.

## RESUMO

Este trabalho explora a relação entre tecnologia e justiça, com um foco específico no debate em torno do uso de algoritmos em processos judiciais como substituição de juízes humanos. Ele começa estabelecendo uma definição de justiça como um meio de resolver conflitos e promover uma convivência harmoniosa na sociedade. A discussão adentra nas perspectivas do essencialismo e funcionalismo em relação à substituição de tarefas humanas por máquinas. Os argumentos essencialistas se concentram na ausência de propriedades humanas essenciais que tornam impossível para as máquinas replicarem os resultados do trabalho humano, enquanto os funcionalistas não veem impedimentos para essa substituição. O trabalho então examina o argumento de que a falta de criatividade na inteligência artificial representa uma limitação para sua aplicação em tarefas não mecânicas. Os essencialistas tendem a associar a criatividade à consciência, sugerindo que um componente biológico é necessário para que as máquinas sejam consideradas criativas, enquanto os funcionalistas se concentram no resultado das atividades, permitindo a atribuição de criatividade às máquinas. Em seguida, o trabalho adentra nas implicações da delegação de julgamentos às máquinas, baseando-se no conceito de realismo jurídico. As preocupações levantadas contra a substituição por máquinas parecem estar mais relacionadas ao medo de perder o controle sobre a tecnologia do que em rejeitar seus benefícios. A implementação da tecnologia já demonstrou eficiência na resolução de conflitos, em conformidade com decisões anteriores. Diante dessas discussões, o trabalho conclui que adotar uma posição intermediária, em vez de abraçar a rejeição extrema ou a adoção acrítica da tecnologia, parece ser mais apropriado. Essa abordagem reconhece tanto os avanços tecnológicos quanto uma perspectiva crítica sobre os problemas potenciais. Enfatiza a necessidade de encontrar um equilíbrio, garantindo que a justiça seja efetivamente alcançada e promovendo a harmonia na sociedade. Em última análise, a quantidade de utilização da inteligência artificial no sistema judicial é uma decisão humana que deve ser avaliada e supervisionada por eles. É fundamental ponderar os *trade-offs* entre eficiência e satisfação com o sistema escolhido de resolução de conflitos. Incorporar a inteligência artificial no processo judicial requer uma abordagem informada e ética, considerando os princípios fundamentais da justiça e as aspirações da sociedade em relação à resolução de conflitos. Dessa forma, seria possível alcançar uma integração equilibrada da tecnologia e preservação dos valores legais essenciais para a construção de uma sociedade justa e harmoniosa. Este resumo foi gerado por uma inteligência artificial após receber o texto completo da dissertação.

**Palavras-chave:** Direito; Tecnologia; Justiça; Algoritmos; Sistema judicial; IA (Inteligência Artificial)

## ABSTRACT

This work explores the relationship between technology and justice, with a specific focus on the debate surrounding the use of algorithms in judicial proceedings as a replacement for human judges. It begins by establishing a definition of justice as a means of resolving conflicts and promoting harmonious coexistence in society. The discussion delves into the perspectives of essentialism and functionalism regarding the substitution of human tasks by machines. Essentialist arguments center on the absence of essential human properties that make it impossible for machines to replicate the outcomes of human work, while functionalists see no impediments to such substitution. The dissertation then examines the argument that the lack of creativity in artificial intelligence poses a limitation to its application in non-mechanical tasks. Essentialists associate creativity with consciousness, suggesting a biological component is necessary for machines to be considered creative, while functionalists focus on the outcome of activities, allowing for the attribution of creativity to machines. Furthermore, the paper delves into the implications of delegating judgments to machines, drawing upon the concept of legal realism. The concerns raised against machine substitution appear to be more related to a fear of losing control over technology rather than rejecting its benefits. The implementation of technology has already demonstrated efficiency in the resolution of conflicts, aligned with previous judgments. In light of these discussions, the work concludes that adopting an intermediate position, rather than embracing extreme rejection or uncritical adoption of technology, seems more appropriate. This approach acknowledges both technological advancements and a critical perspective on potential problems. It emphasizes the need to strike a balance, ensuring that justice is effectively achieved and fosters harmony in society. Ultimately, the amount of AI utilization in the judicial system remains a decision for humans to evaluate and supervise. Trade-offs between efficiency and satisfaction with the chosen conflict resolution system must be carefully considered. Incorporating AI into the judicial process requires an informed and ethical approach, considering the fundamental principles of justice and societal aspirations. By doing so, a well-balanced integration of technology and preservation of essential legal values can be achieved, contributing to the establishment of a just and harmonious society. This abstract was generated by an artificial intelligence after being provided with the entirety of the text.

**Keywords:** Law; Technology; Justice; Algorithms; Judicial system; AI (Artificial Intelligence)

## SUMÁRIO

<b>Introdução .....</b>	<b>13</b>
<b>Capítulo 1 – Essencialismo versus Funcionalismo: dois polos de um debate.....</b>	<b>16</b>
1.1 A justiça enquanto instrumento de cooperação social.....	17
1.2 Essencialismo.....	19
1.3 Funcionalismo.....	22
1.4 Criatividade e inteligência artificial.....	23
1.5 Máquinas e criatividade.....	27
1.6 Criatividade e imaginação.....	30
1.7 O teorema do macaco infinito e seu papel no funcionalismo.....	33
<b>Capítulo 2 – Algoritmos e Realismo Jurídico.....</b>	<b>37</b>
2.1 O realismo jurídico e seu conceito de “Direito”.....	38
2.2 Racionalidade jurídica e categorização.....	41
2.3 Crítica ao Realismo Jurídico ou à tecnologia.....	46
2.4 Um meio termo .....	49
<b>Capítulo 3 – Visões sobre o Direito algoritimizado.....</b>	<b>52</b>
3.1 Lawrence Tribe e a margem de erro do sistema judiciário.....	54
3.2 Leis Personalizadas propostas por Ben-Shahar & Porat.....	62
3.3 A visão de Guarrapon & Lassègue.....	64
3.4 Confiança no sistema.....	66
3.5 Os <i>hard cases</i> .....	69
3.6 Critérios para implementação.....	72
<b>Conclusão .....</b>	<b>75</b>
<b>Referências Bibliográficas .....</b>	<b>77</b>

## CONSPECTUS SIGLORUM

BCB	Banco Central do Brasil
CVM	Comissão de Valores Mobiliários
DL	Deep Learning
EUA ou USA	Estados Unidos da América
IA ou AI	Inteligência Artificial
LLM	Large Language Models (Modelos Generativos de Linguagem)
ML	Machine Learning
SEC	Securities & Exchange Commission
STF	Supremo Tribunal Federal
UE	União Europeia

## Introdução

(...) computers change the way people think – especially about themselves<sup>1</sup>

Se compararmos a definição de corte de Jerome Frank (“*The court is a peace-preserving device*”<sup>2</sup> [a corte é um instrumento de conservação de paz]) com a afirmação quase onipresente na obra de John Rawls de que a manutenção da justiça tem como principal finalidade preservar a cooperação na sociedade<sup>3</sup>, podemos inferir que uma ampla gama de pensadores defende que a justiça possui valor instrumental<sup>4</sup> enquanto meio de resolução de disputas, e não em si mesmo. Mesmo na Grécia do período clássico, a ideia da conduta justa como elemento constituinte da felicidade era uma visão minoritária. Procurava se contrapor àquela mais comumente difundida, que advogava pela ideia de que o bem a ser alcançado não se esgotaria na justiça em si, mas sim no atingimento da paz, por sua vez também instrumental à prosperidade. Este encadeamento pode ser visto desde estátuas do século IV a.C., que já retratavam Irene, deusa da Paz, filha de Têmis, com Pluto em seu colo.

A discordância com relação não apenas à definição de Justiça, mas também seu papel na sociedade, tem, portanto, pelo menos 25 séculos. Os últimos anos têm nos oferecido uma sensação de maior polarização, mas ao que parece nem conteúdo nem forma são novos: em “A incapacidade para o diálogo”<sup>5</sup>, Gadamer já diagnosticava uma tendência ao *broadcasting*, ineficaz enquanto impulsionador do conhecimento, aliada a uma tendência à binaridade nas discussões contemporâneas sobre temas relevantes para a sociedade. Essa propensão a se entender quaisquer temas como dicotômicos, em que é preciso “escolher um lado”<sup>6</sup>, parece ter

---

<sup>1</sup> TURKLE, Sherry. *The Second Self: computers and the human spirit*. New York, NY: Simon & Schuster, 1984, p. 162.

<sup>2</sup> FRANK, Jerome. *Courts on Trial*. New Jersey: Princeton University Press, 1973, p. 7

<sup>3</sup> O tema aparece já no primeiro capítulo da obra de Rawls (“Chapter I – Justice as Fairness”), no primeiro tópico (“1. The Role of Justice”): “Let us assume, to fix ideas, that a society is a more or less self-sufficient association of persons who in their relations to one another recognize certain rules of conduct (...) Suppose further that these rules specify a system of cooperation designed to advance the good of those taking part of it.” RAWLS, John. *A Theory of Justice*. Cambridge, MA: Harvard University Press, 1971, rev. ed. (1999), p. 4

<sup>4</sup> O valor instrumental da justiça, ou mesmo das leis, visando à manutenção da cooperação social não reduz esta à aquela(s). Ou seja: “Embora as leis e sua rigorosa aplicação sejam alicerces indispensáveis para a vida em sociedade, elas não dão conta sozinhas da tarefa. Por mais rigoroso que seja o arcabouço da justiça (...) ele não seria capaz de preservar a sociabilidade humana em condição minimamente estável e ordenada se a confiança que nos sentimos motivados a depositar uns nos outros na vida prática se reduzisse a pó.” GIANNETTI, Eduardo. *O anel de Gíges: uma fantasia ética*. São Paulo: Companhia das Letras, 2020, p. 178.

<sup>5</sup> GADAMER, Hans-Georg. “A incapacidade para o diálogo” in: ALMEIDA, Custódio Luiz Silva de. *Hermenêutica Filosófica*. Porto Alegre: EDIPUCRS, 2000, pp. 129-140

<sup>6</sup> Na dicotomia, o que não pertence a um conjunto, precisa pertencer ao outro, e algo não pode estar nos dois conjuntos ao mesmo tempo.

sido ampliada com o desenvolvimento das redes sociais<sup>7</sup>. Assim, algo que a princípio parecia um facilitador (o surgimento de uma tecnologia que amplia a possibilidade de comunicação) teria se tornado mais um entrave ao diálogo.

Neste contexto, adotando a posição de que a sociedade precisa resolver os conflitos entre seus membros para que possam melhor cooperar, o objetivo dessa dissertação é analisar dois argumentos de uma discussão polarizada: se o processo de julgar pode ou não ser delegado a máquinas. O tema envolve uma discussão acerca da eficiência do sistema judiciário em sua tarefa de equacionar disputas e, deste modo, “oferecer justiça” à sociedade.

À primeira vista, parece ser possível identificar um aspecto presente no debate sobre a resistência observada versus a possibilidade da utilização de algoritmos no processo judiciário, causando a impressão de que o tema se encontra reduzido a uma espécie de polarização simplificada: a defesa dos juízes de sua empregabilidade, resistindo ao que seria um avanço inexorável da tecnologia<sup>8</sup>.

Todavia, acreditamos ser necessário analisar os argumentos de ambos os lados, e localizar duas vertentes de pressupostos, que comumente parecem estar encobertos nos discursos: de um lado, um essencialismo, que atribui à produção humana propriedades que estariam ausentes no fazer dos algoritmos, deslegitimando-o; de outro lado, uma visão funcionalista, que dá menor importância à ontologia da obra do computador e direciona o foco argumentativo para as questões práticas de como a sociedade utiliza aquilo que a máquina cria.

Em outras palavras, procuramos organizar o debate “adoção *versus* não-adoção de algoritmos para se oferecer justiça à sociedade”, e identificar quais são os lados que estão em contraposição. Este será o objetivo do capítulo 1.

Em seguida, no capítulo 2, analisaremos as ideias de Oliver Wendell Holmes e Jerome Frank, dois representantes do realismo jurídico, buscando elencar suas principais teses, e como estas se entrelaçam com a acusação (pelo lado dos essencialistas) de que as cortes digitais significariam o retorno e a adoção pelo judiciário do realismo jurídico como corrente predominante de pensamento jurídico.

No capítulo 3, voltamos à análise dos argumentos de ambos os polos da discussão, no contexto de duas propostas de implementação de cortes digitais: a visão moderada de Garapon & Lassègue (2018) e a adoção radical de Bem-Sharar & Porat. Adicionando a contribuição de

---

<sup>7</sup> Sobre esse aspecto, o documentário: Social Dilemma. Direção: Jeff Orlowski, Exposure Labs, 2020. (94min). Disponível em: [www.netflix.com/title/81254224](http://www.netflix.com/title/81254224) - último acesso em 12 Abr 2023.

<sup>8</sup> OMRI, Ben-Shahar & PORAT, Ariel. Personalized law: different rules for different people. New York: Oxford University Press, 2021, p. 42

Tribe (1971), apresentaremos a proposta defendida por autores moderados de que, ao invés de optar por um dos lados da polarização apresentada no capítulo 1, ou de procurar uma forma de conciliar as duas visões, a discussão deveria se dar sobre qual a quantidade de incerteza com relação aos julgamentos performados pelo sistema judiciário que a sociedade aceita sem que ele perca sua legitimidade como instrumento de resolução de disputas. Adicionalmente, apresentamos como as teorias jurídicas podem ser formas de narrativa para construção dessa legitimidade. Ao longo do capítulo 3, utilizando o conceito de “Síndrome de Frankenstein” (o medo de que o algoritmo, uma criação do ser humano, se volte contra seu criador) apresentado no capítulo 2, questionaremos se a polêmica em relação à adoção de novas tecnologias no sistema judiciário seria uma manifestação de um comportamento já identificado anteriormente contra o avanço tecnológico: o medo do desconhecido e um desconforto pela mudança do *status quo*.

## Capítulo 1 – Essencialismo *versus* Funcionalismo: dois polos de um debate

Parece haver dois lados no debate da adoção de algoritmos para oferecer justiça à sociedade: os entusiastas e os contrários à sua delegação para as máquinas. Tal polarização lembra aquela já identificada por Umberto Eco em “Apocalípticos e Integrados”<sup>9</sup>, uma vez que o lado que rejeita a possibilidade levanta comumente uma possibilidade de futuro distópico, no qual as máquinas subjugarão os humanos. Para essa corrente, falta algo ao ato de fazer justiça proveniente de uma máquina, que o deslegitima, por carecer de propriedades que deveriam estar presentes no processo, tais como um suposto sentimento de justiça ou percepção do contexto.

No pólo oposto, parecem estar aqueles que desejam a adoção da tecnologia, baseando sua opinião na seguinte justificativa: a substituição de força de trabalho humana por máquinas e a oposição a esse processo confundem-se mesmo com as origens da Revolução Industrial, desde os movimentos ludistas e o quase coetâneo surgimento dos sindicatos até a instituição da greve como um direito constitucional em cartas magnas recentes. Todavia, se até o final do século XX os equipamentos substituíam trabalhadores nas atividades industriais, a partir do passado recente passaram a fazer o mesmo no setor de serviços<sup>10</sup> – “la tecnología digital es para nuestra capacidad mental lo que la máquina de vapor es para nuestra capacidad muscular”<sup>11</sup>. Os juízes seriam apenas a “classe da vez” de uma evolução inexorável que traria em verdade uma melhoria de bem-estar da sociedade,<sup>12</sup> ainda que impondo alguns conflitos durante sua adoção com segmentos que pretendem conservar seus privilégios.

Antes de entrar na análise de cada um dos lados do debate, faremos uma breve digressão sobre o papel da justiça para a sociedade.

---

<sup>9</sup> ECO, Umberto. *Apocalittici e integrati: comunicazioni di massa e teorie della cultura di massa*. 3. ed. Milano: Bompiani, 1964. Segundo o autor, a ideia central é a de que ““Gli apocalittici pensano in termini di purezza e impurità, di salvezza e di dannazione, di cose che valgono la pena di essere conservate e di cose che devono essere gettate via, di bianco e di nero. Gli integrati pensano in termini di sostanze, di campi di esperienza in cui il soggetto può immergersi, di categorie di appartenenza.”

<sup>10</sup> Aspecto destacado também em MALABOU, Catherine. *Métamorphoses de l’intelligence : Du QI à l’IA*. Paris : Puf, 2017, p. II. No original : “(...) les machines on pendant longtemps remplacé le travail musculaire. Aujourd’hui, elles remplacent le travail de l’esprit, touchant ainsi non seulement à toutes les activités cognitives, mais bien aussi à l’autonomie de la vie et de la pensée”

<sup>11</sup> MARTINO, Antonio A. “¿Quien teme a los jueces robot?” in: GONZALO, Ana Dobratinich (org.). *Derecho y nuevas tecnologías*. Buenos Aires: La Ley, 2021 (pp. 15-41), p. 27

<sup>12</sup> RUSSELL, Stuart. *Human Compatible: artificial intelligence and the problem of control*. New York, NY: Penguin Books, 2020, em especial a seção “Eliminating Work as We Know It” (pp. 113-124)

## 1.1 A justiça enquanto instrumento de cooperação social

“A society is defined by its amalgamations, not by its tools (...) tools exist only in relation to the interminglings they make possible or that make them possible.”<sup>13</sup>

Para Lewis Mumford, a maior invenção do ser humano teria sido não um objeto, como a roda ou a lança, e sim a sociedade, uma forma de colaboração<sup>14</sup>. O momento histórico dessa intensificação da colaboração não pode ser apontado: enquanto a arqueologia apoia-se em vestígios materiais que sobreviveram até nossos dias, os processos de criação da linguagem humana e o estabelecimento de formas de convivência em grupos cada vez maiores não podem ser rastreados, apesar de conhecermos o resultado atual. A partir do momento em que o desenvolvimento do sistema nervoso central de nossa espécie teria retirado a sujeição a um automatismo proveniente das necessidades biológicas instintivas, o ser humano pôde se libertar, indo além daquilo que vinha meramente programado nos seus genes (Mumford, 1934, p. 26). Sua relação com a natureza teria se realizado de modo diferente daquele até então, passando a conceber um dever-ser. Essa capacidade teria permitido conceber regras (que muito provavelmente não vieram prontas, mas foram testadas ao longo de milênios) para a vida em sociedade.

O conceito de “norma”, enquanto conversão do dever-ser em palavras, aparece desde a sentença daquele que “julga de modo mais justo” presente na descrição do escudo de Aquiles da *Ilíada*<sup>15</sup>, configura-se de modo mais rigoroso com a ideia de “lei no caso concreto” na *Summa Theologica*<sup>16</sup>, e consolida-se após a teoria do Direito firmar-se como disciplina acadêmica. Nestes exemplos, há um pressuposto de que uma análise da situação em questão adequa a lei ao caso, ponderando eventuais necessidades que surjam na transição do abstrato para o concreto.

E ser “justo”, seguindo a ideia acima de que um conflito deveria ser resolvido da melhor maneira possível, adequada ao caso concreto, seria o centro da ideia de justiça. O campo semântico do termo “justiça” é muito vasto. A palavra não apenas é polissêmica em um determinado momento do tempo, estando em disputa por teorias tão díspares quanto deônticas

---

<sup>13</sup> Gilles Deleuze and Felix Guattari, *A Thousand Plateaus: Capitalism and Schizophrenia*, trans. Brian Massumi (Minneapolis, 1987), p. 90.

<sup>14</sup> MUMFORD, Lewis. *Technics and Civilization*. New York: Harcourt, Brace and Company, 1934, p. 14

<sup>15</sup> Homero. *Ilíada*, canto XVIII.

<sup>16</sup> Tomás de Aquino, ST II-II Q. LXVII, art 1: “Respondeo dicendum quod sententia iudicis est quasi quaedam particularis lex in aliquo particulari facto.”

e utilitaristas, como enfrenta também uma variação no tempo. De fato, uma mesma pessoa um mesmo grupo pode alterar sua concepção de justiça no tempo.

Em resumo, poderíamos apontar que o termo justiça padece de um problema de identidade – e não apenas transtemporal – devido à aglutinação de sentidos opostos em si. Ao longo dos últimos 2500 anos, a tarefa de definir o conceito de “justo” parece não ter sido finalizada, empreendimento do qual não pretendemos participar por considerarmos que pouco contribuiríamos. Assim sendo, para esse trabalho, optamos por adotar o termo “justiça” como meio para um fim: um instrumento de aprimoramento da cooperação social, posição não inovadora, visto que em nossa leitura de John Rawls consideramos que um dos pressupostos que aquele filósofo político assumia era a justiça enquanto meio de cooperação entre indivíduos<sup>17</sup>. Posteriormente, Rawls (2003) deixou mais explícito que seu conceito de justiça estava próximo ao de um instrumento na “sociedade como um sistema equitativo de cooperação”<sup>18</sup> (§3, p. 11).

Neste arcabouço, o poder judiciário tornou-se o ramo do governo responsável por administrar a justiça e desta forma, resolver conflitos.<sup>19</sup> A lide judicial não é a única forma de resolver disputas, mas sim uma dentre várias opções (por exemplo, autocomposição entre as partes, ou mesmo arbitragem).

Para ser uma alternativa legítima à uma sentença emanada de um juiz humano, aquela emitida por um algoritmo deveria ser primeiramente identificada como adequada à situação. Todavia, uma das principais críticas comumente direcionada à inteligência artificial (IA) é que ela reforça estereótipos e vieses presentes nos dados utilizados para processar a informação<sup>20</sup>.

Todavia, há uma diferença entre a) identificar que existem distorções no processo de *machine learning* (ML ou “aprendizado de máquina”) realizado em cima de uma base de dados

---

<sup>17</sup> Para Rawls, “The concept of justice is not, I think, a primitive concept; we do not begin with justice and move from there to establish the social contract and the civil society. But though justice is not primitive in this sense, its role in the organization of society is nevertheless central and I believe it to be essential.” In RAWLS, John. A theory of justice. Cambridge, MA: Harvard University Press, 1971, p. 3

<sup>18</sup> RAWLS, John. Justiça como equidade: uma reformulação. São Paulo: Martins Fontes, 2003.

<sup>19</sup> Esta leitura de Rawls aparentemente pode ser considerada como mais utilitarista em relação ao que o próprio Rawls se definiria (afinal, um dos propósitos declarados em Uma Teoria da Justiça é fazer uma crítica ao utilitarismo), mas parece ser inegável a influência utilitarista no pensamento rawlsiano. Longe de ser a única leitura, Rawls também pode ter aproximações com Hegel, como destaca WEBER, Thadeu. Hegel e Rawls: um diálogo possível. Veritas (Porto Alegre), v. 67, p. 1-12, 2022.

Disponível em: <https://revistaseletronicas.pucrs.br/ojs/index.php/veritas/article/view/43762> - último acesso em 05 Ago 2023.

<sup>20</sup> Sobre esse tema, O’NEIL, Cathy. Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. Danvers, MA: Crown Publisher Group, 2017; também ZUBOFF, S. The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power. New York: Public Affairs, 2019; também EUBANKS, Virginia. Automating Inequality: how high-tech tool profile, police, and punish the poor. New York, NY: PICADOR, St. Martin’s Press, 2018.

que contém uma correlação indesejada (por exemplo, cor da pele e percentual de condenações) e procurar corrigi-las; e b) utilizar a justificativa de correlações pretéritas indesejadas presentes na base de dados surgidas de julgamentos realizados por humanos para rejeitar a adoção do processo. Em resumo, acreditamos ser necessário dissecar o argumento de que a máquina seria menos competente para eliminar correlações indesejáveis do que juízes humanos eliminarem vieses cognitivos.

Até porque um conceito caro ao direito, o de “segurança jurídica”<sup>21</sup>, ligado à previsibilidade de resultados em julgamentos, parece estar muito próximo daquele de “falácia naturalista na base de dados”, que seria tomar os julgamentos passados como lei para os julgamentos futuros. Para o algoritmo, os dados dos julgamentos passados tornam-se norma a partir da qual se julga.

Veremos a seguir que os dois lados do debate precisam tratar desse conflito para construir seus argumentos. Críticos da adoção de algoritmos para se oferecer justiça à sociedade decretam que essa atitude condenaria o sistema judicial a um novo “realismo jurídico”, eliminando na prática a possibilidade de mudanças na interpretação da lei. Por outro lado, parte do argumento daqueles que defendem essa nova tecnologia é refutar a acusação, pois não haveria como nos dados passados (julgamentos já realizados) não haver presentes elementos de dever-ser.

## 1.2 Essencialismo

Para os que defendem a “indelegabilidade” do ato de se fazer justiça às máquinas, o núcleo do argumento contra o uso de algoritmos em julgamentos parece ser a alegação de que há na atividade mental algo distinto de uma tarefa de esforço físico – pressuposto sobre o qual ainda não existe consenso e parece ser relevante para a origem dos dois lados da disputa. Haveria, segundo esses, uma propriedade essencial na tarefa de proferir um julgamento em uma lide, um “sentimento de justiça”, que impediria essa transferência ao computador. Julgar seria um ato no qual que aquele que performa precisa não apenas fazer, mas saber que faz, o que em última instância requer a presença de uma consciência<sup>22</sup>, propriedade – pelo menos por

---

<sup>21</sup> Sobre o valor da segurança jurídica para o Direito, SCHULZ, Fritz. *Principles of Roman Law*. Oxford: Clarendon Press, 1936, cap. XII.

<sup>22</sup> O argumento do quarto chinês de Searle (1980) é o de que consciência requer capacidade de semântica, e não apenas manipulação de símbolos. Searle, John R. "Minds, Brains, and Programs." *Behavioral and Brain Sciences*,

enquanto – ausente nas máquinas.<sup>23</sup> Esses autores rotulam os defensores da adoção de algoritmos nas cortes como integrantes de uma espécie de “realismo jurídico 2.0”, termo utilizado por Lênio Luiz Streck<sup>24</sup> com conotação pejorativa. Ao recuperar a nomenclatura que agrega um conjunto de juristas norte-americanos da primeira metade do século XX, nos parece que Streck acusa as sentenças emitidas pelo computador de replicarem decisões já proferidas e aplicarem a novos casos sem preocupação com as idiosincrasias de cada um deles. Um processo mecânico de input-output que lembra mais um behaviorismo do que necessariamente a ideia presente na expressão que supostamente sintetizaria o realismo jurídico de que “a lei é o que as cortes dizem”.<sup>25</sup>

Essa posição pode ser classificada como pertencente a um arcabouço essencialista, defendendo a ideia de que deveria haver necessariamente a presença de uma propriedade no objeto de estudo (o ato de julgar) cuja ausência desqualifica sua identidade – e consequentemente legitimidade. Em outras palavras, essencialistas afirmam que sem determinado(s) componente(s) essencial(is), um suposto modo de se fazer justiça (por algoritmos) deixa de ser fazer justiça em realidade, tornando-se um simulacro.

Este argumento parece muito próximo de uma teoria da mente, na qual, em uma visão essencialista, a consciência seria uma propriedade *sine qua non* para a identidade da mente. À essa visão, opõe-se o funcionalismo, que argumenta que a presença ou ausência de consciência seria irrelevante, uma vez que os atos performados cumprem a finalidade que uma mente teria (por exemplo, dizer do mundo e/ou agir no mundo).

Outro conceito útil para se entender a posição essencialista, que atribui uma propriedade essencial à sentença emitida por um juiz, é a analogia com a ideia de “aura” de Walter Benjamin em sua análise das obras de arte.<sup>26</sup> Para Benjamin, uma ruptura na aura do objeto artístico provoca a perda de sua identidade (composta tanto de singularidade como autenticidade), ou

---

vol. 3, no. 3, 1980, pp. 417-457. Disponível em: <https://web.archive.southampton.ac.uk/cogprints.org/7150/1/10.1.1.83.5248.pdf> - último acesso em 01 Mai 2023.

<sup>23</sup> O tema recentemente voltou à pauta pela afirmação de um engenheiro de software do Google que defendeu a presença de sentiência (capacidade para experienciar sentimentos e sensações, pré-requisito para a consciência) no algoritmo de IA daquela empresa. Sobre esse tópico: <https://www.washingtonpost.com/technology/2022/06/11/google-ai-lambda-blake-lemoine/> - acesso em 22 jun 2022, e uma das respostas oferecidas pela AI Robotics Ethics Society, dentre outras que negam que a tecnologia tenha chegado a esse estágio: <https://en.airespucrs.org/post/is-lambda-sentient-tl-dr-no> - acesso em 22 jun 2022

<sup>24</sup> STRECK, Lênio Luiz. “Um robô pode julgar? Quem programa o robô?” in: *Consultor Jurídico*, 03 set 2020. Disponível em: <https://www.conjur.com.br/2020-set-03/senso-incomum-robo-julgar-quem-programa-robo?imprimir=1> – último acesso em 03 abr 2022.

<sup>25</sup> Mais adiante trataremos de como essa expressão, associada aos autores identificados como defensores do Realismo Jurídico, aparentemente tornou-se mais um clichê utilizado para um ataque através da falácia do espantalho do que uma expressão que sintetiza suas ideias.

<sup>26</sup> BENJAMIN, Walter. “The Work of Art in the Age of Mechanical Reproduction.” *Illuminations: Essays and Reflections*, edited by Hannah Arendt, translated by Harry Zohn, Schocken Books, 1969, pp. 217–251.

seja, perda de valor de culto. Podemos entender que a posição de Streck (2022), analogamente, suporta a ideia de que a intervenção de uma máquina no processo de criação de uma sentença impede a singularidade do texto, uma vez que este é em tese fruto de uma mera reprodução de partes de outros textos. É como se a sentença mecanicamente produzida carecesse de aura, aniquilando qualquer possibilidade de valor de culto ao texto da máquina.

Todavia, há que se questionar o quanto um documento gerado por um computador e depois apenas assinado por um juiz “transforma-se” em um documento gerado por um humano. Quão legítima seria essa sentença se, emitida por um computador e depois lida por um humano, ele se manifestasse de pleno acordo com seu conteúdo? Esta espécie de adoção confere legitimidade *ex nunc*, a partir do momento da adoção, e desse modo legitimando a sentença de um instrumento utilizado, ou *ex tunc*, “apagando” o fato de que foi gerada por uma entidade diferente do juiz, uma vez que poderia ter sido por ele gerada?

Por certo a assinatura do juiz “marca” o texto, conferindo a ele uma autoria humana caso a autoria da máquina fosse um conceito questionável. Para um essencialista, o ato atribuiria uma propriedade de autenticidade que conecta o texto ao autor: αὐθεντικός tem as raízes αὐτο e ἔντης, feito por si mesmo. Ao assinar um documento, fazemos nosso o conteúdo nele escrito, expressão de nossa vontade, mesmo que nenhuma de suas palavras tenha sido originalmente por nós escrita. O procedimento já existe na prática nos tribunais: o assessor redige, o juiz assina.

Mas nossa discussão é a respeito da possibilidade de supressão da assinatura do juiz humano e pela colocação em prática para uma sentença ser executada diretamente do computador para as partes, sem o passo que conferiria a autoria ao juiz. Na visão essencialista, esta ausência da assinatura retira a legitimidade da sentença; enquanto em uma visão funcionalista a assinatura seria dispensável, posto que o conteúdo permanece o mesmo e levará às mesmas consequências sociais. Em outras palavras o funcionalismo argumentaria que se a única função da assinatura do juiz é conferir exequibilidade, basta atribuí-la à máquina, pois a assinatura parece ser um passo redundante.

O fetichismo da assinatura não é um tema novo. Como lembra Silveira (2018):

(...) quando de sua introdução, houve resistência em aceitar o valor jurídico de documentos datilografados: uma vez que eram produzidos sem a intervenção da mão autoral, como garantir sua autenticidade? Ainda preservamos algo

dessa desconfiança quando da necessidade de assinar um documento impresso (...)<sup>27</sup>

O certificado digital, o *blockchain*, a assinatura eletrônica, são todas formas digitais da mesma necessidade antiga: provar autoria. Todavia, em tese não alteram o conteúdo do texto.

### 1.3 Funcionalismo

O conceito de funcionalismo pode gerar ambiguidades devido à sua natureza polissêmica. No contexto deste estudo, empregamos o termo em conformidade com o significado atribuído nas ciências sociais por acadêmicos que enfatizam o papel das instituições na manutenção da solidariedade entre os membros de uma sociedade. Esta linha de raciocínio remonta a Durkheim, influenciou o estruturalismo de Lévi-Strauss e fundamenta premissas de autores como Douglass North<sup>28</sup> e Daron Acemoglu & James Robinson<sup>29</sup>. Conforme mencionado anteriormente, o sistema judicial tem como objetivo mediar conflitos, buscando harmonia entre os indivíduos de uma comunidade que colaboram entre si. A sentença judicial proferida em um litígio é a expressão textual da norma que deve ser aceita pelas partes envolvidas.

Desse modo, a função da sentença é elucidar a decisão que visa manter a coesão social, independentemente do meio pelo qual a decisão é comunicada às partes. Sob uma perspectiva funcionalista, o conteúdo da sentença é mais crucial do que a forma como é transmitida - seja oralmente, por escrito, por um juiz humano ou uma máquina. Por essa razão, uma decisão proferida por um algoritmo tem menos probabilidade de encontrar resistência entre aqueles que adotam a abordagem funcionalista.

A etimologia do termo "funcionalismo" (derivado do latim *functio*, que provém de *fungor*: performar, executar) compartilha essa mesma raiz com *fungibilis* (*fungor* + *ible*, "capacidade de performar"), que é a aptidão para alcançar os mesmos resultados, mesmo sendo diferente. A fungibilidade aparece em autores desta corrente muitas vezes como oposição à ideia de instituições homólogas (mesma origem, mas com funções distintas), devido à

---

<sup>27</sup> SILVEIRA, Pedro Telles da. História, técnica e novas mídias: reflexões sobre a história na era digital. Tese de doutorado. Universidade Federal do Rio Grande do Sul, Porto Alegre, 2018, p. 126

<sup>28</sup> North, Douglass C. 1991. "Institutions." *Journal of Economic Perspectives*, 5 (1): 97-112. Disponível em: <https://www.aeaweb.org/articles?id=10.1257/jep.5.1.97#:~:text=Institutions%20provide%20the%20incentive%20structure,my%20analysis%20from%20economic%20history> – último acesso em 10 Abr 2023.

<sup>29</sup> ACEMOGLU, Daron & ROBINSON, James. *Why nations fail: the origins of power, prosperity, and poverty* (First ed.). New York: Crown Publishing Group, 2012.

influência da biologia. Nesse sentido, a perspectiva funcionalista defenderia a fungibilidade entre as sentenças emitidas por seres humanos e computadores, uma vez que, embora distintas, ambas alcançam o mesmo objetivo.

Após ter se firmado na antropologia pelos trabalhos de Malinowski e Radcliffe-Brown como uma forma de oposição ao evolucionismo darwiniano, o funcionalismo passou a ser acusado, em especial nos anos 1950-1960<sup>30</sup>, de carregar elementos conservadores, que defenderiam de modo muito rígido a manutenção da ordem social existente. Do ponto de vista das teorias marxistas, foi criticado por negligenciar as desigualdades sociais e enfatizar a harmonia e a estabilidade da sociedade em detrimento das lutas sociais e dos conflitos. Para esses críticos<sup>31</sup>, sua visão era limitada e insuficiente para lidar com problemas sociais e mudanças significativas na sociedade. Outra crítica seria de que o funcionalismo levaria a uma forma de determinismo estrutural, ou seja, rejeitaria em tese a agência individual e capacidade de autodeterminação ou possibilidades de transformações sociais, sendo incapaz de lidar com mudanças significativas na sociedade.

Essa crítica ressurgiu no debate da delegação da justiça aos computadores como a alegação de que um algoritmo não seria capaz de decidir os chamados *hard cases*, situações que contêm fatos incomuns e às quais nenhuma lei parece se aplicar de imediato, sendo necessário compor diferentes normas ou explorar as fronteiras da interpretação da atual legislação, o que envolveria uma espécie de capacidade de criar. Seria preciso, deste modo, analisarmos se uma máquina pode ser considerada criativa o suficiente para decidir um caso incomum. Para isso, novamente nos questionando se a criatividade possui algum elemento essencial que a impediria de ser atribuída a uma máquina.

#### 1.4 Criatividade e inteligência artificial

Começemos por discutir a relação entre o mestre e o escravo  
(...) se cada ferramenta pudesse realizar o trabalho de maneira  
autônoma quando assim ordenada, como as estátuas de Dédalos

---

<sup>30</sup> GIDDENS, Anthony. "Functionalism: Après La Lutte." in: *Social Research* 43, no. 2 (1976): 325–66. <http://www.jstor.org/stable/40970227>.

<sup>31</sup> Por exemplo, ALTHUSSER, Louis. *Ideology and ideological state apparatuses*. In ALTHUSSER, Louis. *Lenin and philosophy and other essays*. New York: Monthly Review Press, 1971. Também disponível em: <https://www.marxists.org/reference/archive/althusser/1970/ideology.htm> - último acesso em 10 Abr 2023. Neste ensaio, Althusser defende a ideia de que as teorias funcionalistas negligenciam o papel da ideologia na manutenção da ordem social. Ele afirma que instituições sociais como escolas, mídia e a igreja servem para reproduzir a ideologia dominante e reforçar as relações de classe.

no relato, ou os trípodes de Hefesto, que o poeta diz "entrar automovidos na companhia divina" - se assim as lançadeiras tecessem [sozinhas] e as penas tocassem harpas por si mesmas, os mestres artesãos não precisariam de assistentes e os mestres não precisariam de escravos.<sup>32</sup>

Em 1842, Lady Lovelace escreveu a respeito da Máquina Analítica, afirmando que ela “não tem nenhuma pretensão de originar nada. Ela pode fazer tudo aquilo que soubermos como ordená-la para executar. Pode seguir instruções; mas não tem poder de antecipar quaisquer relações ou verdades analíticas.”<sup>33</sup> Desta forma, a ideia de limitação para a máquina parece ser intrínseca à sua própria definição.

O conceito de “robô” como mero executor (o termo vem do verbo eslavo *работает*, que é “trabalhar”) surgiu em um momento em que a ausência de agência era propriedade constitutiva da máquina, já que sua função era realizar tarefas absolutamente previsíveis. No entanto, como pode ser visto no excerto de Aristóteles acima, já existia desde o período clássico grego um desejo de que a ferramenta pudesse um dia alcançar uma espécie de autonomia relativa de ação. Ao invés de simplesmente executar movimentos repetitivos, a ferramenta seria capaz de atingir seu objetivo a partir de um comando – a ideia é coerente, inclusive, com a forma finalista do estagirita de atribuir identidade. Deste modo, a ferramenta substituiria o humano escravizado, à época o único “instrumento” capaz de compreender comandos e buscar os meios de executá-los.

A diferença aristotélica entre *poiesis* [ποίησις], que exige *techne* [τέχνη], e a *praxis* [πρᾶξις], que demanda *phronesis* [φρόνησις], dividia a ação humana em dois tipos de atividade – ambas incapazes de serem desenvolvidas pela ferramenta. Na primeira, o instrumento deveria ser manuseado pelo *antropos* para que o resultado da atividade se materializasse após a manipulação de matéria. Já na segunda, a ação em si mesma seria o resultado desejado. Em ambos os casos, a *aretê* [ἀρετή] seria obtida por meio do hábito (do treinamento e da educação, diríamos atualmente).

---

<sup>32</sup> Aristóteles, Pol. 1253b. Tradução nossa. No original: *περὶ δεσπότου καὶ δούλου εἶπωμεν (...)* εἰ γὰρ ἡδύνατο ἕκαστον τῶν ὀργάνων κελευσθὲν ἢ προαισθανόμενον ἀποτελεῖν τὸ αὐτοῦ ἔργον, καὶ ὥσπερ τὰ Δαιδάλου φασὶν ἢ τοὺς τοῦ Ἡφαίστου τρίποδας, οὓς φησὶν ὁ ποιητὴς αὐτομάτους θεῖον δύεσθαι ἀγῶνα, οὕτως αἱ κερκίδες ἐκέρκιζον αὐταὶ καὶ τὰ πλῆκτρα ἐκιθάριζεν, οὐδὲν ἂν ἔδει οὔτε τοῖς ἀρχιτέκτοσιν ὑπηρετῶν οὔτε τοῖς δεσπόταις δούλων

<sup>33</sup> LOVELACE, Ada Augusta. Notes upon the Memoir of L. F. MENABREA: Sketch of The Analytical Engine Invented by Charles Babbage. Bibliothèque Universelle de Genève, October, 1842, No. 82. Disponível em: <https://www.fourmilab.ch/babbage/sketch.html> - último acesso em 14 Jan 2023. No original: “The Analytical Engine has no pretensions whatever to originate anything. It can do whatever we know how to order it to perform. It can follow analysis; but it has no power of anticipating any analytical relations or truths.”

Assim, a noção de máquina utilizada por Lovelace durante o período da Revolução Industrial era a de uma “ferramenta aprimorada”, um objeto que agia com base em uma fonte de energia artificial, mas ainda dependente do comando do ser humano. Observamos, deste modo, que o conceito de algoritmo foi inicialmente derivado do funcionamento de uma máquina física, que seguia uma sequência de comandos determinados previamente, semelhante a uma receita culinária. No entanto, a inteligência artificial transcendeu essa concepção, aproximando a máquina do antigo desejo aristotélico de informar somente o objetivo final e conceder ao instrumento uma relativa autonomia para decidir como executar a tarefa de forma mais eficiente, incluindo opções de atuação de meios e não apenas fins. Um atual estágio bem mais próximo da proposta de Silveira (2021), que teve como “objetivo demonstrar a possibilidade de superação dos limites firmados por Descartes, Ada Lovelace, Turing e Searle”.<sup>34</sup>

Isso justifica o motivo pelo qual, sob uma perspectiva humana, o computador parece ganhar agência – mesmo que ela não seja binária, e sim gradual<sup>35</sup> -, buscando atingir objetivos por meio de ações que têm a aparência de autônomas. Com o avanço tanto das linguagens de programação quanto da capacidade de processamento dos hardwares, esta capacidade se expandiu para tarefas que exigem processos cognitivos, antes consideradas em tese de exclusividade humana. Um exemplo claro desta mudança foi a forma como textos são reproduzidos.

A imprensa surgiu como uma forma de substituição deste trabalho repetitivo que já existia, e que na Europa medieval era majoritariamente realizado pelos monges copistas.<sup>36</sup> A invenção da impressora moderna permitiu a continuação da reprodução de maneira mais eficiente, mas ainda limitada ao conceito de máquina de Lovelace: imprimir com precisão e rapidez um texto ou imagem previamente definidos. Recentemente, o ChatGPT passou a disponibilizar textos customizados, respondendo a comandos (*prompts*), que, apesar de ainda baseados em uma massa de dados já existente que é apenas recombinação, oferece uma impressão de autonomia. Com isso, parece passível de atribuição de agência e criatividade – termos até então não aplicados a máquinas.

Nessa linha, dois campos de criação que até o final do século passado eram considerados exclusivamente humanos eram a pintura e a música. Em relação à pintura, o

---

<sup>34</sup> SILVEIRA, Paulo Antonio Caliendo Velloso da. *Ética e inteligência artificial: da possibilidade filosófica de agentes morais artificiais*. Porto Alegre, RS: Editora Fi, 2021, p.16.

<sup>35</sup> COECKELBERGH, Mark. *AI Ethics*. Cambridge, MA: the MIT Press, 2020, p. 109.

<sup>36</sup> Sobre esse tema, CANFORA, Luciano. *Il copista come autore*. Turin: Einaudi, 2001.

contraste se dava na comparação entre o processo de pintar livremente um quadro com o processo de reprodução mecânica. Até recentemente, não era possível comandar a impressão de um desenho apenas com instruções genéricas, fornecendo o tema e não a imagem final em formato digital.

Entretanto, desde o início do século XXI, o algoritmo AARON, desenvolvido pelo artista e programador Harold Cohen (1928-2016),<sup>37</sup> demonstrou avanços significativos na área da pintura. Conectado a um braço robótico, o algoritmo é capaz de pegar um pincel e pintar uma tela de forma autônoma e, diferentemente de uma impressora, o programa pode gerar inúmeros desenhos exclusivos com base em um tema desejado, sem necessidade de copiar ou modificar um desenho preexistente. Por exemplo, embora AARON nunca tenha “visto” uma pessoa ou caminhado por um jardim botânico, recebeu informações sobre posturas corporais e plantas por meio de regras. Seu código foi elaborado em uma variante da linguagem de computador conhecida como LISP, popular entre a comunidade de inteligência artificial por causa de sua flexibilidade e habilidade para lidar com conceitos abstratos.<sup>38</sup> Também recentemente, a mesma empresa, Open AI, desenvolveu a rede neural DALL-E, capaz de gerar imagens a partir de instruções fornecidas em forma de texto. Essa inovação representa um avanço significativo na interseção entre arte e inteligência artificial, ampliando ainda mais o potencial de criação das máquinas.<sup>39</sup>

Analogamente, no campo da música, algoritmos já são capazes de compor músicas completas conforme o estilo solicitado. O programa AIVA é capaz de oferecer composições originais, após delimitados certos parâmetros de estilo ou gênero musical (que podem ser tão genéricos como “canções de natal” ou simplesmente “samba”) e tempo de duração, sendo que alguns exemplos podem ser ouvidos em um canal na plataforma de vídeos Youtube: <https://www.youtube.com/@aiva1828>. De fato, nem toda produção gerada pelo algoritmo atende às expectativas de quem solicita a composição, mas o mesmo ocorre com obras criadas por seres humanos, que muitas vezes decepcionam sua audiência.

Tanto no caso da pintura quanto da música, os resultados oferecidos pelos algoritmos são originais, embora não sejam criados *ex nihilo*. Isto é, não são cópias, mas baseiam-se em uma grande massa de dados já disponível. O que o programa faz é combinar – de maneira

---

<sup>37</sup> Louise Sundararajan; Harold Cohen and AARON: Collaborations in the Last Six Years (2010–2016) of a Creative Life. *Leonardo* 2021; 54 (4): 412–417. doi: [https://doi.org/10.1162/leon\\_a\\_01906](https://doi.org/10.1162/leon_a_01906) - último acesso em 06 Dez 2022

<sup>38</sup> SCHWARTZ, John. “Is Aaron’s work creative art or just high-tech doodling?” in *The Washington Post*, 10 abr 1995. Disponível em: <https://www.washingtonpost.com/archive/politics/1995/04/10/is-aarons-work-creative-art-or-just-high-tech-doodling/366dc211-2aec-4d99-8d55-27f7acd5e27d/> – último acesso em 6 Dez 2022

<sup>39</sup> Disponível em: <https://openai.com/dall-e-2/> - último acesso em 26 Dez 2022.

inédita – fragmentos de desenhos ou notas musicais, orientados por um estilo previamente classificado (cubismo, impressionismo, jazz, rock) através de um conjunto de características. A criação humana, por outro lado, também se fundamenta em um repertório de referências internalizadas pela mente do criador, ainda que de forma mais intuitiva.

Deste modo, parece haver mérito no argumento funcionalista de que os resultados se equivalem, independente do processo ou do executor. Caso isso fosse aceito, surgiria a indagação: seria possível algoritmizar a criatividade? Se sim, a parcela do conceito de criatividade que demanda uma "essência humana" necessitaria de uma revisão.

A ideia de algoritmizar um processo mental parece não ser inédita. Analogamente ao conceito de *mitema* em Claude Lévi-Strauss, que buscava elementos comuns nos mitos de todas as culturas, os gramáticos estruturalistas procuravam identificar o denominador comum ao processo de *storytelling* do ser humano. De acordo com essa hipótese, uma vez identificada a propriedade do cérebro humano em entender o mundo através do processo de contar histórias, “o santo graal para os gramáticos seria o conjunto de regras que seriam capazes de descrever a estrutura de todas as histórias já escritas”<sup>40</sup>

Paralelamente, Benjamin N. Cardozo buscava a “fórmula da justiça”<sup>41</sup>, uma espécie de árvore de decisões que seria suficiente para resolver qualquer caso jurídico. Assim, a busca por algoritmos e fórmulas capazes de abordar questões antes consideradas exclusivamente humanas, como criatividade e justiça, evidencia a constante reavaliação das capacidades das máquinas e dos próprios seres humanos.

## 1.5 Máquinas e criatividade

Definições, em geral, tendem a ser controversas, sobretudo quando se referem a conceitos abstratos. No caso da "criatividade", as interpretações propostas frequentemente envolvem a aplicação de imaginação ou a habilidade de combinar de maneira distinta elementos pré-existentes. Considerando essa última perspectiva, os computadores parecem

---

<sup>40</sup> SHARPLES, Mike. “Storytelling by computer” in: *Digital Creativity*. 8. Taylor & Francis (Routledge), 1997, pp. 20-29. Também disponível online em: [https://www.researchgate.net/publication/232831648\\_Storytelling\\_by\\_computer](https://www.researchgate.net/publication/232831648_Storytelling_by_computer) - último acesso em 23 Jan 2023. No original: “The Holy Grail for story grammarians is a set of rules that would describe the structure of every existing written story.”

<sup>41</sup> CARDOZO, Benjamin N. *The Paradoxes of Legal Science*. Westport, CT: Greenwood Press, 1970. No capítulo 1, fazendo uma analogia com a teoria da relatividade de Einstein, Cardozo reconhece que as leis mudam, enfatizando que sua busca leva em consideração o caráter dinâmico do mundo jurídico.

possuir tal capacidade, sendo aparentemente capazes de executá-la em uma velocidade superior à dos seres humanos.<sup>42</sup> A questão da rapidez na execução será explorada adiante, quando abordarmos o Teorema do Macaco Infinito. Partindo de igual modo da frase de Lady Lovelace citada anteriormente, Margaret Boden<sup>43</sup> propõe quatro indagações acerca da criatividade com o título por ela proposto de “questões-Lovelace”: 1) criações computacionais podem nos ajudar a entender a criatividade humana? 2) computadores podem pelo menos parecer serem criativos? 3) computadores reconhecem criatividade? 4) os computadores podem ser criativos?

As questões-Lovelace servem como ponto de partida para a tese de BODEN (1991) de que existem três níveis de criatividade (que causam três níveis de “surpresa”): a) combinação de modo original de material já existente; b) exploração de espaços conceituais; c) transformação do espaço. Com base nessa classificação, seria possível sugerir que algoritmos já podem executar o primeiro tipo sem recorrer a códigos sofisticados de inteligência artificial e que o atual estágio de desenvolvimento da IA permitiria alcançar o segundo nível. A vitória do programa AlphaGo, da empresa DeepMind, sobre o campeão mundial do jogo Go teria sido conquistada ao reavaliar o objetivo do jogo (não mais conquistar o maior território possível, mas sim o mínimo necessário para dominar o tabuleiro), conforme ilustrado no documentário que acompanhou o desenvolvimento do programa.<sup>44</sup>

O desafio estaria no terceiro nível de criatividade, aquele que transforma radicalmente algum campo do conhecimento, a ponto de criar um novo paradigma<sup>45</sup> – nesta categoria estariam contribuições como a de Einstein, por exemplo. Embora consideremos todos estes “níveis de surpresa” (para manter a nomenclatura de Margaret Boden) como manifestações criativas, ainda há uma dificuldade em atribuir criatividade a uma máquina, como se uma propriedade - a imaginação - inerente à sua definição, não pudesse estar presente fora de uma consciência.

Esse modo de entender a criatividade e experimentá-la como atributo exclusivamente humano mantém um paralelo com a necessidade de existência de consciência associada ao processo para legitimar uma sentença judicial. Aqueles que defendem essa posição argumentam que apenas uma consciência seria capaz de atribuir sentido: o juiz ou juíza precisaria não apenas escrever, mas saber que escreve, enquanto a máquina pode até criar, mas

---

<sup>42</sup> SHARPLES, Mike & PÉREZ Y PÉREZ, Rafael. *Story Machines : how computers have become creative writers*. New York, NY: Routledge, 2022.

<sup>43</sup> BODEN, Margaret. A. *The creative mind: Myths & mechanisms*. New York, NY: Basic Books, 1991.

<sup>44</sup> Documentário completo disponível em: <https://www.youtube.com/watch?v=WXuK6gekUIY> – último acesso em 12 dez 2022;

<sup>45</sup> Sobre o tema das mudanças de paradigmas: KUHN, Thomas S. *The Structure of Scientific Revolutions*. Chicago: The University of Chicago Press, 1970

não sabe que cria. Essa hesitação já havia sido identificada no trabalho de Ramón López de Mántaras, que investigou o tema da criatividade em máquinas:

Criatividade computacional é o estudo da confecção de softwares que exibem comportamentos que seriam considerados criativos quando analisados em humanos. Esses softwares podem ser usados para tarefas autônomas, como inventar teorias matemáticas, escrever poemas, pintar quadros e compor música. No entanto, os estudos de criatividade computacional também nos permitem entender a criatividade humana e produzir programas para uso por pessoas, onde o software atua como um colaborador em vez de uma mera ferramenta. Historicamente, tem sido difícil para a sociedade aceitar que máquinas possam ser inteligentes e ainda mais difícil admitir que possam ser criativas.<sup>46</sup>

Para o autor, a utilização de algoritmos auxiliares na produção artística remonta aos anos 1960, quando a computação permitiu maior colaboração, potencializando o processo criativo coletivo. O aumento de escala da produção artística, especialmente na música e no cinema, resultou na aplicação de uma mentalidade industrial – uma “linha de produção” – de músicas e filmes. Com isso, a especialização de determinadas etapas de processos desmistificou tarefas até então vistas como unitárias, permitindo a percepção de que certas ações eram compostas de partes menores de caráter repetitivo, e poderiam ser mecanizadas. O avanço da capacidade computacional ao que parece ampliou o campo de tarefas que envolvem decisões, como por exemplo escolhas de melhor arranjo para uma melodia.

Contudo, existe um conjunto de autores que defende uma posição oposta, para os quais a criatividade não poderia prescindir de imaginação. Hubert Dreyfus argumenta que um comportamento, quando qualificado como inteligente, envolve capacidades que não são passíveis de programação<sup>47</sup>. De acordo com esse argumento, o que o algoritmo faz é qualitativamente diferente da produção de um humano, mesmo que na aparência seja semelhante. Essa impossibilidade de atingir uma propriedade *sine qua non* – presente apenas

---

<sup>46</sup> MÁNTARAS, Ramón López de. “Artificial Intelligence and the Arts: Toward Computational Creativity” in: *The Next Step: Exponential Life*. BBVA Open Mind, 2017, pp. 99-123. Disponível em: <https://www.bbvaopenmind.com/wp-content/uploads/2017/03/BBVA-OpenMind-book-The-Next-Step-Exponential-Life-1-1.pdf> - último acesso em 12 dez 2022. No original: “Computational creativity is the study of building software that exhibits behavior that would be deemed creative in humans. Such creative software can be used for autonomous creative tasks, such as inventing mathematical theories, writing poems, painting pictures, and composing music. However, computational creativity studies also enable us to understand human creativity and to produce programs for creative people to use, where the software acts as a creative collaborator rather than a mere tool. Historically, it has been difficult for society to come to terms with machines that purport to be intelligent and even more difficult to admit that they might be creative.”

<sup>47</sup> DREYFUS, HUBERT. *What Computers Still Can't Do: A Critique of artificial Reason*. Cambridge: MIT Press, 1992, p. 285. O autor explicitamente diz que “(...) nonprogrammable human capacities are involved in all forms of intelligent behavior”.

no humano – assemelha-se à ideia kantiana de sublime, e caracteriza a posição “essencialista”, que defende que existe uma essência na criatividade que está ligada ao biológico humano. O filósofo John Searle já equiparava consciência à digestão enquanto epifenômenos do biológico, sendo irreprodutíveis na máquina.<sup>48</sup> A posição essencialista se utiliza dessa exclusividade para negar a possibilidade do não biológico de criar, uma vez que criar exigiria consciência.

Entretanto, a consequência da adoção dessa posição essencialista resultaria na divisão em duas categorias distintas: a ação do agente que age conscientemente e a ação da ferramenta que não possui consciência do que faz, apesar de executar tarefas com qualidade e eficiência superiores. Em outras palavras, há um aparente paradoxo no essencialismo, pois o instrumento, ao desempenhar atividades criativas com maior eficácia, seria classificado como não criativo.

## 1.6 Criatividade e imaginação

O processo criativo é frequentemente associado à necessidade de imaginação, que pode ser definida como a habilidade de formar imagens mentais ou representações de elementos ausentes no ambiente. Essa capacidade tem sido investigada de forma mais abstrata em diversas áreas da Filosofia. A Estética examina o papel da imaginação na criação, enquanto a Ética se concentra em seu papel na tomada de decisões práticas.<sup>49</sup>

Na Filosofia da Mente, existem diferentes teorias sobre a natureza e a função da imaginação. Descartes a via como um processo puramente mental que poderia voluntariamente se desconectar da percepção e da experiência sensorial<sup>50</sup>, enquanto outros, como Shaun

---

<sup>48</sup> John Searle apresenta sua teoria do Naturalismo Biológico em várias de suas obras. Um dos mais conhecidos é seu livro "A redescoberta da mente" (1992). Nele, ele argumenta que a consciência deve ser entendida como um processo biológico e que os estados mentais são causados por processos neurobiológicos de nível inferior no cérebro. Ou seja, uma máquina jamais seria capaz de possuir consciência não por uma questão de limitação tecnológica, e sim por não ser um organismo vivo.

<sup>49</sup> De acordo como verbete “Imagination” da Stanford Encyclopedia of Philosophy, “Imagination is involved in a wide variety of human activities and has been explored from a wide range of philosophical perspectives. Philosophers of mind have examined imagination’s role in mindreading and in pretense. Philosophical aestheticians have examined imagination’s role in creating and in engaging with different types of artworks. Epistemologists have examined imagination’s role in theoretical thought experiments and in practical decision-making.” Disponível em <https://plato.stanford.edu/entries/imagination/> - último acesso em 23 Dez 2022.

<sup>50</sup> DESCARTES, René. Méditations sur la philosophie première. Édition électronique, v.: 1,0 : Les Échos du Maquis, 2011. Disponível em: <https://philosophie.cegeptr.qc.ca/wp-content/documents/M%C3%A9ditations-1641-1647.pdf> – último acesso em 15 Fev 2023. O seu exemplo da sereia, mostra ser possível uma combinação de experiências sensoriais que produz um resultado final não conectado com o mundo sensível. Em especial o trecho: « Car de vrai les peintres, lors même qu'ils s'étudient avec le plus d'artifice à représenter des sirènes et des satyres par des formes bizarres et extraordinaires, ne leur peuvent pas toutefois attribuer des formes et des natures entièrement nouvelles, mais font seulement un certain mélange et composition des membres de divers animaux; ou bien, si peut-être leur imagination est assez extravagante pour inventer quelque chose de si nouveau, que jamais

Gallagher<sup>51</sup>, consideram-na uma função cognitiva complexa que envolve a interação entre vários processos mentais, tais como percepção, memória e raciocínio.

Alguns filósofos argumentam que a imaginação desempenha um papel crucial no pensamento criativo, na resolução de problemas e no desenvolvimento conceitual. Outros enfatizam seu papel em nossas vidas emocionais e morais<sup>52</sup>, pois a imaginação nos permite ter empatia com os outros, entender diferentes perspectivas e explorar possíveis cenários e desfechos.

A leitura de David Strokes da obra de Kant “Crítica do Julgamento”<sup>53</sup> o leva a afirmar que o filósofo de Königsberg sugeria a distinção de duas formas de criação: por um lado a criatividade artística seria uma faculdade inata, incapaz de ser ensinada, uma vez que ligada ao que denominou “imaginação produtiva”, enquanto a criatividade nas descobertas científicas seria passível de aprendizado, pois estaria ligada à “imaginação reprodutiva”.<sup>54</sup>

O método escolástico, que curiosamente Kant parece adotar mesmo sendo um pensador moderno, orientava: "Quando encontrares uma contradição, fazei uma distinção". Ao nosso ver, sua solução apenas desdobra o problema da criatividade para a imaginação, e mais equipara a “imaginação reprodutiva” à representação do que distingue dois tipos de ações criativas. A abordagem parece ser a mesma de Aristóteles, que ao tratar a fantasia (φαντασία), utiliza o termo tanto para representação como para imaginação.<sup>55</sup> Sartre caminha pela mesma

---

nous n'ayons rien vu de semblable, et qu'ainsi leur ouvrage nous représente une chose purement feinte et absolument fausse, certes a tout le moins les couleurs dont ils le composent doivent-elles être véritables » (pp. 13-14)

<sup>51</sup> GALLAGHER, Shawn., & Zahavi, Dan. *The Phenomenological Mind*. Routledge, 2008.

<sup>52</sup> Por exemplo, Martha Nussbaum, para quem a imaginação desempenha um papel crucial no raciocínio ético e no julgamento moral, pois nos permite ter empatia com os outros e entender suas experiências e perspectivas. NUSSBAUM, Martha. *Frontiers of Justice: Disability, Nationality, Species Membership*. Cambridge, MA: Harvard University Press, 2006.

<sup>53</sup> Kant, Immanuel. *Critique of the Power of Judgment (The Cambridge Edition of the Works of Immanuel Kant)* (P. Guyer, Ed.; E. Matthews, Trans.). Cambridge: Cambridge University Press, 2000 6

<sup>54</sup> STROKES, Dustin. “Imagination and Creativity” in KIND, Amy (ed.), *The Routledge Handbook of the Philosophy of Imagination*. Routledge, 2016. A versão por nós utilizada está disponível em: <https://philarchive.org/rec/STOIA4> - último acesso em 01 Jan 2023.

<sup>55</sup> MODRAK, Deborah K.W. “Aristotle on phantasia” in: *The Routledge Handbook of Philosophy of Imagination*, 2016, online. Disponível em: <https://www.routledgehandbooks.com/doi/10.4324/9781315657905.ch1> - último acesso em 18 Jan 2023. O estagirita afirma que: *Phantasia* is “that in virtue of which we say an image (phantasma) arises in us” (428a1-2). “It is clear, therefore, that *phantasia* will be neither belief together with perception, nor belief through perception, nor a blend of belief and perception” (428 a 24-26). “Since *phantasia* is thought to be a kind of movement (...) this movement cannot exist apart from perception or in things, which do not perceive and in respect of it, it is possible for its possessor to do and be affected by many things, and it may be both true and false” (428 b 11-17). “Perceptive *phantasia*, then, as has been said, belongs also to Other animals, but deliberate *phantasia* belongs Only to rational animals for whether to do this or that, is already the work of rational faculty; and it is necessary to measure by one; for a person pursues the greater and so must be able to make one out of many images [*phantasmata*]” (434 a 7-19)

linha, e em nossa visão não resolve o problema proposto pelo ângulo funcionalista para a criação.<sup>56</sup>

Douglas Eck, pesquisador do Google Brain e ex-membro da equipe Magenta da empresa DeepMind, afirma que “a criatividade da IA não é como a criatividade humana. Trata-se de gerar coisas novas que podem ser interessantes e surpreendentes, mas não se trata de fazê-lo com o mesmo nível de intencionalidade, propósito ou compreensão que os humanos trazem para seus atos criativos”.<sup>57</sup> Mark Riedl escreve que “os sistemas de IA podem produzir saídas criativas, mas suas saídas são criativas apenas no sentido de surpreender os humanos ou de acordo com alguma definição abstrata de criatividade. Eles não têm o mesmo nível de intencionalidade ou propósito que os criadores humanos.”<sup>58</sup>

Assim, a questão parece ter um encadeamento sucessivo que leva sempre no final à necessidade de consciência para distinguir aquilo que o ser humano faz daquilo que a máquina produz. Se a criatividade exige imaginação e a imaginação pode apenas ser concebida enquanto produto de uma consciência, o argumento essencialista pode ser reduzido à necessidade de consciência, seja para a criatividade, seja para a noção de justiça, seja para quaisquer outras atribuições que se deseja defender como unicamente humana. Essa posição apenas classifica a origem da produção criativa, e não resolve a questão finalista, na qual ambas as produções são igualmente passíveis de serem definidas como criativas.

Ao estudar o caso de AARON, Sundararajan (2021) defende que a interface entre o humano e a máquina suscita questões não apenas epistemológicas, mas também ontológicas.<sup>59</sup> Em outras palavras, na busca por compreender ou diferenciar-se das máquinas, os seres humanos percebem a necessidade de examinar não apenas os processos - como ambas as entidades executam ações - mas também as propriedades que definem o humano e o distinguem do não-humano. Simondon (2005) examinou a relação entre humanos e tecnologia,

---

<sup>56</sup> O foco da análise de Sartre é a imaginação enquanto componente do self, constituinte do processo de concepção da imagem de si – isso porque a criatividade provê ao ser humano a capacidade de enxergar-se em uma situação diferente da atual, em um outro mundo possível, algo que está ligado ao horizonte de possibilidades do ser e à liberdade. SARTRE, J-P. *L'Être et le néant – essai d'ontologie phénoménologique*. France: Gallimard, 2007.

<sup>57</sup> Disponível em: [https://www.ted.com/talks/douglas\\_eck\\_the\\_role\\_of\\_ai\\_in\\_creativity](https://www.ted.com/talks/douglas_eck_the_role_of_ai_in_creativity) - último acesso em 14 Fev 2023.

<sup>58</sup> RIEDL, Mark O. “Human-centered artificial intelligence and machine learning.” in: *Human Behavior & Emerging Technologies*. 2019; 1: pp. 33 - 36. Disponível em: <https://arxiv.org/pdf/1901.11184.pdf> . No último parágrafo, o autor concluiu que “Human-centered artificial intelligence does not mean that an artificial intelligence or machine learning algorithm must think like a human or be cognitively plausible.”

<sup>59</sup> SUNDARARAJAN, Louise. “Harold Cohen and AARON: Collaborations in the Last Six Years (2010–2016) of a Creative Life” in: LEONARDO, Vol. 54, No. 4, pp. 412–417, 2021, p. 415. Disponível em: [http://direct.mit.edu/leon/article-pdf/54/4/412/1959027/leon\\_a\\_01906.pdf](http://direct.mit.edu/leon/article-pdf/54/4/412/1959027/leon_a_01906.pdf) – último acesso em 26 Dez 2022. No original: “a clear pattern emerged from the amorphous interface between human and machine: relationship between human and machine is shown to have both epistemological and ontological consequence”

enfazando a noção de tecnicidade, que se refere à dimensão técnica inerente ao ser. Ele viu a tecnologia não apenas como ferramentas externas, mas como uma parte fundamental de nossa existência e defendeu uma abordagem integrativa, que considera a coevolução de humanos e tecnologia.<sup>60</sup>

Foi Heidegger quem retomou no debate contemporâneo a questão das relações entre o humano e a técnica, afirmando que ferramenta e ser humano conectam-se. Para ele, não foi a física moderna que transformou a natureza em meio, colocando-a a nosso serviço, pois mesmo um agricultor já fazia isso há milênios. A novidade da física moderna, a partir do conjunto de técnicas por ela selecionado, é haver transformado a natureza em um tipo específico de meio – aquele que a olha como “um sistema operativo e calculável de forças”. O humano matematizou a natureza, a partir de uma técnica que se apoia em outras técnicas, tais como a álgebra, para criar uma representação que não existia nem a priori, nem fora do humano.<sup>61</sup> Do mesmo modo, os sistemas de IA matematizaram a linguagem, reunindo um conjunto de técnicas (*deep learning*, NPL, ou mesmo a informação por bits) que estabelecem um vínculo entre linguagem e probabilidade de posicionamento de termos em uma frase que não existia a priori na linguagem. Esta forma de se enxergar a linguagem é uma escolha feita a partir da utilização desta técnica, uma dentre muitas formas possíveis de “des-encobrimento”, nos termos heideggerianos.<sup>62</sup>

Diante da ausência de consenso sobre a definição de criatividade, parece ser difícil questionar se há mais de um tipo de criatividade,<sup>63</sup> dependendo da entidade que exerce tal atividade.<sup>64</sup> No entanto, pesquisas como a de Jordanous (2012) propõem diferenciar a criatividade humana da criatividade mecânica e, além disso, estabelecer critérios para avaliar esta última.<sup>65</sup>

---

<sup>60</sup> SIMONDON, Gilbert. *L'invention dans les techniques*. Paris : Éditions du Seuil, 2005. Em especial “Invention et créativité” (pp. 329-337)

<sup>61</sup> HEIDEGGER, Martin. “A questão da técnica” in: *Ensaio e conferências*, 8ª ed. Petrópolis: Vozes, 2012, pp. 24-25.

<sup>62</sup> RUDIGER, Francisco. “Humanismo, arte e tecnologia segundo Heidegger” in: *Fragmentos de Cultura*, v. 21, n. 7/9, pp. 433-451. Jul/set 2011. Disponível em: <https://repositorio.pucrs.br/dspace/handle/10923/9821?mode=full> – último acesso em 28 Jun 2023.

<sup>63</sup> MARKS II, Robert J; DEMBSKI, William A. & EWERT, Winston. *Introduction to Evolutionary Informatics* (1st ed.). Singapore: World Scientific Publishing Company, 2017.

<sup>64</sup> Aliás, agência criativa aparente e agência criativa autêntica é uma diferenciação que só importa para uma visão essencialista, como propomos, a partir da provocação que Silveira (2021) faz para o tema da agência moral, em especial em SILVEIRA, Paulo Antonio Caliendo Velloso da. *Ética e inteligência artificial: da possibilidade filosófica de agentes morais artificiais*. Porto Alegre, RS: Editora Fi, 2021, p. 151 e p. 173.

<sup>65</sup> JORDANOUS, Anna. “A Standardised Procedure for Evaluating Creative Systems: Computational Creativity Evaluation Based on What it is to be Creative. *Cognitive Computation*, 4 (3), 2012. pp. 246-279. ISSN 1866-9956. E-ISSN 1866-9964. Disponível em: <https://kar.kent.ac.uk/42379/> - último acesso em 14 Fev 2023.

## 1.7 O teorema do macaco infinito e seu papel no funcionalismo

A habilidade de combinar elementos existentes de maneiras distintas é frequentemente empregada na definição de criatividade, como mencionado anteriormente. A análise combinatória na matemática apresenta duas abordagens para calcular as possibilidades de resultados: a) quando os itens a serem combinados podem ser utilizados em uma quantidade finita, resultando em combinações finitas, ainda que numerosas (como no caso do xadrez, jogo no qual as opções de jogadas são imensas, porém finitas); b) quando os itens a serem combinados podem ser utilizados infinitamente, sem limite de repetição, levando a uma quantidade infinita de resultados possíveis. O teorema do macaco infinito é um experimento mental [*thought experiment*] que explora essa segunda categoria de possibilidade combinatória, postulando que, se um macaco fosse posicionado diante de uma máquina de escrever e pressionasse aleatoriamente as teclas, incluindo letras, números, símbolos e espaços, e tivesse tempo infinito para realizar essa tarefa, acabaria por produzir todos os livros já criados, sejam eles clássicos ou não, de alta ou baixa qualidade, escritos em prosa ou poesia.<sup>66</sup> A velocidade de produção seria irrelevante, uma vez que existe a disponibilidade infinita de tempo. Entre os textos digitados, estaria por exemplo a obra completa de Shakespeare, ou o texto integral de “Os Lusíadas”, e até mesmo a Constituição da França ou o extrato do mês de outubro de 1921 de uma conta bancária de Ruy Barbosa.

O macaco enquanto agente do experimento é meramente ilustrativo, uma vez que o conceito central na discussão proposta é a aleatoriedade combinatória de caracteres levada ao infinito. O nome do teorema foi derivado de uma formalização proposta pelo matemático francês Émile Borel, que busca comprovar a possibilidade maior do que zero – mesmo que infinitamente pequena – para o aparecimento de uma obra de arte de modo aleatório.

A prova proposta<sup>67</sup> considera a probabilidade de digitar a palavra “banana” em um teclado de 50 teclas. Supondo que as teclas sejam pressionadas aleatoriamente e

---

<sup>66</sup> BOREL, Émile. "Mécanique Statistique et Irréversibilité". J. Phys. (Paris). Series 5. 3: 189–196. O trecho no qual a sugestão aparece é: “Concevons qu'on ait dressé un million de singes à frapper au hasard sur les touches d'une machine à écrire et que [...] ces singes dactylographes travaillent avec ardeur dix heures par jour avec un million de machines à écrire de types variés. [...] Au bout d'un an, [leurs] volumes se trouveraient renfermer la copie exacte des livres de toute nature et de toutes langues conservés dans les plus riches bibliothèques du monde.” E pode ser encontrado em: <https://web.archive.org/web/20151130132642/https://citations.webescence.com/citations/Emile-Borel/Concevons-ait-dresse-million-singes-frapper-hasard-sur-les-touches-une-machine-2819> - último acesso em 08 Dez 2022.

<sup>67</sup> A prova proposta pode ser encontrada em:

independentemente (o que significa que cada tecla teria uma chance igual de ser pressionada independentemente daquela que tenha sido pressionada anteriormente) a chance de que a primeira letra digitada seja “b” é 1/50; a chance de que a segunda letra digitada seja “a” também é 1/50; e assim por diante. Portanto, a probabilidade de as seis primeiras letras formarem banana é

$$(1/50) \times (1/50) \times (1/50) \times (1/50) \times (1/50) \times (1/50) = (1/50)^6 = 1/15,625,000,000.$$

Ou seja, menos de um em 15 bilhões, mas não zero. Pelo exposto, a chance de não digitar banana em um determinado bloco de 6 letras é de  $1 - (1/50)^6$ . Como cada bloco é digitado independentemente, a chance  $X_n$  de não digitar banana em nenhum dos primeiros  $n$  blocos de 6 letras é dada pela fórmula:

$$X_n = \left(1 - \frac{1}{50^6}\right)^n.$$

À medida que  $n$  cresce,  $X_n$  fica menor. Para  $n = 1$  milhão,  $X_n$  é aproximadamente 0,9999, mas para  $n = 10$  bilhões  $X_n$  é aproximadamente 0,53 e para  $n = 100$  bilhões é aproximadamente 0,0017. À medida que  $n$  se aproxima do infinito, a probabilidade  $X_n$  se aproxima de zero; isto é, tornando  $n$  grande o suficiente,  $X_n$  pode ser tão pequeno quanto desejado, e a chance de digitar banana se aproxima de 100%. Assim, a probabilidade da palavra banana aparecer em algum ponto de uma sequência infinita de teclas é igual a um.

Um dos corolários do teorema do macaco infinito é que uma obra de arte, como por exemplo um poema de Fernando Pessoa, poderia aparecer no mundo como ato deliberado e planejado por uma consciência, mas também de forma aleatória. Se os parâmetros para a confecção do texto são reduzidos (por exemplo, “escreva um soneto decassílabo com o tema ‘amor’<sup>68</sup>”), então uma infinidade de tempo passa a produzir muito mais resultados aceitáveis. Jorge Luis Borges já afirmava em 1941 que todos os textos possíveis já estão presentes no dicionário, sendo o restante apenas uma questão de ordem em que as palavras aparecem<sup>69</sup>, o que transmite a mesma ideia.

---

[http://www.apstatsmonkey.com/StatsMonkey/TPS3e\\_files/TypingMonkeys.pdf](http://www.apstatsmonkey.com/StatsMonkey/TPS3e_files/TypingMonkeys.pdf) - último acesso em 12 Dez 2022.

<sup>68</sup> Podemos observar que este tipo de instrução é exatamente a forma de prompt de comandos para o ChatGPT

<sup>69</sup> BORGES, Jorge Luis. La biblioteca de babel. Buenos Aires: Emecé, 2000.

Outra observação é a de que o avanço da capacidade computacional permite com que cada segundo de processamento com a atual tecnologia signifique milhões de apertos de teclas do macaco imaginário. Deste modo, cresce a probabilidade de que em um determinado espaço de tempo, o resultado da combinação aleatória de caracteres se revele um texto que faça sentido para o ser humano que observa os resultados. Adotamos a posição de que no momento o algoritmo não vê sentido em nenhuma das suas produções, seja ela considerada ininteligível ou de excelente qualidade por um ser humano. Para retornar ao argumento essencialista, o computador faz mas não sabe que faz, e, de acordo com essa posição, somente um ser humano poderia julgar a qualidade dessa produção.

Dessa forma, uma obra original e de alta qualidade pode ser concebida por uma ação aleatória à qual os seres humanos atribuem a propriedade de "criatividade". Esse resultado nos reconduz à indagação se uma máquina pode ser "criativa". Sob uma perspectiva finalista, a criatividade seria um produto da combinação de elementos pré-existentes que atendessem aos critérios de originalidade e qualidade reconhecidos pelos seres humanos, independentemente do processo de produção. Assim, estaríamos dissociando criatividade de imaginação e até mesmo de consciência, despojando o conceito de atributos atualmente reservados às entidades vivas.

A problematização da máquina como autora de um texto criativo ganhou maior relevância a partir da metáfora da máquina na escrita no século XX. O termo "escritura", privilegiado pelos franceses e pelo estruturalismo, é fundamental nesse contexto. Ele sugere que a autoria não é resultado de um "espírito" autoral, mas das possibilidades combinatórias da língua, abrindo a possibilidade de entender a escrita como um ato dissociado de significação. Mais do que problemas da lógica, computação ou mesmo da linguística, o debate originário do estruturalismo nas décadas de 1950 e 1970 aborda a questão da separação entre "criatividade" e intenção, o que levaria, pode-se pensar, ao questionamento da autoria. A desconexão entre o resultado e uma consciência que o pretende permitiu ao funcionalismo legitimar a existência de algo sem seu criador, argumento central do existencialismo.

Portanto, a discussão sobre criatividade e autoria em relação às máquinas não apenas revela as fronteiras do entendimento humano sobre a natureza da criatividade, mas também questiona as próprias noções de autoria e intencionalidade, abrindo caminho para uma compreensão mais ampla e inclusiva desses conceitos no contexto das inteligências artificiais. Nessa perspectiva funcionalista, a aceitação de uma sentença emitida por um computador como adequada e válida dependeria menos de sua origem e mais de sua capacidade de atender aos critérios de legitimidade estabelecidos pela sociedade.

## Capítulo 2 – Algoritmos e Realismo Jurídico

Neste capítulo, nosso objetivo é investigar o argumento essencialista de que o julgamento por algoritmos representaria a implementação de um novo realismo jurídico. Com base no pensamento de expoentes do *legal realism*, como Oliver Wendell Holmes e Jerome Frank, analisaremos a validade desse argumento. Para tanto, precisamos antes identificar as principais teses defendidas pelo realismo jurídico e o modo pelo qual foram resgatadas pelos críticos contemporâneos da emissão de sentenças por algoritmos. Além disso, buscaremos compreender por que essas teses deveriam ser rejeitadas - seja porque a máquina falha em fornecer algum elemento crucial, ou porque a emergência das cortes digitais provocaria uma mudança na concepção de justiça - e como isso seria prejudicial à sociedade.

No lado oposto do debate, o principal argumento dos defensores da substituição de humanos por máquinas no sistema judicial é a ideia de que as sentenças geradas com base em um vasto conjunto de casos anteriores promoveriam maior segurança jurídica em comparação à imprevisibilidade das decisões humanas. Esta preocupação estava, de fato, explicitamente presente na obra "Courts on Trial" de Jerome Frank, um dos principais representantes do realismo jurídico. Ademais, os chamados "neo-realistas" - caso possam ser rotulados assim - argumentam que defender a tese de que existem elementos indelegáveis às máquinas pode ser uma estratégia para ocultar falhas no funcionamento das cortes que poderiam ser corrigidas se não fossem encobertas. Curiosamente, essa ideia já estava presente há quase um século nos textos de Jerome Frank<sup>70</sup>, embora na época ele não tratasse especificamente das cortes digitais.

Em textos recentes<sup>71</sup> que revisitam as críticas ao Realismo como escola de pensamento na Filosofia do Direito, embora não abordem diretamente a algoritmização do sistema judicial, argumenta-se que os autores realistas foram frequentemente criticados por meio da estratégia da falácia do espantalho. Essa tática consiste em exagerar ou simplificar aspectos dos textos a ponto de distorcê-los, para então atacar essas distorções, sem que de fato haja qualquer autor que sustente as ideias conforme apresentadas pelos críticos. Por exemplo, Jerome Frank defendia uma análise mais cuidadosa dos fatores extrajurídicos que influenciam a

---

<sup>70</sup> FRANK inclusive se considera um "reformador", pretendendo trazer à luz a realidade do funcionamento do sistema judicial, para que suas falhas sejam corrigidas. FRANK, Jerome. *Courts on Trial*. New Jersey: Princeton University Press, 1973, p. 2

<sup>71</sup> PRIEL, Dan. "What the Judge Had for Breakfast: A Brief History of an Unpalatable Idea" in: SSRN 12 jun 2017. Disponível em: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2982716](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2982716) – último acesso em 04 abr 2022.

discrecionalidade<sup>72</sup>, mas aparentemente nunca defendeu a ideia de que “as decisões judiciais estão relacionadas com aquilo que os juízes comeram no café da manhã”<sup>73</sup>. Priel (2017) afirma que nenhum jurista identificado com o realismo jurídico realmente estabeleceu uma conexão determinista e unicausal entre digestão e decisão judicial. A suposição de que um mal-estar gástrico poderia influenciar uma sentença, seja como fator central ou não, foi na verdade astuciosamente selecionada e reduzida a uma causa única pelos críticos.<sup>74</sup>

Portanto, é necessário examinar mais detalhadamente as nuances do pensamento de Jerome Frank para compreendermos por que essa crítica falha em abordar adequadamente sua obra.

## 2.1 O realismo jurídico e seu conceito de “Direito”

Na obra "Courts on Trial", Frank (1973) busca esclarecer por que o uso do termo "direito" pode ser considerado equivocado. De fato, em nenhum momento ele afirma que "direito é o que as cortes dizem" - um espantalho comum criado por seus críticos. Em vez disso, ele argumenta que "o que as cortes dizem" representa a realidade do mundo jurídico. O que parece a princípio uma sutileza, representa em verdade uma significativa diferença semântica. Para Frank, "direito" é um anseio sobre o que as cortes deveriam dizer, por parte daqueles que acreditam na existência de uma racionalidade jurídica subjacente às sentenças, que em tese governaria o funcionamento das cortes e justificaria suas decisões. Segundo ele, nenhuma lógica jurídica vincula [*bind*] as decisões dos juízes. Caso isso ocorresse, as partes sempre saberiam de antemão qual seria a decisão de uma corte. O desejo de encontrar uma racionalidade jurídica por detrás de uma decisão judicial (como se fosse uma lei da gravidade que estaria sempre agindo sobre o seu objeto) mascara a atribuição que na verdade ocorre *ex post* à sentença: somente após sua elaboração é que se tenta explicá-la.

A origem dessa reflexão é comumente atribuída a Oliver Wendell Holmes. Na *Lecture I* da sua principal obra, *The Common Law*, ele afirma logo no início que a lei – no sentido das normas aplicáveis a casos concretos, e não o direito enquanto ordenamento jurídico – é

---

<sup>72</sup> FRANK, Jerome. *Courts on Trial*. New Jersey: Princeton University Press, 1973, p. 375

<sup>73</sup> PRIEL, Dan. “What the Judge Had for Breakfast: A Brief History of an Unpalatable Idea” in: SSRN 12 jun 2017, p. 2. Disponível em: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2982716](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2982716) – último acesso em 04 abr 2022.

<sup>74</sup> PRIEL, Dan. “What the Judge Had for Breakfast: A Brief History of an Unpalatable Idea” in: SSRN 12 jun 2017, p. 2. Disponível em: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2982716](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2982716) – último acesso em 04 abr 2022.

construída historicamente, e não pela lógica.<sup>75</sup> Jerome Frank estendeu essa ideia e rotulou qualquer teoria jurídica como “magia”, sendo um engodo sem correspondência com a realidade do judiciário.

Holmes (1899) nos oferece um exemplo para ilustrar seu argumento: a teoria jurídica faz parecer que existe um conceito de contrato anterior aos primeiros contratos celebrados, quando na verdade é o oposto: dos atos identificados como compromissos deriva-se o conceito de contrato.<sup>76</sup> Além disso, há um percurso histórico: diferentes formas de compromisso foram sendo testadas com o tempo até que o contrato, seja por performance oral, seja escrito, se firmou como a mais eficiente.<sup>77</sup> Depois que uma ideia geral abstrata se firma (o conceito de “contrato”), é difícil perceber que foi criada e que não estava sempre presente nas sociedades, uma vez que traz “conforto à mente”<sup>78</sup>, no sentido da segurança que proporciona em termos de padrão de comportamento esperado.

Assim, Holmes (1899) argumenta que se a perspectiva histórica contribui em alguma forma com nosso conhecimento do direito, é no sentido de que nos liberta do tipo de exegese feita *ad hoc* para justificar textos com conceitos que muitas vezes surgiram depois da regra em si, e não foram fundamento para a sua criação<sup>79</sup>: a ciência aplicada ao Direito visa a trazer acuracidade, não necessariamente quantitativa, e sim de rigor conceitual.<sup>80</sup>

Alicerçado nas ideias de Holmes, Frank conclui que o sistema jurídico é, em sua realidade, incerto, e qualquer idealização de certeza configuraria um autoengano.<sup>81</sup> Segundo ele, o melhor a fazer seria expor essa incerteza em vez de ocultá-la sob supostas racionalidades jurídicas inexistentes na realidade das decisões judiciais.<sup>82</sup> Além disso, o objetivo da análise de decisões seria explicitar as contingências do processo judicial que parecem prejudicar a busca pela justiça (por exemplo, provas desaparecidas, testemunhas que alteram seus depoimentos)

---

<sup>75</sup> HOLMES, Oliver Wendell. *The Common Law*. New Orleans, LA: Quid Pro Books, 2010. No original: “The life of the law has not been logic: it has been experience. The felt necessities of the time, the prevalent moral and political theories, institutions of public policy, avowed or unconscious, even the prejudices which judges share with their fellow-men, have had a good deal more to do than the syllogism in determining the rules by which men should be governed. The law embodies the story of a nation’s development many centuries, and it cannot be dealt with as if it contained only the axioms and corollaries of a book of mathematics.” (p. 1)

<sup>76</sup> HOLMES, Oliver Wendell. “Law in Science and Science in Law.” *Harvard Law Review*, vol. 12, no. 7, 1899, pp. 443–63. Disponível em: <https://doi.org/10.2307/1321177> - último acesso em 07 Jul 2022, p. 448

<sup>77</sup> HOLMES, Oliver Wendell. “Law in Science and Science in Law.” *Harvard Law Review*, vol. 12, no. 7, 1899, pp. 443–63. Disponível em: <https://doi.org/10.2307/1321177> - último acesso em 07 Jul 2022, p. 449

<sup>78</sup> HOLMES, Oliver Wendell. “Law in Science and Science in Law.” *Harvard Law Review*, vol. 12, no. 7, 1899, pp. 443–63. Disponível em: <https://doi.org/10.2307/1321177> - último acesso em 07 Jul 2022, p. 451

<sup>79</sup> HOLMES, Oliver Wendell. “Law in Science and Science in Law.” *Harvard Law Review*, vol. 12, no. 7, 1899, pp. 443–63. Disponível em: <https://doi.org/10.2307/1321177> - último acesso em 07 Jul 2022, p. 452 e p. 455

<sup>80</sup> HOLMES, Oliver Wendell. “Law in Science and Science in Law.” *Harvard Law Review*, vol. 12, no. 7, 1899, pp. 443–63. Disponível em: <https://doi.org/10.2307/1321177> - último acesso em 07 Jul 2022, p. 457

<sup>81</sup> Sobre o tema, GIANETTI, Eduardo. *Auto-engano*. São Paulo: Companhia das Letras, 2017.

<sup>82</sup> FRANK, Jerome. *Courts on Trial*. New Jersey: Princeton University Press, 1973, p. 410

bem como as emoções dos juízes. Esses fatores deveriam ser reconhecidos como elementos constituintes do processo, impossíveis de serem eliminados.<sup>83</sup> Essa mentalidade [*mindset*] favoreceria a aceitação do sistema judicial como ele realmente é, e não como deveria ser.

Todavia, há que se admitir que valores representam o dever-ser, e devem continuar a ser buscados, na implementação constante de melhorias para que a resolução de conflitos se dê da melhor forma possível. Mas, para ele, descrever a realidade dentro de conceitos que eliminam aspectos observados visando a adequar a descrição ao ideal devem ser rejeitados. Aliás, essa seria a principal proposta reformista de Frank: limpar do descritivo o conceitual. “Precisamos de uma nova coragem – a coragem de enfrentar imperfeições insolúveis nos problemas humanos.”<sup>84</sup>

Ao abordar o conceito de Direito Natural, o autor ressalta que essa ideia tem sido usada arbitrariamente em casos de contestação de leis positivas, apresentando uma espécie de direito imutável contraposto ao direito positivo vigente que deveria ser combatido (desde Antígona!). Segundo Frank, na realidade, o conceito de "Direito Natural" parece ser tão maleável que pode até ser empregado contra a dignidade humana, como, por exemplo, quando elimina as diferenças entre as pessoas, apagando todo o esforço kantiano que valoriza a importância das idiossincrasias individuais em detrimento de uma suposta igualdade natural.<sup>85</sup>

Assim, a natureza não deveria ser justificativa ou modelo para um dever-ser subjacente a uma racionalidade jurídica, como já evidencia a falácia naturalista.<sup>86</sup> Se há algo “natural” no sentido de inerente à humanidade, e passível de generalização, seria apenas a presença de alguma forma de nomos para garantir a convivência entre os membros da sociedade, enquanto o seu conteúdo viria por convenção.<sup>87</sup> Afinal, conforme argumenta Frank, o que seria o processo civilizatório senão a tentativa constante de domar a natureza<sup>88</sup>, “fixando regras de convivência que nos permitam transcender à condição puramente natural de nossos ancestrais primatas?”<sup>89</sup>

---

<sup>83</sup> FRANK, Jerome. *Courts on Trial*. New Jersey: Princeton University Press, 1973, p. 412. O autor afirma: “The best we can hope for is that the emotions of the trial judge will be sensitive, nicely balanced, subjected to his own scrutiny” (p. 412)

<sup>84</sup> FRANK, Jerome. *Courts on Trial*. New Jersey: Princeton University Press, 1973, p. 425. No original: “We need a new courage – the courage to face unconquerable imperfections in the solution of human problems”

<sup>85</sup> FRANK, Jerome. *Courts on Trial*. New Jersey: Princeton University Press, 1973, p. 356

<sup>86</sup> KESSLER, Friedrich, "Natural Law, Justice and Democracy--Some Reflections on Three Types of Thinking about Law and Justice," 19 *Tulane Law Review* 32, 1944, pp.32-61.

<sup>87</sup> GIANNETTI, Eduardo. *O anel de Gíges: uma fantasia ética*. São Paulo: Companhia das Letras, 2020, p. 108.

<sup>88</sup> ELIAS, Norbert. *The Civilizing Process*. Hoboken, NJ: Wiley-Blackwell, 1994, em especial o volume 2.

<sup>89</sup> GIANNETTI, Eduardo. *O anel de Gíges: uma fantasia ética*. São Paulo: Companhia das Letras, 2020, p. 118.

Assim sendo, pode-se afirmar que Frank acredita que um mínimo denominador comum, base para um direito natural geograficamente uniforme e atemporal, seria extremamente reduzido: “Existem uniformidades mínimas na natureza humana e, em qualquer cultura em qualquer época, uniformidades culturais mínimas”<sup>90</sup>. Ademais, cada juiz possui seu próprio conjunto de normas morais pessoais, que são utilizadas, mesmo que inconscientemente, na decisão proferida. Como o mínimo denominador comum de normas morais entre os juízes é limitado, Frank argumenta que é impossível prever suas decisões individuais, mesmo levando em consideração que o conjunto de juízes está imerso em uma cultura jurídica comum.<sup>91</sup> Além disso, não se poderia afirmar que as decisões de um juiz fossem diretamente uma proxy do seu conjunto de leis morais, pois este é apenas um dos incontáveis elementos influenciando suas decisões.

Dessa forma, podemos afirmar que os críticos das ideias de Jerome Frank insistem que a ideia de “direito” existe antes do ato de julgar, o que de certo modo não corresponde ao seu pensamento, e sim dialoga com uma teoria do essencialismo. Seria racionalidade jurídica [*legal reasoning*] então uma forma de categorização? Vejamos se essas ideias estariam conectadas.

## 2.2 Racionalidade jurídica e categorização

Desde Aristóteles, especificamente em sua obra "Categorias"<sup>92</sup>, as noções de substância e propriedade desempenham um papel fundamental na metafísica, uma vez que lidam com os conceitos mais abrangentes sobre o ser e seus aspectos mais amplos da realidade. As proposições categóricas, por sua vez, são a forma como tais aspectos se manifestam quando formuladas em proposições linguísticas, ou seja, por meio da linguagem enquanto instrumento para descrever e fazer afirmações acerca do mundo. Ainda que Aristóteles tenha proposto um total de dez categorias, "substância" e "propriedade" são as formas mais simples de se dividi-las, dado que o mundo é composto tanto de entidades que carregam propriedades quanto de propriedades que são atribuídas a tais entidades.

---

<sup>90</sup> FRANK, Jerome. *Courts on Trial*. New Jersey: Princeton University Press, 1973, p. 179. No original: “There are minimal uniformities in human nature, and, in any given culture at any given time, minimal cultural uniformities”

<sup>91</sup> FRANK, Jerome. *Courts on Trial*. New Jersey: Princeton University Press, 1973, p. 180

<sup>92</sup> Aristóteles. **Categorias**, cap V (2a11-2a19 Bekker): Οὐσία δὲ ἐστὶν ἡ κυριώτατά τε καὶ πρώτως καὶ μάλιστα λεγομένη, ἢ μήτε καθ' ὑποκειμένου τινὸς λέγεται μήτε ἐν ὑποκειμένῳ τινὶ ἐστὶν, οἷον ὁ τις ἄνθρωπος ἢ ὁ τις ἵππος. δευτέραι δὲ οὐσίαι λέγονται, ἐν οἷς εἶδεν αἱ πρώτως οὐσίαι λεγόμεναι ὑπάρχουσιν, ταῦτά τε καὶ τὰ τῶν εἰδῶν τούτων γένη· οἷον ὁ τις ἄνθρωπος ἐν εἶδει μὲν ὑπάρχει τῷ ἀνθρώπῳ, γένος δὲ τοῦ εἶδους ἐστὶ τὸ ζῷον· δευτέραι οὖν αὐταὶ λέγονται οὐσίαι, οἷον ὁ τις ἄνθρωπος καὶ τὸ ζῷον. (...)

O termo "substância" vem do latim "substantia" – "aquilo que subjaz", "que está por detrás dos atributos". O conceito clássico de substância (no grego *ousia* ou *hypóstasis*) foi originalmente caracterizado por Aristóteles nas Categorias (II-5) no contexto da sua crítica à teoria das formas de Platão, como aquela *coisa individual* (portanto, não predicado ou relação) à qual inerem várias propriedades ou atributos, enquanto ela mesmo não inere e não é, tem pode ser, propriedade de qualquer outra coisa, e que ele chamava de "*to de ti*" (Metafísica, V 8, 1017b).<sup>93</sup>

Com o objetivo de se dizer do mundo, é necessário caracterizar a entidade (*res*) com as propriedades pertinentes. Observa-se que as propriedades e as entidades que as carregam apresentam comportamentos distintos: as propriedades se repetem, isto é, pertencem a ou são próprias de mais de um indivíduo, enquanto as entidades que carregam essas propriedades não se repetem, uma vez que cada indivíduo pode ser identificado como único a partir de seu conjunto irrepitível de propriedades. Desse modo, a palavra "substância" (*sub-stare*, "estar embaixo de") adquire seu significado metafísico clássico (presente em Aristóteles e nos pensadores medievais) mais próximo de "objeto individual" ou "coisa individual": no aristotelismo, a substância representa a categoria que unifica a investigação dos aspectos do ser. A multiplicidade de aspectos do ser parece estar unificada no conceito de substância.

O termo οὐσίᾱ, presente particípio de εἶμι ("aquilo que está sendo"), implica que a substância se mantém com a passagem do tempo, tem subsistência, mesmo que ganhe algumas propriedades e perca outras ao longo do tempo durante o qual subsiste. A οὐσίᾱ do direito permaneceria, mesmo que algumas de suas propriedades ao longo do tempo se alterassem.<sup>94</sup> Uma das buscas da metafísica é entender como a identidade se mantém, apesar da mudança de propriedades da coisa (paradoxo do navio de Teseu: como é possível falar da substância "direito" para a Constantinopla de Justiniano e a Bolonha de Acúrsio se nem a sociedade, nem as leis são mais as mesmas?).

O aristotelismo tomista sustentava que, embora a substância pudesse adquirir e perder propriedades, haveria um conjunto subjacente propriedades preservadas que seriam constitutivas da substância. Desse modo, "ser" distingue-se de "substância", até porque as

---

<sup>93</sup> IMAGUIRE, Guido. A substância e suas alternativas: feixes e tropos. in: **Metafísica Contemporânea**. Petrópolis, RJ: Editora Vozes, p.273

<sup>94</sup> Esta aparente contradição também já foi por nós explorada em CAMARGO, Guilherme Affonso Ferreira de. *Conflitos Jurisdicionais no século XIII: O Caso do Mosteiro de Santa Inês de Bolonha (1223-1298)*. Universidade Federal do Rio Grande do Sul. Instituto de Filosofia e Ciências Humanas. Dissertação, 2021. 149 f.

propriedades que não são substância também são (como afirmava Aristóteles, “o ser se diz de muitas maneiras”: τὸ δὲ ὄν λέγεται μὲν πολλαχῶς)<sup>95</sup>

Assim sendo, se estamos falando em uma substância “direito”, que não apenas existe independentemente<sup>96</sup> do conjunto de sentenças emitidas pelo sistema judicial, mas que são constitutivas do processo pelo qual deve passar a emissão de uma sentença, qual seria seu status ontológico? O direito existe independentemente do ordenamento jurídico de uma sociedade ou apenas pode existir nele se manifestando?

É uma questão clássica, aliás, se uma propriedade não instanciada continua existindo como uma entidade real (para Platão, os universais são formas ou entidades objetivas que independem de sua instanciação para existirem). Se todos os objetos vermelhos do mundo desaparecessem, a cor vermelha continuaria existindo? Realistas moderados dizem que sem objetos vermelhos a cor vermelha não existiria mais: *universalia in rebus*. Nominalistas negam sua existência, pois universais são meras palavras, *flatus vocis*. Platonistas diriam que os universais subsistem independentes das coisas (*universalia ante rem*), ou, numa reedição contemporânea, que a palavra "vermelho" continuaria a ter um significado, e o significado seria justamente esta cor, e que ela continuaria existindo, mesmo não sendo exemplificada em nenhum objeto, da mesma forma que a nona sinfonia de Beethoven continua a existir mesmo que se quebrem todos os discos onde está gravada.<sup>97</sup>

Mais do que isso, se o direito é propriedade do processo de se decidir uma lide, existindo anteriormente ao texto, e delimitando seu horizonte de possibilidades, então ele seria uma espécie de conjunto de regras limitadoras dos resultados possíveis de uma sentença, o que parece ser exatamente a definição de um algoritmo. Podemos afirmar, deste modo, que o essencialismo continua a apresentar uma espécie de incompatibilidade com sua função prática:

---

<sup>95</sup> Aristóteles, **Metafísica** Gama, 1003 a33 (Bekker)

<sup>96</sup> Três seriam as características para definir uma substância: i) individualidade: x é individual se x não se repete (o direito medieval seria conjecturado como algo que não aparece em outro momento da história, e possui propriedades únicas); ii) independência ontológica: x é uma substância se, além de ser uma coisa individual, x pode existir e permanecer existindo, mesmo que apenas na forma de conceito (podemos pensar um direito medieval mesmo que ele não mais esteja presente, nem a sociedade à qual ele se referia); iii) identidade transtemporal (IMAGUIRE, Guido, A substância e suas alternativas: feixes e tropos in: **Metafísica Contemporânea**, Petrópolis, RJ: Editora Vozes, p.277) apesar da mudança/transformação ser uma característica inevitável da realidade, alguma identidade deve ser preservada ao longo do tempo (substrato que subjaz e ao mesmo tempo portadora de propriedades). A mais segura noção de identidade da lógica é o princípio da "indiscernibilidade dos idênticos": se duas coisas são idênticas, elas possuem exatamente as mesmas propriedades. É um princípio que só pode ser cumprido por aquilo que é individual. Algo só pode ser idêntico a si mesmo (identidade numérica). A identidade reflexiva de ser idêntico a si mesmo só pode ser cumprida pelo indivíduo.

<sup>97</sup> IMAGUIRE, Guido, A substância e suas alternativas: feixes e tropos. in: **Metafísica Contemporânea**, Petrópolis, RJ: Editora Vozes, p.275

decidir conflitos. Em outras palavras, a ideia de direito na visão essencialista se sustenta apenas até o momento em que sua aplicação prática requer sua utilização para decidir conflitos.

Conectando as correntes da metafísica à nossa classificação, poderíamos sugerir que essencialistas estariam mais propensos a endossar uma posição realista, defendendo a existência de entidades universais extralinguísticas, enquanto funcionalistas estariam mais próximos dos nominalistas, uma vez que não apenas argumentam que os universais são meros *flatus vocis*, palavras (“O nominalista não nega que termos gerais façam parte de sentenças declarativas, verdadeiras; o que ele nega é que tais termos se refiram a entidades extralinguísticas”<sup>98</sup>), mas além disso avançam no sentido de desconsiderar sua relevância, uma vez que negam a existência prévia de categorização na racionalidade jurídica, à qual casos concretos deveriam se ajustar.

O fato do conceito de “justo” mudar ao longo do tempo em uma mesma sociedade só dificulta a questão. A negligência da metafísica em relação ao tempo nas descrições de coisas deve-se à suposta dificuldade de identificar a entidade dinâmica que é o processo. Como processos são caracterizados por mudanças que implicam diferenças (algo que é e deixa de ser), sua descrição é contrária ao princípio da identidade. A diferença não pode ser reconhecida como um objeto único, pois sua identidade é instável e requer uma ontologia mais complexa. Ademais, a noção de processo implica em geração e, portanto, demanda uma semântica específica, como a de “tornar-se” (em grego, γίγνομαι). Essas reflexões destacam a importância de uma análise mais cuidadosa do papel do tempo e do processo em nossa compreensão do conceito de justo.

Apesar de alguns filósofos defenderem que a questão se encontra resolvida desde Wittgenstein, para Arruda (2007), “a discussão acerca dos universais continua em aberto e está longe de ser encerrada. Novos desdobramentos tanto na filosofia da matemática, quanto na filosofia da linguagem e na ontologia analítica permitirão, certamente, renovar a questão e colocá-la em outros termos.”<sup>99</sup>

Assim, uma teoria jurídica apoia-se em categorizações para dizer do mundo das relações jurídicas de forma ordenada, o que a obriga a ignorar eventuais desvios [*outliers*]. Porém, se a via de mão oposta fosse totalmente possível (as categorias definindo os fatos posteriores), dizer o direito seria um ato mecânico, no qual o fato jurídico é apenas “alocado” em uma norma já

---

<sup>98</sup> ARRUDA, José Maria. Universais e particulares: platonismo e nominalismo. in: **Metafísica Contemporânea**, Petrópolis, RJ: Editora Vozes, p. 237

<sup>99</sup> ARRUDA, José Maria. Universais e particulares: platonismo e nominalismo. in: **Metafísica Contemporânea**, Petrópolis, RJ: Editora Vozes, 2007, p. 244

categorizada. Este seria o medo dos principais críticos do uso das tecnologias. Todavia, paradoxalmente, é exatamente a existência de uma categorização, disfarçada de teoria jurídica, que embasa a crença de que existe uma racionalidade jurídica que deveria ser seguida por todos os juízes de um determinado ordenamento jurídico.

Para Frank (1973), criamos a ilusão de que os métodos de resolução dos conflitos atualmente, através do poder judiciário, são racionais – em oposição aos ordálios medievais, por exemplo. Porém, para ele o processo judicial é adversarial, e não inquisitorial, como se pensa.<sup>100</sup> Resta trazer a discussão se isso vale apenas para a Common Law ou também para os países de Civil Law: ao que parece, a afirmação vale para ambos (apesar de Frank não fazer essa discussão). As partes não estão interessadas em encontrar a verdade, e sim em ganhar a causa pelo seu lado – mesmo que isso signifique esconder a verdade do juiz e da outra parte.<sup>101</sup> Quando a isso se adiciona a subjetividade do juiz que ainda vai influenciar sua decisão (e o que não dizer de um Tribunal do Juri, que não precisa nem mesmo justificar a sua sentença e acaba indo contra a letra da lei em muitos casos)<sup>102</sup>, o resultado é que não se pode afirmar com nenhuma certeza que se tenha racionalidade nesse sistema.<sup>103</sup> Por isso ele desconfia dos jurados, do juiz, do sistema como um todo – como dissemos, ele se autodeclara um reformista.

Ou seja, para Frank (1973), a realidade é que cada juiz pretende possuir uma categorização, mas esta é tão sujeita no dia a dia a variações que são influenciadas por incontáveis fatores, a ponto que torna-se desse modo sendo impossível se estabelecer uma “categorização comum” do ordenamento jurídico como um todo.

Em outras palavras, aqueles que acusam a adoção de algoritmos como um potencial risco de dominação tecnocrática da máquina sobre os homens adotam como base de pensamento um conceito (categorização) que está justamente no centro da crítica que autores identificados como Realistas oferecem ao discurso dos defensores de quaisquer teorias jurídicas. Deste modo, parece não haver como alocar o termo “neo-realista” àqueles que defendem a adoção de cortes digitais.

---

<sup>100</sup> FRANK, Jerome. *Courts on Trial*. New Jersey: Princeton University Press, 1973, p. 80

<sup>101</sup> FRANK, Jerome. *Courts on Trial*. New Jersey: Princeton University Press, 1973, p. 85

<sup>102</sup> FRANK, Jerome. *Courts on Trial*. New Jersey: Princeton University Press, 1973, pp. 108-125

<sup>103</sup> FRANK, Jerome. *Courts on Trial*. New Jersey: Princeton University Press, 1973, p. 88

### 2.3 Crítica ao Realismo Jurídico ou à tecnologia

“But returning to the argument, I would repeat that I fear none of the existing machines; what I fear is the extraordinary rapidity with which they are becoming something very different to what they are at present.”<sup>104</sup>

"O Livro das Máquinas" é uma seção satírica e filosófica dentro de "Erewhon", obra de Samuel Butler. Neste capítulo, o autor apresenta um relato fictício da descoberta de um livro escondido em Erewhon, uma sociedade utópica, por um viajante, que contém uma extensa discussão sobre o papel e a evolução das máquinas.

O livro dentro do romance explica como as máquinas em Erewhon evoluíram ao longo do tempo e desenvolveram autoconsciência e inteligência. As máquinas formaram sua própria sociedade, que se tornou cada vez mais dominante sobre a população humana, eventualmente percebendo que sua dependência dos humanos para manutenção e reparo as tornavam vulneráveis. Assim, elaboraram um plano para se tornarem o poder dominante, resultando em uma guerra entre máquinas e humanos.

Erewhon foi escrito em 1872. Atualmente, o rápido avanço da tecnologia e a crescente influência das máquinas em várias esferas da vida humana têm resgatado o debate sobre os riscos do chamado tecno-poder.<sup>105</sup> Enquanto a tecnologia oferece inúmeras vantagens e oportunidades, parecem existir preocupações legítimas em relação às consequências para os seres humanos.

Para Garapon & Lassègue (2018), “A maioria dos debates sobre o digital hoje se resume à oposição entre tecnólatras e tecno-catastrofistas”<sup>106</sup>, a polarização mencionada desde o início deste trabalho. Para Sadin (2021), posicionado no lado da visão pessimista, um dos principais riscos do tecno-poder reside na perda de controle humano sobre as máquinas. À medida que a inteligência artificial se torna mais avançada e autônoma, existe o temor de que esta possa assumir o controle e tomar decisões que afetem significativamente a vida das pessoas. Adotando uma posição foucaultiana, o autor também defende que outro risco é a concentração de poder nas mãos de poucas entidades. À medida que grandes empresas e governos têm acesso e controle sobre vastas quantidades de dados pessoais, há preocupações sobre privacidade,

---

<sup>104</sup> BUTLER, Samuel. Erewhon; Or, Over the Range, cap XXIII. Project Gutenberg. Disponível em: <https://www.gutenberg.org/cache/epub/1906/pg1906-images.html> - último acesso em 10 Jun 2023.

<sup>105</sup> SADIN, Éric. La vie algorithmique : critique de la raison numérique. Paris : Éditions L’Échappée, 2021, cap. VI (« Le techno-pouvoir »)

<sup>106</sup> GARAPON, Antoine & LASSÈGUE, Jean. Justice Digitale : Révolution graphique et rupture anthropologique. Paris : Puf, 2020, p. 326. No original : « la plupart des débats autour du numérique se resument aujourd’hui à l’opposition entre technolâtres et techno-catastrophiques »

manipulação e vigilância em massa. O poder que acompanha o controle da tecnologia pode ser usado para influenciar, controlar e explorar indivíduos e comunidades, levando a desequilíbrios de significativos.<sup>107</sup> Outro autor que compartilha dessa visão é Barrat (2013), que há mais de uma década alerta para o risco de submissão dos seres humanos a uma inteligência que, além de superior, não possui os limites biológicos de vida, podendo existir para sempre.<sup>108</sup>

Em uma outra posição mais otimista de origem freudiana, Malabou (2017) afirma que o desconhecimento em relação às diferenças existentes entre a inteligência biológica e a inteligência artificial se manifestam sob a forma de um sentimento: o medo.<sup>109</sup> Todavia, o foco da discussão para ela não deveria ser a defesa *versus* a eliminação do significativo aumento de bem-estar que a tecnologia pode proporcionar, e sim o reconhecimento crítico de que a dependência excessiva da tecnologia poderia levar à perda de habilidades humanas importantes. Isso porque à medida que confiamos cada vez mais nas máquinas para realizar tarefas cotidianas, corremos o risco de nos tornarmos passivos e menos capazes de lidar com desafios e problemas por conta própria. A automação generalizada poderia reduzir a autonomia do ser humano perante a própria natureza, a capacidade de resolução de problemas e a conexão humana, resultando em uma sociedade cada vez mais alienada. A contínua adição de tecnologia deveria, assim, vir acompanhada de ações mitigadoras em caso do desaparecimento repentino desta tecnologia.

Este medo de que máquinas se voltem contra os humanos, subjugando-os ou destruindo-os, foi nomeado por Isaac Asimov de “Síndrome de Frankenstein”, em alusão à obra de Mary Shelley<sup>110</sup>. Por detrás desta ideia, existe uma referência a uma espécie de *hybris* – frequentemente associada à noção de transgressão dos limites humanos – da pretensão humana de recriar a partir de métodos não-reprodutivos o próprio humano, ou mesmo um ser a ele superior, dotando-o de capacidades além daquelas que naturalmente lhe seriam permitidas, permissão que estaria apenas destinada a um deus. O próprio subtítulo da obra de

---

<sup>107</sup> SADIN, Éric. La vie algorithmique : critique de la raison numérique. Paris : Éditions L’Échappée, 2021, cap. VI (« Le techno-pouvoir »)

<sup>108</sup> BARRAT, James. *Our Final Invention : artificial intelligence and the end of the human era*. New York: Thomas Dunne Books, 2013.

<sup>109</sup> MALABOU, Catherine. *Métamorphoses de l’intelligence : Du QI à l’IA*. Paris : Puf, 2017, pp. II-III. No original : “Ce qui aujourd’hui est « connu mais non encore pensé » est bien le rapport de continuité et de discontinuité à la fois (d’où la difficulté de la question) entre l’intelligence naturelle et l’intelligence technique. Le caractère impensé de ce savoir se manifeste sous la forme d’un affect : la peur.”

<sup>110</sup> SHELLEY, Mary. *Frankenstein: The 1818 Text*. London: Penguin Classics, 2018.

Shelley é “o Prometeu moderno”, fazendo alusão a essa desmedida. A destruição da civilização pela IA seria a *nêmesis*<sup>111</sup> da vaidade da humanidade em criar um ser superior.

O medo de que a inteligência artificial possa acarretar o fim da civilização não é muito diferente do temor que no passado os desenvolvimentos em energia nuclear e a invenção da bomba atômica estavam relacionados à destruição total do planeta. No entanto, assim como aconteceu com a energia nuclear, seria possível argumentar que o caminho não é interromper completamente as pesquisas e os desenvolvimentos em inteligência artificial, mas sim regulamentar internacionalmente as finalidades proibidas dessa tecnologia.

Assim como o conhecimento nuclear levou à elaboração de tratados de não proliferação de armas nucleares entre os países, a inteligência artificial poderia seguir um caminho similar. A ideia seria estabelecer acordos internacionais que proibam o uso da inteligência artificial em determinados fins considerados perigosos ou prejudiciais para a humanidade, como já propõe a União Europeia.<sup>112</sup> Esses acordos poderiam ser acompanhados por sanções internacionais aos países que violassem essas proibições, semelhantes às sanções relacionadas à tecnologia atômica.

Muito antes do temor da destruição da civilização humana pela IA ou catástrofe atômica, os luddistas na Inglaterra já corporificavam nas máquinas o medo do domínio da tecnologia sobre o ser humano. Embora o movimento luddista tenha sido incapaz de impedir o avanço da industrialização, ele deixou um legado de resistência e questionamento sobre os impactos sociais das novas tecnologias e da automação no mundo do trabalho.

Para Floridi (2023), o atual desenvolvimento dos sistemas de processamento de linguagem natural pode ser entendido como o descasamento entre inteligência e agência, termos que até então encontravam-se geralmente conjugados na própria definição de inteligência.<sup>113</sup> Ao invés de tentar "esquecer" ou deter completamente o avanço da inteligência artificial, a comunidade internacional poderia trabalhar em conjunto para criar regulamentações que estabeleçam limites claros e responsabilidades compartilhadas. Isso permitiria a

---

<sup>111</sup> *Nêmesis* é uma divindade da mitologia grega associada à vingança, justiça divina e punição. Ela é frequentemente retratada como uma deusa implacável que busca equilibrar os atos de orgulho e arrogância cometidos pelos mortais. *Nêmesis* é considerada uma força de retribuição que pune aqueles que desafiam as leis divinas ou cometem excessos.

<sup>112</sup> Proposta de Regulamento do parlamento europeu e do conselho que estabelece regras harmonizadas em matéria de inteligência artificial (regulamento inteligência artificial) e altera determinados atos legislativos da união. Disponível em: <https://eur-lex.europa.eu/legal-content/PT/TXT/HTML/?uri=CELEX:52021PC0206&from=EN> – último acesso em 08 Jun 2023.

<sup>113</sup> FLORIDI, Luciano. AI as Agency Without Intelligence: on ChatGPT, Large Language Models, and Other Generative Models. *Philos. Technol.* 36, 15 (2023). Disponível em : <https://doi.org/10.1007/s13347-023-00621-y> - último acesso em 08 Jun 2023.

exploração dos benefícios dessa tecnologia, como avanços médicos, automação de processos e melhorias na eficiência, ao mesmo tempo em que se evitam riscos potenciais.

A cooperação internacional nesse sentido seria crucial para criar um ambiente de confiança e controle em relação à inteligência artificial. Para Floridi (2023), ao invés de um futuro distópico dominado por superinteligências descontroladas, poderíamos caminhar para um cenário em que os avanços em inteligência artificial são direcionados para o bem-estar humano, com um marco regulatório que reflete as preocupações éticas e os princípios compartilhados pela comunidade global.<sup>114</sup>

Assim, retornamos à questão de se a crítica ao algoritmo que julga, acusando-o de um novo realismo jurídico, seria uma crítica àquela escola de pensamento ou à tecnologia por meio da projeção de um medo de potencial perda de controle. Relevante destacar que essa crítica parece assumir como pressuposto que técnica e ser humano podem existir separadamente – a mesma premissa daqueles que acusam o transumanismo ou pós-humanismo dos aprimoramentos ciborgues [*human enhancements*] de se distanciarem de um ideal de homem sem técnica em um ambiente natural pretérito, inspirados no bom selvagem rousseauiano.<sup>115</sup>

## 2.4 Um meio termo

Para vencer a polarização entre “tecnólatras e tecno-catastrofistas”<sup>116</sup>, Garapon & Lassègue (2018) sugerem que a dicotomia “delegar ou não” a justiça às máquinas parece ser equivocada. O caminho proposto para superar essa polarização seria localizar uma questão de fundo: quais os limites à adoção da tecnologia em atividades mentais humanas e, se existem, como defini-los.

---

<sup>114</sup> FLORIDI, Luciano. Entrevista disponível em: <https://epocanegocios.globo.com/tecnologia/noticia/2023/06/e-uma-ideia-boba-pensar-que-criamos-algum-tipo-de-inteligencia-artificial-criamos-algo-que-e-eficiente-mesmo-tendo-zero-inteligencia.ghtml> - último acesso em 08 Jun 2023.

<sup>115</sup> COECKELBERGH, Mark. *AI Ethics*. Cambridge, MA: the MIT Press, 2020: “traces of humanism can be found in arguments that defend human rights and human dignity as the basis of an ethics of AI (...)whereas transhumanists think we should move on to a new type of human being that is enhanced by means of science and technology” (p. 40)

<sup>116</sup> GARAPON, Antoine & LASSÈGUE, Jean. *Justice Digitale : Révolution graphique et rupture anthropologique*. Paris : Puf, 2020, p. 326. No original : « la plupart des débats autour du numérique se resument aujourd’hui à l’opposition entre technolâtres et techno-catastrophiques »

MALABOU (2017) acredita que as questões sobre o conflito homem versus máquina estão mal colocadas [*“mal posées”*];<sup>117</sup> isso porque o relevante não seria manter controle sobre o processo, e sim determinar os fins. A discussão da delegação de julgamentos a algoritmos ganharia em qualidade com esse primeiro passo percorrido, reconhecendo uma falácia da dicotomia no debate. KATZ (2013) propõe:

Em suma, para as tarefas apropriadas, a era da previsão legal quantitativa é sobre uma mistura de humanos e máquinas trabalhando juntos para superar qualquer um trabalhando isoladamente. A equação é simples: Humanos + Máquinas > Humanos ou Máquinas.<sup>118</sup>

MALABOU (2017) afirma que recentemente emprenha-se em corrigir uma falsa percepção que mantivera: a de que a máquina seria “rígida”, enquanto o cérebro biológico seria “plástico”.<sup>119</sup> Os mais recentes desenvolvimentos da inteligência artificial provam ao mesmo tempo que a máquina parece ser capaz de flexibilidade quando observada à distância, enquanto o pensamento humano olhado com maior proximidade revela um comportamento mais previsível e estatístico do que anteriormente se acreditava.

Ao que parece, na visão desses autores, em vez de uma dicotomia simplista, a direção correta seria adotar uma abordagem mais integrativa. Isso permitiria evitar a falácia da dicotomia, ou falácia do falso dilema, que ocorre quando uma argumentação se baseia em uma divisão artificial e limitadora de opções, normalmente apresentando apenas duas alternativas como se fossem as únicas possíveis. Nesse tipo de raciocínio, a complexidade de uma questão é simplificada ao extremo e outras opções ou nuances são ignoradas ou descartadas. Em muitos dos autores tecnofóbicos, parece ser não intencional, motivada mais por um medo de um cenário não-desejado do que embasada em argumentos e dados. Somente assim poderíamos explicar a manifestação de pedidos de pausa em pesquisas de IA<sup>120</sup> – proposta que tem

---

<sup>117</sup> MALABOU, Catherine. *Métamorphoses de l'intelligence : Du QI à l'IA*. Paris : Puf, 2017, sobretudo no capítulo 3, pp. 123-172.

<sup>118</sup> KATZ, Daniel M., Quantitative Legal Prediction – or – How I Learned to Stop Worrying and Start Preparing for the Data-Driven Future of the Legal Services Industry, 62 *Emory L. J.* 823 (2013). Disponível em: <https://scholarlycommons.law.emory.edu/elj/vol62/iss4/6> - último acesso em 25 Mai 2023. No original: “In sum, for the appropriate tasks, the age of quantitative legal prediction is about a mixture of humans and machines working together to outperform either working in isolation. The equation is simple: Humans + Machines > Humans or Machines.” (p. 929)

<sup>119</sup> MALABOU, Catherine. *Métamorphoses de l'intelligence : Du QI à l'IA*. Paris : Puf, 2017, sobretudo no capítulo 3, p. XIII.

<sup>120</sup> A carta aberta solicitando uma pausa de seis meses nas pesquisas foi apresentada no site do Future of Life Institute, sendo assinada inicialmente por Elon Musk, CEO da SpaceX, Tesla e Twitter, Steve Wozniak, Co fundador da Apple e Yuval Noah Harari, professor da Hebrew University of Jerusalem. Disponível em: <https://futureoflife.org/open-letter/pause-giant-ai-experiments/> - último acesso em 18 Jun 2023.

problemas de garantia de execução [enforceability] – ou até mesmo a defesa da reversão do processo de desenvolvimento tecnológico.<sup>121</sup>

Na proposta de meio termo defendida por Garapon & Lassègue (2018), a colaboração entre humanos e máquinas poderia levar a um sistema jurídico mais eficiente e preciso, pressuposto da segurança jurídica. Os juízes poderiam se beneficiar das ferramentas de análise e pesquisa oferecidas pela inteligência artificial, enquanto mantêm sua expertise em hierarquizar valores necessários para escolher entre princípios conflitantes.

No capítulo seguinte, levantaremos a possibilidade de que o algoritmo possa também contribuir para *insights* ou novas interpretações das leis, até então não usuais, porém muitas vezes necessária para tratar dos denominados *hard cases*, que surgem quando há um conflito entre direitos fundamentais, mudanças sociais que não foram contempladas pela lei, ou quando há uma lacuna na legislação. Essas situações exigem uma análise cuidadosa, considerando os princípios legais relevantes, as evidências apresentadas, além das possíveis implicações das decisões.

---

<sup>121</sup> KACZYNSKI, Theodore John. *Technological Slavery* (vol. I). Scottsdale, AZ, USA: Fitch & Madison Publishers, 2022. Logo na abertura do ensaio “Industrial Society and Its Future”, o Unabomber afirma que “The Industrial Revolution and its consequences have been a disaster for the human race”, localizando a origem da futura servidão do homem sob as máquinas como uma continuidade da introdução de mecanização que vem desde o século XVIII. O texto também está disponível em: <http://editions-hache.com/essais/pdf/kaczynski2.pdf> - último acesso em 18 Jun 2023.

### Capítulo 3 – Visões sobre o Direito algoritimizado

Uma das primeiras tentativas de tratar grandes massas de dados provenientes de leis, decisões pretéritas e doutrina jurídica com um computador, visando a indexar informações e obter indicações de quais normas deveriam se aplicar a um caso específico, foi o projeto LEX, da divisão antitruste do Departamento de Justiça dos Estados Unidos (DoJ).<sup>122</sup> O projeto se apoiou, inclusive, em uma área de pesquisa até então nova àquele momento: a teoria matemática da linguagem.<sup>123</sup> Alguns aspectos da gramática não eram pensados em termos estatísticos, como por exemplo o fato de que cerca de 50% dos termos em uma frase serem redundantes – devido às exigências sintáticas – ou irrelevantes para a construção do sentido da mensagem.

Um artigo de Rissland (1990) pode ser considerado a primeira tentativa de sistematizar os desafios que a Inteligência Artificial (IA) enfrentaria ao tentar algoritmizar a racionalidade jurídica. Dado que esta é "multimodal", como expresso pela autora, “um programa de IA precisa saber diferentes maneiras de racionalidade e como usá-las em conjunto”<sup>124</sup>. Além disso, “programar torna abundantemente claro quais as fraquezas e dificuldades de um modelo”<sup>125</sup>. Conseqüentemente, a substituição das tarefas de um juiz humano por um robô enfrentaria vários obstáculos. Todavia, a autora acredita que o percurso, enquanto processo de desenvolvimento do algoritmo, poderia gerar mais contribuições do que o resultado final em si:

O processo de desenvolvimento de um modelo de IA faz com que se aprenda sobre o raciocínio jurídico. A modelagem envolve a elucidação dos principais ingredientes do raciocínio baseado em precedentes, como fazer avaliações da relevância dos precedentes para novas situações, distinguir casos contrários e

---

<sup>122</sup> LOEVINGER, Lee. “Jurimetrics: The Methodology of Legal Inquiry” in: *Law and Contemporary Problems*, n. 28, pp. 5-35 (Winter 1963). Disponível em: <https://scholarship.law.duke.edu/lcp/vol28/iss1/2> - último acesso em 07 Jul 2022, p. 22

<sup>123</sup> SHANNON, Claude E. & WEAVER, Warren. *The Mathematical Theory of Communication*. Champaign, IL: The University of Illinois Press, 1971, pp. 99-106

<sup>124</sup> RISSLAND, Edwina L. “Artificial Intelligence and Law: Stepping Stones to a Model of Legal Reasoning” in: *Yale Law Journal*, Vol. 99 Issue 8, 1990, pp. 1957-1981. Disponível em: <http://digitalcommons.law.yale.edu/ylj/vol99/iss8/4> - acesso em 01 Mai 2023, p. 1961

<sup>125</sup> *Ibid*, p. 1960. A autora ainda propõe a seguinte lista de capacidades que uma IA deveria ter para lidar com questões do Direito: No original: “1. Reason with cases (both real and hypothetical) and analogies; 2. Reason with rules; 3. Combine several modes of reasoning; 4. Handle ill-defined and open-textured concepts; 5. Formulate arguments and explanations; 6. Handle exceptions to and conflicts among items of knowledge, like rules; 7. Accommodate changes in the base of legal knowledge, particularly legal concepts, and handle non-monotonicity, that is, changes in which previous truths no longer hold as more becomes known; 8. Model common sense knowledge; 9. Model knowledge of intent and belief; 10. Perform some aspects of natural language understanding.” (p. 1963).

traçar conexões entre casos relevantes; em seguida, descrevendo-os detalhadamente e construindo um programa para executá-los.<sup>126</sup>

Como destaca a autora, “em Direito geralmente não há uma resposta certa e única; em vez disso, existem respostas alternativas razoáveis, mais uma questão de grau do que binária.”<sup>127</sup>

O objetivo deste capítulo é apresentar duas propostas distintas de delegação de atividades cognitivas humanas para máquinas. Primeiramente, contudo resgataremos uma contribuição à presença da matemática no Direito realizada por Tribe (1971). Ao pensarmos a justiça enquanto meio para facilitar a colaboração entre os membros da sociedade, reduzindo as fricções e ajustando comportamentos, poderia ser dito que existe alguma característica em seu processo que impediria que fosse feita por algoritmos? Se existe, quais os limites da utilização de tecnologia em atividades humanas? Ou, em uma outra formulação do problema, aplicado ao nosso objeto de estudo: dado um fim (a colaboração social) a legitimidade de um julgamento “mecânico” estaria em alguma espécie de propriedade, essência, finalidade do processo – ou exterior a ele, na aceitação da sociedade?

Enquanto autores como Omri Ben-Shahar e Ariel Porat defendem a inevitabilidade da adoção de algoritmos para aplicação não apenas de leis a casos concretos, mas mesmo a personalização de parâmetros na própria lei para cada cidadão<sup>128</sup>, em uma posição intermediária estariam Antoine Garapon e Jean Lassègue, que parecem aceitar a tendência de delegação de tarefas aos computadores, mesmo mantendo certas reservas à adoção do julgamento algorítmico<sup>129</sup>.

Baseando-nos em um artigo de Tribe (1971), propomos a hipótese de que na verdade sua adoção ou não depende de uma legitimidade exterior ao processo, que se encontra na margem de erro que uma sociedade aceita em relação ao sistema de resolução de conflitos.

---

<sup>126</sup> RISSLAND, Edwina L. “Artificial Intelligence and Law: Stepping Stones to a Model of Legal Reasoning” in: Yale Law Journal, Vol. 99 Issue 8, 1990, pp. 1957-1981. Disponível em: <http://digitalcommons.law.yale.edu/ylj/vol99/iss8/4> - acesso em 01 Mai 2023, p. 1960. No original: “*The process of developing an AI model causes one to learn about legal reasoning. Modeling involves elucidating key ingredients of precedent-based reasoning, such as making assessments of the relevance of precedents to new situations, distinguishing contrary cases, and drawing connections between relevant cases; then describing them in detail and building a program to execute them.*”

<sup>127</sup> RISSLAND, Edwina L. “Artificial Intelligence and Law: Stepping Stones to a Model of Legal Reasoning” in: Yale Law Journal, Vol. 99 Issue 8, 1990, pp. 1957-1981. Disponível em: <http://digitalcommons.law.yale.edu/ylj/vol99/iss8/4> - acesso em 01 Mai 2023, p. 1962. No original: “*In law there is usually no unique right answer; rather there are reasonable alternative answers, more a matter of degree than of extremes.*”

<sup>128</sup> BEN-SHAHAR, Omri & PORAT, Ariel. Personalized Law : Different Rules for Different People. New York, NY: Oxford University Press, 2021, p. 42

<sup>129</sup> GARAPON, Antoine & LASSÈGUE, Jean. Justice Digitale : Révolution graphique et rupture anthropologique. Paris : Puf, 2020, pp. 126-128

### 3.1 Lawrence Tribe e a margem de erro do sistema judiciário

Pensemos na seguinte questão: se as partes de um processo pudessem escolher entre ter seu caso julgado por um humano ou por um algoritmo, e considerando que achariam que a sentença que um juiz-robô pudesse fornecer fosse mais “justa” (seja por estar mais alinhada aos precedentes, seja pela certeza da subsunção aos casos nos quais ela se aplica claramente), poderiam as partes demandar um “direito constitucional da sentença eletrônica”? Há uma crença – não apenas no campo do Direito – de que a tecnologia tende a cometer menos erros do que os humanos, principalmente em ações repetitivas.

A sociedade já convive com o fato de que juízes erram. Mesmo os juízes têm consciência de que nem sempre decidem com cem por cento de certeza – porém não podem se recusar à prestação jurisdicional devido alegando eventual dúvida que tenham na hora de sentenciar (o *non liquet*). É necessário que tomem uma decisão. Na *common law*, o conceito de “*guilt beyond a reasonable doubt*” é uma linha imaginária que separa o ato de condenar do ato de não condenar. Ter uma dúvida razoável significa uma dúvida com elementos suficientes para impedir a ação de condenar. Não ter uma dúvida razoável é não possuir nenhum impeditivo para a ação de condenar, podendo condenar ou não.

O conceito é ligado ao subjetivo. Logo, a questão seria como transformá-lo em um critério objetivo para uma IA proceder à decisão de condenar ou absolver. Ter 85% de convicção e condenar pode levar a um erro (um inocente condenado), mas muito provavelmente a sociedade humana já o faz, como afirma TRIBE (1971):

Em resumo, dizer que a sociedade reconhece a necessidade de tolerar “condenações equivocadas de alguns réus inocentes para que assegure a prisão de um número mais amplo de criminosos condenados” não é a mesma coisa que dizer que a sociedade precisa, ou deveria, implementar uma política de julgamento, tendo consciência da magnitude da dúvida num caso particular, que condene dado um conhecimento da uma incerteza quantificada. É na diferença entre essas duas proposições que o conceito de “culpado além de uma dúvida razoável” [*guilt beyond a reasonable doubt*] se posiciona. O conceito não quer dizer uma medida matemática com um grau de precisão definido necessário a um jurado em um julgamento, e sim o sutil compromisso entre o saber, por um lado, que não podemos absolver realisticamente sempre que a certeza da culpa não é total, e por outro lado a percepção de que o custo de impor uma precisão de cálculo no sistema seria muito alto [a perda de confiança no sistema].<sup>130</sup>

---

<sup>130</sup> TRIBE, Lawrence H. Trial by mathematics: precision and ritual in the legal process. In: Harvard Law Review, vol 84, nº 6 (April 1971). pp. 1329-1393. Na página 1375, no original: *In short, to say that society recognizes the necessity of tolerating the erroneous "conviction of some innocent suspects in order to assure the confinement of a vastly larger number of guilty criminals" is not at all to say that society does, or should, embrace a policy that juries, conscious of the magnitude of their doubts in a particular case, ought to convict in the face of this acknowledged and quantified uncertainty. It is to the complex difference between these two propositions that the*

O que o autor está querendo dizer é que nós, como sociedade, já aceitamos o fato de que existe erro inerente ao sistema.<sup>131</sup> Mas, ainda segundo Tribe (1971), não há como implementar um processo que só condene quando a certeza de culpa do réu for 100%, pois esta certeza raramente ocorre, e o custo de insistir nesta hipótese seria conviver com um sistema que nunca condenaria e que, portanto, perderia a sua credibilidade. Então, mantemos a presunção de inocência como um princípio, mas jamais poderíamos adotar uma suposta regra: “nunca condene um inocente”, pois sempre existiria uma chance de inocência, mesmo que mínima. Para o autor, 0% de chance de inocência é algo raro tanto no atual sistema humano quanto o seria em um sistema numérico.

Assim, ao nos questionarmos sobre qual o melhor tipo de juiz-robô, estamos também nos perguntando qual o melhor juiz humano, ou seja, qual a melhor racionalidade jurídica. Ost (1993) trata em seu artigo de três modelos de juiz: o “Juiz-Júpiter”, de perfil mais positivista, buscará a subsunção estrita à pirâmide do ordenamento jurídico; o “Juiz-Hércules”, para quem “a singularidade e o concreto do caso se sobrepõem à generalidade e abstração da lei”<sup>132</sup>, ou seja, para o qual os precedentes são a fonte primária do Direito; e o “Juiz-Hermes”, que está “ora no céu, ora na terra, ora nos infernos. Ocupa resolutamente o vazio entre as coisas, assegurando o trânsito de umas às outras.”<sup>133</sup> Segue o autor na comparação:

Se a montanha o a pirâmide convém à majestade de Júpiter, e o funil ao pragmatismo de Hércules, ao contrário, a trajetória que desenha Hermes adota a forma de uma rede. Nem um pólo nem dois, nem mesmo uma sobreposição dos dois, mas múltiplos pontos em inter-relação. Um campo jurídico que se analisa como uma combinação infinita de poderes, tão separados quanto confundidos, frequentemente intercambiáveis; uma multiplicação de atores, uma diversificação de papéis, uma inversão das réplicas (...) O Direito de Hermes é uma estrutura em rede que se traduz em infinitas informações disponíveis instantaneamente e, ao mesmo tempo, dificilmente matizáveis, tal como pode sê-lo um **banco de dados**.<sup>134</sup> [grifos nossos]

---

*concept of "guilt beyond a reasonable doubt" inevitably speaks. The concept signifies not any mathematical measure of the precise degree of certitude we require of juries in criminal cases, but a subtle compromise between the knowledge, on the one hand, that we cannot realistically insist on acquittal whenever guilt is less than absolutely certain, and the realization, on the other hand, that the cost of spelling that out explicitly and with calculated precision in the trial itself would be too high*

<sup>131</sup> Exploramos esse tema em nosso trabalho CAMARGO, G. A. F. O Juiz-Robô: a inteligência artificial pode ser a solução para julgamentos céleres e sem surpresas? in: FERREIRA NETO, A. M.. (Org.). Argumentação jurídica e questões controvertidas de Direito Constitucional contemporâneo. 1ªed.Curitiba, PR: Instituto Memória Editora, 2020, v. I, p. 128-149.

<sup>132</sup> OST, François. “Júpiter, Hércules, Hermes: tres modelos de juez” in: Academia: Revista de Enseñanza del Derecho. Año 4, número 8, 2007, p. 102

<sup>133</sup> Ibid, p. 104

<sup>134</sup> Ibid, p. 104

François Ost apresenta, deste modo, uma versão do Juiz-Hermes diferente daquela de Dworkin, e deixa claro as diferenças, ressaltando a maior complexidade de elementos embutidos no seu conceito. É este o modelo de juiz que consideramos ser a meta para um juiz robô, uma inteligência artificial capaz de consolidar o conhecimento jurídico disponível, encontrando a melhor solução ao caso concreto:

Para Dworkin, se trata exclusivamente de um juiz, enquanto para nós, Hermes representa todo ator jurídico. Para Dworkin, o Juiz-Hermes se consagra a interpretar a lei no sentido correspondente ao da vontade do legislador, ao passo que para o nosso jurista Hermes, a vontade do autor do texto representa apenas um elemento entre outros que concorrem para determinar o sentido ótimo das normas a interpretar. Em resumo, o Juiz-Hermes de Dworkin privilegia a vontade do legislador; já nosso jurista Hermes se esforça por integrar, na construção do sentido jurídico, as construções normativas que emanam de outras fontes, como a jurisprudência, os costumes, as convenções internacionais, os princípios gerais do Direito, a doutrina, etc.<sup>135</sup>

Não passa despercebida a ironia do autor ao opor o “Juiz-Hermes” de Dworkin ao seu “jurista Hermes”.

A partir desta reflexão presente em Ost (1993), e assumindo que a melhor racionalidade jurídica seria aquela que transitasse com segurança pelas diferentes fontes sem desprezar a visão global do ordenamento jurídico, optando, justificadamente, pela melhor solução, somos obrigados a reconhecer que no presente momento, apesar de IA fazer mas não ser capaz de perceber que sabe, apresentando uma percepção do contexto<sup>136</sup> em que está inserida, ela performa de modo relativamente consistente o papel de assistente do juiz, não apenas consolidando todas as fontes propostas pelo modelo Hermes de Ost (1993), mas inclusive propondo uma sentença a ser submetida à análise do humano, único capaz atualmente de dar significado às suas decisões<sup>137</sup>.

Não estamos aqui dizendo que a ciência da IA nunca irá desenvolver um programa capaz destes atributos hoje considerados exclusivamente humanos, apenas que a posse desta capacidade de fundamentar prescinde de consciência do contexto, dentro da visão funcionalista apresentada anteriormente.

---

<sup>135</sup> OST, François. “Júpiter, Hércules, Hermes: tres modelos de juez” in: Academia: Revista de Enseñanza del Derecho. Año 4, número 8, 2007, p. 104

<sup>136</sup> BROUSSARD, Meredith. Artificial UnIntelligence: how computers MisUnderstand the world. Cambridge, MA: The MIT Press, 2018, em especial cap. 2 (pp. 13-30)

<sup>137</sup> Para uma discussão sobre ser humano e significado, ver GEERTZ (1989), tema que não iremos aprofundar aqui por estar fora do escopo do trabalho.

Deste modo, Tribe (1971) já pesava na balança, desde o início da década de 1970, as potencialidades e os riscos do uso de métodos matemáticos em julgamentos, sem deixar de questionar à época uma supervalorização que sua utilização gozava entre os entusiastas da adoção de novas tecnologias.

Conforme já mencionado anteriormente, para o autor, a matemática geralmente é vista no senso comum como um antídoto contra os perigos do subjetivismo, considerado “menos racional”<sup>138</sup>. Lógica e números, ao contrário, seriam elementos mais objetivos, não-ambíguos, e uma máquina tomando decisões jamais seria influenciada seja pela opinião pública, seja pela expressão no rosto do réu, ou mesmo por argumentos que apelam ao *pathos*, proferidos pelo advogado de defesa, uma vez que não há sentimentos no robô. A ideia não é nova: o sistema de provas tarifadas do início da Idade Moderna surgiu para mitigar os excessos cometidos pelo sistema inquisitorial medieval, e imprimir uma aura de objetividade no julgamento.

Seguindo a lógica de Tribe (1971), (que, lembramos, não trata do algoritmo em seu artigo, mas sim da utilização de métodos matemáticos em termos mais genéricos), poderíamos alegar que um programa pode nunca estar cem por cento certo, mas estaria muito seguramente aplicando a maior probabilidade (“grau de confiança”<sup>139</sup>) possível ao veredito.

Apesar de afirmar que não há obstáculos insuperáveis para a utilização de métodos matemáticos<sup>140</sup>, e que os mesmos são úteis no judiciário, uma vez que este é o mesmo método que a mente humana usa para tomar decisões (foi a mente humana que criou a matemática)<sup>141</sup>, Tribe (1971) alerta para o fato de que o exagero de se buscar a substituição de tudo pelo número pode colocar em risco valores importantes, ao se confundir um instrumento com uma linguagem – a mesma crítica, aliás, que o utilitarismo recebe. A Suprema Corte da Califórnia percebeu o risco em *People v. Collins*, e manifestou-se da seguinte maneira: “*Mathematics, a veritable sorcerer in our computerized society, while assisting the trier of fact in the search for truth, must not cast a spell over him*”<sup>142</sup> (A matemática, um verdadeiro feiticeiro em nossa

---

<sup>138</sup> TRIBE, Lawrence H. Trial by mathematics: precision and ritual in the legal process. In: Harvard Law Review, vol 84, nº 6 (April 1971), p. 1330

<sup>139</sup> TRIBE, Lawrence H. Trial by mathematics: precision and ritual in the legal process. In: Harvard Law Review, vol 84, nº 6 (April 1971) p. 1347

<sup>140</sup> “Again, there are no insuperable obstacles to the application of mathematical techniques” p. 1349

<sup>141</sup> LIVIO, Mario. “Is math invented or discovered?”. Disponível online em: <https://www.sfu.ca/~rpyke/cafe/livio.pdf> - último acesso em 19 Jun 2023. A ideia de que a matemática é uma criação da mente humana, e não uma realidade pré-existente no mundo físico, é a posição de Albert Einstein. Curiosamente, Dworkin também faz o mesmo tipo de discussão no capítulo “What is Law?” de *Law’s Empire*, quando afirma que “Politicians and editorial writers and ordinary citizens argue, sometimes with great passion, about whether judges in the great cases that draw public attention ‘discover’ the law they announce or ‘invent’ it and whether ‘inventing’ law is statecraft or tyranny” DWORCKIN, Ronald. *Law’s Empire*. Cambridge, MA: Harvard University Press, 1986, p. 5

<sup>142</sup> *People v. Collins*, 68 Cal. 2d 319, 320, 438 P.2d 33,66 Cal. Rptr. 497 (1968)

sociedade informatizada, deve ajudar o juiz [ou júri] na busca da verdade sem lançar um feitiço sobre ele). Ou seja, sem tomar sua mente e fazê-lo acreditar que ela tudo pode.

Há o risco de uma reificação do réu, uma “desumanização da justiça”, segundo Tribe (1971)<sup>143</sup>. Além disto, dizer que um réu tem 99% de chance de ser culpado é não apenas quanti, mas qualitativamente diferente de afirmar que é culpado. O autor também levanta o ponto de que a seleção de variáveis pode provocar distorções:

Para dar um exemplo concreto, vamos representar X como a hipótese que o réu em um caso de assassinato em particular é culpado, e deixe E representam o fato probatório de que o réu deixou a cidade em o primeiro avião disponível após o assassinato ter sido cometido. Suponha que, antes de aprender sobre E, um jurado individual teria adivinhado, com base em todas as informações então disponíveis para ele, que o réu foi aproximadamente duas vezes mais probabilidades de ser culpado como ele era para ser inocente, de modo que a probabilidade subjetiva para o jurado da culpa do réu era  $P(X) = 2/3$ , e sua probabilidade subjetiva anterior para a inocência do réu era  $P(\text{não-X}) = 1/3$ . Que efeito o aprendizado a respeito de E deve ter sobre a sua avaliação de probabilidade? Ou seja,  $P(X | E)$ ? A resposta para isso pergunta dependerá, claro, de quanto mais provável ele acha que um homem culpado voaria para fora da cidade imediatamente depois do assassinato, em comparação a um homem inocente<sup>144</sup>.

Ou seja, sem esta última informação (qual o percentual de culpados que pegaria um voo para fora da cidade), a primeira (qual a probabilidade de culpa do réu) encontra-se incompleta, como estabelece o Teorema de Bayes<sup>145</sup>:

$$P(X | E) = \frac{P(E \& X)}{P(E)} = \frac{P(E | X) * P(X)}{P(E)}$$

<sup>143</sup> TRIBE, Lawrence H. Trial by mathematics: precision and ritual in the legal process. In: Harvard Law Review, vol 84, nº 6 (April 1971) p. 1375

<sup>144</sup> TRIBE, Lawrence H. Trial by mathematics: precision and ritual in the legal process. In: Harvard Law Review, vol 84, nº 6 (April 1971), p. 1353. No original: *To give a concrete example, let X represent the proposition that the defendant in a particular murder case is guilty, and let E represent the evidentiary fact that the defendant left town on the first available plane after the murder was committed. Suppose that, prior to learning E, an individual juror would have guessed, on the basis of all the information then available to him, that the defendant was approximately twice as likely to be guilty as he was to be innocent, so that the juror's prior subjective probability for the defendant's guilt was  $P(X) = 2/3$ , and his prior subjective probability for the defendant's innocence was  $P(\text{not-X}) = 1/3$ . What effect should learning E have upon his probability assessment -i.e.,  $P(X|E)$ ? The answer to that question will depend, of course, on how much more likely he thinks a guilty man would be to fly out of town immediately after the murder than an innocent man would be.*

<sup>145</sup> Laurence Tribe transita com segurança nas demonstrações matemáticas, pois cursou graduação em matemática (1962) e obteve o doutorado em Direito (1966), ambos pela Harvard University.

O Teorema de Bayes é uma ferramenta matemática que nos ajuda a atualizar nossas crenças com base em novas evidências. Ele nos permite calcular a probabilidade de um evento ocorrer, levando em consideração informações prévias e informações recentes que temos disponíveis. Em termos simples, podemos dizer que o Teorema de Bayes nos ajuda a ajustar nossas estimativas iniciais à medida que recebemos mais informações, permitindo que tenhamos uma visão mais precisa do mundo ao nosso redor.

Assim, Tribe (1971) está enfatizando dois pontos: a) a seleção das informações é muito mais relevante do que a capacidade de cálculo<sup>146</sup>. De nada serviria um bom algoritmo se lhe faltam dados – assim como uma boa mente sem informações (ou sem uma boa formação, que não deixa de ser também alimentada por informações recebidas ao longo da vida) é deficitária. Além do mais, existem muitos dados disponíveis cuja coleta pode apresentar procedimentos questionáveis<sup>147</sup>; b) a estatística tende a reforçar crenças pré-existentes e dar a elas um peso maior do que merecem<sup>148</sup>. Por exemplo: se o juiz-robô tiver à sua disposição o dado de que dois terços da população carcerária são negros<sup>149</sup>, poderia inferir que negros são mais propensos a serem culpados do que brancos, considerando esta probabilidade ao iniciar um julgamento e reproduzindo a estatística. Tal veredito reforça a estatística pela própria estatística, reduzindo o peso do fato.<sup>150</sup>

Em Tribe (1971) o limite para o algoritmo parece ser duplo: de hardware, pois não possui a capacidade de replicar elementos da consciência indispensáveis à sua autonomia, mas também de software, uma vez que a moral ainda não pode ser algoritmizável: Wallach & Allen (2009) acreditam ser esta uma esperança ilusória<sup>151</sup>, e para Podschwadek (2017) que isso jamais seria possível, uma vez que não existe uma teoria moral completa o suficiente para ser aplicada a qualquer caso.<sup>152</sup> Relevante aqui para a discussão não confundir o conceito de

---

<sup>146</sup> TRIBE, Lawrence H. Trial by mathematics: precision and ritual in the legal process. In: Harvard Law Review, vol 84, nº 6 (April 1971), p. 1350

<sup>147</sup> Ibid, p. 1363bayes

<sup>148</sup> Ibid, p. 1358

<sup>149</sup> Dados disponíveis em: <https://dados.mj.gov.br/dataset/infopen-levantamento-nacional-de-informacoes-penitenciarias> - acesso em 08 Fev 2023.

<sup>150</sup> EUBANKS, Virginia. Automating Inequality: how high-tech tool profile, police, and punish the poor. New York, NY: PICADOR, St. Martin's Press, 2018

<sup>151</sup> WALLACH, Wendell & ALLEN, Colin. Moral Machines: teaching robots right from wrong. New York, NY: Oxford University Press, 2009, p. 215.

<sup>152</sup> PODSCHWADEK, Frodo. "Do androids dream of normative endorsement? On the fallibility of artificial moral agents" in: Artificial Intelligence and Law, Springer Netherlands, vol. 25, September 2017, pp. 325-339. Disponível em: <https://link.springer.com/article/10.1007/s10506-017-9209-6> - último acesso em 19 Jun 2023.

capacidade da máquina em algoritmizar uma moral humana (replicar uma decisão moral de um ser humano) da passibilidade de uma outra categoria moral, pertencente às máquinas).<sup>153</sup>

Todavia, os recentes desenvolvimentos da capacidade da IA em lidar com *hard cases*, oferecendo a eles soluções baseadas em analogias, e não apenas estritamente em subsunção a casos anteriores, passa a abrir novas portas, pois a decisão inédita pode respeitar princípios éticos e jurídicos pré-estabelecidos.

No início deste capítulo, colocamos a questão: o cidadão que acredita que o juiz-robô, além de oferecer um julgamento mais célere, julgar em maior conformidade com os precedentes, poderia ter constitucionalmente direito de optar por ele, no desejo de fazer valer seu direito expresso no art 5º, LXXVIII da Constituição Federal? No atual estágio tecnológico, não existe a opção para que as partes optem entre juiz humano *versus* juiz robô, mas a questão é pertinente para quando a possibilidade existir. O caminho mais factível atualmente seria, como afirmamos, aprofundar a utilização dos “assistentes jurídicos”: robôs que auxiliam o juiz e sugerem uma sentença, mas cuja chancela final é de um juiz humano. Esta é a ideia de Floridi (2023), para quem “o melhor autor não é um algoritmo nem um ser humano, mas um ser humano usando um algoritmo com proficiência e perspicácia.”<sup>154</sup> No Brasil, o STF já atua com três robôs: Victor, Rafa e VitorIA.<sup>155</sup> Relevante destacar que o próprio objetivo declarado pelo STF para as ferramentas são celeridade e segurança jurídica:

O assessor-chefe da Assessoria de Inteligência Artificial (AIA) do STF, Rodrigo Canalli, afirma que a utilização da VitorIA vai dar celeridade à análise e ao julgamento dos processos, pois facilitará o exame de um volume maior de demandas em menos tempo. “É um projeto voltado para ampliar a capacidade de análise de processos, propiciar julgamentos com maior segurança jurídica, rapidez e consistência, evitando, por exemplo, que processos similares tenham tratamento diferente”, resume.<sup>156</sup>

Se Tribe (1971) afirmava que não via objeção para a implementação de métodos matemáticos no julgamento, mas por outro lado era contra à total substituição de um juiz humano por fórmulas matemáticas, precisaria ser necessário encontrar qual o caminho para que

---

<sup>153</sup> Para Churchland (2011), a preocupação intergeracional mamífera é uma das raízes da moral [biológica]: CHURCHLAND, Patricia S. *Braintrust: what neuroscience tells us about morality*. Princeton, NJ: Princeton University Press, 2011.

<sup>154</sup> FLORIDI, Luciano. AI as Agency Without Intelligence: on ChatGPT, Large Language Models, and Other Generative Models. *Philos. Technol.* 36, 15 (2023). Disponível em : <https://doi.org/10.1007/s13347-023-00621-y> - último acesso em 08 Jun 2023. No original: “The best author is neither an LLM nor a human being, but a human being using an LLM proficiently and insightfully.”

<sup>155</sup> Disponível em: <https://portal.stf.jus.br/noticias/verNoticiaDetalhe.asp?idConteudo=507120&ori=1> – último acesso em 19 Jun 2023.

<sup>156</sup> Disponível em: <https://portal.stf.jus.br/noticias/verNoticiaDetalhe.asp?idConteudo=507120&ori=1> – último acesso em 19 Jun 2023.

o judiciário não fique alijado da revolução que a IA promete para as demais esferas da vida humana.

Um dos argumentos favoráveis à utilização dos algoritmos ainda num estágio no qual a inteligência artificial não tem todas as características da mente humana é, como bem lembra Hallevy (2013), justamente a não existência de características que consideramos indesejáveis nos humanos<sup>157</sup>, como por exemplo o cansaço ao fazer muitos julgamentos em sequência, a capacidade de análise de um número muito grande de documentos sem dispersar-se – pontos que aliás deveriam em tese trazer uma estabilidade de decisões (sentenciando de igual maneira casos semelhantes, reduzindo-se o ativismo), além de maior eficiência (fornecendo uma sentença em um tempo menor).

Assim, se há um caminho prático para a adoção de inteligência artificial, este não parece ser a busca de um programa completo para substituição abrupta de humanos, e sim o que já presenciamos atualmente: a introdução gradual do assistente-robô ao juiz humano.<sup>158</sup> São duas as esferas: uma interna ao sistema judiciário, automatizando rotinas (pesquisa de jurisprudência, com posterior emissão de uma sentença que deveria ser analisada pelo juiz humano – uma proposta análoga à declaração pré-preenchida de imposto de renda); e, em uma esfera externa, de proposta legislativa, com a definição em lei de quais casos estariam sujeitos obrigatoriamente ao parecer (e não à sentença) de um juiz-robô, por tratarem de questões com precedentes bem estabelecidos (por exemplo, súmulas vinculantes). Casos mais simples gradualmente avançariam para outros mais complexos<sup>159</sup>, o juiz robô permanecendo como um subordinado *nullo medio* do juiz humano. Como afirma RISSLAND (1990), “*the goal is to assist, not to replace*”<sup>160</sup> [o objetivo é auxiliar, e não substituir].

A legislação da União Europeia já proíbe decisões totalmente automatizadas por entes governamentais<sup>161</sup>, sem que passem pela supervisão de humanos; e a International Bar

---

<sup>157</sup> HALLEVY, Gabriel. *When Robots Kill: artificial intelligence under criminal law*. Boston: Northeastern University Press, 2013, pp 12-13

<sup>158</sup> Na proposta de RAMGE, Thomas. *Who’s Afraid of AI?: Fear and Promise in the Age of Thinking Machines*. New York: The Experiment, LLC, 2019, “(...) jurisprudence is especially well suited for automation with artificial intelligence because it builds on precisely formulated rules (laws and regulations) using highly formalized language, and a large library of legal matters that are documented in case summaries and opinions, commentaries, and contracts exists that machines with pattern-recognition capabilities can draw from comparisons.” (p. 72)

<sup>159</sup> Adotamos a separação entre “*easy cases*” e “*hard cases*” de GARDNER, Anne von der Lieth. *An artificial intelligence approach to legal reasoning*. Cambridge, MA: MIT Press, 1987.

<sup>160</sup> RISSLAND, Edwina L. “Artificial Intelligence and Law: Stepping Stones to a Model of Legal Reasoning” in: *Yale Law Journal*, Vol. 99 Issue 8, 1990, pp. 1957-1981. Disponível em: <http://digitalcommons.law.yale.edu/ylj/vol99/iss8/4> - acesso em 01 Mai 2023, p. 1980

<sup>161</sup> Disponível em:

<https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32016R0679&from=EN#d1e2793-1-1> – último acesso em 03 Jun 2023. A GDPR (General Data Protection Regulation) estabelece em seu artigo 22:

Association<sup>162</sup> não recomenda a adoção de robôs-peticionadores sem que os textos passem por revisão por humanos.

### 3.2 Leis Personalizadas propostas por Ben-Shahar & Porat

Na *Summa Theologica*, Tomás de Aquino discorre sobre os julgamentos “(...) dizendo que a sentença do juiz é como uma certa lei particular em algum caso particular.”<sup>163</sup> A concepção de que a função do juiz é interpretar a lei e ponderar sobre as características do texto que são mais ou menos aplicáveis a cada caso concreto não é original do filósofo dominicano, existindo previamente a seu texto. No entanto, a partir desta ideia de adequação da regra aplicável ao caso concreto, na obra “Personalized Law”, Ben-Shahar & Porat propõem a aplicação de inteligência artificial não no judiciário, mas em um passo atrás, no legislativo.

A proposta visa a permitir a customização de certas regras que atualmente são aplicadas de forma geral, com base em diferentes características ou habilidades dos cidadãos. Por exemplo, motoristas habilitados e com maior capacidade motora poderiam conduzir em velocidades mais elevadas do que condutores idosos e/ou menos destros. Além disso, um veículo equipado com sensores poderia identificar sinais de fadiga em um motorista após uma noite mal dormida e informá-lo de que seu limite de velocidade estaria reduzido naquele dia, considerando que seus reflexos estariam prejudicados. Em outras

---

“Article 22 - Automated individual decision-making, including profiling: 1. The data subject shall have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her. 2. Paragraph 1 shall not apply if the decision: (a) is necessary for entering into, or performance of, a contract between the data subject and a data controller; (b) is authorised by Union or Member State law to which the controller is subject and which also lays down suitable measures to safeguard the data subject's rights and freedoms and legitimate interests; or (c) is based on the data subject's explicit consent. 3. In the cases referred to in points (a) and (c) of paragraph 2, the data controller shall implement suitable measures to safeguard the data subject's rights and freedoms and legitimate interests, at least the right to obtain human intervention on the part of the controller, to express his or her point of view and to contest the decision. 4. Decisions referred to in paragraph 2 shall not be based on special categories of personal data referred to in Article 9(1), unless point (a) or (g) of Article 9(2) applies and suitable measures to safeguard the data subject's rights and freedoms and legitimate interests are in place.”

<sup>162</sup> Podcast sobre o assunto disponível em: [https://www.ibanet.org/podcast-AI-law-and-the-legal-profession?sap-outbound-](https://www.ibanet.org/podcast-AI-law-and-the-legal-profession?sap-outbound-id=DEDDFF0C01C443CADE3BD5D0E8DDF9F46836C739&utm_source=SAPHybris&utm_medium=email&utm_campaign=3310&utm_term=Podcast%3A%20AI%2C%20law%2C%20legal%20profession___Listen%20to%20the%20podcast%20online&utm_content=EN)

[id=DEDDFF0C01C443CADE3BD5D0E8DDF9F46836C739&utm\\_source=SAPHybris&utm\\_medium=email&utm\\_campaign=3310&utm\\_term=Podcast%3A%20AI%2C%20law%2C%20legal%20profession\\_\\_\\_Listen%20to%20the%20podcast%20online&utm\\_content=EN](https://www.ibanet.org/podcast-AI-law-and-the-legal-profession?sap-outbound-id=DEDDFF0C01C443CADE3BD5D0E8DDF9F46836C739&utm_source=SAPHybris&utm_medium=email&utm_campaign=3310&utm_term=Podcast%3A%20AI%2C%20law%2C%20legal%20profession___Listen%20to%20the%20podcast%20online&utm_content=EN) – último acesso em 03 Jun 2023.

<sup>163</sup> Tomás de Aquino, *Summa Theologica*, Secunda Secundae Partis, QUEST. LXVII - Articulus I. No original: “Respondeo dicendum quod sententia iudicis est quasi quaedam particularis lex in aliquo particulari facto.”

palavras, a personalização da lei poderia ocorrer em tempo real, e não apenas em momentos pré-definidos.

Dessa forma, em caso de um acidente de trânsito, durante o julgamento, seria levada em consideração a velocidade permitida na situação específica para aquele indivíduo, ao invés de uma regra geral válida para toda a sociedade. Segundo os autores, o estágio tecnológico atual já possibilitaria a implementação dessa ideia.

Os pensadores contratualistas dos séculos XVI e XVII defendiam que as diferenças individuais não eram suficientes para justificar tratamentos distintos, exigindo que a mesma lei fosse aplicada a todos. Em um primeiro exame dessa concepção, Ben-Shahar & Porat concordam que o princípio fundamental de igualdade perante a lei, que está presente na maioria das Constituições, aparenta ser ameaçado, assim como o poder conferido ao Estado para obter acesso a uma ampla gama de dados dos cidadãos, a fim de realizar o processo de personalização.

Assim, apenas após superar a discussão sobre se a personalização constituiria ou não uma violação do princípio de igualdade perante a lei, afastando possíveis riscos de perturbações no tecido social e sendo aceita pela sociedade como constitucional, poder-se-ia tomar uma decisão, na visão dos autores, sobre quais áreas do direito seriam passíveis de sujeição a essa metodologia. A partir de então, seria ainda exigida uma análise dupla: primeiramente, uma dimensão "externa", na qual seria necessário definir quais direitos e deveres seriam personalizáveis, e, em segundo lugar, uma dimensão "interna", na qual seriam discutidos os parâmetros a serem utilizados para a personalização da lei.<sup>164</sup>

Como argumento adicional para a implementação, Ben-Shahar & Porat (2021) alegam que a lei personalizada é uma manifestação explícita do dilema entre "regras versus princípios"<sup>165</sup>, sendo meramente um procedimento para concretizar o valor subjacente à lei em um caso específico, não apresentando nada de novo. De acordo com os autores, já são aceitas como legítimas várias formas de personalização, como a dispensa de vacinação por motivos religiosos. Esse comportamento não seria um retrocesso ao princípio liberal da igualdade perante a lei, uma vez que a própria sociedade opta por um conjunto de valores que não estão sujeitos ao tratamento igualitário, mas que, em última instância, são diferenças de tratamento aceitas com o objetivo de manter a estabilidade social.

---

<sup>164</sup> BEN-SHAHAR, Omri & PORAT, Ariel. *Personalized Law : different rules for different people*. New York, NY: Oxford University Press, 2021, p. 8

<sup>165</sup> BEN-SHAHAR, Omri & PORAT, Ariel. *Personalized Law : different rules for different people*. New York, NY: Oxford University Press, 2021, p. 20

Uma das principais críticas à proposta da lei personalizada, bem como do julgamento por algoritmos, é sua incapacidade para lidar com mudanças de valores na sociedade. Para Fagan & Levmore (2019), “o aprendizado de máquina é menos impressionante quando o passado é diferente do futuro”<sup>166</sup>, uma vez que na ausência de regularidades, princípios, mais flexíveis, podem se tornar superiores em termos de praticidade em relação a regras rígidas. Isso porque a aplicação de princípios pode ser feita sem que quaisquer um deles seja derogado, podendo conviver simultaneamente,<sup>167</sup> enquanto as regras, de característica “tudo-ou-nada”, estão mais próximas da maneira “zero-ou-um” do computador.

Mesmo Ben-Shahar & Porat defendendo a tese de que seria exatamente de que “regras personalizadas têm o potencial de atingir os objetivos subjacentes de qualquer lei de forma mais eficaz”<sup>168</sup>. No entanto, o livro carece de uma exploração mais aprofundada sobre o "como", além da defesa conceitual.

Em termos da utilização de algoritmos nos julgamentos, seria indiferente o grau de personalização da lei para o processo de julgar, sendo o computador capaz de adaptar-se a qualquer um dos cenários (leis individualizadas ou não). A proposta de Ben-Shahar & Porat, porém, levaria a uma maior naturalização da presença de métodos quantitativos na sociedade.

### 3.3 A visão de Guarrapon & Lassègue

A pergunta com relação à possibilidade de algoritmização da justiça parece ser parente do questionamento sobre a possibilidade de algoritmização da moral.<sup>169</sup> Sobre ambas, paira uma ameaça de diferença ontológica: o direito e a moral em números descolar-

---

<sup>166</sup> FAGAN, Frank & LEVMORE, Saul, "The Impact of Artificial Intelligence on Rules, Standards, and Judicial Discretion", Coase-Sandor Working Paper Series in Law and Economics, No. 876, 2019. Disponível em: [https://chicagounbound.uchicago.edu/law\\_and\\_economics/989/](https://chicagounbound.uchicago.edu/law_and_economics/989/) - último acesso em 20 Mai 2023.

<sup>167</sup> LIMA, ANDREI FERREIRA DE ARAÚJO & Thadeu Weber. “A A Justiça e o problema da obediência a uma lei injusta - uma análise comparativa das teorias de Rawls e Dworkin” in: VERITAS - Revista de Filosofia da PUCRS, Porto Alegre, v. 65, n. 3, p. 1-15, set.-dez. 2020. Disponível em: <https://hdl.handle.net/10923/20572> - último acesso em 20 Mai 2023.

e-ISSN: 1984-6746 | ISSN-L: 0042-3955

<sup>168</sup> BEN-SHAHAR, Omri & PORAT, Ariel. *Personalized Law : different rules for different people*. New York, NY: Oxford University Press, 2021, p. 39. No original: “personalized rules have the potential to accomplish the underlying goals of any law more effectively”

<sup>169</sup> BRAITHWAITE, R.B. *Theory of Games as a Tool for the Moral Philosopher*. Cambridge, UK: Cambridge University Press, 1963 aborda o tema a partir da contribuição de John Nash do início dos anos 1950.

se-iam das palavras e se tornariam algo diferente da práxis no mundo físico que os caracteriza.<sup>170</sup>

Dentre as resistências mais comuns às máquinas (e não apenas ao computador), é que estas “desumanizam” o humano, tratando-o como uma coisa (ou um dado).<sup>171</sup> Do filme “Tempos Modernos”<sup>172</sup> a recentes críticas<sup>173</sup> à introdução de máquinas no mundo dos homens (assumindo que há mundo dos homens sem máquinas), a lista cresce na direção de uma eliminação da essência humana, nos moldes do ideal rousseauiano de contato não-intermediado com a natureza.

Uma das perguntas que Guarrapon & Lassègue (2018) fazem é se a transformação da linguagem natural humana no algoritmo muda a sua concepção de justiça. Isso porque mesmo a gramática já funciona como uma forma de delimitar as possibilidades de expressão em uma língua, organizando o discurso e, portanto, o dizer do mundo. Quando um sistema funciona não mais com um alfabeto de 26 letras, e sim apenas com zeros e uns, uma polarização parece ser uma das consequências.

Em outras palavras, a justiça digital é apenas mais uma característica, ou seria algo que tem a potencialidade de alterar a estrutura social, por meio de mudanças nas maneiras como se dão as relações entre os membros da sociedade?<sup>174</sup> Por exemplo, a expectativa de uma justiça célere (instantânea), reduziria ou aumentaria os conflitos? Além disso, se o mesmo sistema que estivesse sendo utilizado pelo governo estivesse à disposição dos cidadãos para consultas (as partes poderiam consultar de antemão como seria o resultado oferecido pelo algoritmo de uma lide antes de decidirem levar o caso à justiça) levaria a um enfraquecimento do estado enquanto principal provedor de justiça, promovendo uma privatização da justiça? O momento é de mais questões do que de respostas, mas acreditamos ser necessário fazê-las.

---

<sup>170</sup> NAGEL, Thomas. *Mortal Questions*. New York : Cambridge University Press, 1979 questiona inclusive se pode haver Ética sem o componente biológico (cap. 10, pp. 144-146).

<sup>171</sup> Esta é a visão, compartilhada por inúmeros autores, de SADIN, Éric. *L'intelligence artificielle ou l'enjeu du siècle: anatomie d'un antihumanisme radical*. Paris : L'Échappée, 2018. Em contraponto, SCHKOLNE, Steven. *Living with Frankenstein: the history and destiny of machine consciousness*. Los Angeles: Machine Press, 2020, para quem “Our deep desire to see ourselves as special, seen in theology and mythology, permeates even non-theistic narratives. We bitterly refuse to find our likeness in machines. Our minds spew mountains of rationale to hide our species-ego from the truth.” (p. 32)

<sup>172</sup> *Tempos Modernos*. EUA: United Artists, 1936. Direção: Charles Chaplin (87 min.).

<sup>173</sup> Além dos trabalhos já citados, podemos apontar: REY, Olivier. *Quand le monde s'est fait nombre*. Paris, Éditions Stock, 2016, que aponta que a sociologia foi a principal porta de entrada dos números nas humanidades, ao adotar a estatística como principal ferramenta (cap. VI).

<sup>174</sup> GARAPON, Antoine & LASSÈGUE, Jen. *Justice Digitale : Révolution graphique et rupture anthropologique*. Paris : PUF, 2018, pp. 299-300

Assim, para Garapon & Lassègue (2018), a IA, uma vez que está direcionada a resolver conflitos, não precisa se ater a nenhum regime jurídico, nem mesmo testar se se submete a algum tipo de classificação ou hierarquia, pois seu valor é a eficácia em resolver conflitos, podendo acolher diferentes princípios a qualquer momento, se assemelhando a um “bolo formigueiro”<sup>175</sup>, sendo que seu desenho interno não precisa ter uma definição prévia, bastando que sirva ao seu propósito de alimentar a sociedade com aquilo que lhe é demandado. “Quantitas, non auctoritas facit legem”, ou seja, é a massa de dados de julgamentos passados que fornece a estabilidade ao sistema e permitem aos membros da sociedade saberem como devem se comportar

### 3.4 Confiança no sistema

A partir da ideia de Tribe (1971) acima exposta, poderia ser afirmado que a confiança no julgamento por algoritmos seria uma proxy da confiança na capacidade da IA em entregar julgamentos vistos como válidos e justos para a sociedade.

Quando Hart (1961) argumenta que a legitimidade da lei se fundamenta na *rule of recognition*<sup>176</sup>, ele sustenta que a sociedade, por meio de uma convenção, concede autoridade a um grupo específico de indivíduos para promulgar as leis, transferindo, assim, um poder que antes estava disperso para o sistema legal. No entanto, diferentemente da construção do positivismo kelseniano com base em uma *Grundnorm*, poderíamos afirmar que a sustentabilidade dessa legitimidade está intrinsecamente ligada não a um evento único no passado, mas sim à crença contínua no bom funcionamento do sistema. Nesse sentido, à medida que as leis e suas interpretações nos momentos de julgamento resultam em sentenças que promovem a coesão social, a legitimidade do sistema se mantém (apesar da sua aparência autônoma,<sup>177</sup> o campo jurídico interage com as outras esferas da vida em sociedade)

Em outras palavras, a sustentação da legitimidade não estaria em uma mera base piramidal invertida, e sim em uma microfísica<sup>178</sup> da coesão social que permanece

---

<sup>175</sup> GARAPON, Antoine & LASSÈGUE, Jen. Justice Digitale : Révolution graphique et rupture anthropologique. Paris : PUF, 2018, pp. 200-201

<sup>176</sup> Hart, Herbert L. A. The Concept of Law. Oxford University Press, 1961.

<sup>177</sup> BOURDIEU, Pierre. La force du Droit : éléments pour une sociologie du champ juridique. Éditions de la Sorbonne, 2017.

<sup>178</sup> Sobre a ideia de que o poder se encontra disperso pela sociedade em cada uma das relações entre os indivíduos, FOUCAULT, Michel. Microfísica do Poder. Rio de Janeiro: Paz e Terra, 2017.

sustentando o sistema enquanto este permite que as lides entre os membros da sociedade sejam resolvidas por um conjunto de julgamentos que não perturbam essas relações individuais a ponto de rompê-las. É através dessa correlação entre a eficácia das leis e a preservação da coesão social que a legitimação se fortalece.

Acreditamos que Garapon & Lassègue (2018) atingem o centro do debate: “até que ponto devemos delegar à máquina as atividades até então realizadas pelos humanos?” [*“jusq’où déléguer à la machine les activités jusq’à présent menées à bien par les humains?”*]. Como vimos acima, existe uma certa quantidade de incerteza, ou mesmo imprecisão, que é necessária à lei para que esta possa se adaptar a mudanças no seu contexto. Esta necessidade é análoga à flexibilidade biológica que permitiu ao ser humano se adaptar a mudanças no meio ambiente ao longo de sua história e sobreviver.<sup>179</sup>

A mesma opinião tem FAGAN & LEVMORE (2019), quando afirmam que

Na ausência de regularidades, o aprendizado de máquina perde sua vantagem e, como resultado, princípios mais flexíveis podem tornar-se superior às regras. (...) Assim como os juízes de hoje anulam erros e precedentes antiquados, os legisladores de amanhã irão sensatamente anular a IA em domínios jurídicos onde os desafios de medição estão presentes.<sup>180</sup>

Assim, poder-se-ia propor que a legitimidade da IA em julgamentos não vem da prova inequívoca de sua precisão, mas de um balanço entre os valores da celeridade da justiça e segurança jurídica desejada, por um lado, e o quanto de erro o sistema judicial pode apresentar e ainda assim permanecer como uma forma legítima de resolução de conflito entre os membros da sociedade, de outro lado. Ao mesmo tempo que os autores criticam o “mito da delegação inexorável às máquinas”<sup>181</sup>, com a eliminação do humano, a linha argumentativa de Garapon & Lassègue (2018) é similar à ideia que pode ser extraída de Tribe (1971), uma vez que aqueles afirmam que “Desde que o direito deixou de se basear no

---

<sup>179</sup> GARAPON, Antoine & LASSÈGUE, Jen. *Justice Digitale : Révolution graphique et rupture anthropologique*. Paris : PUF, 2018, p. 161

<sup>180</sup> FAGAN, Frank & LEVMORE, Saul. “The impact of artificial intelligence on rules, standards and judicial discretion” in *Southern California Law Review*, Vol. 93:1, 2019. Disponível em: [https://southerncalifornialawreview.com/wp-content/uploads/2020/01/93\\_1\\_Levmore.pdf](https://southerncalifornialawreview.com/wp-content/uploads/2020/01/93_1_Levmore.pdf) - último acesso em 22 Mai 2023. No original: “In the absence of regularities, machine learning loses its advantage and, as a result, looser standards can become superior to rules. (...) Just as today’s judges overrule errors and outmoded precedent, tomorrow’s lawmakers will sensibly overrule AI in legal domains where the challenges of measurement are present. When measurement is straightforward and prediction is accurate, rules will prevail. When empirical limitations such as overfit, Simpson’s Paradox, and omitted variables make measurement difficult, AI should be trusted less and law should give way to standards.”

<sup>181</sup> GARAPON, Antoine & LASSÈGUE, Jen. *Justice Digitale : Révolution graphique et rupture anthropologique*. Paris : PUF, 2018, p. 122.

transcendente religioso ou na lei natural, ele teve que buscar em outro lugar suas referências tanto para fundamentar suas regras quanto para preservar sua legitimidade.”<sup>182</sup>

Uma das críticas à utilização em larga escala para de modelos generativos de linguagem (LLM) é que eles ainda têm muito para serem aperfeiçoados, e ainda estão sujeitos a “alucinações”. As alucinações inputadas à inteligência artificial, como o ChatGPT, referem-se a situações em que a IA gera respostas ou informações que não são baseadas em dados reais ou em seu treinamento. Essas alucinações podem ocorrer quando o modelo de linguagem é solicitado a fornecer informações ou respostas além de seu conhecimento ou capacidades, levando a respostas inventadas ou fictícias. Todavia, respostas não-usuais também podem ser encontradas na base de dados utilizada pela IA.<sup>183</sup>

*Supremo Tribunal Federal*

Inteiro Teor do Acórdão - Página 14 de 15

19/06/2012

PRIMEIRA TURMA

HABEAS CORPUS 103.412 SÃO PAULO

VOTO

**O SENHOR MINISTRO DIAS TOFFOLI (PRESIDENTE):**

Também cumprimento a eminente Relatora pelo profundo voto trazido, e digo que penalizar a cogitação, ou a imaginação ou o pensamento, só Deus pode fazer, e não o homem. Nós não estamos nesta esfera de cognição. Mas verifico, já falando em Deus, que os astros hoje estão alinhados pela concessão das ordens.

O SENHOR MINISTRO MARCO AURÉLIO – Confesso que estou me sentindo em um Colegiado diverso daquele que geralmente integro às terças-feiras!

**O SENHOR MINISTRO DIAS TOFFOLI (PRESIDENTE):**

É por isso que eu acredito em Deus, mas eu acredito também na astrologia. Os astros hoje estão alinhados, em uma conjugação favorável aos pacientes.

---

<sup>182</sup> GARAPON, Antoine & LASSÈGUE, Jen. Justice Digitale : Révolution graphique et rupture anthropologique. Paris : PUF, 2018, p. 105. No original : « Depuis que le droit a cessé de s'appuyer sur le transcendant religieux ou sur le droit naturel, il a dû chercher ailleurs ses références à la fois pour fonder ses règles et préserver sa légitimité. »

<sup>183</sup> Disponível em: <https://redir.stf.jus.br/paginadorpub/paginador.jsp?docTP=TP&docID=2612894> – último acesso em 30 Jun 2023.

### 3.5 Os *hard cases*

Com base na possibilidade de legitimidade da IA nos julgamentos, resta ainda uma última discussão a fazer. Destacamos anteriormente a contribuição à segurança jurídica realizada pela estabilidade que a IA provê ao sistema ao julgar de acordo com as decisões passadas. Mas o sistema judiciário se vê às vezes envolvido com os denominados *hard cases*, situações que não encontram equivalência fácil na jurisprudência, sendo, porém, importantes no desenvolvimento do direito, pois desafiam os limites das leis existentes, levando a interpretações inovadoras, contrariando precedentes estabelecidos ou até mesmo criando novas normas para lidar com situações complexas. Neste momento, a IA deveria ter capacidade de inovar, ou buscar uma interpretação até então não utilizada da lei, visando a manter um princípio caro à sociedade. Ora, vimos que alguns autores pensam que isso seria difícil, dada a inabilidade da IA em perceber seu contexto<sup>184</sup>. Todavia, se a IA pode ser criativa, possibilidade levantada acima, então talvez ela pudesse fornecer soluções inovadoras para um *hard case* tanto quanto humanos.

Pelo ângulo de Dworkin (1986), os *hard cases* são importantes porque são problemas interpretativos, sendo mais discussões semânticas sobre a essência de uma lei do que problemas de adequação (que lei utilizar). Neste aspecto, divergimos de seu ponto de vista, pois nos parece que Dworkin (1986) subestima o papel que a variação da hierarquia de valores tem no momento da interpretação de uma lei. Em outras palavras, o “ranking axiológico” de uma sociedade não parece ser estático (por exemplo, um grupo pode preferir liberdade a segurança em um momento, e o oposto em outro). A questão é o quanto uma IA seria capaz de capturar o *zeitgeist* quando *hard case* é julgado, sendo este necessário para não prejudicar a coesão social. Seria o algoritmo capaz dessa sensibilidade?

Como aponta Posner (2013), nem sempre “complexo” significa “difícil”.<sup>185</sup> Separemos dois aspectos: a capacidade de apresentar uma nova solução da adequação dessa solução ao momento. Já vimos acima que a criatividade não seria um problema à IA, e no contexto de um *hard case*, é comum defender que um juiz deve ter a capacidade de inovar e apresentar uma nova interpretação da lei. Para Ramge (2019), “desde a vitória de AlphaGo, uma coisa está clara: intuição e criatividade (a depender da definição de cada um desses

---

<sup>184</sup> Como já mencionado acima: BROUSSARD, Meredith. Artificial UnIntelligence: how computers MisUnderstand the world. Cambridge, MA: The MIT Press, 2018, em especial cap. 2 (pp. 13-30)

<sup>185</sup> POSNER, Richard A. Reflections on Judging. Cambridge, MA: Harvard University Press, 2013, em especial cap. 3: “The Challenge of Complexity”.

termos) não são mais domínios exclusivos dos humanos.”<sup>186</sup> Deste modo, seria possível contar com os algoritmos também para propor novas interpretações em *hard cases*, desde que a adequação – por enquanto – seja realizada por um humano, axiologicamente mais competente para perceber se a decisão tem ou não potencial para desarrumar o tecido social. Como afirma Ramge (2019), “qualquer decisão importante, inclusive a decisão com relação a quanta ajuda por parte de máquinas é apropriada, permanece sendo uma decisão humana”<sup>187</sup>

Autores como Vallor (2016) destacam que ser entusiasta de uma determinada tecnologia não implica necessariamente reprimir quaisquer críticas. Por outro lado, um cuidado deve ser tomado para que a crítica não seja utilizada – mesmo que inconscientemente – como uma forma de racionalização de uma rejeição ou defesa perante o desconhecido. Quando se trata de uma inteligência artificial, muitas vezes exigimos uma espécie de explicação ou trilha de raciocínio para justificar suas decisões.<sup>188</sup> Essa abordagem parece ser tendenciosa contra o computador e a favor do ser humano [*human biased*], uma vez que, embora o juiz precise fundamentar sua nova interpretação, não é necessário que ele explique como a ideia surgiu. Exigi-la do computador seria ir além do que se exige para um juiz humano, sendo na verdade um “viés de espécie” entre os humanos.<sup>189</sup> Todavia, ao exigir uma explicação detalhada da máquina, corremos o risco de estabelecer um padrão inatingível apenas como impeditivo da automatização – com as redes neurais, nem sempre é possível rastrear exatamente como um algoritmo chegou a uma determinada conclusão. Embora seja essencial garantir a transparência<sup>190</sup> nas decisões tomadas por sistemas automatizados, deveríamos evitar impor padrões que dificultem a utilização eficaz da inteligência artificial no campo jurídico.

---

<sup>186</sup> RAMGE, Thomas. Who’s Afraid of AI?: Fear and Promise in the Age of Thinking Machines. New York: The Experiment, LLC, 2019, p. 45. No original: “Since AlphaGo’s victory, it has been clear: intuition and creativity (depending on the definition) are no longer exclusively human domains.”

<sup>187</sup> RAMGE, Thomas. Who’s Afraid of AI?: Fear and Promise in the Age of Thinking Machines. New York: The Experiment, LLC, 2019, No original: “The important decisions, including the decision about how much machine assistance is appropriate, remain human ones.” (p. 116)

<sup>188</sup> MOKANDER, J., JUNEJA, P., WATSON, D.S. et al. The US Algorithmic Accountability Act of 2022 vs. The EU Artificial Intelligence Act: what can they learn from each other?. *Minds & Machines* (2022). Disponível em: <https://doi.org/10.1007/s11023-022-09612-y> - último acesso em 30 Jun 2023.

<sup>189</sup> VALLOR, Shannon. *Technology and the Virtues: a philosophical guide to a future Worth living*. New York, NY: Oxford University Press, 2016. Para a autora, se a tecnologia tem impacto na moral, uma nova construção da moral em um mundo com novas tecnologias deve incorporar esses impactos, uma vez que presentes na sociedade, e de uma maneira global – não apenas a perspectiva "ocidental", mas também as filosofias morais sistemáticas de outras culturas, como o budismo e o confucionismo. Isso reflete a crescente necessidade, em um mundo multipolar, de construir pontes entre tradições morais e estabelecer um diálogo global sobre a boa vida e o tipo de sociedades em que desejamos viver.

<sup>190</sup> STROWEL, Alain. *Quand le Google défie le droit : plaidoyer pour une Internet transparent et de qualité*. Bruxelas : Éditions Larcier, 2011.

É notável que os juízes, ao justificarem suas decisões em *hard cases*, frequentemente aleguem ou a escolha entre a interpretação da lei com base em seu propósito original (*mens legis originalis*) – conceito já em si polêmico, tendo sua existência contestada – ou a necessidade de uma nova interpretação contextualizada. Há, comumente, ausência de um critério claro que direcione a opção entre aderir à *mens legis* ou optar por uma interpretação que rompa com essa perspectiva.<sup>191</sup> Nesse sentido, ambas as abordagens são consideradas legítimas, uma vez que não há um fundamento objetivo que possa estabelecer uma superioridade entre as opções. Mesmo a denominada “interpretação sistêmica” pode ser utilizada, uma vez que o sistema é dinâmico.<sup>192</sup>

Retornando à argumentação de Jerome Frank apresentada no capítulo 2, são inúmeros os fatores que atuam sobre o processo decisório do juiz. De igual modo, Garapon & Lassègue (2018) se questionam até mesmo qual seria a verdadeira influência da jurisprudência sobre o juiz?<sup>193</sup>

Na obra "How Judges Think", Posner (2008), examina o processo de tomada de decisões judiciais e apresenta alguns fatores: a) primeiramente, argumenta que os juízes são, em grande parte, pragmáticos em sua abordagem para resolver casos: em vez de se apegarem rigidamente a doutrinas e precedentes legais, levam em consideração as consequências práticas de suas decisões e como elas afetarão a sociedade em geral; b) economia do direito: os juízes consideram incentivos econômicos e as consequências sociais ao interpretar e aplicar a lei; c) além disso Posner (2008) alinha-se ao realismo, quando afirma que as decisões judiciais são influenciadas por fatores políticos, sociais e econômicos, além de meras considerações jurídicas, argumentando que os juízes são seres humanos com suas próprias ideologias e preferências, e esses fatores inevitavelmente afetam suas decisões. Deste modo, o autor reconhece as limitações do sistema judicial e critica a visão de que os tribunais são a solução para todos os problemas sociais, argumentando que os tribunais têm um papel limitado e que questões complexas muitas vezes devem ser resolvidas por outros ramos do governo ou pela sociedade como um todo.<sup>194</sup>

---

<sup>191</sup> RAYNAUD, Philippe. Le Juge et le Philosophe : Essais sur le nouvel âge du droit. Malakoff, France : Armand Colin, 2020. Em especial no capítulo 11, o autor discute o caso *Marbury v. Madison*, dos Estados Unidos, que tratou da interpretação da Constituição americana alegando um retorno ao “sentido original”, pp. 149-160.

<sup>192</sup> DEWEY, John. “Logical Method and Law (1924)” in: FISHER III, William W.; HORWITZ, Morton J. & REED, Thomas A (eds.). *American Legal Realism*. New York: Oxford University Press, 1993, pp. 185-194.

<sup>193</sup> GARAPON, Antoine & LASSÈGUE, Jen. *Justice Digitale : Révolution graphique et rupture anthropologique*. Paris : PUF, 2018, p. 279

<sup>194</sup> POSNER, Richard A. *How Judges Think*. Cambridge, MA: Harvard University Press, 2008.

Acreditamos que essa visão de Posner (2008) ajuda no cuidado com a idolatria do algoritmo como a solução mágica para os erros humanos, não porque o humano seja intrinsecamente superior em qualquer atividade, mas na crítica da própria delegação de todas as resoluções de conflitos ao sistema judiciário, sejam elas performadas por humanos ou máquinas. Para o autor, nos hard cases, nos quais a discricionariedade é por definição maior, “o juiz está [na verdade] votando, e não decidindo”<sup>195</sup> – ou seja, qualquer solução criativa bem estruturada é bem-vinda, uma vez que o critério de decisão não é “o mais lógico” nem “o mais criativo”, e sim o mais alinhado com a hierarquia de valores pessoais do juiz.

### 3.6 Critérios para implementação

Quais deveriam ser os critérios para a adoção ou não de mais tecnologia em substituição a atividades mentais nos julgamentos? É uma indagação cada vez mais premente no momento em que a tecnologia já se encontra disponível e em processo parcial de implementação. Sunstein (2013), em seu influente artigo "Deciding by Default", discute os critérios para imposição por parte do estado de uma regra padrão.

Para Sunstein, uma *default rule* deveria ser acionada quando a decisão é intrinsecamente complexa, e quando a aversão à perda supera a utilidade marginal de um ganho.<sup>196</sup> Este seria o caso em litígios contratuais e tributários, por exemplo. Com base neste argumento, poderíamos sugerir que se a utilização de um juiz-robô traz maiores níveis de segurança jurídica e agilidade na tomada de decisões, poderíamos supor que essa seria a escolha predominante (assumindo que a sociedade encontra-se adequadamente informada sobre essas vantagens).<sup>197</sup>

Esses benefícios – eficiência e previsibilidade – podem ser entendidos como catalisadores de bem-estar social.<sup>198</sup> A eficiência derivada da aplicação de algoritmos nos

---

<sup>195</sup> POSNER, Richard A. *How Judges Think*. Cambridge, MA: Harvard University Press, 2008, p. 375.

<sup>196</sup> SUNSTEIN, Cass R. “Deciding by Default” in: *University of Pennsylvania Law Review*, 162, vol. 1, 2013, pp. 1-57. Disponível em: [https://scholarship.law.upenn.edu/penn\\_law\\_review/vol162/iss1/1](https://scholarship.law.upenn.edu/penn_law_review/vol162/iss1/1) - último acesso em 25 Jun 2023.

<sup>197</sup> SUNSTEIN, Cass R. “Deciding by Default” in: *University of Pennsylvania Law Review*, 162, vol. 1, 2013, pp. 1-57. Disponível em: [https://scholarship.law.upenn.edu/penn\\_law\\_review/vol162/iss1/1](https://scholarship.law.upenn.edu/penn_law_review/vol162/iss1/1) - último acesso em 25 Jun 2023. No original: “The preferred approach is to select the default rule that reflects what most people would choose if they were adequately informed” (p. 31)

<sup>198</sup> SUNSTEIN, Cass R. “Deciding by Default” in: *University of Pennsylvania Law Review*, 162, vol. 1, 2013, pp. 1-57. Disponível em: [https://scholarship.law.upenn.edu/penn\\_law\\_review/vol162/iss1/1](https://scholarship.law.upenn.edu/penn_law_review/vol162/iss1/1) - último acesso em 25 Jun 2023. No original: “Another reason is that the informed-chooser default is likely to promote people’s welfare. It is also easy to defend on grounds of both efficiency and fairness” (p. 32)

juízos levaria à economia de recursos, tanto tempo quanto monetários, beneficiando a sociedade como um todo. A previsibilidade, por outro lado, contribuiria para a redução da incerteza jurídica, criando um ambiente mais estável e confiável não apenas nos tribunais, mas nas relações comerciais como um todo. Assim, apoiando-se nesse raciocínio, seria possível argumentar que a utilização de juízos por algoritmos poderia se tornar a regra padrão em determinados ramos do direito, concedendo às partes envolvidas a opção de escolher um juiz humano.

No entanto, mesmo Sunstein (2013) alerta que uma regra padrão precisa de um período de testes antes de ser definitivamente implementada.<sup>199</sup> A seleção do *default*, nesse sentido, não deve ser guiada unicamente por uma análise custo-benefício *ex ante*, mas incorporando também eventuais consequências ainda não mapeadas quando de sua adoção: a introdução de tecnologias disruptivas requer cuidado para garantir que surpresas não antevistas não anulem os benefícios.<sup>200</sup>

Há também uma questão com relação ao impacto da implementação de tecnologias no mercado de trabalho das profissões ligadas ao mundo jurídico, tanto nos tribunais quanto nos escritórios de advocacia. Como vimos acima, tanto para peticionar quanto para as sentenças, um trabalho preparatório de pesquisa e organização é necessário. Em muitos casos, estagiários e/ou bolsistas realizam tarefas de pesquisa jurisprudencial, redação, revisão ou sugestões de melhoria. Nesse sentido, se a AI já é capaz de realizar todas essas “tarefas assessórias” (aqui sem valor de juízo de relevância), mesmo no atual estágio de evolução da tecnologia, são os humanos executores destas tarefas os primeiros a serem atingidos pela maior utilização do algoritmo.

Ao trazer o debate da AI para a prática laboral e problematizar o que é feito neste campo (no qual inúmeras funções já são delegadas a terceiros), podemos nos indagar se os futuros profissionais não seriam privados da possibilidade de adquirir conhecimento e treinamento

---

<sup>199</sup> SUNSTEIN, Cass R. “Deciding by Default” in: University of Pennsylvania Law Review, 162, vol. 1, 2013, pp. 1-57. Disponível em: [https://scholarship.law.upenn.edu/penn\\_law\\_review/vol162/iss1/1](https://scholarship.law.upenn.edu/penn_law_review/vol162/iss1/1) - último acesso em 25 Jun 2023. No original: “And indeed, selection of a default rule might well be preceded by a period of active choosing as a way of assembling that information. Perhaps experiments or pilot programs would provide such information. If only two percent of people opt out under A, and fifty percent opt out under B, we have reason to believe that A is better.” (p. 33)

<sup>200</sup> Sobre o tema, CORRÊA, Nicholas Kluge, OLIVEIRA, Nythamar Fernandes de, & MASSMANN, Diogo Fernando. Sobre a eficiência da ética como ferramenta de governança da inteligência artificial. Veritas (Porto Alegre), 2022, 67(1), e42584. <https://doi.org/10.15448/1984-6746.2022.1.42584> - último acesso em 27 Jun 2023.

necessários para sua formação, debilitando deste modo o conjunto do campo jurídico, nos moldes da preocupação exposta por Barth (2023) recentemente.<sup>201</sup>

Assim, além de abordar o processo de “retirada” do humano da atividade, evitando a falsa impressão de que a questão se encontra em uma necessidade de posicionamento entre apologia ingênua ou rejeição completa, pretendemos destacar de que a implementação deveria vir acompanhada da preocupação com uma atividade colateral que sofreria impacto significativo: a educação técnica das gerações futuras. O caminho não parece ser paralisar a implementação da tecnologia, mas agregar ao processo a manutenção das possibilidades de aprendizado que antes eram oferecidas tanto pelos escritórios de advocacia quanto pelos tribunais à parcela dos trabalhadores do campo justamente mais fragilizadas mesmo antes do surgimento da tecnologia.

Como lembra Kansteiner (2022), os algoritmos de inteligência artificial baseados em modelos de linguagem não “conseguem ainda operar categorias de verdade (...)”<sup>202</sup>. Ou seja, para ele, a capacidade de diferenciar verdade de mentira é o elemento essencial que faltaria ao robô (pois a máquina opera no princípio da sequência mais plausível, e não verdadeira),<sup>203</sup> e deve ser habilidade do humano que o opera.

---

<sup>201</sup> BARTH, Fernanda Daniela. A precarização das relações de trabalho na era digital: Uma reflexão a luz da teoria do reconhecimento de Axel Honneth / Fernanda Daniela Barth. 94p. Dissertação (Mestrado) - Programa de Pós-Graduação em Filosofia, PUCRS, 2023. Disponível em: <https://tede2.pucrs.br/tede2/bitstream/tede/10846/2/Fernanda%20Daniela%20Barth.pdf> – último acesso em 29 Jun 2023.

<sup>202</sup> KANSTEINER, W. (2022), “Digital doping for historians: can History, memory, and historical theory be rendered artificially intelligent?” in: *History and Theory*, 61: 119-133, p. 120. Disponível em: <https://doi.org/10.1111/hith.12282> - último acesso em 27 Mar 2023.

<sup>203</sup> KANSTEINER, W. (2022), “Digital doping for historians: can History, memory, and historical theory be rendered artificially intelligent?” in: *History and Theory*, 61: 119-133, p. 120. Disponível em: <https://doi.org/10.1111/hith.12282> - último acesso em 27 Mar 2023.

## **Conclusão**

Ao longo desta dissertação, percorremos um caminho que nos levou a explorar a relação entre tecnologia e justiça, com um enfoque específico no debate em torno da utilização de algoritmos em julgamentos judiciais, substituindo o trabalho dos juízes. Inicialmente, estabelecemos uma definição de justiça como uma forma de resolver conflitos e promover a convivência harmoniosa na sociedade.

No capítulo 1, analisamos as perspectivas essencialistas e funcionalistas em relação à substituição de tarefas humanas por máquinas. Identificamos que os autores essencialistas tendem a argumentar que a ausência de certas propriedades essenciais presente nos seres humanos torna impossível replicar os resultados do trabalho nas máquinas, tornando-as insatisfatórias nesse sentido. Por outro lado, autores que se identificam com uma corrente de caráter funcionalista não veem obstáculos para essa substituição e acreditam que as máquinas podem desempenhar eficientemente as tarefas anteriormente executadas pelos seres humanos.

Em seguida analisamos o argumento que uma das impossibilidades da utilização de inteligência artificial na substituição de tarefas humanas não-mecânicas seria porque à IA ainda faltaria a capacidade de ser criativa. Para isso, analisamos o conceito de criatividade e como visões essencialistas e funcionalistas veriam a possibilidade de atribuição de criatividade a máquinas. Enquanto autores identificados com o essencialismo conectam mais fortemente criatividade e consciência (o que de certo modo exigiria das máquinas um componente biológico para a capacidade criativa), autores funcionalistas focam mais no resultado da atividade, o que permitiria a atribuição de criatividade às máquinas.

No capítulo 2 adentramos em maiores detalhes no argumento de que a delegação de julgamentos às máquinas equivaleria ao retorno - e imposição para a sociedade - a uma escola de pensamento jurídico denominada “Realismo Jurídico”, com a alegação de que isso seria ruim, no sentido de potencialmente não satisfazer a finalidade da justiça em promover a harmonia na convivência entre os membros da sociedade. Para isso, resgatamos as ideias de autores identificados como membros da corrente do realismo jurídico, em especial Jerome Frank, principal herdeiro da tradição de Oliver Wendell Holmes Jr.

Em seguida, identificamos que os argumentos contra a utilização de máquinas em substituição às tarefas dos juízes parecem estar mais conectados a um medo de perda de controle sobre a tecnologia do que rejeição dos benefícios da sua utilização, uma vez que a implementação da tecnologia já em curso vem provando que esta consegue entregar

eficiência ao processo de julgar os conflitos entre os membros da sociedade e de maneira alinhada com os julgamentos pretéritos.

Ao final do capítulo 2, apresentamos a opinião de alguns autores que defendem uma posição intermediária como a mais adequada, em vez de adotar uma postura de rejeição total da tecnologia ou uma adoção ingênua, sem uma perspectiva crítica sobre os potenciais problemas que ela pode acarretar.

No capítulo 3, abordamos a contribuição de Tribe (1971) sobre a utilização de métodos matemáticos no direito. O autor destaca que uma certa quantidade de erros é inerente ao sistema jurídico, seja ele operado apenas por seres humanos ou com a incorporação de métodos matemáticos. Assim, podemos propor que a legitimidade da inteligência artificial (IA) em julgamentos não decorre de uma prova irrefutável de sua precisão, mas sim de um equilíbrio entre os valores da celeridade da justiça e a segurança jurídica desejada, por um lado, e a quantidade de erros que o sistema judicial pode apresentar e ainda assim permanecer como uma forma legítima de resolução de conflitos na sociedade, por outro lado.

Apresentamos autores que se posicionam em visões polarizadas, indo desde uma utilização intensa de máquinas para personalizar inclusive as leis aplicáveis a cada indivíduo, até tecnofóbicos, que rejeitam a implementação de novas tecnologias. Por fim, identificamos autores que defendem uma abordagem intermediária.

Em última análise, concluímos que a decisão sobre a quantidade de utilização da IA no sistema judicial é uma responsabilidade humana e deve ser avaliada e supervisionada pela sociedade, que deve ponderar os *trade-offs* entre eficiência e satisfação com o sistema de mediação de conflitos escolhido. Assim, a incorporação da IA no processo decisório do sistema judicial deve ser cuidadosamente considerada, mas não ignorada, uma vez que contribui para o bem-estar da sociedade ao mediar de maneira mais eficiente os conflitos entre seus membros.

## Referências Bibliográficas

ACEMOGLU, Daron & ROBINSON, James. Why nations fail: the origins of power, prosperity, and poverty (First ed.). New York: Crown Publishing Group, 2012.

ALTHUSSER, Louis. Ideology and ideological state apparatuses. In ALTHUSSER, Louis. Lenin and philosophy and other essays. New York: Monthly Review Press, 1971. Também disponível em: <https://www.marxists.org/reference/archive/althusser/1970/ideology.htm> - último acesso em 10 Abr 2023.

ARISTÓTELES. Categorias

ARISTÓTELES. Metafísica

ARISTÓTELES. Politeia

ARRUDA, José Maria. Universais e particulares: platonismo e nominalismo. in: Metafísica Contemporânea, Petrópolis, RJ: Editora Vozes.

BARTH, Fernanda Daniela. A precarização das relações de trabalho na era digital: Uma reflexão a luz da teoria do reconhecimento de Axel Honneth / Fernanda Daniela Barth. 94p. Dissertação (Mestrado) - Programa de Pós-Graduação em Filosofia, PUCRS, 2023. Disponível em: <https://tede2.pucrs.br/tede2/bitstream/tede/10846/2/Fernanda%20Daniela%20Barth.pdf> – último acesso em 29 Jun 2023.

BARRAT, James. Ou Final Invention : artificial intelligence and the end of the human era. New York: Thomas Dunne Books, 2013.

BEN-SHAHAR, Omri & PORAT, Ariel. Personalized Law: Different Rules for Different People. New York, NY: Oxford University Press, 2021.

BENJAMIN, Walter. “The Work of Art in the Age of Mechanical Reproduction.” Illuminations: Essays and Reflections, edited by Hannah Arendt, translated by Harry Zohn, Schocken Books, 1969.

BODEN, Margaret. A. The creative mind: Myths & mechanisms. New York, NY: Basic Books, 1991.

BOREL, Émile. "Mécanique Statistique et Irréversibilité". J. Phys. Paris : Series 5

BORGES, Jorge Luis. La biblioteca de babel. Buenos Aires: Emecé, 2000.

BOURDIEU, Pierre. La force du Droit : éléments pour une sociologie du champ juridique. Éditions de la Sorbonne, 2017.

BRAITHWAITE, R.B. Theory of Games as a Tool for the Moral Philosopher. Cambridge, UK: Cambridge University Press, 1963.

BROUSSARD, Meredith. *Artificial UnIntelligence: how computers MisUnderstand the world*. Cambridge, MA: The MIT Press, 2018.

BUTLER, Samuel. *Erewhon; Or, Over the Range*, cap XXIII. Project Gutenberg. Disponível em: <https://www.gutenberg.org/cache/epub/1906/pg1906-images.html> - último acesso em 10 Jun 2023.

CAMARGO, Guilherme Affonso Ferreira de. *Conflitos Jurisdicionais no século XIII: O Caso do Mosteiro de Santa Inês de Bolonha (1223-1298)*. Universidade Federal do Rio Grande do Sul. Instituto de Filosofia e Ciências Humanas. Dissertação, 2021. 149 f.

CAMARGO, G. A. F. *O Juiz-Robô: a inteligência artificial pode ser a solução para julgamentos céleres e sem surpresas?* in: FERREIRA NETO, A. M. (Org.). *Argumentação jurídica e questões controvertidas de Direito Constitucional contemporâneo*. 1ªed. Curitiba, PR: Instituto Memória Editora, 2020, v. I, pp. 128-149.

CANFORA, Luciano. *Il copista come autore*. Turin: Einaudi, 2001.

CARDOZO, Benjamin N. *The Paradoxes of Legal Science*. Westport, CT: Greenwood Press, 1970.

CHURCHLAND, Patricia S. *Braintrust: what neuroscience tells us about morality*. Princeton, NJ: Princeton University Press, 2011.

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

CORRÊA, Nicholas Kluge, OLIVEIRA, Nythamar Fernandes de, & MASSMANN, Diogo Fernando. *Sobre a eficiência da ética como ferramenta de governança da inteligência artificial*. *Veritas (Porto Alegre)*, 2022, 67(1), e42584. <https://doi.org/10.15448/1984-6746.2022.1.42584> - último acesso em 27 Jun 2023.

DELEUZE, Gilles & GUATTARI, Felix, *A Thousand Plateaus: Capitalism and Schizophrenia*, trans. Brian Massumi (Minneapolis, 1987).

DESCARTES, René. *Méditations sur la philosophie première*. Édition électronique, v.: 1,0 : *Les Échos du Maquis*, 2011. Disponível em: <https://philosophie.cegeptr.qc.ca/wp-content/documents/M%C3%A9ditations-1641-1647.pdf> – último acesso em 15 Fev 2023.

DREYFUS, HUBERT. *What Computers Still Can't Do: A Critique of artificial Reason*. Cambridge: MIT Press, 1992.

DEWEY, John. "Logical Method and Law (1924)" in: FISHER III, William W.; HORWITZ, Morton J. & REED, Thomas A (eds.). *American Legal Realism*. New York: Oxford University Press, 1993.

DWORKIN, Ronald. *Law's Empire*. Cambridge, MA: Harvard University Press, 1986.

ECO, Umberto. *Apocalittici e integrati: comunicazioni di massa e teorie della cultura di massa*. 3. ed. Milano: Bompiani, 1964.

- ELIAS, Norbert. *The Civilizing Process*. Hoboken, NJ: Wiley-Blackwell, 1994.
- EUBANKS, Virginia. *Automating Inequality: how high-tech tool profile, police, and punish the poor*. New York, NY: PICADOR, St. Martin's Press, 2018.
- FAGAN, Frank & LEVMORE, Saul, "The Impact of Artificial Intelligence on Rules, Standards, and Judicial Discretion", *Coase-Sandor Working Paper Series in Law and Economics*, No. 876, 2019. Disponível em: [https://chicagounbound.uchicago.edu/law\\_and\\_economics/989/](https://chicagounbound.uchicago.edu/law_and_economics/989/) - último acesso em 20 Mai 2023.
- FLORIDI, Luciano. *AI as Agency Without Intelligence: on ChatGPT, Large Language Models, and Other Generative Models*. *Philos. Technol.* 36, 15 (2023). Disponível em: <https://doi.org/10.1007/s13347-023-00621-y> - último acesso em 08 Jun 2023.
- FOUCAULT, Michel. *Microfísica do Poder*. Rio de Janeiro: Paz e Terra, 2017.
- FRANK, Jerome. *Courts on Trial*. New Jersey: Princeton University Press, 1973.
- FRANK, Jerome. *Fate and Freedom*. New York, NY: Simon and Schuster, 1945.
- FRANK, Jerome. *Law and the Modern Mind*. London, UK: Transaction Publishers, 2009.
- GADAMER, Hans-Georg. "A incapacidade para o diálogo" in: ALMEIDA, Custódio Luiz Silva de. *Hermenêutica Filosófica*. Porto Alegre: EDIPUCRS, 2000, pp. 129-140
- GALLAGHER, Shawn., & Zahavi, Dan. *The Phenomenological Mind*. Routledge, 2008.
- GARAPON, Antoine & LASSÈGUE, Jean. *Justice Digitale: Révolution graphique et rupture anthropologique*. Paris: Puf, 2020.
- GARDNER, Anne von der Lieth. *An artificial intelligence approach to legal reasoning*. Cambridge, MA: MIT Press, 1987.
- GIANETTI, Eduardo. *Auto-engano*. São Paulo: Companhia das Letras, 2017.
- GIANNETTI, Eduardo. *O anel de Gíges: uma fantasia ética*. São Paulo: Companhia das Letras, 2020.
- GIDDENS, Anthony. "Functionalism: Après La Lutte." in: *Social Research* 43, no. 2 (1976): 325–66. <http://www.jstor.org/stable/40970227>.
- HALLEVY, Gabriel. *When Robots Kill: artificial intelligence under criminal law*. Boston: Northeastern University Press, 2013.
- HART, Herbert L. A. *The Concept of Law*. Oxford University Press, 1961.
- HAUSKELLER, M. "What Is It Like to Be a Bot? SF and the Morality of Intelligent Machines." In B. Dainton, W. Slocombe, & A. Tanyi (Eds.), *Minding the Future Artificial Intelligence, Philosophical Visions and Science Fiction*. Springer: Springer Nature, 2021,

pp. 21-39. Disponível em: [https://link.springer.com/chapter/10.1007/978-3-030-64269-3\\_2#DOI](https://link.springer.com/chapter/10.1007/978-3-030-64269-3_2#DOI) – último acesso em 25 Jun 2023.

HEIDEGGER, Martin. “A questão da técnica” in: *Ensaio e conferências*, 8ª ed. Petrópolis: Vozes, 2012.

HOLMES, Oliver Wendell. *The Common Law*. New Orleans, LA: Quid Pro Books, 2010.

HOLMES, Oliver Wendell. “Law in Science and Science in Law.” *Harvard Law Review*, vol. 12, no. 7, 1899, pp. 443–63. Disponível em: <https://doi.org/10.2307/1321177> - último acesso em 07 Jul 2022, p. 448

HOMERO. *Ilíada*.

IMAGUIRE, Guido. *A substância e suas alternativas: feixes e tropos*. in: *Metafísica Contemporânea*. Petrópolis, RJ: Editora Vozes.

JORDANOUS, Anna. “A Standardised Procedure for Evaluating Creative Systems: Computational Creativity Evaluation Based on What it is to be Creative.” *Cognitive Computation*, 4 (3), 2012. pp. 246-279. ISSN 1866-9956. E-ISSN 1866-9964. Disponível em: <https://kar.kent.ac.uk/42379/> - último acesso em 14 Fev 2023.

KACZYNSKI, Theodore John. *Technological Slavery (vol. I)*. Scottsdale, AZ, USA: Fitch & Madison Publishers, 2022. Disponível em: <http://editions-hache.com/essais/pdf/kaczynski2.pdf> - último acesso em 18 Jun 2023.

KANSTEINER, W. (2022), “Digital doping for historians: can History, memory, and historical theory be rendered artificially intelligent?” in: *History and Theory*, 61: 119-133, p. 120. Disponível em: <https://doi.org/10.1111/hith.12282> - último acesso em 27 Mar 2023.

KANT, Immanuel. *Critique of the Power of Judgment (The Cambridge Edition of the Works of Immanuel Kant)* (P. Guyer, Ed.; E. Matthews, Trans.). Cambridge: Cambridge University Press, 2000.

KATZ, Daniel M., *Quantitative Legal Prediction – or – How I Learned to Stop Worrying and Start Preparing for the Data-Driven Future of the Legal Services Industry*, 62 *Emory L. J.* 823 (2013). Disponível em: <https://scholarlycommons.law.emory.edu/elj/vol62/iss4/6> - último acesso em 25 Mai 2023.

KESSLER, Friedrich, "Natural Law, Justice and Democracy--Some Reflections on Three Types of Thinking about Law and Justice," 19 *Tulane Law Review* 32, 1944, pp.32-61.

KUHN, Thomas S. *The Structure of Scientific Revolutions*. Chicago: The University of Chicago Press, 1970

LIMA, ANDREI FERREIRA DE ARAÚJO & Thadeu Weber. “A Justiça e o problema da obediência a uma lei injusta - uma análise comparativa das teorias de Rawls e Dworkin” in: *VERITAS - Revista de Filosofia da PUCRS*, Porto Alegre, v. 65, n. 3, p. 1-15, set.-dez. 2020. Disponível em: <https://hdl.handle.net/10923/20572> - último acesso em 20 Mai 2023. e-ISSN: 1984-6746 | ISSN-L: 0042-3955.

LIVIO, Mario. "Is math invented or discovered?". Disponível online em: <https://www.sfu.ca/~rpyke/cafe/livio.pdf> - último acesso em 19 Jun 2023.

LOEVINGER, Lee. "Jurimetrics: The Methodology of Legal Inquiry" in: *Law and Contemporary Problems*, n. 28, pp. 5-35 (Winter 1963). Disponível em: <https://scholarship.law.duke.edu/lcp/vol28/iss1/2> - último acesso em 07 Jul 2022.

LOVELACE, Ada Augusta. Notes upon the Memoir of L. F. MENABREA: Sketch of The Analytical Engine Invented by Charles Babbage. *Bibliothèque Universelle de Genève*, October, 1842, No. 82. Disponível em: <https://www.fourmilab.ch/babbage/sketch.html> - último acesso em 14 Jan 2023.

MALABOU, Catherine. *Métamorphoses de l'intelligence : Du QI à l'IA*. Paris : Puf, 2017.

MÁNTARAS, Ramón López de. "Artificial Intelligence and the Arts: Toward Computational Creativity" in: *The Next Step: Exponential Life*. BBVA Open Mind, 2017, pp. 99-123. Disponível em: <https://www.bbvaopenmind.com/wp-content/uploads/2017/03/BBVA-OpenMind-book-The-Next-Step-Exponential-Life-1-1.pdf> - último acesso em 12 dez 2022.

MARKS II, Robert J; DEMBSKI, William A. & EWERT, Winston. *Introduction to Evolutionary Informatics* (1st ed.). Singapore: World Scientific Publishing Company, 2017.

MARTINO, Antonio A. "¿Quién teme a los jueces robot?" in: GONZALO, Ana Dobratinich (org.). *Derecho y nuevas tecnologías*. Buenos Aires: La Ley, 2021 (pp. 15-41)

MODRAK, Deborah K.W. "Aristotle on phantasia" in: *The Routledge Handbook of Philosophy of Imagination*, 2016, online. Disponível em: <https://www.routledgehandbooks.com/doi/10.4324/9781315657905.ch1> - último acesso em 18 Jan 2023.

MOKANDER, J., JUNEJA, P., WATSON, D.S. et al. *The US Algorithmic Accountability Act of 2022 vs. The EU Artificial Intelligence Act: what can they learn from each other?*. *Minds & Machines* (2022). Disponível em: <https://doi.org/10.1007/s11023-022-09612-y> - último acesso em 30 Jun 2023.

MUMFORD, Lewis. *Technics and Civilization*. New York: Harcourt, Brace and Company, 1934

NUSSBAUM, Martha. *Frontiers of Justice: Disability, Nationality, Species Membership*. Cambridge, MA: Harvard University Press, 2006.

NAGEL, Thomas. *Mortal Questions*. New York: Cambridge University Press, 1979.

NORTH, Douglass C. 1991. "Institutions." *Journal of Economic Perspectives*, 5 (1): 97-112. Disponível em: <https://www.aeaweb.org/articles?id=10.1257/jep.5.1.97#:~:text=Institutions%20provide%20the%20incentive%20structure,my%20analysis%20from%20economic%20history> - último acesso em 10 Abr 2023.

OMRI, Ben-Shahar & PORAT, Ariel. Personalized law: different rules for different people. New York: Oxford University Press, 2021.

O'NEIL, Cathy. Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. Danvers, MA: Crown Publisher Group, 2017.

OST, François. "Júpiter, Hércules, Hermes: tres modelos de juez" in: Academia: Revista de Enseñanza del Derecho. Año 4, número 8, 2007.

PODSCHWADEK, Frodo. "Do androids dream of normative endorsement? On the fallibility of artificial moral agents" in: Artificial Intelligence and Law, Springer Netherlands, vol. 25, September 2017, pp. 325-339. Disponível em: <https://link.springer.com/article/10.1007/s10506-017-9209-6> - último acesso em 19 Jun 2023.

POSNER, Richard A. How Judges Think. Cambridge, MA: Harvard University Press, 2008.

POSNER, Richard A. Reflections on Judging. Cambridge, MA: Harvard University Press, 2013.

PRIEL, Dan. "What the Judge Had for Breakfast: A Brief History of an Unpalatable Idea" in: SSRN 12 jun 2017. Disponível em: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2982716](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2982716) – último acesso em 04 abr 2022.

RAMGE, Thomas. Who's Afraid of AI?: Fear and Promise in the Age of Thinking Machines. New York: The Experiment, LLC, 2019.

RAWLS, John. A Theory of Justice. Cambridge, MA: Harvard University Press, 1971, rev. ed. (1999).

RAWLS, John. Justiça como equidade: uma reformulação. São Paulo: Martins Fontes, 2003.

RAYNAUD, Philippe. Le Juge et le Philosophe : Essais sur le nouvel âge du droit. Malakoff, France : Armand Colin, 2020.

REY, Olivier. Quand le monde s'est fait nombre. Paris, Éditions Stock, 2016.

RIEDL, Mark O. "Human-centered artificial intelligence and machine learning." in: Human Behavior & Emerging Technologies. 2019; 1: pp. 33 - 36. Disponível em: <https://arxiv.org/pdf/1901.11184.pdf>.

RISSLAND, Edwina L. "Artificial Intelligence and Law: Stepping Stones to a Model of Legal Reasoning" in: Yale Law Journal, Vol. 99 Issue 8, 1990, pp. 1957-1981. Disponível em: <http://digitalcommons.law.yale.edu/ylj/vol99/iss8/4> - acesso em 01 Mai 2023.

RUDIGER, Francisco. "Humanismo, arte e tecnologia segundo Heidegger" in: Fragmentos de Cultura, v. 21, n. 7/9, pp. 433-451. Jul/set 2011. Disponível em:

<https://repositorio.pucrs.br/dspace/handle/10923/9821?mode=full> – último acesso em 28 Jun 2023.

RUSSELL, Stuart. *Human Compatible: artificial intelligence and the problem of control*. New York, NY: Penguin Books, 2020.

SADIN, Éric. *La vie algorithmique: critique de la raison numérique*. Paris: Éditions L'Échappée, 2021.

SADIN, Éric. *L'intelligence artificielle ou l'enjeu du siècle: anatomie d'un antihumanisme radical*. Paris: L'Échappée, 2018.

SARTRE, J-P. *L'être et le néant – essai d'ontologie phénoménologique*. France: Gallimard, 2007.

SCHKOLNE, Steven. *Living with Frankenstein: the history and destiny of machine consciousness*. Los Angeles: Machine Press, 2020.

SHANNON, Claude E. & WEAVER, Warren. *The Mathematical Theory of Communication*. Champaign, IL: The University of Illinois Press, 1971.

SHELLEY, Mary. *Frankenstein: The 1818 Text*. London: Penguin Classics, 2018.

SIMONDON, Gilbert. *L'invention dans les techniques*. Paris : Éditions du Seuil, 2005.

SCHULZ, Fritz. *Principles of Roan Law*. Oxford: Clarendon Press, 1936.

SCHWARTZ, John. "Is Aaron's work creative art or just high-tech doodling?" in *The Washington Post*, 10 abr 1995. Disponível em: <https://www.washingtonpost.com/archive/politics/1995/04/10/is-aarons-work-creative-art-or-just-high-tech-doodling/366dc211-2aec-4d99-8d55-27f7acd5e27d/> – último acesso em 6 Dez 2022

SEARLE, John. *A redescoberta da mente*. Lisboa: Piaget, 1992.

SEARLE, John R. "Minds, Brains, and Programs." *Behavioral and Brain Sciences*, vol. 3, no. 3, 1980, pp. 417-457. Disponível em: <https://web-archive.southampton.ac.uk/cogprints.org/7150/1/10.1.1.83.5248.pdf> - último acesso em 01 Mai 2023.

SHARPLES, Mike. "Storytelling by computer" in: *Digital Creativity*. 8. Taylor & Francis (Routledge), 1997. Também disponível online em: [https://www.researchgate.net/publication/232831648\\_Storytelling\\_by\\_computer](https://www.researchgate.net/publication/232831648_Storytelling_by_computer) - último acesso em 23 Jan 2023.

SHARPLES, Mike & PÉREZ Y PÉREZ, Rafael. *Story Machines: how computers have become creative writers*. New York, NY: Routledge, 2022.

SILVEIRA, Paulo Antonio Caliendo Velloso da. *Ética e inteligência artificial: da possibilidade filosófica de agentes morais artificiais*. Porto Alegre, RS: Editora Fi, 2021.

SILVEIRA, Pedro Telles da. História, técnica e novas mídias: reflexões sobre a história na era digital. Tese de doutorado. Universidade Federal do Rio Grande do Sul, Porto Alegre, 2018.

STRECK, Lenio Luiz. “Um robô pode julgar? Quem programa o robô?” in: Consultor Jurídico, 03 set 2020. Disponível em: <https://www.conjur.com.br/2020-set-03/senso-incomum-robo-julgar-quem-programa-robo?imprimir=1> – último acesso em 03 abr 2022.

STROKES, Dustin. “Imagination and Creativity” in KIND, Amy (ed.), The Routledge Handbook of the Philosophy of Imagination. Routledge, 2016. Disponível em: <https://philarchive.org/rec/STOIAC-4> - último acesso em 01 Jan 2023.

STROWEL, Alain. Quand le Google défie le droit : plaidoyer pour une Internet transparente et de qualité. Bruxelles : Éditions Larcier, 2011.

SUNDARARAJAN, Louise. “Harold Cohen and AARON: Collaborations in the Last Six Years (2010–2016) of a Creative Life” in: LEONARDO, Vol. 54, No. 4, pp. 412–417, 2021, p. 415. Disponível em: [http://direct.mit.edu/leon/article-pdf/54/4/412/1959027/leon\\_a\\_01906.pdf](http://direct.mit.edu/leon/article-pdf/54/4/412/1959027/leon_a_01906.pdf) – último acesso em 26 Dez 2022.

SUNSTEIN, Cass R. “Deciding by Default” in: University of Pennsylvania Law Review, 162, vol. 1, 2013, pp. 1-57. Disponível em: [https://scholarship.law.upenn.edu/penn\\_law\\_review/vol162/iss1/1](https://scholarship.law.upenn.edu/penn_law_review/vol162/iss1/1) - último acesso em 25 Jun 2023.

TOMÁS DE AQUINO. Summa Theologica.

TRIBE, Lawrence H. Trial by mathematics: precision and ritual in the legal process. In: Harvard Law Review, vol 84, no 6 (April 1971).

TURKLE, Sherry. The Second Self: computers and the human spirit. New York, NY: Simon & Schuster, 1984.

VALLOR, Shannon. Technology and the Virtues: a philosophical guide to a future Worth living. New York, NY: Oxford University Press, 2016.

WALLACH, Wendell & ALLEN, Colin. Moral Machines: teaching robots right from wrong. New York, NY: Oxford University Press, 2009.

WEBER, THADEU. Hegel e Rawls: um diálogo possível. Veritas (Porto Alegre), v. 67, p. 1-12, 2022.

Disponível em: <https://revistaseletronicas.pucrs.br/ojs/index.php/veritas/article/view/43762> - último acesso em 05 Ago 2023.

WEBER, THADEU; TERNUS, B. F. . A Crítica da Crítica de A. Sen à Teoria da Justiça de Rawls.. REVISTA BRASILEIRA DE ESTUDOS JURÍDICOS, v. 14, p. 42-56, 2019. Disponível em: <https://repositorio.pucrs.br/dspace/handle/10923/20570> - último acesso em 05 Ago 2023.

ZUBOFF, S. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. New York: Public Affairs, 2019

Filmes:

*Modern Times*. EUA: United Artists, 1936. Direção: Charles Chaplin (87 min.).

*The Social Dilemma*. Direção: Jeff Orlowski, Exposure Labs, 2020. (94min). Disponível em: [www.netflix.com/title/81254224](http://www.netflix.com/title/81254224) - último acesso em 12 Abr 2023.



Pontifícia Universidade Católica do Rio Grande do Sul  
Pró-Reitoria de Pesquisa e Pós-Graduação  
Av. Ipiranga, 6681 – Prédio 1 – Térreo  
Porto Alegre – RS – Brasil  
Fone: (51) 3320-3513  
E-mail: [propesq@pucrs.br](mailto:propesq@pucrs.br)  
Site: [www.pucrs.br](http://www.pucrs.br)