



# TEXTO PARA DISCUSSÃO

ISSN 0103-9466

453

**Apontamentos e digressões sobre o manifesto  
nooscópico:  
Inteligência Artificial e Complexidade**

**Eduardo Barros Mariutti**

**Julho 2023**



UNICAMP

**ie** Instituto de  
economia

# Apontamentos e digressões sobre o manifesto nooscópico: Inteligência Artificial e Complexidade

Eduardo Barros Mariutti \*

**Resumo** – O *manifesto nooscópico* foi publicado em 2020 por Matteo Pasquinelli e Vladan Joler. No texto os autores apresentam uma cartografia dos limites da Inteligência Artificial, um mapa-manifesto que visa “desafiar as mistificações sobre a inteligência artificial”, particularmente a percepção de que ela tende a se autonomizar e encarnar uma espécie de “mente alienígena”, isto é, totalmente estranha às formas humanas de percepção e cognição. Este artigo exploratório parte da problemática explicitada pelos autores, reforçando o imbricamento entre homens e máquinas que viabiliza a IA.

**Palavras-chave:** Inteligência artificial, Machine learning, Nooscópico.

## Abstract

### *Notes and digressions on the nooscope manifested: Artificial Intelligence and Complexity*

The nooscopic manifesto was published in 2020 by Matteo Pasquinelli and Vladan Joler. In the text, the authors present a cartography of the limits of Artificial Intelligence, a map-manifesto that aims to "challenge the mystifications of artificial intelligence", particularly the perception that it tends to autonomize itself and embody a kind of "alien mind", that is, totally foreign to human forms of perception and cognition. This exploratory article starts from the problematic stated by the authors, reinforcing the intertwining between men and machines that makes AI possible.

**Keywords:** Artificial intelligence, Machine learning, Nooscope.

**JEL code:** A19.

## Introdução: o que é o nooscópico?

De acordo com Pasquinelli e Joler, o nooscópico é um *diagrama* que explicita os agenciamentos ligados às aplicações da Inteligência Artificial. O termo foi criado para possibilitar “uma cartografia dos limites da inteligência artificial (IA), que se apresenta como uma provocação tanto para a ciência da computação quanto para as humanidades”. Estabelecendo uma analogia com as próteses óticas desenvolvidas na renascença, os autores visam combater as mistificações que gravitam em torno da IA: não estamos defronte “máquinas inteligentes” ou “mentes alienígenas” que não guardam nenhuma relação com a cognição humana, mas sim a um novo instrumento do conhecimento que, partindo de um conjunto de dados tão vasto e heterogêneo que está além do alcance da mente dos homens, identifica e constrói padrões de forma automática. Logo, tal como um telescópio, ele ajuda a enxergar e a navegar em um novo *espaço do conhecimento* que tende a assumir a forma de um espaço vetorial *multidimensional*. Mas o nooscópico faz mais do que navegar. Por ter sido gerado prioritariamente pelas pressões da concorrência militar e pelas *big techs*, ele expressa a mutação do projeto moderno de *mecanização da razão humana* que tomou forma no século XXI: a constituição de um novo “regime corporativo” calcado no *extrativismo do conhecimento* e no *colonialismo epistêmico*.

---

\* Professor Associado do Instituto de Economia da Unicamp e do Programa de Pós-Graduação *San Tiago Dantas*. Pesquisador do INCT/Ineu e membro da Rede de Pesquisa em Autonomia Estratégica, Tecnologia e Defesa (PAET&D). E-mail: [mariutti@unicamp.br](mailto:mariutti@unicamp.br). ORCID: [0000-0002-3674-3194](https://orcid.org/0000-0002-3674-3194).

A primeira coisa que se nota ao ler o *manifesto nooscópico*<sup>1</sup> é uma certa flutuação terminológica. Embora os autores afirmem que o *machine learning* é seu principal elemento, este termo é utilizado muitas vezes como um sinônimo de IA e até mesmo de *deep learning* que, como se sabe, são coisas diferentes. A confusão parece ser proposital, pois reflete uma recusa da arraigada percepção de que a IA é uma forma “alienígena” de inteligência (no sentido de ser radicalmente diferente da humana) e uma entidade divorciada da sociedade. Pelo contrário. Todos os processos ligados a este termo são *maquímicos*, isto é, são baseados em uma tensa  *fusão* entre homens, máquinas e demais objetos técnicos. De acordo com os autores, o nooscópio é a atual manifestação do processo de *automação da percepção* posto em marcha pelo menos desde a segunda metade do século XX.

Além desta introdução, o artigo está subdividido em 5 partes. A primeira envolve uma discussão sobre o modo como o nooscópio amplia o alcance do conhecimento, acelera as transformações na concepção de *espaço* e, desse modo, *intensifica* a crise do perspectivismo cartesiano, o regime escópico que preponderou na renascença e ajudou a modelar a modernidade. A ideia de um espaço homogêneo e isotrópico *extrínseco* aos objetos nele contidos tende a ser dissolvida com a generalizações de aplicações práticas derivadas de um espaço vetorial multidimensional, isto é, uma *topologia* em que o espaço é definido *ex post*, como um jogo de *possibilidades* definido pelas distâncias matemáticas entre os objetos e suas propriedades. A segunda parte retoma e desenvolve a analogia estabelecida por Leibniz entre o *calculus ratiocinator* e os instrumentos óticos de ampliação da visão. Tal como as próteses da visão que ganharam proeminência na renascença, o nooscópio deve ser visto predominantemente como um *instrumento de ampliação do conhecimento* que, contudo, também está sujeito a distorções. A terceira parte tematiza o modo como as correlações produzidas pelos agenciamentos maquímicos favorecem a consolidação de uma nova racionalidade, uma racionalidade *difratada* que, ao contrário da episteme centrada na causalidade, tem como fundamento básico a detecção e a criação automatizada de padrões em um universo gigantesco de informações, que está muito além do alcance da mente humana. A seção quatro tem como eixo a discussão sobre a visão maquímica e a inteligência adversarial, isto é, as possibilidades de *sabotar* o processo de normalização das condutas em curso atualmente. Esta temática prepara o caminho para a próxima seção, que aborda a dificuldade dos sistemas preditivos automatizados de lidar com o *novo* e com as anomalias com que se deparam. A última parte – as considerações finais – não deve ser encarada como uma conclusão, mas como a explicitação de uma tensão que é recorrente em todo o texto: de um lado o ímpeto de tentar aprisionar as virtualidades da IA em um regime disciplinar centrado no extrativismo do conhecimento e, de outro, as resistências e, sobretudo, os novos horizontes e as possibilidades inéditas que ela pode viabilizar.

### **Transformando o terreno do conhecimento: o espaço vetorial multidimensional.**

Analogias sempre são um pouco traiçoeiras. Mas a estabelecida entre o nooscópio e as próteses da visão que começaram a se proliferar na renascença é particularmente frutífera. Por exigir técnicas geométricas apuradas que possibilitam representar cenas tridimensionais em uma superfície bidimensional, o desenvolvimento da perspectiva linear deu um grande impulso à matematização do campo visual e, sobretudo, a sua conversão em uma dimensão isotrópica (Ivins, 1938, p. 9; 12). Este

---

(1) Estou seguindo o modo como o termo tem sido majoritariamente traduzido para o português. Mas o título escolhido pelos autores é *The Nooscope Manifested*, que passa a ideia de *manifestações do nooscópio*, muito mais próxima do teor do texto.

impulso foi suplementado pelo desenvolvimento do *perspicillum* – rebatizado depois como telescópio – por Galileu que, como se sabe, afirmou que o instrumento que criou não deveria ser visto apenas como uma extensão da visão, mas sobretudo como uma espécie de *prótese da razão* capaz de corrigir as falhas do olho humano. A combinação entre os estímulos provenientes da arte e da proliferação de novos instrumentos óticos acelerou uma modificação na concepção do espaço, que passou a ser entendido como um campo geometricamente isotrópico, abstrato, no qual os objetos podem ser localizados de acordo com suas *relações exteriores* a partir de um ponto focal. Esse *regime escópico* foi denominado posteriormente como *perspectivismo cartesiano* (Jay, 2020, p. 333; Bousquet, 2018, p. 17).

Na medida em que se baseia na matematização e na racionalização do campo visual, o nooscópio pode ser visto como uma *atualização* desta herança, mas que a transforma significativamente. O espaço passa a ser visualizado em múltiplas dimensões e os objetos nele contidos – e suas relações – são representados por *vetores*. Logo, a ideia de um campo visual formado a partir de um único ponto de vista – um olhar focado – que era central no regime escópico anterior é abandonada. A clara separação entre o *espaço* – uma dimensão extrínseca - e os objetos nele contidos começa também a ficar cada vez mais nuançada. O desenvolvimento da topologia e da teoria dos grafos favoreceu o surgimento de aplicações práticas ligadas a um espaço n-dimensional. No manifesto, a questão é abordada da seguinte forma:

*Multi-dimensional vector space is another reason why the logic of machine learning is difficult to grasp. Vector space is another new cultural technique, worth becoming familiar with. The field of Digital Humanities, in particular, has been covering the technique of vectorialisation through which our collective knowledge is invisibly rendered and processed. William Gibson's original definition of cyberspace prophesized, most likely, the coming of a vector space rather than virtual reality: 'A graphic representation of data abstracted from the banks of every computer in the human system. Unthinkable complexity. Lines of light ranged in the nonspace of the mind, clusters and constellations of data. Like city lights, receding (Pasquinelli; Joler, 2020, p. 12).*

A questão decisiva não é a *virtualidade*, mas a vetorialização e a constituição de um espaço apreendido predominantemente pela topologia. As formas automatizadas de percepção – redes neurais, *deep learning* etc. – operam mediante representações vetoriais puramente digitais do mundo, que rompem com o *perspectivismo cartesiano*.

Este campo se desenvolveu fortemente associado ao problema do reconhecimento de imagens no âmbito da *visão computacional*. Porém, com a formalização crescente da topologia, foi possível extravasar as aplicações para outras tarefas. Um dos casos mais conspícuos já faz parte da experiência cotidiana de qualquer pessoa que possua um smartphone: o famigerado corretor (preditor) automático. *Large Language Models* como o chat GPT são outro exemplo. Sistemas de reconhecimento e geração de texto operam encapsulando as palavras em um espaço vetorial multidimensional de acordo com a sua *vizinhança*. Rei e Rainha, por exemplo, têm a mesma distância de homem e mulher em uma das dimensões. Casa e Moradia são próximas. É possível montar sentenças e prever o que será digitado a partir de similaridades derivadas desta topologia. Por fim, ao longo do tempo, as interações com os humanos e a capacidade de analisar quantidades gigantescas de textos aumenta a eficiência destes modelos.

Em um instigante artigo, Nuria Rodríguez-Ortega destaca como essas operações produzem transformações ontológicas:

*The current relevance of the concept of n-dimensional space for cultural analysis and interpretation must be related, in the first place, to the «ontological» transformation that cultural objects have been undergoing for decades as a result of their digitization or direct digital production: cultural objects, regardless of their nature (images, words, sounds, etc.), in their digital mode of existence are essentially matrices or sets of numerical data. It is precisely this transformation that makes their computation possible. From a computational point of view, images or texts are nothing more than a spatial surface of numerical information from which it is possible to extract (also) numerical characteristics using computational systems. Therefore, this ontological transformation also implies an epistemological transformation insofar as cultural objects that are transformed into digital forms become a problem of a computational and mathematical order (2022, p. 2).*

Esse *modo digital de existência* se manifesta em um espaço radicalmente distinto daquele privilegiado pelo perspectivismo cartesiano. Parte de sua peculiaridade reside no fato deste espaço não possuir nenhum correlato empírico, pois ele consiste em uma “generalização intelectual” derivada da possibilidade lógica de existência de objetos multidimensionais, uma *formalização* de diferentes *tipos de geometria e de suas formas espaciais* em uma *base comum*: um modelo matemático que pode ser manipulado digitalmente. Trata-se, portanto, de um conceito matemático e abstrato de espaço, configurado por múltiplas dimensões (cada coordenada é uma dimensão) e que possibilita encontrar comensurabilidades entre objetos multidimensionais distintos. A diferenciação euclidiana entre o espaço e os elementos nele contidos é erodida: signo e espaço passam a constituir uma unidade indissolúvel (cf. Rodríguez-Ortega, 2022, p. 9-11).

Esta forma de se conceber o espaço e a relação entre os seus elementos é extremamente contraintuitiva:

*As I pointed out earlier, making a mental image of a high-dimensional space is very complicated since it is practically impossible to process using our human brain, which is used to thinking – at least until recently – in four-dimensional parameters. The counterintuitive character of these vector spaces lies not only in the number of dimensions but also in some of their properties and in the algorithmic rationale that operates on and produces them, which are sometimes unintelligible to our forms of understanding (Rodríguez-Ortega, 2022, p. 7).*

Só podemos acessar esse espaço por meio do nooscópio e, pelo menos por enquanto, ela permanece pouco familiar. Novas percepções geralmente demoram muito tempo para serem assimiladas. Isto seguramente colabora para a equivocada percepção da IA como algo dotado de uma aura mágica, alheia a qualquer experiência ou forma humana de compreensão.

### **Gottfried Wilhelm Leibniz: *característica universalis e calculus ratiocinator***

A despeito das desavenças e da rivalidade entre seus patronos, a publicação de *Cybernetics* (1948) e o prestigiado *Dartmouth Summer Research Project on Artificial Intelligence* (1956) são marcos significativos na constituição da Inteligência Artificial como um ramo do pensamento

científico e tecnológico<sup>2</sup>. Norbert Wiener assinala a importância de Leibniz na constituição da cibernética:

*If I were to choose a patron saint for cybernetics out of the history of science, I should have to choose Leibniz. The philosophy of Leibniz centers about two closely related concepts - that of a universal symbolism and that of a calculus of reasoning. From these are descended the mathematical notation and the symbolic logic of the present day. Now, just as the calculus of arithmetic lends itself to a mechanization progressing through the abacus and the desk computing machine to the ultra-rapid computing machines of the present day, so the calculus ratiocinator of Leibniz contains the germs of the machina ratiocinatrix, the reasoning machine. Indeed, Leibniz himself, like his predecessor Pascal, was interested in the construction of computing machines in the metal. It is therefore not in the least surprising that the same intellectual impulse which has led to the development of mathematical logic has at the same time led to the ideal or actual mechanization of processes of thought (1985, p. 12).*

Desde os seus primórdios a cibernética aspirava *automatizar* e controlar os processos de produção, coleta e distribuição das informações por meio de sistemas de *feedback*. Um projeto muito próximo às aspirações de Leibniz, particularmente seu ambicioso projeto de criar uma *ciência geral* capaz de unificar todas as diversas áreas do conhecimento e, com isso, resolver definitivamente as disputas filosóficas em curso.

Esta “ciência geral” leibniziana teria como fundamento uma lógica matemática – uma “álgebra do pensamento” – dividida em dois esforços que se amparavam mutuamente: a *characteristica universalis* (ou *lingua characteristic*) e o *calculus ratiocinator* (Hintikka, 1977, p. ix). Esta *démarche* se fundamenta na ideia de que o universo é uma espécie de *máquina lógica* regida por princípios universais e, também, com opacia afirmação de Leibniz de que a complexidade do pensamento humano tinha como base um número limitado de “pensamentos primitivos” irredutíveis (Peckhaus, 2004, p. 599). Em conformidade com esta percepção do universo e da natureza do pensamento, ele aventou a possibilidade de construir uma linguagem universal (a *characteristica universalis*) onde todo pensamento primário seria representado por um símbolo, em clara analogia com o modo como a matemática introduz letras e símbolos para designar objetos matemáticos (Breger, 2005, p. 487). A *characteristica universalis* expressaria uma espécie de *álgebra do pensamento*, isto é, a formalização de um código capaz de indexar todas as variantes possíveis de pensamento. Por sua vez, o *calculus ratiocinator* era apresentado por Leibniz como um sistema capaz de expressar todos os pensamentos possíveis em uma *estrutura lógica*, permitindo a realização de deduções e inferências de forma automática. Precisamente por conta disto o *calculus ratiocinator* poderia ser instanciado em uma *máquina*, uma máquina calculadora capaz de identificar e manipular todas as conexões possíveis entre os conceitos elementares, resolvendo de forma mecânica *qualquer* problema.

---

(2) John McCharty foi o principal organizador do *Dartmouth Summer Research Project on AI*. Ele odiava Norbert Wiener, especialmente depois que ele publicou *The Human Use of Human Beings* (1950), um livro considerado muito ácido e antagonista ao interesse de comercializar as novas tecnologias de processamento de informações que começavam a emergir. McCharty se negou a utilizar a expressão cibernética, preferindo cunhar a expressão Inteligência Artificial, termo que rapidamente se ligou ao posterior desenvolvimento dos computadores digitais (Brockman, 2019, p. 15).

O *manifesto nooscópico* tem como uma de suas referências principais o *calculus ratiocinator* e a analogia estabelecida pelo próprio Leibniz entre a sua eventual concretização e o papel que as próteses óticas – microscópio e telescópio – desempenharam na ampliação do conhecimento:

*Once the characteristic numbers are established for most concepts, mankind will then possess a new instrument which will enhance the capabilities of the mind to a far greater extent than optical instruments strengthen the eyes, and will supersede the microscope and telescope to the same extent that reason is superior to eyesight* (Leibniz apud Pasquinelli; Joler, 2020, p. 2)

A analogia é bastante ilustrativa. O telescópio e o microscópio permitiram que o pensamento penetrasse em zonas da realidade que estavam fora do escrutínio da humanidade pelo fortalecimento da *visão*. Por conta disto, o pensamento humano foi capaz de acessar novas dimensões. O desenvolvimento de uma “álgebra do pensamento” automatizada traria efeitos ainda mais revolucionários, pois ela incrementaria a *razão* que, para Leibniz, é superior à vista humana. Os evangelistas da IA prometem uma *ampliação* e indexação do conhecimento ainda maior do que a imaginada por Leibniz.

### ***Racionalidade Difratada e Racionalização das condutas***

Todo instrumento de ampliação da percepção produz algum tipo de distorção. O telescópio, por exemplo, dentre diversas outras anomalias óticas, é suscetível à *aberração cromática*, isto é, uma variação na palheta de cores derivada da curvatura da lente e de seus materiais que distorce a imagem. As cores possuem distintos comprimentos de onda tendem a se desviar com distintas variações quando atravessam as lentes. O resultado é uma imagem ligeiramente diferente do que seria esperado em um sistema ótico ideal. Essas micro variações podem gerar halos coloridos ou franjas no contorno dos objetos avistados, comprometendo a sua nitidez. Mesmo usando filtros e correções eletrônicas é sempre esperada alguma perturbação na imagem. Esta é uma das características que levaram a Pasquinelli e Joler recorrer à analogia entre o nooscópio e as próteses óticas:

*Instruments of measurement and perception always come with inbuilt aberrations. In the same way that the lenses of microscopes and telescopes are never perfectly curvilinear and smooth, the logical lenses of machine learning embody faults and biases. To understand machine learning and register its impact on society is to study the degree by which social data are diffracted and distorted by these lenses. This is generally known as the debate on bias in AI, but the political implications of the logical form of machine learning are deeper. Machine learning is not bringing a new dark age but one of diffracted rationality, in which, as it will be shown, an episteme of causation is replaced by one of automated correlations. More in general, AI is a new regime of truth, scientific proof, social normativity and rationality, which often does take the shape of a statistical hallucination* (Pasquinelli; Joler, 2020, p. 2).

Essa passagem é decisiva, pois ela sintetiza elementos importantes da discussão e remete a outros momentos do manifesto nooscópico. Por conta disto, é necessário desagregar e desenvolver as ideias que a fundamentam.

Como já foi apontado, o *machine learning* é o princípio elementar do nooscópio e, como todo instrumento perceptivo, ele não somente amplia, mas também *distorce* o conhecimento e as representações da realidade que produz. Para entender melhor o efeito das suas “lentes lógicas”, é importante analisar o seu *diagrama*, um agenciamento que combina três elementos. i) um *objeto* a ser

observado (os dados de treinamento); ii) um instrumento de observação (o algoritmo de aprendizado) e iii) uma representação final, que assume a forma de um *modelo estatístico*. Reiterando a analogia com nos instrumentos óticos, Pasquinelli e Joler afirmam que o fluxo de informação que perpassa o nooscópio pode ser visto como um feixe de luz que é projetado pelos dados de treinamento, *comprimido* pelo algoritmo e difratado pelas “lentes” do modelo estatístico. Cada uma destas etapas produz anomalias peculiares que precisam ser analisadas para que se compreenda como o nooscópio funciona e, simultaneamente, como ele falha. Este duplo jogo, sustentam os autores, levanta sérias dúvidas sobre o entendimento da Inteligência Artificial e seus agenciamentos como um *paradigma monolítico de racionalidade*, uma “mente alienígena” indecifrável. Parece mais correto pensá-la como uma “arquitetura espúria” que amplia os vieses humanos e opera predominantemente por meio de truques (tricks) e improvisações (Pasquinelli; Joler, 2020, p. 4), de forma análoga ao desenvolvimento das máquinas a vapor, que precedeu a descoberta das leis que explicam e controlam o seu funcionamento interno.

Como destacam os autores, existem redes neurais dedicadas ao reconhecimento de imagens, por exemplo, que funcionam relativamente bem, mas não há nenhuma *teoria do aprendizado* (*theory of learning*) capaz de explicar como elas acertam e, também, como elas erram tanto. Em parte, isto ajuda a entender por que Pasquinelli e Joler salientam que, ao invés de uma nova era das trevas, estamos testemunhando a manifestação de uma nova forma de racionalidade, uma “racionalidade difratada (*diffracted rationality*), na qual (...) a episteme baseada na causalidade é substituída por outra, constituída por correlações automatizadas”. O nooscópio é, essencialmente, um agenciamento capaz de detectar *padrões* em um conjunto de dados gigantesco e heterogêneos em uma velocidade e escala que está muito além do alcance da mente humana. E é precisamente neste ponto que reside a sua força e, também, a sua fraqueza. Qualquer pessoa minimamente educada sabe que correlação não implica causalidade. A grande dificuldade envolve separar as correlações efetivas – que reflete alguma causalidade ou jogo de causalidades subjacentes – das arbitrárias, que são tomadas como reais exatamente porque não é possível explicar como os sistemas “inteligentes” chegam aos seus resultados. Neste caso, uma correlação equivocada se converte em uma profecia autorrealizável que se tornará “verdade” por conta da nossa confiança nestes sistemas. Quanto mais generalizados, maior a probabilidade de construção de “verdades” que, no fim das contas, nunca passaram e *alucinações estatísticas*. Profecias desta natureza podem prender e até mesmo eliminar pessoas, dando vazão a um aparato biopolítico baseado na *preempção* (Pasquinelli; Joler, 2020, p. 16).

Neste ponto em particular, Pasquinelli e Joler se baseiam em um artigo de Dan McQuillan publicado em 2018. As duas passagens transcritas abaixo ilustram o argumento básico.

*The mode of operation of machine learning is to “learn”; that is, when supplied with a large amount of input data and a corresponding set of targets it finds a function to map the features of the input data to the desired target outputs. Machine learning finds reproducible patterns in the data. Moreover, these patterns have predictive power in that they can predict the target value for new and unknown input. Thus, machine learning is a form of numerical pattern finding with predictive power, prompting comparisons with science. But rather than being universal and objective, it produces knowledge that is irrevocably entangled with specific computational mechanisms and the data used for the training* (2018, p. 1).

O aprendizado de máquina detecta padrões reproduzíveis nos dados que o abasteceram, seja por interpolação (definição de novos padrões dentro do intervalo de dados) ou por *extrapolação*. Mas o aspecto decisivo é a capacidade de *prever* novas entradas:

*The predictive nature of machine learning promotes preemption, that is, action that attempts to anticipate or prevent the predicted outcome. In a world overflowing with data, machine learning and deep learning are powerfully attractive. For any context where a lot is riding on uncertain outcomes, the lure of being able to peer through the fog of data and read off the probable future through computational pattern analysis comes to be almost irresistible. At the current time, it would be hard to think of an area of life that is not under consideration for treatment by algorithmic methods (2018, p. 3).*

O fato é que a quantidade e a variedade de dados produzidos na vida social estão aumentando de forma extremamente acelerada, favorecendo com isso a penetração dos métodos preditivos baseados em algoritmos “aprendizes” em todas as dimensões da realidade social.

Para melhor compreender o que está em jogo aqui, contudo, é necessário destacar uma indagação presente no manifesto nooscópico e que também se insinua no artigo supracitado: o que é um padrão? Ou, de forma alternativa, padrões são descobertos ou *criados*? Isso é importante pois a Inteligência Artificial “não é um autômato pensante, mas um algoritmo que performa o reconhecimento de padrões” em praticamente todas as formas e tipos de dados que podem ser detectados.

*Today, in the case of self-driving cars, the patterns that need to be recognised are objects in road scenarios. In the case of automatic translation, the patterns that need to be recognised are the most common sequences of words across bilingual texts. Regardless of their complexity, from the numerical perspective of machine learning, notions such as image, movement, form, style, and ethical decision can all be described as statistical distributions of pattern. In this sense, pattern recognition has truly become a new cultural technique that is used in various fields. For explanatory purposes, the Nooscope is described as a machine that operates on three modalities: training, classification, and prediction. In more intuitive terms, these modalities can be called: pattern extraction, pattern recognition, and pattern generation (Pasquinelli; Joler, 2020, p. 7).*

A qualidade (e o tamanho) dos *datasets* de treinamento são elementos decisivos para a performance do algoritmo. Os dados de treinamento nunca são brutos: há sempre um grau de *construção* envolvido na geração dos padrões. Um conjunto de dados de treinamento envolve uma intensa participação humana que, geralmente combina profissionais altamente qualificados com trabalhos precários (*ghost workers*), que simplesmente rotulam as imagens ganhando muito pouco por uma tarefa estressante.

Embora existam diversos sistemas de terceirização de trabalho para classificar imagens, o mais famoso é o *Amazon Mechanical Turk*, que funciona como uma espécie de marketplace (a Amazon é apenas uma intermediária entre o ofertante do serviço e os “turkers”) no qual são oferecidas diversas micro tarefas de difícil automação, conhecidas como *Human Intelligence Tasks* (Hits). Além da classificação de imagens para formar datasets, é bastante comum a oferta de trabalho para treinar os assistentes de voz, jogar videogames e formar datasets para os sistemas de moderação de conteúdo. Neste caso, os trabalhadores, que ganham poucos centavos de dólares para rotular centenas de imagens por minuto, são expostos a cenas agressivas – gore, pornografia infantil etc. – para treinar os algoritmos. Enquanto os engenheiros e cientistas da computação aparecem com pompa na mídia, os

*ghost workers* permanecem anônimos, embora o seu trabalho seja vital não só para treinar, mas também para corrigir as eventuais falhas dos sistemas automatizados:

*The human labor powering many mobile phone apps, websites, and artificial intelligence systems can be hard to see – in fact, it’s often intentionally hidden. We call this opaque world of employment ghost work. Think about the last time you searched for something on the web. Maybe you were looking for a trending news topic, an update on your favorite team, or fresh celebrity gossip. Ever wonder why the images and links that the search engine returned didn’t contain adult content or completely random results? After all, every business, illicit or legitimate, advertising online would love to have its site ranked higher in your web search. Or think about the last time you scrolled through your Facebook, Instagram, or Twitter feed. How do those sites enforce their no-graphic-violence and no-hate-speech policies? On the internet, anyone can say anything, and, given the chance, people certainly will. So how do we get such a sanitized view? The answer is people and software working together to deliver seemingly automated services to customers like you and me (Gray; Suri, 2019, p. 7).*

O fato é que, pelo menos em seu estágio atual, a “inteligência artificial” não consegue funcionar sem a participação massiva do trabalho humano. Quando os sistemas automatizados encontram dificuldades – um pedido de pizza por aplicativo mais complicado, em que o cliente muda muito os ingredientes, por exemplo – muitas empresas recorrem a trabalhadores fantasmas que precisam estar disponíveis a qualquer momento para “concluir silenciosamente a tarefa”.

Como sugerem os autores, isso configura uma espécie de linha de montagem digital:

*This new digital assembly line aggregates the collective input of distributed workers, ships pieces of projects rather than products, and operates across a host of economic sectors at all times of the day and night. In fact, the rise of this shadow workforce is part of a larger, more profound reorganization of employment itself. This yet-to-be-classified form of employment done on demand is neither inherently good nor bad. But left without definition and veiled from consumers who benefit from it, these jobs can easily slip into ghost work (Gray, Suri, 2019, p. 7).*

Os homens são requisitados para resolver tarefas muito simples que, contudo, confundem as máquinas. No caso, a intervenção humana é pontual e baseada no *common sense*, no sentido que Hume atribui ao termo. Essa habilidade, contudo, exatamente por ser praticamente ubíqua, é extremamente mal remunerada. Dentre vários *insights* poderosos, Mary Gray e Siddhart Suri destacam o grande paradoxo da automação: as tentativas de eliminar o trabalho humano acabam gerando novas tarefas que só os humanos conseguem desempenhar. O problema é que, na maioria dos casos, essas necessidades são contingentes e momentâneas e, portanto, são incompatíveis com a concepção convencional de jornada de trabalho.

Frente ao que foi exposto podemos tirar pelo menos duas conclusões provisórias. A primeira é que, como sustentam Pasquinelli e Joler, a inteligência artificial não é uma *mente alienígena*, pois seus agenciamentos envolvem ações humanas tanto para gerar os dados – afinal, eles são um produto da vida social – quanto para estruturar os *datasets*. A segunda, ligada à primeira, é que a classificação dos dados performada por humanos é uma das fontes do *viés* que os algoritmos ampliam e explicitam. No manifesto o primeiro tipo de *viés* é classificado como “histórico”, pois já era aparente na sociedade antes das transformações tecnológicas aqui discutidas. O segundo (“*dataset bias*”) é introduzido pelos homens na fase de estruturação dos dados de treinamento do algoritmo. Nestes dois casos, a peso do fator humano é saliente. Já o *viés* do algoritmo é um pouco diferente:

*Algorithmic bias (also known as machine bias, statistical bias or model bias, to which the Nooscope diagram gives particular attention) is the further amplification of historical bias and dataset bias by machine learning algorithms. The problem of bias has mostly originated from the fact that machine learning algorithms are among the most efficient for information compression, which engenders issues of information resolution, diffraction and loss. Since ancient times, algorithms have been procedures of an economic nature, designed to achieve a result in the shortest number of steps consuming the least amount of resources: space, time, energy and labour. The arms race of AI companies is, still today, concerned with finding the simplest and fastest algorithms with which to capitalise data. If information compression produces the maximum rate of profit in corporate AI, from the societal point of view, it produces discrimination and the loss of cultural diversity (Pasquinelli; Joler, 2020, p. 4).*

O nooscópio classifica, estabelece correlações e *comprime* os dados para poder produzir representações. Nesta compressão os vieses e as discriminações tendem a ser ampliados ao mesmo tempo que muita informação é perdida ou deliberadamente descartada.

O tema da redução da diversidade cultural se entrelaça com a questão da criação de padrões. Figuras e imagens mais populares – as “celebridades”, os cenários icônicos, imagens que se tornam virais etc. – vão se tornando os padrões que norteiam os sistemas automatizados de identificação, organização, rotulagem e difusão das imagens e informações. Quem – ou o quê – está muito fora destes padrões tende a ficar cada vez menos visível, pois seu rastro digital e suas imagens não são consideradas relevantes para parametrizar os datasets. Este processo infla a exposição do que já era visível e tende a condenar ao desaparecimento o que é menos visível.

Por fim, resta discutir a conexão entre o terceiro componente do nooscópio - o modelo estatístico – e a formação dos padrões:

*The business of these neural networks, however, was to calculate a statistical inference. What a neural network computes, is not an exact pattern but the statistical distribution of a pattern. Just scraping the surface of the anthropomorphic marketing of AI, one finds another technical and cultural object that needs examination: the statistical model. What is the statistical model in machine learning? How is it calculated? What is the relationship between a statistical model and human cognition? These are crucial issues to clarify (Pasquinelli; Joler, 2020, p. 7).*

O modelo estatístico, embora não possa ser dissociado do algoritmo,<sup>3</sup> pode ser visto como a etapa final do processo de classificação e formação de padrões desempenhado pelo nooscópio.

Quem tem alguma familiaridade com a obra de Foucault já deve ter notado que a biopolítica, dentre várias outras características, expressa a influência dos modelos estatísticos sobre a cultura e a política. Logo, como destacam Pasquinelli e Joler, o peso destes modelos não emerge com o *machine learning*, pois o uso da estatística para controle da população remonta pelo menos ao século XVIII. Devemos conceber o *machine learning* como uma forma de *automatizar* as modelagens estatísticas

---

(3) “The product of such machine processes is a statistical model (more accurately termed an ‘algorithmic statistical model’). In the developer community, the term ‘algorithm’ is increasingly replaced with ‘model.’ This terminological confusion arises from the fact that the statistical model does not exist separately from the algorithm: somehow, the statistical model exists inside the algorithm under the form of distributed memory across its parameters. For the same reason, it is essentially impossible to visualise an algorithmic statistical model, as is done with simple mathematical functions. Still, the challenge is worthwhile” (Pasquinelli; Joler, 2020, p. 7).

que, contudo, opera com uma forma peculiar de opacidade. Após contrastar esta técnica com as modelagens produzidas pelas ciências do clima (alegadamente menos opacas), os autores afirmam:

*Machine learning models, on the contrary, are opaque and inaccessible to community debate. Given the degree of myth-making and social bias around its mathematical constructs, AI has indeed inaugurated the age of statistical science fiction. Nooscope is the projector of this large statistical cinema* (Pasquinelli; Joler, 2020, p. 8).

A ideia é interessante, mas um pouco exagerada. Boa parte dos modelos climáticos são extremamente opacos, muitos deles tão opacos quanto a geração de textos no chat GPT. O fato de a comunidade científica compartilhar os modelos e os dados que utiliza – algo que os algoritmos proprietários não fazem por sigilo comercial – não reduz a sua opacidade das previsões. Não faria sentido tentar provar a acurácia do modelo refazendo manualmente os cálculos que geram as previsões. No entanto, creio que essa não seja a questão mais importante. Nenhum modelo “resolve” questões do debate público, pois para cada modelo, geralmente há um contra-modelo que espelha os interesses de seus financiadores. Algum grau de opacidade<sup>4</sup> é um elemento incontornável de qualquer processo de automação da percepção e, inclusive, de qualquer forma de vida complexa. Creio que não há mais lugar para o mito de uma sociedade transparente.

A despeito destes floreios retóricos, os autores do manifesto são certos em um aspecto: as implicações da captura da Inteligência Artificial pelas grandes corporações e seu imbricamento com o aparato securitário dos Estados. Essa é uma questão central, pois ela afeta significativamente a relação entre a classificação e a geração de padrões performada pelas aplicações do aprendizado de máquina. No primeiro caso, trata-se de *classificar* de forma automatizada um alvo (um signo, um objeto, um rosto etc.) respeitando os parâmetros incorporados ao modelo estatístico pelos dados de treinamento. As taxonomias ou convenções sociais são indexadas a certas distribuições estatísticas. Cada objeto que se enquadrar nesta distribuição recebe um rótulo que o identifica ou regista alguma das suas propriedades. Já a *geração de padrões amplia* as informações com base em uma amostra ou um pedaço da informação. Pasquinelli e Joler incluem nesta categorização dois procedimentos distintos que, contudo, possuem um princípio bastante similar. Quando se restaura um arquivo de áudio corrompido, por exemplo, o que se faz é *completar* as informações faltantes seguindo a distribuição estatística do modelo. O segundo procedimento, cada vez mais utilizado pelas forças armadas e policiais, é a dimensão preditiva: com base na trajetória e nos comportamentos registrados no passado, os algoritmos projetam *tendências futuras*. Tanto a fase de classificação dos dados quanto

---

(4) No manifesto os autores fazem uma referência às reservas de Judea Pearl – um dos pioneiros das redes bayesianas aplicadas à IA – com relação à opacidade do *machine learning*. O que ele faz é explicitar os *limites teóricos* das inferências estatísticas dos modelos de aprendizado de máquina para que possamos utilizá-los com mais eficácia. “*I view machine learning as a tool to get us from data to probabilities. But then we still have to make two extra steps to go from probabilities into real understanding—two big steps. One is to predict the effect of actions, and the second is counterfactual imagination. We cannot claim to understand reality unless we make the last two steps*” (Pearl, 2019, p. 33). Essa passagem dos dados para as probabilidades é opaca e ele não acha que isso seja um problema. O que se deve ser questionado é tomar a IA como um oráculo ou como a via para uma inteligência capaz de *suplantar* a humana. “*Model-blind approaches impose intrinsic limitations on the cognitive tasks that Strong AI can perform. My general conclusion is that human-level AI cannot emerge solely from model-blind learning machines; it requires the symbiotic collaboration of data and models. (...) Data science is a science only to the extent that it facilitates the interpretation of data – a two-body problem, connecting data to reality*” (2019, p. 34).

o processo de geração de padrões são sobredeterminados pela *finalidade* a que foram desenhados, o que já afasta de antemão qualquer ideia de neutralidade.

Logo, quando capturada pela rivalidade interestatal e pela concorrência entre as *big techs*, a inteligência artificial e suas aplicações tende a acentuar o processo de *normalização* de condutas:

The normative power of AI in the 21<sup>st</sup> century has to be scrutinised in these epistemic terms: what does it mean to frame collective knowledge as patterns, and what does it mean to draw vector spaces and statistical distributions of social behaviours? According to Foucault, in early modern France, statistical power was already used to measure social norms, discriminating between normal and abnormal behaviour. AI easily extends the ‘power of normalisation’ of modern institutions, among others bureaucracy, medicine and statistics (originally, the numerical knowledge possessed by the state about its population) that passes now into the hands of AI corporations. The institutional norm has become a computational one: the classification of the subject, of bodies and behaviours, seems no longer to be an affair for public registers, but instead for algorithms and datacenters (Pasquinelli; Joler, 2020, p. 15).

Esse processo deliberado de normalização das condutas é faz com que o a opacidade (*black box*) assuma uma feição legitimadora.

*While the social consequences of AI are popularly understood under the issue of bias, the common understanding of technical limitations is known as the black box problem. The black box effect is an actual issue of deep neural networks (which filter information so much that their chain of reasoning cannot be reversed) but has become a generic pretext for the opinion that AI systems are not just inscrutable and opaque, but even ‘alien’ and out of control. The black box effect is part of the nature of any experimental machine at the early stage of development (it has already been noticed that the functioning of the steam engine remained a mystery for some time, even after having been successfully tested). The actual problem is the black box rhetoric, which is closely tied to conspiracy theory sentiments in which AI is an occult power that cannot be studied, known, or politically controlled (Pasquinelli; Joler, 2020, p. 4).*

Toda forma de vida social complexa baseia-se em uma distribuição do conhecimento que nos obriga a operar com máquinas e dispositivos que não podemos dominar. As redes neurais radicalizam ainda mais esta tendência, mas me parece cedo demais para dizer se o seu advento marca alguma ruptura radical. À exceção dos especialistas, quando usamos algum dispositivo complexo o que geralmente compreendemos são os *princípios* que governam a sua operação. Seria simplesmente impossível dominarmos completamente *todos* os equipamentos que usamos. É verdade que, de um ponto de vista mais operacional, todos os agenciamentos ligados à IA são necessariamente opacos. E isto não decorre *apenas* do caráter experimental destas técnicas. Se pudéssemos compreender as cadeias causais operadas pelo nooscópio ele não seria necessário. O problema é tratá-lo como um oráculo. Neste aspecto, a crítica ao efeito *black box* é certa. Contudo, a dimensão *retórica* deste efeito precisa ser dissociada do aspecto operacional. A IA não é uma “mente alienígena” autônoma que pode operar de forma dissociada dos homens. Mas a sua existência cria uma zona de opacidade suplementar que *nunca* poderemos dominar completamente.

### ***Visão Maquinica e Inteligência Adversarial***

A ideia de que a IA não totalmente estranha à cognição humana, é recorrente no manifesto. Contudo, há uma dimensão da percepção e da inteligência das máquinas que lhes garante uma certa

peculiaridade, uma *diferença* que possibilita a manifestação do nooscópio como um instrumento do conhecimento e de intervenção social. As máquinas captam estímulos que são imperceptíveis ao sensorio humano de forma quase instantânea e, com isto, criam padrões inusitados. Outro aspecto importante é que, ao contrário dos usuários humanos, a percepção que as máquinas formam pelos sensores e outras formas de captação de dados não necessita de uma saída videográfica ou de qualquer outra forma de projeção de imagens (Virilio, 1994, p. 60). O fato do “imaginário” criado pelas máquinas e para as máquinas não precisar retornar à dimensão do visível<sup>5</sup> ajuda a consolidar o mito de que a IA é autônoma.

Tendo isto em mente, John Johnston propõe uma abordagem frutífera, onde ele refuta as *oposições* entre o humano e o técnico, explorando *o terreno comum* entre os homens e as máquinas utilizando a noção de *visão maquínica*:

*Machinic vision, as I shall use the term, presupposes not only an environment of interacting machines and human-machine systems but a field of decoded perceptions that, whether or not produced by or issuing from these machines, assume their full intelligibility only in relation to them* (Johnston, 1999, p. 27).

O que está em pauta é um campo de *percepções* codificadas e recodificadas que pressupõe um imbricamento efetivo entre o orgânico e o eletromecânico, onde as fronteiras entre homem e máquina – ou, se preferir, do humano e do não-humano – são reiteradamente dissolvidas e reconstituídas.

O que me parece particularmente interessante na abordagem de Johnston é sua insistência nas possibilidades que se desdobram do processo de *codificação e recodificação* subjacente à digitalização que torna possível o funcionamento das máquinas perceptivas.

*The digital image is a decoding because it frees the cinematic image from its material support, mobilizing it within a communicational network wherein it can be transmitted anywhere instantaneously; but it is also a recoding because, instead of being inscribed directly on a chemically treated surface, light is converted into information, mathematical data whose infinitesimal discreteness allows the real to be synthesized or recomposed* (Johnston, 1999, p. 39).

A “libertação” da imagem do suporte material possibilita a sua inscrição no campo da informação matematicamente manipulável por meios eletrônicos, o terreno por excelência da operação do nooscópio. Com isto, termos como *observador e representação* perdem a sua solidez anterior. Como destaca Jonathan Crary, até a década de 1970 o cinema, a fotografia e a televisão ainda produziam *representações miméticas* que correspondiam ao mundo analógico, percebido mediante um ponto de vista estimulado pela faixa visível do espectro da luz, de forma congruente com o perspectivismo cartesiano. Técnicas videográficas como hologramas sintéticos, mapas de textura, ressonância magnética, óculos de realidade virtual e diversas outras tecnologias de *produção de*

---

(5) Paul Virilio foi muito sagaz ao dizer no final da década de 1980 que, quando nos referimos às representações internas que as máquinas de visão geram, sequer poderíamos falar de imagens: *Don't forget, though, that 'image' is just an empty word here since the machine's interpretation has nothing to do with normal vision (to put it mildly!). For the computer, the optically active electron image is merely a series of coded impulses whose configuration we cannot begin to imagine since, in this 'automation of perception', image feedback is no longer assured. That being, of course, the whole idea* (Virilio, 1994, p. 73). A “visão” destas máquinas é, portanto, matemática e instrumental e só precisa retornar ao plano do visível para estabelecer uma comunicação com os homens.

*imagens* em espaços visuais fabricados estão possibilitando uma *cisão* entre as imagens eletronicamente constituídas e a posição do sujeito que observa situado em um mundo “real”. Dito de outro modo: a visão maquínica dissolve aquele mundo oticamente estruturado que marcou a modernidade. Ao serem constituídas por dados, as imagens ajudam a instaurar um regime de visualidade radicado em “um terreno cibernético e eletromagnético em que elementos abstratos, linguísticos e visuais coincidem, circulam, são consumidos e trocados em escala global” (Crary, 1992, p. 2).

O fato de as máquinas perceptivas não precisarem de uma saída videográfica para entrarem em sinergia gera um campo que, enquanto homens, não podemos penetrar diretamente. Uma zona de opacidade que, de forma um pouco exagerada, Virilio qualificou como um “enigma insondável”, um *imaginário mecanizado*.<sup>6</sup> Mas isto também ocorre no sentido contrário. É possível ludibriar a percepção das máquinas, explorando os limites da visão computacional e do *machine learning*. Um exemplo corriqueiro ilustra isso: o teste CAPTCHA (*Completely Automated Public Turing test to tell Computers and Humans Apart*), desenhado para distinguir homens de *bots* quando se navega na Internet.

Os CAPTHA “jogam como uma espécie de fronteira da visibilidade computacional: visíveis para os humanos, mas não para as máquinas” (Mintz, 2015, p. 26). As obstruções e sobreposições das imagens também dificultam os processos de reconhecimento por depender muito das inferências para completar as informações que faltam. Dois objetos sobrepostos com texturas similares podem confundir o classificador.

Pasquinelli e Joler exploram estas fronteiras por meio da oposição entre *inteligência adversarial e artificial*. As máquinas perceptivas reconhecem e classificam imagens levando em consideração as bordas e a distribuição estatística do contraste entre pixels claros e escuros, isto é, geram valores baseados no brilho e na proximidade. Por isso é difícil reconhecer as figuras do teste CAPTCHA acima mencionado, que utiliza imagens distorcidas.

*The algorithm is programmed to record only the dark edge of a profile (that is to fit that desired pattern) and not all the pixels across the image (that would result in overfitting and repeating the whole visual field). A statistical model is said to be trained successfully when it can elegantly fit only the important patterns of the training data and apply those patterns also to new data ‘in the wild’. If a model learns the training data too well, it recognises only exact matches of the original patterns and will overlook those with close similarities, ‘in the wild’. In this case, the model is overfitting, because it has meticulously learnt everything (including noise) and is not able to distinguish a pattern from its background. On the other hand, the model is underfitting when it is not able to detect meaningful patterns from the training data (Pasquinelli; Joler, 2020, p. 10).*

Calibrar o equilíbrio entre *underfitting* e *overfitting* é a chave para garantir acurácia dos modelos. Logo, pelo menos neste caso, o aprendizado de máquina não aprende nada no sentido humano do termo. Ele opera mapeando estatisticamente a distribuição de valores numéricos e

---

(6) Esta tradução é sujeita a controvérsias. O autor usa o termo *imaginaire machinique*, que foi traduzido para o inglês como *mechanized imaginary*. A princípio, seria possível usar o termo *imaginário maquínico*. Contudo, o termo *mecanizado* me parece mais fiel ao sentido que Virilio dá à expressão. Longe de ressaltar o imbricamento homem-máquina que caracterizam as relações maquínicas, Virilio está ressaltando a *autonomização* das “imagens virtuais instrumentais” que as máquinas perceptivas constroem em sua interação, um imaginário no qual, enquanto homens, estamos excluídos.

inferindo funções matemáticas que se aproximem ao máximo da compreensão humana. Esta característica, contudo, exatamente por marcar uma diferença “lança uma nova luz sobre o modo como os seres humanos compreendem”.

É possível, entretanto, pensar essa questão pelo avesso. Uma vez definida essa fronteira, é possível explorar a zona de opacidade da “visão computacional” para resistir às pressões dos sistemas de vigilância. A vigilância baseada no reconhecimento facial engendrou a *contra-vigilância*, isto é, a transposição das técnicas de ofuscação para a dimensão matemática, com o objetivo de tornar os humanos e suas ações ininteligíveis para as máquinas. Nesse caso, o homem vira uma *black box*. Os autores do manifesto usam como exemplo o *HyperFace*, uma camuflagem têxtil criada por Adam Harvey que explora os *glitches* dos algoritmos de visão computacional para ludibriá-los. Os sistemas de reconhecimento “enxergam” múltiplos rostos humanos quando não há nenhum. Só é possível usar este truque explorando a *diferença* entre o modo como homens e máquinas reconhecem rostos.

É nessa fronteira que se materializam os *ataques adversariais*, isto é, o ato de explorar as vulnerabilidades matemáticas dos modelos usando estímulos que confundem e inviabilizam as classificações automatizadas de objetos.

*Such attacks are often instantiated by adversarial examples: legitimate inputs altered by adding small, often imperceptible, perturbations to force a learned classifier to misclassify the resulting adversarial inputs, while remaining correctly classified by a human observer* (Papernot et al., 2017, p. 1).

No artigo, os autores reproduzem duas imagens praticamente idênticas de placas de pare. Em uma das imagens foi adicionada uma perturbação imperceptível ao olho humano que, contudo, induz os sistemas de classificação a identificá-la como uma placa de atenção. Logo, adesivos e objetos 3D podem confundir os sistemas de navegação de carros autônomos, provocando acidentes. Esse jogo de visibilidade e ocultamento tem sido a norma na rivalidade estratégica de ponta. Mísseis de cruzeiro, por exemplo, perseguem o seu alvo de forma automatizada. Defensivamente, é possível “enganar” o míssil confundido seus sensores utilizando *decoys* e contra-medidas. Não existe poder sem contrapoder, sem desobediência. “O limite lógico e ontológico do aprendizado de máquina é o sujeito rebelde (unruly) ou o evento anômalo que escapa à classificação e ao controle” (Pasquinelli; Joler, 2020, p. 19).

### ***A captura do futuro?***

Creio que, dentre as várias teses apresentadas no manifesto, a ideia de que os sistemas preditivos baseados no aprendizado de máquina são incapazes de detectar um fenômeno genuinamente novo é a mais importante, a ponto de estruturar todas as demais. Minha proposta aqui é relativizar e qualificar melhor esta afirmação, recorrendo à uma sutil distinção entre complexidade e caos. No imaginário popular a teoria do caos é fortemente associada ao *efeito borboleta*: um evento aparentemente banal como o voo de uma borboleta desencadeia uma reação em cascata que pode resultar em um furacão do outro lado do planeta. A imagem foi criada para destacar a alta sensibilidade às condições iniciais em sistemas não-lineares. Contudo, embora exista uma disputa intensa sobre o seu significado, especialmente em sua forma matemática, a teoria do caos tende a expressar a noção de que existe uma *ordem oculta* em todo fenômeno, por mais complexo que ele possa parecer. Essa ordem é relativamente instável e, portanto, só pode ser aproximada por meio de um quadro de

possibilidades no curto prazo (*ensemble forecast*). O que se busca é, portanto, *domar* a imprevisibilidade. Já os estudos sobre a complexidade tendem a caminhar no sentido contrário, especialmente por conta do conceito de *emergência*. Sistemas complexos congregam uma grande quantidade e diversidade de elementos com uma cadeia múltipla de interações entre eles, a qual expressa um incessante intercâmbio de matéria, energia e informação com o meio circundante. Nestas condições é possível a emergência de fenômenos totalmente inesperados, que geram uma *nova ordem* capaz de se autorreplicar por conta de sua eficiência. Embora sejam frutos do longo processo de *erosão do determinismo* que desembocou na ideia de *leis estatísticas* (Hacking, 2002, p. 10; 160 e segs.; Hacking, 2006), os sistemas complexos e a teoria do caos possuem formas diferentes de lidar com a aleatoriedade.

Sistemas preditivos se baseiam em formas automatizadas de *classificação* dos dados que processam. Como estas técnicas estão se generalizando como formas de vigilância e governo, Pasquinelli e Joler sugerem que estamos testemunhando a consolidação de um novo regime disciplinar, centrado em uma *governança estatística* que tenta aprisionar o futuro com base na tentativa sistemática de *regenerar o velho*:

*A logical limit of machine learning classification, or pattern recognition, is the inability to recognise a unique anomaly that appears for the first time, such as a new metaphor in poetry, a new joke in everyday conversation, or an unusual obstacle (a pedestrian? a plastic bag?) on the road scenario. The undetection of the new (something that has never 'been seen' by a model and therefore never classified before in a known category) is a particularly hazardous problem for self-driving cars and one that has already caused fatalities. Machine learning prediction, or pattern generation, show similar faults in the guessing of future trends and behaviours. As a technique of information compression, machine learning automates the dictatorship of the past, of past taxonomies and behavioural patterns, over the present. This problem can be termed the regeneration of the old – the application of a homogenous space-time view that restrains the possibility of a new historical event (Pasquinelli; Joler, 2020, p. 16).*

Esta é uma questão fundamental. Os agenciamentos de vigilância e de gestão automatizada de riscos incorporaram as técnicas de aprendizado de máquina como uma tática de lidar com a contingência. Mas sua orientação dominante é tentar converter o contingente no provável (Hui, 2019, p. 40) e, com isto, pela via da *repetição do passado* tentar controlar as virtualidades para normalizar as condutas. Os futuros indesejáveis devem ser eliminados no nascedouro.

O que Pasquinelli e Joler chamam de *regeneração do velho* pode ser definido como o aprisionamento das virtualidades dentro dos limites lógicos inerentes aos modelos estatísticos que governam o nooscópio. Essa passagem é particularmente esclarecedora:

*Interestingly, in machine learning, the logical definition of a security issue also describes the logical limit of its creative potential. The problems characteristic of the prediction of the new are logically related to those that characterise the generation of the new, because the way a machine learning algorithm predicts a trend on a time chart is identical to the way it generates a new artwork from learnt patterns. The hackneyed question 'Can AI be creative?' should be reformulated in technical terms: is machine learning able to create works that are not imitations of the past? Is machine learning able to extrapolate beyond the stylistic boundaries of its training data? The 'creativity' of machine learning is limited to the detection of styles from the training data and then random improvisation within these styles. In other words, machine learning can*

*explore and improvise only within the logical boundaries that are set by the training data. For all these issues, and its degree of information compression, it would be more accurate to term machine learning art as **statistical art***” (Pasquinelli; Joler, 2020, p. 16).

Alimentado pelos dados de treinamento, os algoritmos inferem *padrões* que são transpostos para o espaço vetorial n-dimensional que discutimos no começo desse artigo. Os produtos que eles geram são improvisações aleatórias de estilos que dão a aparência de criatividade. Não se trata, portanto, de uma imitação estática do passado, mas de recombinações que respeitam os limites lógicos das “múltiplas lentes do nooscópio”.

Este argumento não é consensual mesmo entre os críticos da IA, pois é comum apontar que, a contrapelo dos oligopólios e dos aparelhos estatais que dão substância a um presumido “capitalismo de vigilância”, a captura do futuro pelo passado nem sempre funciona. É possível utilizar a própria complexidade ensejada por estes agenciamentos para possibilitar a eclosão de novas e formas de pensamento e, principalmente, um novo modo de abordar a realidade. Sem questionar frontalmente a *démarche* do manifesto nooscópico, Nuria Rodríguez-Ortega indica essa possibilidade.

*The opposite approach is represented by those positions that understand machine learning as a method that provide us with augmented and magnified knowledge of the reality we observe due to their capacity to make intelligible implicit characteristics, patterns and correlations across a huge space of data. (...) In this regard, my position is based on a strategic and pragmatic stance. I consider that a scenario that is defined by the ubiquity and prevalence of AI compels us to seek productive frameworks to expand our comprehension of cultural phenomena and processes. I consider it essential to explore the possibilities of «seeing» that AI offers while bearing in mind that it is a diffracted vision. In short, I think it is more interesting to think that AI helps us to see differently than to think that it helps us to see more and better. It is in this different seeing where, in my view, lies the heuristic and epistemological potential of these technologies (2022, p. 9).*

Deste ponto de vista, que também recorre à analogia ótica do nooscópio, o *machine learning* nos permite ver muito mais, embora de forma indireta. Isso também ajuda a mudar a nossa compreensão da realidade e, desse modo, a recolocar no centro da discussão a questão ontológica.

A despeito dessas objeções, o me parece mais interessante no diagnóstico é o modo como Pasquinelli e Joler encaminham a discussão sobre o modo como as difrações estatísticas criam uma ilusão perigosa: fica parecendo que a dimensão técnica – as teorias científicas e os dispositivos computacionais – é que ilumina a vida social e, na prática, a conversão da realidade em uma *fonte de dados* processada incessantemente pelo nooscópio é um produto da “técnica”. Essa concepção “datacêntrica” do mundo precisa ser frontalmente atacada. Os autores nos convidam a pensar este problema dissolvendo as fronteiras entre as abordagens externalistas e externalistas da IA: existe uma dimensão subjetiva na “matemática do controle” que anima a partir de dentro a “inteligência de máquina”. Ecoando a visão de Guattari sobre as máquinas em geral, eles afirmam que “a inteligência das máquinas é constituída também de ‘formas hiper-desenvolvidas e hiper-concentradas de certos aspectos da subjetividade humana” (Pasquinelli; Joler, 2020 p. 17). A IA não é, portanto, uma externalidade absoluta que não guarda nenhuma comensurabilidade com a mente e as formas humanas de percepção e cognição. São nestas interseções – nos agenciamentos maquinais – que a luta contra o poder normativo da IA pode ser travada, sem cair na batida oposição homem-máquina ensejada por alguns humanistas extemporâneos.

## Considerações finais

A máquina-ferramenta parecia encarnar os princípios fundamentais da tradição newtoniana: um mecanismo linear baseado na transmissão do movimento por meios puramente mecânicos, em que cada peça possui uma função específica cujo sentido é dado pela operação do conjunto. Contudo, a absorção das ferramentas pelo maquinário e sua subordinação à acumulação incessante de capitais fomentou a busca de novas fontes de energia. A generalização da máquina a vapor como forma de propulsão foi a resposta mais imediata. Além do seu impacto no cotidiano, a generalização das máquinas-ferramentas movidas a vapor precipitou transformações no plano ontológico. Algo se perdia definitivamente quando se convertia calor em movimento. O saber prático entrou em contradição com a tradição dos saberes ancorados na ontologia dominante. A simetria temporal passou a ser desafiada pela seta do tempo e pela entropia, assim como a ideia de *energia* assumiu um papel central, ao lado das forças e partículas que até então dominavam a cena. Este clima intelectual favoreceu o surgimento da cibernética, que colocou em primeiro plano o problema da informação e da auto-organização dos sistemas, precipitando novas transformações ontológicas que, por conta da aceleração da digitalização e dos processos de automação da percepção, ficaram mais explícitas no final do século XX.

O surgimento de máquinas perceptivas articuladas por redes de comunicação que atravessam os homens e a “natureza” abalou o que restava do humanismo renascentista. A perspectiva do homem começou a ser complementada – e até mesmo ameaçada – pelas percepções maquinicas. Esta démarche gerou uma forte sinergia com os estudos sobre a complexidade, colocando mais pressão sobre o determinismo subjacente à tradição newtoniana e ao naturalismo. Quando se constata que o futuro não é integralmente determinado pelo passado, o papel da contingência muda. Ela deixa de ser encarada como um efeito derivado da falta de dados e de capacidade de processamento pelo observador para ser concebida como um elemento constitutivo da realidade. O paradoxo é que o reconhecimento desta característica tornou ainda mais premente a questão do *controle*: para garantir alguma margem de interferência voluntária na realidade é necessário quantificar, circunscrever e, na medida do possível, *domar a aleatoriedade* (Hacking, 2002, p. 10).

Esta tensão entre determinismo e contingência está no cerne dos estudos sobre a complexidade. Boa parte dos seus entusiastas acentuam a abertura para o novo que essa perspectiva possibilita. Longe de ser temida, a incerteza tende a ser vista como um elemento *criativo*, indispensável para a emergência do inusitado e do genuinamente novo. Mas há quem caminhe no sentido contrário. A teoria matemática do caos, por exemplo, tende a enfatizar a dimensão determinista do caos, que configura uma espécie de *ordem oculta* manipulável no curto prazo por modelos estatísticos. Esta ordem não é estática, pois ela encarna campo de *possibilidades* em constante transformação. Nisto reside o âmbito do problema, especialmente se levarmos em conta a captura da IA e de seus agenciamentos pelos grandes oligopólios privados e pelas agências estatais de vigilância e repressão. Neste caso, a tendência dominante é a instauração de um regime disciplinar orientado pela regeneração sistemática do passado, onde o “futuro” é modelado por uma perspectiva securitária e instrumental que tenta preservar ao máximo a institucionalidade vigente.

O *manifesto nooscópio* explora as fissuras deste regime disciplinar, particularmente a sua dificuldade em lidar com as anomalias inerentes aos acoplamentos homem-máquina e, também, com a vulnerabilidade aos ataques adversariais que exploram os limites da visão maquinica. Embora parte

da carga retórica do texto agrade a muitos catastrofistas, creio que este não é o tom dominante da narrativa. O encerramento do manifesto parece deixar isso bem claro. Como salientam os autores, a IA vai muito além de personificar um dispositivo de controle, pois ela também é um *aparelho produtivo* que combina o trabalho altamente especializado de cientistas e engenheiros com os *ghost workers*. A explicitação do laço indissolúvel entre homens e máquinas possibilita desmistificar a IA: ela não é algo totalmente externo aos homens, uma “mente alienígena” que tende a fugir completamente do controle social. O nooscópio, portanto, sinaliza o ressurgimento em novas bases da “questão das máquinas” que agitou o século XIX. Hoje, contudo, ela se expressa por meio de uma tensão entre pelo menos duas orientações antagônicas. O manifesto nooscópico foi elaborado para ressaltar as tentativas de capturar as virtualidades da IA reforçando o seu papel no *controle normativo* sobre a sociedade e, também, a sanha extrativista das *big techs*. Entretanto, se a entendermos como um desdobramento do crescente processo de *digitalização* do mundo, a IA ajuda a descortinar novas formas de cognição e, principalmente, de percepção da realidade. Como foi aqui apontado, Judea Pearl afirma que o *machine learning* apenas nos leva dos dados para a estatística. Mas a partir daí é possível fazer muita coisa interessante. Mesmo se aceitarmos a ideia de que há um *limite* nas possibilidades da IA que está inscrito nos dados de treinamento, levando em conta as possibilidades de combinação e recombinação, esse limite é praticamente infinito (ou finito ilimitado, para usar uma imagem deleuziana) e pode ser continuamente remodelado pela interação homem-máquina.

### Referências bibliográficas

- BOUSQUET, Antoine. *The eye of war*. Minneapolis: University of Minnesota Press, 2018.
- BREGER, Herbert. God and Mathematics in Leibniz’s Thought. In: KOETSIER, T.; BERGMANS, L. (Org.). *Mathematics da the Divine: a historical study*. Amsterdam: Elsevier, 2005.
- BROCKMAN, John (Ed.). *Possible minds*. New York: Penguin Press, 2019.
- GRAY, Mary L.; SURI, Siddhart. *Ghost Work: how to stop Silicon Valley from Building a new global underclass*. Boston & New York: Houghton Mifflin Harcourt, 2019.
- HACKING, Ian. *The Taming of Chance*. Cambridge: Cambridge U. Press, 2002.
- HACKING, Ian. *The Emergence of Probability*. Cambridge: Cambridge U. Press, 2006.
- HINTIKKA, Jaakko. *Lingua universalis vs. calculus ratiocinator: An Ultimate Presupposition of Twentieth-Century Philosophy*. Dordrecht: Kluwer, 1997.
- HUI, Yuk. *Recursivity and Contingency*. London: Rowman & Littlefield, 2019.
- IVINS JR., William M. *On the rationalization of sight*. New York: Metropolitan Museum of Art, 1938.
- JAY, Martin. Regimes escópicos da modernidade. *ARS*, São Paulo, v. 18, n. 38, 2020.
- MINTZ, André Góes. *Visão computacional e visualidades contemporâneas: composições do ver e do visível entre a técnica, a ciência e a arte*. Dissertação (Mestrado)–UFMG, Belo Horizonte, 2015.
- PAPERNOT, Nicolas; McDANIEL, Patrick; GOODFELLOW, Ian; JHA, S.; CELIK, Z. B.; SWAMI, A. Practical Black-Box Attacks against Machine Learning. In: PROCEEDINGS of the 2017 ACM

on Asia Conference on Computer and Communications Security (ASIA CCS' 17). New York, NY Association for Computing Machinery, 2017, p. 506-519.

PEARL, Judea. The limitations of Opaque Learning Machines. In: BROCKMAN, John (Ed.). *Possible Minds*. New York: Penguin Press, 2019

PECKHAUS, Volker. Schröder's Logic. In: GABBAY, Dov M.; WOODS, John (Org.). *Handbook of the History of Logic*. North-Holland: Elsevier, 2004.

RODRÍGUEZ-ORTEGA, Nuria. Techno-Concepts for the Cultural Field: n-Dimensional Space and Its Conceptual Constellation. *Multimodal Technologies and Interactions*, 6, 96, 2022.

VIRILIO, Paul. *The vision Machine*. Indianapolis: Indiana U. Press, 1994.