

2º CICLO DE ESTUDOS
MESTRADO EM FILOSOFIA - RAMO DE ÉTICA E FILOSOFIA POLÍTICA

**O DIÁLOGO COMO PRODUTO
A LIMINARIDADE DA INTELIGÊNCIA ARTIFICIAL
CONVERSACIONAL**

Leandro Ortolan

M

2025



Leandro Ortolan

O DIÁLOGO COMO PRODUTO A LIMINARIDADE DA INTELIGÊNCIA ARTIFICIAL CONVERSACIONAL

Dissertação realizada no âmbito do Mestrado em Filosofia - Ramo de Ética e Filosofia Política, orientada pela Professora Doutora Maria João Couto e pelo Professor Doutor Domingos José Matos Sousa Faria.

[*Status: aprovada*] em Prova de Mestrado realizada em 11/abril/2025 no Salão Nobre da Faculdade de Letras da Universidade do Porto, com o júri presidido pela Professora Doutora Paula Isabel do Vale Oliveira e Silva e com os vogais: Doutor Tiago Bruno Borges Rodrigues Mesquita de Carvalho (arguente) e Professor Doutor Domingos José Matos Sousa Faria.

Faculdade de Letras da Universidade do Porto

2025

*Dedico esta dissertação a Portugal,
por tudo o que pude por cá aprender, receber e dar.*

Sumário

Declaração de honra / Declaration of Honour	6
Agradecimentos	7
Resumo	8
Abstract.....	9
Índice de Figuras	10
Introdução.....	11
1. Em busca da verdadeira problematização – a historicidade	16
1.1. Como chegamos até “aqui”?	16
1.2. A emergência relacional pela perspectiva filosófica.....	20
1.2.1. Aristóteles e a emergência do objeto	20
1.2.2. A emergência dos parmenidianos.....	26
1.2.3. Descartes e a emergência da subjetividade	26
1.2.4. Kant e a emergência da intersubjetividade.....	30
1.2.5. Husserl e a emergência da intencionalidade.....	39
1.2.6. Heidegger e a emergência do dasein	41
1.2.7. Ricoeur e a emergência da narrativa	43
1.2.8. Uma síntese com a emergência do “hífen” como relacional	44
1.2.9. A emergência da IA como relações (ou correlações)	47
1.3. A emergência ontológica pela perspectiva tecnológica	50
1.3.1. Sobre a Ontologia e as ontologias.....	51
1.3.2. Os modelos de representações do conhecimento	56
1.3.3. Sobre o que ensinamos aos algoritmos	60
1.4. Os muitos problemas éticos	65
1.4.1. O mapa dos problemas éticos.....	66
1.4.2. Uma necessária revisão ontológica para o mapa dos problemas éticos	71
1.4.3. Sobre os artefactos e a emergência das dissonâncias teleológicas.....	75
1.5. O que é este “aqui”?.....	80
2. O diálogo como produto – as possibilidades para a IA.....	83
2.1. Onde e quando queremos chegar?	83
2.2. A quarta dimensão: o capital como agente intencional	84
2.3. A IA do futuro pela transitividade dos problemas éticos	94

2.4. A IA do futuro e a privacidade	97
2.5. A IA do futuro e a validação das evidências	99
2.6. Como queremos chegar ao futuro da IA?	113
3.A liminaridade tecnológica – as oportunidades atuais	115
3.1. Por que a liminaridade?	115
3.2. Whitehead e seus contributos filosófico-relacionais	117
3.2.1. A Concrecência como complexidade existencial alcançada pela narrativa	119
3.2.2. A Preensão como dado(s) passíveis de serem qualificado(s)	120
3.2.3. O Contraste como narrativa a partir dos metadados	120
3.2.4. O modo de síntese da concrecência	121
3.2.5. Os Nexos como agrupamentos ontológicos dotados de narrativas	122
3.2.6. As Formas subjetivas como subjetividade, em si, ou mesmo perspectivas individuais	
123	
3.3. Entidades reais como circularidades hermenêuticas	124
3.4. A hipótese alternativa de uma dinâmica metaontológica	126
3.5. A saída da liminaridade pela superação dos problemas	132
3.6. A limitação dos modelos psicológicos – O caso Big Five	133
3.7. O caso do Projeto Signo: <i>design</i> , valores e ética	135
3.8. Que quer dizer isto tudo, afinal? De Tales de Mileto à IA... ..	140
Conclusão e Considerações Finais	143
Referências Bibliográficas	147

Declaração de honra / Declaration of Honour

Declaro que a/o presente tese/dissertação/relatório é de minha autoria e não foi utilizada/o previamente noutro curso ou unidade curricular, desta ou de outra instituição. As referências a outros autores (textos, trabalhos, ideias) respeitam escrupulosamente as regras de atribuição de autoria e encontram-se devidamente indicadas no texto e nas referências bibliográficas, de acordo com as normas de referência. Tenho consciência de que a prática de plágio e auto-plágio constitui um ilícito académico.

I hereby declare that this thesis/dissertation/report is of my authorship and has not been used previously in another course, degree, curricular unit or subject, at this or any other institution. References to other authors (statements, ideas, thoughts) scrupulously respect the rules of attribution and are duly indicated in the text and bibliographical references, in accordance with the rules of referencing. I am aware that the practice of plagiarism and self-plagiarism is an academic offence.

Declaro, ainda, que não utilizei ferramentas de inteligência artificial generativa (*chatbots* baseados em grandes modelos de linguagem) para realização de parte(s) da/o presente tese/dissertação/relatório.

I further declare that I have not used generative artificial intelligence tools (chatbots based on large language models) to carry out part(s) of this thesis/dissertation/report.

Portugal, Porto, 24 de março de 2025

Leandro Ortolan

Agradecimentos

Na inadequação de agradecer a todos os filósofos, desde Tales de Mileto, agradeço primeiro à Filosofia – uma consoladora, para mim, tal qual foi para Boécio. Inadequação, pois, ao agradecer a todos os filósofos, diminuiria a importância aos que abaixo direcionarei os meus mais sinceros agradecimentos:

Aos meus orientadores, Prof.^a Dr.^a **Maria João Couto** e pelo Prof. Dr. **Domingos José Matos Sousa Faria** que, ao me permitirem uma liberdade dissertativa sem igual, me levaram a seguir por caminhos que, de facto, me surpreenderam e engrandeceram;

Aos **colegas do Projeto Signo**, pelo imenso contributo em ensinamentos, além da liberdade ofertada para expor minhas ideias, sempre consideradas e respeitadas:

Ao Prof. Dr. **José Francisco Preto Meirinhos**, pela generosidade que lhe é peculiar em compartilhar os seus saberes e me permitir uma oportunidade única de aprendizado e de autovalorização pessoal ao ingressar no Projeto Signo: foi, para mim, “a” diferença;

Ao Prof. Dr. **João Carlos Martins Rebalde** com seus comentários nas discussões de trabalho, que sempre trouxeram imensos *insights* às minhas investigações; aos Investigadores Doutor **Mário João Rosas Rebelo Correia** e Doutora **Celeste Maria Lourenço da Silva de Oliveira Pedro**, que sempre propuseram artigos imensamente relevantes, com alguns destes referenciados aqui, inclusive; à colega doutoranda **Irândina Afonso** que, mesmo às vésperas de sua defesa de tese de doutoramento, sempre esteve a contribuir de forma produtiva;

Ao Prof. Dr. **Ricardo Manuel Coelho de Melo**, investigador sénior no Fraunhofer Portugal AICOS e principal do Projeto Signo, pela valorosa receptividade e acolhimento que recebi; também agradeço a toda **equipa Fraunhofer** do Projeto Signo;

E, por fim, (e sempre) a **todos os professores da FLUP** que estiveram em meu percurso a darem seus contributos e originarem em mim, além do carinho, uma gratidão profunda e incondicional.

Pelas circunstâncias da vida, dedicar-me a tal empreitada só me foi possível devido aos apoios financeiros da **Fundação para a Ciência e a Tecnologia (FCT)**, pela bolsa de investigação no Projeto Signo e da **Direção-Geral do Ensino Superior (DGES)**, pela bolsa de estudos no primeiro ano investigativo. Portanto, meu compromisso em entregar algo de valor à sociedade, a retribuir o investimento realizado, virou um imperativo de primeira ordem para o que estive a buscar.

Assim, deixo o meu MUITO OBRIGADO; e, portanto, espero estar a devolver mais do que recebi, e isto não acabará por aqui: meu esforço foi imenso e a entrega foi total – e assim continuarei a pensar e a fazer, a dar o meu melhor dentro do tempo que me for possível continuar.

Resumo

Esta dissertação propõe um novo modelo metaontológico como forma de superar os problemas éticos atuais mais relevantes e desafiadores. Argumenta que há uma dissonância teleológica entre o que se espera de uma IA e o que esta pode oferecer, construtivamente. Propõe que tal dissonância se dá a partir de uma tradição filosófica na qual há uma predileção pela extração de um fragmento da realidade, em detrimento de prezar pela análise da própria complexidade, como dada. Assim, o que resulta de todo esse movimento é a liminaridade da atualidade tecnológica, na qual se atravessa um período de complexa transformação e potencialidade, quando muito do que era estável e previsível está a tornar-se o oposto: instável e imprevisível.

A primeira parte da dissertação aborda o percurso do passado ao presente. A segunda parte considera como será o futuro da IA, se tomado a partir da superação dos problemas atuais, sem configurar-se como projeções ficcionais. Faremos tal projeção com base nos recursos tecnológicos existentes ou outros que sejam dados como viáveis e exequíveis de serem utilizados nos constructos algorítmicos. E, por fim, a terceira parte trata justamente da liminaridade, deste presente que não se refere somente à temporalidade, mas também à espacialidade – em termos de limites e alcances.

A conclusão defende que uma IA ética e responsável, isenta dos atuais problemas, exige uma profunda compreensão da complexidade humana – e a respetiva representação do modal existencial – para que seja possível a aplicação dos valores mais convenientes em todas as etapas do desenvolvimento tecnológico. A consideração final, que sintetiza toda a argumentação, é sobre a elaboração de um protocolo metaontológico dinâmico, capaz de capturar e emular a complexidade relacional da realidade. Tal feito é apresentado como necessário, mas talvez não suficiente, para a criação desta IA do futuro que ultrapassará as atuais demandas.

Palavras-chave: inteligência artificial; metaontologia; representação do conhecimento; ética relacional; *design* sensível aos valores.

Abstract

This dissertation proposes a new metaontological model as a way to overcome the most relevant and challenging ethical problems of today. It argues that there is a teleological dissonance between what is expected of an AI and what it can offer, constructively. It proposes that such dissonance arises from a philosophical tradition in which there is a predilection for extracting a fragment of reality, to the detriment of valuing the analysis of complexity itself, as given. Thus, what results from this whole movement is the liminality of the current technological era, in which we are going through a period of complex transformation and potentiality, when much of what was stable and predictable is becoming the opposite: unstable and unpredictable.

The first part of the dissertation addresses the path from the past to the present. The second part considers what the future of AI will be like, if taken from the perspective of overcoming current problems, without configuring them as fictional projections. We will make such projections based on existing technological resources or others that are considered viable and feasible to be used in algorithmic constructs. And, finally, the third part deals precisely with liminality, this present that refers not only to temporality, but also to spatiality – in terms of limits and scope.

The conclusion argues that an ethical and responsible AI, free from current problems, requires a deep understanding of human complexity – and the respective representation of the existential modal – so that it is possible to apply the most appropriate values in all stages of technological development. The final consideration, which summarizes the entire argument, is about the elaboration of a dynamic metaontological protocol, capable of capturing and emulating the relational complexity of reality. This feat is presented as necessary, but perhaps not sufficient, for the creation of this AI of the future that will surpass current demands.

Key-words: artificial intelligence; metaontology; knowledge representation; relational ethics; value-sensitive *design*.

Índice de Figuras

FIGURA 1 - “EVOLUÇÃO” DO TERNÁRIO MACRO-ONTOLÓGICO DA EPISTEMOLOGIA: SUJEITO, HÍFEN E OBJETO E SEUS RESPETIVOS “EFEITOS COLATERAIS”, MAS APRESENTADOS COMO “CAUSAIS”.....	46
FIGURA 2 - MAPA DOS PROBLEMAS ÉTICOS, SEGUNDO MITTELSTADT ET AL. (2016)	67
FIGURA 3 - MAPA DOS PROBLEMAS ÉTICOS (ONTOLOGICAMENTE REVISADO E RENOMEADO), A PARTIR DA FIGURA 2.....	72
FIGURA 4 - FLUXO ALGORÍTMICO ESQUEMATIZADO A PARTIR DE MITTELSTADT ET AL. (2016).....	73
FIGURA 5 - REPRESENTAÇÃO DO CONHECIMENTO SOBRE O MODELO IDEAL DE IA, ESQUEMATIZADO FUNCIONALMENTE CONFORME O ARTIGO REFERENCIADO. (MITTELSTADT ET AL., 2016)	74
FIGURA 6 - REPRESENTAÇÃO DO CONHECIMENTO DE ALGUNS DOS PROBLEMAS ÉTICOS MAIS COMUMENTE CITADOS NAS MÉDIAS, ALOCADOS DENTRO DA CLASSIFICAÇÃO ADOPTADA NA FIGURA ANTERIOR.	95
FIGURA 7 - QUADRO UTILIZADO COMO SUPORTE PARA A ELICITAÇÃO DOS VALORES RELEVANTES, NAS DINÂMICAS ENTRE OS STAKEHOLDERS.	139
FIGURA 8 - CARDS DOS POSSÍVEIS VALORES RELEVANTES UTILIZADOS COMO RECURSOS, A PARTIR DAS QUESTÕES COLOCADAS PELOS DESIGNERS.	139

Introdução

Esta dissertação tem como pretensão ultrapassar os limites teóricos da Filosofia ao propor-se servir, como apropriação e *práxis* para as diversas Ciências afins às tecnologias da informação, principalmente aquelas que estejam a lidar os desafios impostos pelas inteligências artificiais (IA), e, mais especificamente, a conversacional. O termo "IA conversacional" está a ser preferido em detrimento de "IA generativa", pois este último é um termo procedimental que descreve a forma como a inteligência artificial gera linguagem natural. Em contraste, a IA conversacional não só integra a IA generativa, mas também acrescenta intencionalidades que promovem a interação e a construção de uma relação com o utente. Para além dessa pretensa ultrapassagem, também se pretende que a Filosofia possa capturar o que as Ciências da Informação e da Tecnologia tenham a oferecer, como elementos que servirão de base a uma teorização mais filosófica, como contributo ao engrandecimento da IA como meio de progresso.

Para tanto, adotamos uma medida de suplementação informacional nas muitas notas que seguirão e aparecerão, quando oportunas, sempre a buscarem colaborar para que tais escritos possa transitar pela interdisciplinaridade – tão necessária ao que esta dissertação procurará propor e defender, que é uma metaontologia estruturada nos saberes das ciências das humanidades, a incluir as psicológicas, filosóficas, sociológicas e tantas outras.

Por que chamamos “meta”? Por ser uma ontologia da ontologia, basicamente. Pois, em síntese, podemos considerar “meta” como processo de redução. Assim, por exemplo, a teoria da metalinguagem tem desenvolvido um processo de redução da linguagem ao buscar estabelecer quais palavras sejam, elas próprias, irredutíveis a outras. E, este conjunto de palavras irredutíveis formam a metalinguagem. O “meta”, como no caso da linguagem, portanto, é sempre um processo complexo e que transcende o próprio objeto de estudo, como no caso da linguagem, dado como cross-linguístico, ao abordar análise de valores e demais questões de significações (Goddard, 2018; Goddard & Wierzbicka, 2011). Ontologia, como iremos desenvolver ao longo desta dissertação, pode ser considerada, a partir daqui, de todo o esforço e organização para classificar ou

representar o que seja existente em uma dimensão digital: representar o conhecimento, as coisas, as pessoas, as diversas manifestações ou dimensões subjetivas, objetivas, intersubjetivas (as relações) ou mesmo a ideologia ou ordem simbólica, e tudo o mais, como se o modal existencial fosse colocado como objeto de representação para que, em forma de dados e metadados, “existissem” na dimensão digital.

No caso da linguagem, este processo de investigação reducionista do significado se deu como relevante a partir do Iluminismo, especialmente com a Filosofia de Leibniz, que argumentava que “se nada pode ser compreendido em si mesmo, nada pode ser compreendido.” (Goddard, 2018). Por isso, buscaremos propor os mesmos argumentos *leibnizianos*, na qual toda a ontologia possa e deva ser reduzida a pressupostos metaontológicos que não apenas as definam, mas as expliquem e, quiçá, possam dar significados. Assim, acreditamos poder – nós ou as próximas gerações – superar os problemas graves das faltas de teorias das correlações de dados.

Por isso, e em nome de uma necessária interdisciplinaridade, existem as muitas notas de rodapé explicativas e a operarem como uma interface entre filósofos e cientistas da informação. Tais notas, assim consideradas, se ignoradas, não retirarão a essência argumentativa da dissertação, mas considerá-las implicará em novas dimensões, aplicações, casos, desdobramentos e, por que não, um certo grau de provocação em temas correlatos e que sempre apresentam uma perspectiva diferenciada. Ao propormos o “dinamismo”, prezamos por uma escrita igualmente dinâmica e, ao máximo, relacional: pragmática, atual e, também, “imaginativa”.

Dinamismo, aqui, significa não apenas evitar isolar um problema de seu contexto em busca de uma situação controlada e paralisada para se aplicar uma analítica em condições ideais de “temperatura e pressão”, mas sim buscar decifrar a complexidade da qual o problema emerge. Se há dinamismo, todavia, poderá haver diversos níveis dinâmicos, ou até mesmo o caos. Mas, a considerar que o próprio caos tenha em si tanto potencial e ato, haverá de ter elementos que possam ser comuns a tudo o que, algures, também tenha sido caótico. Afinal, a hipótese de que “o caos é uma ordem por decifrar” (Saramago, 2002) poderá ser colocada como questão.

Imaginativa, aqui, empregada na concepção de Castoriadis (2000), como elemento fundamental para a compreensão da história e da sociedade. Uma imaginação radical, dada como capacidade de a sociedade criar novas formas de ser e de significar, quando esta seja capaz de romper com a mera repetição do passado para construir, conscientemente, a sua história – seria uma vitória inédita da razão, e não mais de um espírito racional, ou mesmo irracional, mas da consciência, portanto – uma vitória nossa. A corroborarmos com o autor, de que a imaginação não é um processo individual, mas social, por estar intrinsecamente ligada à dimensão social através da criação e compartilhamento de significações imaginárias ou simbólicas que dão os formatos à sociedade, às instituições e à própria experiência dos indivíduos (Castoriadis, 2000, pg. 268), apostamos que tais leituras possam servir aos propósitos de ruptura com os grilhões da mesmice, ainda que com todos os riscos inerentes aos desdobramentos.

Assim, com base no exposto, estruturamos esta dissertação da seguinte forma, a seguir: Na Secção 1, consideramos o que nos trouxe até aqui (do passado ao presente) e, por isso, trabalhamos de forma a problematizar a análise corrente dos impactos éticos da IA na sociedade. Argumentamos que a abordagem tradicional, focada nos objetos e ignorando as relações, resulta em uma compreensão limitada e inconsistente da problemática que estamos a buscar perceber e, assim, nem sequer se consegue, dada a dissonância teleológica não considerada entre o que há entre os problemas e a IA da atualidade.

Percorremos uma parcialidade da História da Filosofia – especificamente da Epistemologia – para demonstrar como a ênfase em elementos estáticos e imutáveis, que chamamos de parmenidianos, dificulta a análise de um sistema dinâmico e relacional como a IA. Defendemos que estamos, portanto, anacrônicos em termos analíticos, a carecer de uma necessária compreensão da IA como narrativa dada a partir de um conjunto de relações.

Portanto, na Secção 1, realizamos uma necessária crítica à herança "parmenidiana", enaltecemos a elicitación da relação como foco de análise; defendemos a necessidade de uma ontologia relacional capaz de capturar a dinâmica e a complexidade das relações

entre os dados, algoritmos e utentes; relevamos a importância dos metadados como dador de camadas que poderá levar a formas de se capturar a complexidade da informação e as nuances das relações entre os diferentes elementos; atribuímos valor à intencionalidade do capital como fator determinante ao desenvolvimento da IA; analisamos e verificamos, com base no exposto, a inconsistência do mapa dos problemas éticos, quando este ignora a teleologia da IA e a intencionalidade do capital; e, por fim, reivindicamos a necessidade de um ajuste ontológico entre o atual e o digital, a integrar as dimensões objetiva, intersubjetiva e subjetiva – a considerar a teleologia da IA e a intencionalidade do capital, como quarta dimensão presente e influente.

Na Secção 2, consideramos o que pretendemos a partir daqui, com base na superação dos problemas éticos das IAⁱ (da atualidade ao futuro sem problemas) que estará ainda mais e inevitavelmente (se não tivermos as ferramentas de defesa) moldada pela intencionalidade expansionista e influenciadora do capitalⁱⁱ.

Portanto, na Secção 2, realizamos as dimensões existenciais do diálogo ser adotado como produto final e como meio, em um duplo papel como fim de entretenimento, informação ou companhia, quanto como um meio para o capital atingir outros fins, como a coleta de dados, a manipulação do comportamento ou a venda de produtos; acusamos uma transitividade nos problemas éticos mais citados sobre as IA, por frustrarem quaisquer tentativas de atribuição de uma origem a certos fatores isolados ou meramente estáticos; argumentamos a necessidade de se projetar uma IA que opere

ⁱ O uso de IA no plural indicará uma referência que transcende a IA conversacional, a abranger outros “tipos” de IA, não apenas restrita às generativas, mas principalmente as que se comportam como agentes ou decisoras em processos diversos como automações, previsões ou em quaisquer tarefas específicas.

ⁱⁱ Capital, aqui, e doravante, na lúdica tentativa de personificá-lo como agente intencional do capitalismo. Por isso, vale estabelecer que embora os problemas que serão aqui atribuídos ao capitalismo, através da menção ao capital, possam fazer com que algumas críticas parecerão condenatórias, não será essa a intenção e nem mesmo parte do objetivo da dissertação. Buscaremos estabelecer uma análise *stricto sensu* de alguns de seus efeitos, que defendemos que emergem da fusão entre os interesses humanos e a ordem simbólica que os expressa. Defendemos que o capitalismo não seja um ente autônomo, mas um reflexo — e dos mais autênticos — da intencionalidade humana. O capital (ou o capitalismo) não é um criador, mas uma “criatura”: uma de tantas que concebemos, combatemos e às quais atribuímos agência, sem perceber que, no fundo, apenas estamos a combater o que há de demasiadamente humano em nossas próprias projeções imaginárias ou simbólicas.

com base em dados classificados em tempo real, a interagir com estes em constantes operações ontológicas e respondendo às demandas dos usuários de forma instantânea; abordamos sobre a questão da privacidade sob uma perspectiva na qual esta perderá, gradativamente, a relevância; reiteramos a necessidade da validação e da responsabilização das evidências geradas pelos algoritmos.

Na Secção 3, consideramos como estamos, e defendemos que seja uma liminaridade, que consiste em ser uma dimensão intermediária, indefinida espaciotemporalmente, e dada entre duas outras dimensões que, mesmo que definidas e ainda “presentes”, já não se integram mais como antes (o passado analógico e o futuro digital), o que faz esta liminaridade resultar em uma fase de incertezas mas também de transformação e oportunidades únicas com um futuro mais previsível sobre os impactos tecnológicos.

Portanto, na Secção 3, defendemos a nossa liminaridade tecnológica a se configurar nesta fase de transição; nos apoiamos nos conceitos de Withehead (2010) para nos apropriarmos da sua filosofia do organismo e promover um arcabouço teórico para uma modelagem metaontológica dinâmica e relacional; e, assim, chegamos a uma conclusão que justifica a necessidade de uma metaontologia dinâmica, capaz de integrar, ao menos, as dimensões objetiva, intersubjetiva e subjetiva do utente de uma IA conversacional, bem como a intencionalidade do capital; combatemos a relevância atribuída à opacidade da IA; enaltecemos a importância do design sensível aos valores para garantir que a IA seja desenvolvida de forma ética e responsável, que envolve a participação de diferentes *stakeholders* no processo de *design*, tal como citaremos aplicações realizadas nas experiências práticas realizadas pelo Projeto Signoⁱⁱⁱ, que bem pode servir como *benchmark* ao exposto; ao ser dado como exemplo de desenvolver uma abordagem de design sensível aos valores e centrado nos humanos.

ⁱⁱⁱ Signo. Sistemas de valor em tecnologias digitais para a saúde. Programa: FCT / Carnegie Mellon Portugal – 2022. Ref. 2022.15724.CMU. Instituições participantes: Instituto de Filosofia da Universidade do Porto, Fraunhofer Portugal Research — Center for Assistive Information and Communication Solutions e Carnegie Mellon University. (No sítio <https://ifilosofia.up.pt/projects/signo>, acedido em 19/mar/2025)

E, por fim, apresentamos nossas conclusões e consideração final, com base nos desenvolvimentos dissertativos.

1. Em busca da verdadeira problematização – a historicidade

1.1. Como chegamos até “aqui”?

Nesta primeira parte, reconstituiremos sinteticamente alguns dos passos mais influentes que nos permitirão compreender e buscar uma ou mais posições (ou respostas) em relação aos problemas atualmente levantados acerca das IA.

Iniciaremos por perceber como estamos a analisar os problemas éticos das IA, com base no desenvolvimento histórico do pensamento filosófico “analógico”: poderoso, eficaz, que nos serviu muito bem: nos “trouxe” até aqui e ainda nos servirá. Mas, recorrer exclusivamente a tais formas analíticas nos leva a crer que sejam estas insuficientes para resolverem muitas das atuais questões sobre IA, visto que se baseiam em teorias e causalidades, na epistemologia e na lógica maioritariamente de primeira ordem. Por isso, se mostram insuficientes frente a um cenário eminentemente digital no qual tudo, ou quase tudo, se dá por meio de correlações, sem as consagradas questões causais das teorizações – o analógico e o digital são mundos e dimensões diferentes, portanto, cada qual com os seus respectivos ônus e bônus, mas igualmente resistentes em termos de profunda cognoscibilidade acerca de suas realidades últimas.

Em síntese, exploraremos as distintas relações possíveis de como um utente percebe e se relaciona com uma IA, e vice-versa. Pois, as relações, sejam como causas ou efeitos dos “males” ou dos “benefícios” são sempre necessárias. Para tanto, primeiro adentraremos às explorações conceituais acerca de tais “relações”, a partir de uma hermenêutica epistemológica e ontológica para, então, explorar o conceito fenomenológico da IA do ponto de vista tecnológico-constutivo dos algoritmos, e seus impactos, especialmente os de cunhos éticos.

Assim, precisamos declarar que trataremos dos problemas mapeados (Mittelstadt et al., 2016) e reivindicados, pontualmente, um a um, e integrando-os, se necessário. Apesar

disso, adotaremos uma suspensão dos preconceitos^{iv} qualitativos morais sobre os supostos efeitos ou impactos que uma IA pode (ou ainda poderá) causar à humanidade, como mais comumente têm sido difundidos nas esferas políticas e nas mídias em geral, recentemente, a considerarem debates sobre questões como privacidade, responsabilidade e outras temáticas amplamente discutidas. E faremos assim a considerar uma própria instrumentalidade de uma IA. Para tudo que seja uma nova tecnologia, ou um novo desenvolvimento de aplicações, tanto poderá haver benefícios quanto malefícios. E por isso, nos limitaremos a analisar os malefícios sem, todavia, considerar que a causa possa ser suprimida, pois a IA não é algo que será extinto e, portanto, precisaremos desconsiderar que resistir seja uma opção viável de ser realizada, pois não é, como tudo o mais. Mesmo em termos de legislação, há aí algumas

^{iv} Pois, a mesma suposta capacidade de que uma IA possa, como exemplo, adquirir poderes estratégicos para influenciar comportamentos, tais possibilidades já são (preconceituosamente, talvez pelo medo gerado pela ignorância acerca das mesmas) consideradas maléficas (tal qual implicassem fatidicamente a atuação da IA como a de um gênio maligno).

Deduz-se isto, por exemplo, não apenas ao se considerar certos aspetos de eventuais perdas de autonomia da agência humana, ao ser tolhida pelas IA, em geral, mas também por se ficar à mercê de forças intencionais desconhecidas, ocultas e imprevisíveis.

Deduz-se também, preconceituosamente, acerca da própria liberdade da ação, como se todas as ações humanas e comportamentos sejam mesmo autônomos, em nome da sobrevalorização do que se considera vulgarmente como a autônoma liberdade da agência, ou livre-arbítrio, já a adentrar em outras discussões que não se enquadrariam em nossos objetivos aqui definidos.

Mas, ainda assim, ao se insistir em considerar toda e qualquer suposta influência algorítmica como “má”, há casos (e nem tão específicos) que podem servir de contraexemplos em que esta preconceituosa generalidade possa estar em suspeição.

Podemos citar usar os eventuais poderes influenciadores das IA para estimularem “boas” atitudes (como a abertura à adesão a tratamentos médicos, benefícios à saúde mental, interações motivacionais para melhoria da performance e do bem-estar, intervenções educacionais / instrucionais nos campos éticos ou cívicos, por exemplo).

Ou até mitigar “maus” ou nocivos comportamentos (como nos casos de identificações de crises oriundas de estados mentais patológicos ou mediações em ideias suicidas, por exemplo).

Assim, a mesma suposta capacidade de influência negativa de uma IA seria claramente considerada positiva (ainda mais se aplicada proactivamente, ou preditivamente).

Tudo isto, contudo, é uma outra discussão que também não caberia aqui: sobre os limites e direções de supostas influências algorítmicas.

Por isso, a questão investigativa será mesmo sobre os fundamentos acerca da suposta capacidade de influência de uma IA sobre seus potenciais de causar efeitos, impactos, e não sobre a predicação relativa ou absoluta destes, nem mesmo sobre a sua usabilidade.

complicações já percebidas sobre como se legislar^v sobre as IA. Colaborar é criar ideias para se legislar e conscientizar, prevenir e monitorar, e é nesta direção que iremos.^{vi}

Partiremos de uma revisão conceitual introdutória, inspirada e ajustada a partir de similar consideração argumentativa realizada sobre a hermenêutica digital, pelo filósofo Romele (2019), certos de que dotará os leitores dos elementos críticos necessários, sem pressupor que estes sejam já conhecedores dos conceitos que serão evocados, ou não. Por isso, e pelo intento de cultivar uma interdisciplinaridade, estamos a investir em um apêndice robusto com as devidas complementações e referências aos que pretenderem ampliar os conhecimentos, sem que precisem ter conhecimento em Filosofia. Desta forma, discorreremos sobre os critérios mais relevantes das conceituações que serão pautadas nas páginas vindouras, ainda que panoramicamente, a traçar uma crescente do pensamento filosófico através de uma sintética linha temporal que se dá desde os primórdios das buscas pelo conhecimento. Tais critérios têm correlação com a inata necessidade humana em conhecer, desde sempre, as coisas, em facto e em abstrações racionalizadas.

Não é difícil percebermos isto, a partir das nossas próprias experiências, ações e comportamentos, assim como Vergílio também percebeu: “Feliz aquele que logrou das coisas conhecer as causas”, pois o conhecimento é fomentador da ânsia humana de se atingir um nível (transcendente) sempre para além de atual condição (imane) limítrofe de possibilidades, ou mesmo da sua precária situação finita existencial: e assim estaremos inusitadamente a adentrar (intencionalmente, e ontologicamente) à

^v As discussões estão nas diversas esferas de poder legislativas, judiciárias e executivas, para além da imprensa. O Parlamento Europeu tem feito progressos relevantes, a nível mundial, mas ainda assim muito aquém de se atingir algum patamar consistente de ser seguido. Referência acedida em 19/mar/2025 no sítio <https://www.europarl.europa.eu/topics/pt/article/20200918STO87404/inteligencia-artificial-oportunidades-e-desafios>.

^{vi} O que se buscará, portanto, será um conjunto suposto de elementos característicos – em termos ontológicos e/ou adjetivos – que marcam o ponto de vista, ou a perspetiva, que um certo utente mediano tem (ou pode ter, ou adquirir) em relação a uma IA, seja passivamente, a interagir com as tecnologias de forma indireta ou inconsciente, ou mesmo ativamente, em ação conversacional, quando passa a interagir de forma direta e consciente. Afinal, uma ou outra destas situações podem se alternarem para um mesmo utente e, por isso, não são excludentes entre si.

Epistemologia, uma parte pioneira e relevante da Filosofia que trata, à primeira impressão, do conhecimento e, mais intimamente, trata de atingir tal sociedade humana tão secretamente sonhada: o saber transcendental. E adentraremos, portanto, com o objetivo de perceber como se pode obter o conhecimento a partir de instrumentos, se assim pudermos considerar uma IA, instrumentalmente, em relação ao conhecimento. Estamos a considerar um conhecimento transcendental, pois, não se trata apenas do conhecimento como fim, mas como meio de se atingir outros fins. E a IA, defenderemos, reflete justamente isso enquanto tentativa de ultrapassar nossas humanas limitações.

Tanto que Vergílio atribuiu à tal felicidade do conhecimento o subjugar “sob seus pés todos os medos, o implacável destino e o estrépito do ganancioso Aqueronte”; ou seja, o conhecimento tido como um meio, e não um fim, como visto. Veremos que Foucault também o perceberá assim.

Um fim para o “bem”, na maioria das vezes, seja lá o que este “bem” possa significar em termos práticos, mas que nos trouxe aonde estamos, enquanto humanidade, com progressos inimagináveis em todos os campos do desenvolvimento tecnológico, social, econômico, político, cultural, humano e muitos outros, tais como as atuais e ainda surpreendentes IA. Mas, nem sempre em todas as evoluções tivemos apenas o “bem” como resultado, pois há sempre os efeitos colaterais, ou secundários, e nem sempre desejáveis.

Tal discussão utilitarista (e moral), reiteramos, não será o tema central desta dissertação, todavia, como já anunciado. Não estamos intentar promover ou rebaixar a IA, visto que esta não é um projeto futuro, mas sim um facto, atual, e já inserido em nossas vidas, direta ou indiretamente. A IA é uma resultante dos nossos esforços, presentes desde sempre, desde a descoberta do fogo, ou da invenção da roda, ou do desenvolvimento da agricultura, sem nunca termos parado de nos desenvolvermos instrumentalmente rumo a um *optimum*.

Provavelmente deixaremos frustrados tanto os otimistas quanto os pessimistas das tecnologias, ansiosos por encontrarem pares a corroborarem com suas crenças e unirem forças contra ou a favor das IA. Se há culpa no progresso, é preciso também “culpar” a

Filosofia, visto que esta está a alguns milênios a dar relevantes contributos para escalar o saber. Basta-nos perceber, para isso, que a Epistemologia, veterana da Filosofia, desta forma, foi o primórdio (e o medium) de um movimento (que assumimos desde já como intencional) que culminou em todo o conhecimento que temos hoje, e que não para de prosseguir adiante: veio dela a fundação das Ciências, em suas múltiplas formas de ser, o que não pode ser negado e nem desprezado.

A IA não está alheia a esta progressão histórica, e adveio dela, deste nosso empreendimento humano pelo saber, pelo ultrapassar os limites existentes, e é o que desbravaremos doravante. Assim, na próxima secção perceberemos como a IA se relaciona com esta onipresente busca filosófica: epistemológica e “transcendental”.

1.2. A emergência relacional pela perspectiva filosófica

Que quer dizer tudo isso? Indagou emblematicamente Thomas Nagel no título de seu livro^{vii}, pelo qual ele busca dirimir sobre o que consiste ser a Filosofia, panoramicamente. E, o título do livro é, provavelmente, justamente a questão primordial que desde sempre faz identificarmo-nos como humanos, enquanto racionais: somos questionadores sobre tudo e ávidos pelo conhecimento como forma de transcender às nossas próprias condição e natureza. Mas, tal avidez não é apenas por um “isso”, mas por um “tudo isso”. E é sobre este longo caminho entre o “isso” e o “tudo isso” que iremos reconstruir. E faremos esta reconstrução através de uma hermenêutica histórico-filosófica.

1.2.1. Aristóteles e a emergência do objeto

Desde a antiguidade, e impulsionada pela tradição aristotélica, as primeiras preocupações filosóficas acerca do conhecimento se dirigiram prioritariamente aos objetos: buscava-se conhecer o que eram os objetos em suas manifestações particulares, ou seja, nas coisas, em si, como a madeira e a cadeira de madeira, como exemplo. Buscava-se prioritariamente conhecer o que era o mundo ou universo, como um todo, mas quase que exclusivamente a partir de seus constituintes causais

^{vii} Nagel, T. (1995). *Que quer dizer tudo isto? Uma iniciação à Filosofia*. Gradiva. ISBN 978-972-662-421-9.

“essenciais” ou particulares (Shields, 2007), suas classificações, organizações, origens, princípios, características e muitas outras questões que externavam ao filósofo^{viii}, este a se posicionar como um distinto observador totalmente alheio e independente ao objeto, e ao mundo, completamente deslocado, autônomo e inquisidor do que conseguia perceber sensivelmente, e sem assumir nenhum tipo de mediação, responsabilidade ou protagonismo em relação ao que lhe era (ou parecia ser) meramente “exterior”.

Aristóteles^{ix}, a respeito de sua teoria acerca da matéria-prima mais “nobre” do ser, ou o nível mais alto da sua hierarquia ontológica, tida como a substância essencial, atribuiu à mesma uma determinada forma a considerá-la como um modo de ser, e de existência, como que princípio inalterável e incorruptível em termos existenciais (Shields, 2007, p. 49, tradução nossa). Uma substância essencial seria, para ele, sempre a mesma substância essencial, não importa como esta se apresente, lá ela estará inserida a partir de certas considerações^x.

^{viii} Contudo, desde os clássicos, há a referência de filósofos à alma humana, e suas formas de relevância. Em Aristóteles, suas considerações sobre a alma se deram através de estudos sobre a própria, e também sobre a psicologia, “onde ele se concentra principalmente na alma e suas capacidades de percepção e pensamento” (Shields, 2007, p. 33, tradução nossa). O que está a se estabelecer em relação aos conceitos ao longo da História da Filosofia, pressupõe-se, não é uma linha específica e limitada de temáticas atribuída a cada filósofo citado, mas sim a centralidade das mesmas, em termos de relevâncias, na abordagem conceptual ao longo dos tempos.

Aristóteles, inclusive, pode ser considerado como o precursor (dentre tantas outras ciências, como a própria Psicologia) ou fundador da Ontologia (Shields, 2007, p. 146, tradução nossa), ao estabelecer teorias, definições e classificações estruturadas sobre o ser, a partir de certas categorias e hierarquias estabelecidas nas quais ele pôde representar o conhecimento acerca do ser.

^{ix} “Os seres humanos começaram a fazer filosofia, diz Aristóteles, assim como fazem agora, por causa da admiração, a princípio porque se admiravam das coisas estranhas bem na frente deles, e depois, avançando pouco a pouco, porque passaram a achar coisas maiores intrigantes.” (Meta. i 2, 982b12 apud Shields, 2007, p. 36, tradução nossa).

^x Por exemplo, um homem sempre será um homem, de essência racional e com determinadas formas e, mesmo que suas qualidades possam variar, sempre será um homem. Pois esta substância aristotélica, todavia, pode se apresentar de forma essencial (com o acréscimo ou decréscimo de suas qualidades) e/ou de forma accidental (Shields, 2007, p. 99, tradução nossa).

O que difere ontologicamente tais apresentações substanciais, para um observador, é que sua essencialidade é inerente à identidade da própria substância: a essência formal é algo que a faz ser o que é, de forma indissociável a ela própria. O acidente é tudo o que lhe é supérfluo (como a cor, por exemplo), e que não impacta na sua forma (ou modo) de ser, mas apenas na forma (ou modo) de existir (quando a existência do ser pode se dar de um certo jeito, ou de outro). Ou seja, em grosso modo, os acidentes são seus atributos formais: que explicam, descrevem, diferenciam, mas que não causam, definem, fundamentam ou normatizam.

Aristóteles, portanto, ainda segundo Shields (2007), se dedicou a perceber e a classificar a relação íntima dos objetos, e isto envolvia a priorizar a substancialidade, em uma busca de causalidades. Se há uma causa para algo, haverá de ter uma causa para tal causa? E, tal questionamento o levou a teorizar sobre uma causa primeira que seria a origem de todas as relações causais, pois lhe interessava estabelecer como tudo se relacionava até que algo fosse o que é. Esta causa primeira, pela lógica aristotélica, deveria consistir em termos absolutos, em pleno ato sem nenhuma potencialidade a lhe ser acrescida ou decrescida. Esta causa primeira “era”, de facto, a maior e absoluta dimensão existencial. O ser, portanto, é ontologicamente superior^{xi} à aparência, ao seu próprio parecer ser (Romele, 2019, p. 55). A aparência accidental já era considerada menos importante, portanto, pela prioridade formal que dava à verdadeira dimensão substancial do ser. “Ser” firmou-se como um estatuto ontológico central, desde então. Aristóteles (Lear, 1988) conceptualizou o motor imóvel como causa primeira, que foi apresentado como o que a tudo causa, ainda que indiretamente (pela sua própria imobilidade, por estar em pleno ato e, portanto, levar a tudo o que tenha alguma potencialidade – que não esteja em ato, a se movimentar em sua direção).

^{xi} Poder-se-ia considerar, então, que os atributos eram formas fracas de ser, ao menos para Aristóteles (Romele, 2019, p. 55). Dada a imensa influência aristotélica ou também aos próprios interesses que podem ser subversivos, seletivamente, ainda é assim para muitos pensadores ao considerarem que os atributos sejam menos relevantes e secundários, frente à substancialidade essencial e primária do objeto.

Assim, esta referência primeira em Aristóteles servirá de paradigma, no nosso esboço de traçar uma ontologia ^{xii} da ontologia do conhecimento.

Por que Aristóteles ainda é atual e relevante para a nossa temática de IA?

1. Pelo modal: necessário (essencial, substancial) e contingencial (acidentes, atributos).

A distinção modal entre o necessário (essencial, substancial) e o contingencial (acidentes, atributos) remete a uma tradição filosófica que, em Aristóteles, permitia compreender o que constitui a essência de um ser em contraposição aos seus atributos acidentais. No contexto da IA, essa diferenciação revela-se crucial para interpretar a relação entre o existente e a sua representação. Enquanto os elementos necessários correspondem àquilo que define a identidade e a robustez dos dados – a "essência" subjacente que confere significado e estrutura –, os aspectos contingentes referem-se às características variáveis que, embora influenciem a aparência e a funcionalidade, não alteram o núcleo informacional. Assim, ao modelar a realidade, os sistemas de IA devem considerar ambos os aspectos para garantir que as representações digitais capturem não apenas a substância dos fenômenos, mas também a sua mutabilidade e os detalhes contextuais que enriquecem a compreensão global da "realidade".

^{xii} Na próxima seção, abordaremos o que estamos a considerar sobre o termo ontologia que, em síntese, buscará mesclar tanto o que é considerado na Filosofia quanto nas Ciências das Informações.

2. Pela desenvolvimento da adoção da ontologia como forma de conhecimento, o que será relevante na abordagem que faremos sobre os dados^{xiii} em geral, ao *big data*^{xiv} e, mais especialmente, sobre os metadados^{xv}.

A adoção da ontologia como forma de conhecimento é central para a organização e interpretação de dados na IA, especialmente no contexto do big data e dos metadados. Aristóteles foi pioneiro na estruturação do conhecimento por categorias, e essa abordagem continua relevante para dar significado aos vastos volumes de informação gerados digitalmente. A ontologia permite criar modelos conceituais que estruturam dados de maneira hierárquica e relacional, possibilitando inferências mais precisas e uma melhor compreensão das conexões entre os elementos, o que é essencial para a tomada de decisão algorítmica.

^{xiii} A antecipar o que estará por vir nesta dissertação, a inversão que ocorre a partir da teoria de Aristóteles, de se promover gradativamente os atributos ao protagonismo (e perceber as dinâmicas de como isto se dá) é justamente o que está em jogo desde que se iniciou a mais recente epopeia filosófica para se compreender o que se passa na dimensão digital, ainda mais quando se aborda o tema da IA, que apenas “existe” devido às correlações que realiza entre *inputs* e *outputs*. Pois, para se perceber o *modus operandi* com os dados, em síntese, em toda a sua crescente exponencial quantitativa que alimenta as IA, e que culminou no gigantismo e na incomensurabilidade que faz jus ao estatuto ontológico chamado de *big data*, passou-se a percorrer os dados não pelos dados, em si, mas pelos seus atributos, pelas suas formas relacionais de classificação a partir de certas representações que são feitas, também, com base em atributos, que são chamados de metadados.

^{xiv} Grosso modo, podemos imaginar o *big data*, em seu limite, como o “universo” acessível dos dados:

“Apesar da importância do fenômeno, não está claro o que exatamente o termo ‘big data’ significa e, portanto, a que se refere (...). O termo ‘big data’ se refere a conjuntos de dados grandes, diversos, complexos, longitudinais e/ou distribuídos gerados a partir de instrumentos, sensores, transações na Internet, e-mail, vídeo, fluxos de cliques e/ou todas as outras fontes digitais disponíveis hoje e no futuro. Você não precisa ser um filósofo analítico para achar isso obscuro e vago.” (Floridi, 2012)

^{xv} Metadados são, portanto, os atributos dos dados, ou os seus acidentes, que, atualmente, e opostamente às teorias de Aristóteles, passaram a definir os dados, mas que não os explicam, necessariamente. Pois, para explicar, tal qual preza a epistemologia formal, é preciso não apenas correlacionar, mas perceber as relações causais existentes e prová-las como verdadeiras, justificando-as. E não há, quase sempre, nas correlações que geram os *outputs* das IA, uma relação de causalidade, como veremos mais à frente. *Outputs* são meras replicações do que se está a ser considerado como o mais provável, dentro de tantas possibilidades de probabilidade existentes dentre os dados disponíveis.

3. Pela atualidade digital, que está a promover a primazia dos atributos e, com isso, a amplificar^{xvi} massivamente antigos problemas éticos.

Pois a era digital tem promovido uma inversão (e até mesmo uma subversão) na hierarquia entre essência e atributos, privilegiando características superficiais e contextuais em detrimento da substância. Algoritmos de IA categorizam e interpretam dados com base em padrões derivados de atributos contingentes, muitas vezes ignorando a profundidade ontológica dos fenômenos que representam. Essa primazia dos atributos intensifica problemas éticos já existentes, como viés algorítmico, despersonalização e manipulação da informação. A falta de uma abordagem ontológica mais rigorosa na IA pode levar a decisões injustas e à perpetuação de desigualdades estruturais.

4. Pela superação dos parmenidianos (na secção a seguir), visto que a IA se dá a partir do dinamismo^{xvii} das (cor)relações e, por isso, as análises estáticas falham.

^{xvi} Há problemas práticos derivados de tal primazia da substancialidade em detrimento de seus atributos, todavia, sobre as aplicações destas premissas hierárquico-classificatórias. Tais como quando cada humano percebe claramente o que seja um humano, por exemplo, mas alguns destes humanos se percebem mais humanos do que outros, o que configura a base do racismo, que é um dos mais relevantes problemas éticos atuais. Nestes casos, os atributos dos racistas, como a cor da pele, ficam assim mais próximos de uma essencialidade particular na qual julgam fazer parte, do que da acidentalidade geral. Há sempre, portanto, a possibilidade da ocorrência de uma subversão seletiva, quando conveniente a certos grupos ou perspectivas.

E, por assim dizer, há diferentes valores, pesos e medidas também para cada essencialidade e acidentalidade, para diferentes juízos, em que tudo precisa ser (e será), então, convencionado. O que não quer dizer que ficará resolvido, todavia. Logo mais, ainda neste capítulo, abordaremos tais aspetos em Kant. Assim, o que buscamos acusar é que a relação das categorias primárias e secundárias, em níveis de representações hierárquicas, é conflituosa e difícil de explicar pelas dificuldades de se perceber os limites entre elas. E é o que já preocupava Aristóteles, provavelmente, ao fazê-lo considerar na conceituação da *anima* as formas psicológicas de percepção e impressões.

^{xvii} Ao final, cada um percebe conforme seus próprios interesses, capacidades, conveniências ou ideologias. Ou pelas crenças, desejos ou vontades, por exemplo. A relação interna das qualificações, meramente mentais ou empíricas, sem que existam referenciais externos, como padrões convencionados ou modelos estabelecidos, é ainda mais problemática. Por isso, nestes labirintos epistemológicos, desde sempre, as relações foram aliciadas a ficarem subordinadas a algo imutável, fixo e sempre presente como referencial. O valor passou a estar na referência “segura” e “confiável”, dada pelas supostas estabilidade ou imobilidade.

Os sistemas de inteligência artificial não apenas analisam dados isolados, mas identificam padrões dinâmicos, ajustam-se a novos contextos e evoluem com o tempo. Assim, compreender a IA exige um modelo filosófico que reconheça o caráter relacional e transformador da realidade.

1.2.2. A emergência dos parmenidianos

Para que possamos fundamentar uma crítica sobre as formas inconsistentes de como os problemas com as IA estão a serem desenvolvidos, formaremos um conjunto com todos os elementos teóricos que remetem à imobilidade nas diversas teorias que apresentaremos aqui. Assim, o motor imóvel de Aristóteles será deslocado, por enquanto, e ainda isoladamente, para este conjunto. Outros elementos logo lhe farão companhia, também por terem uma mesma função referencial de imobilidade, e na centralidade fundacional das teorias das quais participam, visto que se fazem necessários para que todo o resto da teoria se mantenha consistente.

Agruparemos todas estas referências imóveis, uma a uma, neste conjunto inaugurado pelo motor imóvel, e que bem poderíamos denominar como os dos parmenidianos, em justa homenagem ao suposto precursor (ou promotor) da imobilidade na Filosofia: Parmênides^{xviii}.

1.2.3. Descartes e a emergência da subjetividade

^{xviii} A fundamentação filosófica de Parmênides está na disparidade entre a "via da verdade" (*aletheia*) e a "via da opinião" (*doxa*). A primeira conduz à compreensão do ser verdadeiro (que ele qualifica como uno, imutável, imóvel e indivisível) – e daqui tiramos a nossa "homenagem" a ele. Parmênides desenvolveu a teoria de que o movimento e a mudança iriam pressupor uma espécie de não-ser, o que seria inconcebível para ele, pois implicaria em algo deixar de ser o que era para se tornar algo diferente: o que, obviamente, não era. E como o não-ser era inconcebível e inexistente para Parmênides, concluiu que o movimento e a mudança seriam ilusórios, pois nunca ocorreriam, de facto. A imobilidade do ser foi deduzida como uma consequência lógica de sua unidade e imutabilidade. O ser deve ser "todo igual", sem quaisquer diferenciações, e, portanto, não pode se alterar ou mover. A "via da opinião" é a compreensão ilusória da realidade, meramente empírica e perdida na multiplicidade incognoscível do mundo sensível. (Heidegger, 2008)

Quase dois milênios após Aristóteles^{xix}, René Descartes (2005) consolidou a transição de foco que decorreu progressivamente desde então, no limiar em que a consciência passou a assumir alguma relevância existencial e relacional, subjetivamente, a alterar o protagonismo epistemológico do objeto para o do sujeito, a nível psicológico, ou mental, como considera Newman (2023):

A certeza/indubitabilidade para Descartes é de caráter psicológico, embora não meramente psicológico – não é simplesmente um sentimento inexplicável. Ela também tem um caráter distintamente epistêmico envolvendo um tipo de insight racional. Durante momentos de certeza, é como se a percepção fosse guiada por “uma grande luz no intelecto” (Med. 4, AT 7:59, CSM 2:41).

Descartes assim considerou, ao adotar a dúvida (visceralmente existencial) como um método de questionamento investigativo, até mesmo da própria existência do cogito subjetivo questionador. A dúvida foi, então, assim legitimada ao permitir que o cogito “continuasse” a existir mesmo que, radicalmente, nada mais pudesse existir para além dele próprio: ou seja, a teoria de Descartes até ponderou considerar, extremadamente, que tudo o que “existe” seria apenas a mente pensante, e nada mais (Fodor, 1980), embora ele próprio tenha combatido, ainda que de forma inusitada, o “mal” radical que sua teoria trouxe, a partir de um dos conceitos (mesmo que oriundo de efeito colateral) que tem sido mais altamente combatido^{xx} e desprezado em toda a História da Filosofia: o solipsismo.

^{xix} Nesta nossa primeira exposição e nas demais, há que se considerar que os *gaps*, tal como supostamente ocorrido entre Aristóteles e Descartes são meramente ilusórios. Apenas existiriam se, obviamente, nada mais tivesse sido teorizado entre os dois momentos filosóficos considerados, o que não ocorreu, pois foram muitos os filósofos a contribuírem nestes aparentes gaps. Portanto, não é um facto evocado a existência de tais gaps, mas apenas uma ilusão que possa vir a ocorrer.

^{xx} E, progressivamente aos pensamentos aristotélicos, cartesianos e todos os demais, a Filosofia evolui, como tudo, embora nem sempre assertivamente, em termos de avanços incontestáveis. Pois, há as funcionais e necessárias objeções que surgem e, assim, a cada nova proposta ou consideração, a complexidade acerca do entendimento se incrementa, ao ponto de, por vezes, parecer que está mais a se afastar de alguma conclusão do que a se aproximar da mesma. E tudo se dá tal como foi com Descartes, que trouxe avanços aos pensamentos filosóficos, mas também deixou graves efeitos colaterais após o escrutínio de seus objetores, que não foram poucos. E isto é peculiar a todas as propostas de teorias inovadoras, e não é algo necessariamente deletério aos filósofos. Os problemas ocorrem, a partir de nossa

Não seria precipitado supor que a tentativa realizada por Descartes foi, em última análise, isolar a mente de todas as relações externas, exceto com ela própria, sem buscar perceber como tais relações se dão e, por isso, ao não conseguir resolver o problema de uma dinâmica incontrollável e imprevisível, ou melhor, caótica, buscou um referencial fixo para tal^{xxi}. E isto, em síntese, culminou no mesmo problema aristotélico de considerar apenas as relações internas aos objetos, sem conseguir estabelecer como tais relações se davam^{xxii}, a partir delas próprias, que o fez buscar uma causa primeira

hipótese, por serem também devido aos factos de que as relações nas teorias ou são geralmente desconsideradas ou assumem uma posição marginal.

Os problemas da negligência das relações não ocorrem apenas na Filosofia, mas também nas investigações matemáticas, segundo Eugenia Cheng (2020, cap. 6, tradução nossa):

“(...) vimos como é crucial considerar os sistemas como um todo, com todas as suas interações, em vez de apenas pessoas ou eventos isolados. A ideia de considerar as coisas umas em relação às outras é um dos princípios básicos da matemática moderna. Mas nem sempre foi assim, pois somente nas investigações relativamente mais recentes, a partir de meados do século XX que as colocaram na vanguarda. Vemos que observar como as coisas ou pessoas se relacionam é muitas vezes a chave para compreender uma situação, em vez de olhar para as características intrínsecas dessas coisas ou pessoas. Isto é verdade em muitos níveis e em diferentes escalas, desde a forma como os países interagem no mundo até à forma como as pessoas interagem nos relacionamentos.”

^{xxi} Sobre tais efeitos colaterais ou dialéticos que se agregam aos anteriores, todos oriundos de cada nova teoria filosófica apresentada, como o já citado solipsismo cartesiano, por exemplo, que fez concentrar o foco na subjetividade para resolver os problemas existentes. Faz parecer que uma suposta substancialidade da coisa em si (ou o “x” da equação, alegoricamente) fica sempre mais distante quando se está a se movimentar em sua busca. Percebe-se a priori, abstratamente, por vezes, mas nunca de forma satisfatória, saciadora. Pois faz parecer que o abstrato racional resultante é sempre um esforço filosófico inerente ao entendimento idealizado como estático, acerca de um mundo que é puro movimento e transformação, por vezes de aparência caótica, a quebrar qualquer teoria linear paradigmática que se queira estabelecer sobre o mesmo.

As relações são geralmente subvalorizadas, evitadas e negligenciadas, pois são instáveis, não tão fáceis de serem ontologizadas. Estaremos, todavia, a perceber que os problemas atuais em IA são essencialmente relacionais: de não se perceber que (e como) “tudo” se dá devido às relações existentes entre suas “partes”. É como se fosse uma negação da circularidade pelo mero desprezo a ela, como se assim fosse desaparecer o problema da circularidade. Mas, desprezar algo não implica necessariamente em sua extinção, todavia, pois o desprezo dado não torna algo desprezível, necessariamente. E, se o torna, não o anula a ser o que é, a atuar como atua: a velha questão entre a necessidade e a contingência, portanto.

^{xxii} E, assim, passa-se a se considerar um todo sem suas partes todas. Um “todo” que não é um todo, e partes que não são suficientes para justificarem o “todo”. E aí está a faceta mais circumspecta da circularidade, a mesma que tem levado a hermenêutica às críticas constantes e a um papel que, defenderemos, deveria ser mais relevante, e não como problema, mas como protagonista. A circularidade evidenciada e exposta, tal como na hermenêutica, portanto, poderá vir a ser um de seus maiores méritos,

externa, dando vida ao seu equivalente funcional do motor imóvel aristotélico. Descartes logo tratou de consertar o seu problema, mas sem precisar criar “algo” novo, pois usou de um conceito já existente e o fixou não como causa primeira, necessariamente, mas como relação primeira, todavia.

Há, portanto, uma limitação para uma releitura atual especificamente sobre as IA, tanto na filosofia aristotélica quanto na cartesiana, que é a tendência de focar nas propriedades internas dos objetos sem compreender plenamente como suas relações externas ocorrem e influenciam sua dinâmica. Vimos que Aristóteles analisava os entes a partir de suas essências e acidentes, mas sua teoria carecia de um mecanismo que explicasse as interações entre eles sem que precisasse recorrer ao motor imóvel. Da mesma forma, Descartes, ao buscar um referencial fixo para a mente, inicialmente isolou-a do mundo externo sem considerar plenamente como essa relação se dava. Assim, ambos acabaram criando soluções que dependiam de um princípio externo e absoluto para sustentar a coerência de seus sistemas.

E assim fez Descartes, ao dramaticamente buscar deus como referência relacional e diferenciada do cogito. O deus cartesiano é diferenciado significativamente do motor imóvel aristotélico, mas nem tanto assim, em termos funcionais aos conceitos apresentados^{xxiii}. Então, no conjunto dos parmenidianos, o deus cartesiano junta-se ao até então solitário motor imóvel aristotélico.

e não o oposto, como veremos: “é importante examinar as interações como círculos viciosos” (Cheng, 2020, tradução nossa).

^{xxiii} Os problemas cartesianos são os mesmos das teorias aristotélicas, ao não focarem nas multiplicidades das relações sem que estas possuam referenciais externos imóveis. A diferença em Descartes – e, cremos, um progresso interessante – é que este alterou o domínio investigativo do objeto para o sujeito – emergiu a subjetividade, afinal, pelo cogito. Assim posto, e a passar direto pelos ainda mais problemáticos efeitos cartesianos gerados (e que parecem mais “quentes”, talvez justamente pela maior aproximação com a fonte do problema), pelo seu cogito, reificado, ainda que através de um indigesto dualismo derivado, também derivou outros problemas para além do solipsismo, mas que não serão aqui abordados, por transcenderem aos objetivos dissertativos assumidos como centrais.

1.2.4. Kant e a emergência da intersubjetividade

A corroborar destacadamente com todos os esforços para uma evolução filosófica, surge Immanuel Kant (2008), que pode ser considerado, ao menos nesta nossa jornada, como o mais relevante divisor entre o foco filosófico orientado ao objeto para o foco no sujeito, quando apresenta os conceitos de númeno e fenômeno, a expandir e a ultrapassar completamente o pensamento cartesiano para além do psicologismo cético-radical cartesiano:

Tese da "objetividade dos objetos e conteúdos": nem os objetos nem os conteúdos de nossas cognições têm identidades que são esgotadas pelos atos individuais de representação que os envolvem, de modo que tanto os objetos de nossas cognições quanto as maneiras de se relacionar com tais objetos (ou seja, os conteúdos de nossas cognições) têm uma posição que não é (em geral) dependente de nenhuma de nossas atividades mentais reais. (Tolley, 2012, tradução nossa)

A impossibilidade kantiana atribuída ao conhecimento da coisa em si – do númeno, portanto, é relevante, pois reduz e atualiza as verdadeiras possibilidades de se conhecer algo (existente, atual) e dota o sujeito da responsabilidade (e exclusividade) de ele próprio ter as estruturas racionais mentais suficientes e necessárias para perceber a coisa que se apresenta sensivelmente para si, ao próprio sujeito capaz de conhecer: não mais como a busca do númeno, mas sim do fenômeno. Começa-se a dar, portanto, significativa relevância atribuída às relações, desde então. A partir de uma hermenêutica contemporânea, se um objeto existe (e é), portanto, tal existência kantiana se dá exclusivamente devido ao sujeito percebê-lo de tal e tal maneira, muito mais similar à Aristóteles, pelas essencialidades e acidentalidades das formas, do que em relação à Descartes, pela autoconsciência (Wood, 1975)^{xxiv}. O autor, embora não nomeie ou classifique a hermenêutica considerada, nos leva a crer que se refira a uma hermenêutica fenomenológica, a mais próxima à tradição de Heidegger e Gadamer. Pois

^{xxiv} “Kant rejeita a concepção cartesiana de autoconsciência porque ele sustenta que pensar não é uma auto-intuição intelectual espontânea, mas uma atividade que conecta as representações (sejam de origem interna ou externa) que são dadas em uma intuição sensível e receptiva.” (Wood, 1975, p. 603, tradução nossa).

é a que enfatiza que o conhecimento não é uma captação direta da realidade objetiva (como no realismo clássico), mas uma interpretação mediada pelo sujeito e suas estruturas cognitivas. Essa abordagem se alinha à ideia kantiana de que só podemos conhecer os fenômenos (e não os númenos) e reforça a primazia das relações subjetivas na constituição do significado, contrastando com o racionalismo cartesiano e aproximando-se da ontologia aristotélica das formas e acidentes. Portanto, o protagonismo existencial-epistemológico do objeto dá lugar, assim, ao protagonismo das relações subjetivas.

Mas Kant (2008) não se ateve apenas aos objetos, e nem poderia. Surgem sempre novas questões (ou conjeturas) a cada teorização. Neste caso, todavia, foram acerca da multiplicidade de sujeitos: Como é possível que algo que um sujeito conheça possa ser o mesmo (exatamente, em sentido de identidade) que um outro sujeito também conheça, visto que as percepções podem se diferir entre si, assim como juízos e capacidades? Ainda que sejam comuns a todos os seres humanos, há que se verificar tal comunhão perceptiva, pois, ainda que não divirjam entre si, há que se perceber que sejam convergentes. Pois não há, à priori, formas de “aferição” ou “equalização” possíveis das subjetividades, afinal, para dá-las como ajustadas, garantidamente. Kant (2008) resolve (ou ao menos se esforça para resolver) tal questão ao recorrer a duas conceituações distintas, mas conexas entre si, relacionalmente, o que já é um imenso avanço no filosofar. A primeira é sobre a objetividade. A segunda é, ainda que indiretamente, sobre a universalidade. E, como visto, e dito, quanto mais se avança, mais complexidade há e mais longe parece se estar de alguma cabalidade, mas, neste caso, Kant insere algumas relações que trarão luz ao entendimento. (Tolley, 2012)^{xxv}

A objetividade (o que “existe”, de forma significativa), portanto, a partir de uma leitura mais atenta, ou menos ingênua, passa a ser entendida como uma resultante de

^{xxv} “O que Kant parece ter em mente ao enfatizar a ‘objetividade’ das cognições é o facto de que os objetos que elas ‘têm’ (ou são intencionalmente direcionados para) são objetos que estão disponíveis para serem objetos de outros atos de cognição também. Agora, se isso fosse verdade, então, uma vez que um objeto de cognição em si não poderia ser idêntico a nenhum dos atos de cognição que o representam, seguir-se-ia que Kant endossa a distinção ato/objeto (...) (Tolley, 2012, p.9, tradução nossa)”

subjetividades que se relacionam e convencionam certos resultados como tais, tais como citamos na parte sobre Aristóteles. As coisas kantianas acabam por ser, simplifadamente, o que são, por serem assim convencionadas entre os que as experienciam e as racionalizam, conjuntamente ou não. E isto é algo que substitui as tais das “aferições” ou “equalizações” inviáveis das subjetividades. A subjetividade, portanto, não é considerada tal qual um artefacto, que pode ser operada controladamente. Há uma clara separação entre a objetividade e a subjetividade, em Kant, diferentemente do que havia em Aristóteles, em sua anima, sem que este conseguisse separar claramente os limites entre a essencialidade e a causalidade. Assim, tais convenções sobre a objetividade passam pelos processos subjetivos kantianos, que consistem, mais ou menos, em um crivo racional das categorias do conhecimento, que são dadas pela sensibilidade através das experiências citadas, e tudo a priori (Tolley, 2012, p.10, tradução nossa).

Há uma inegável estética nesta teoria kantiana fundada sobre a racionalidade que a tudo pode se aplicar, até mesmo sobre a estética, para além da ética e da própria epistemologia. Pois,

“o intelecto não se move sem o desejo, sem o exercício da faculdade desejante. Guiamo-nos pelo desejável, procuramos, através do pensamento, identificar-nos com ele. (...) O telos, o objetivo final do pensamento é a intuição intelectual, uma figura, como vimos, do pensamento não-discursivo” (Wood, 1975)

O conceito do a priori kantiano é mesmo sedutor, pois eleva a racionalidade a um nível quase sobre-humano, a fundir-se com o desejável, ou mesmo com o divinal, ao ser capaz de atingir um entendimento cabal. Mas, ainda que seja algo mesmo apriorístico, continua a ser uma convenção que requer relacionamentos para que “algo” seja nomeado ou referenciado como “algo”.^{xxvi}

^{xxvi} Pois, para cada subjetividade há também certos relacionamentos estruturados e necessários, obviamente, visto que não há, em Kant, admissão da ocorrência do refutado solipsismo cartesiano: “o solipsismo de uma razão subjetivamente centrada deve ser dissolvido nas estruturas intersubjetivas da razão comunicativa” (Keul, 2002).

São as relações (a se firmarem como nossas protagonistas conceituais) das subjetividades que resultam em algumas formas de significação acerca (da objetividade) do mundo, e das coisas, a partir de tais interações. Isto, então, nos leva à intersubjetividade, que é esta espécie de convenção compartilhada e aceita consensualmente entre as diversas subjetividades que se relacionam acerca de algo, ou melhor, acerca das próprias subjetividades que objetificam o mundo.^{xxvii}

A objetividade kantiana com tal, portanto, não poderá estar conclusa antes destas relações intersubjetivas (que operam como significadoras), pois o status objetivo resulta daí, como tal, visto que é a intersubjetividade que dá vida e garante a subsistência aos significados. Assim posto, a subjetividade kantiana requer e pressupõe, em toda a sua Filosofia, uma capacidade intrínseca de perceber na outra subjetividade, na alteridade, e o que há de comum em termos intencionais, através dos caminhos racionais e, portanto, supostamente isentos de males.^{xxviii}

O momento intersubjetivo mostra-se na declaração de Kant no início da Crítica da Razão Pura (2018, A XI, p. 5) de que “o juízo amadurecido da época, que já não se deixa seduzir por um saber aparente”, ou seja, sem significados, pois complementa que sua proposta constitui dar primazia à razão para “de novo empreender a mais difícil de suas tarefas, a do conhecimento de si mesma e da constituição de um tribunal que lhe assegure as

^{xxvii} Mais recentemente, os filósofos da linguagem, dentre eles Gottlob Frege, já referenciado por Ricoeur, chamariam, por outros caminhos, que tais processos seriam parte da significação linguística, por exemplo. Pois a linguagem está intrinsecamente associada às questões intersubjetivas das significações, mas não apenas, mas relevantemente linguísticas. Assim, “a condição de possibilidade de compreensão pode agora ser entendida concretamente como a condição para uma comunicação genuína.” (Keul, 2002, tradução nossa). Um tema que logo voltaremos a abordar na secção 2.

^{xxviii} Pois a racionalidade kantiana é revestida (ou percebida) com certo nível de pureza não-infetável (ou até mesmo com poderes assépticos, desinfetantes) no exercício da intersubjetividade. Embora Kant não recorra diretamente ao termo intersubjetivo, esta presunção do processo intersubjetivo é central em sua Filosofia que sempre aspira ao comum, ao compartilhado e consensual, ao melhor que é originado pela racionalidade em seu grau máximo e, portanto, ao bem, que passa a ser ao mesmo tempo imanente e transcendental, pois é antes de tudo, o racional intrínseco e extrínseco. E a racionalidade humana é um atributo apresentado como o melhor que emerge de cada subjetividade interativa pelo senso comum colmatado pelas mesmas.

pretensões legítimas e, em contrapartida, possa condenar-lhe todas as presunções infundadas”. O que seria, afinal, um tribunal para além da busca de um consenso compartilhado acerca de algo?

Como escreveu Kant (2018, B766, p. 596), na *Doutrina do Método*:

“É mesmo sobre esta liberdade que repousa a existência da razão; esta não tem autoridade ditatorial alguma, mas a sua decisão outra coisa não é que o acordo de cidadãos livres, cada um dos quais deve poder exprimir as suas reservas e mesmo exercer seu veto sem impedimentos”^{xxix}.

E, por resultante do senso comum tido como incontestado, a própria dimensão formada e compartilhada pelas subjetividades, passa a ser percebida como um universal. Ou melhor, tal universalidade passa^{xxx}, ou precisa passar, necessariamente, pelo crivo do senso comum. A universalidade em Kant, embora pareça independente e autônoma de quaisquer restrições, só é assim “reconhecida” quando submetida às relações de intersubjetividade^{xxxi}.

^{xxix} A culminação da ideia é encontrada na Seção 40 da Crítica do Julgamento, onde Kant introduz o *sensus communis* como um “sentido comum a todos” que introduz as “máximas da compreensão humana comum”, desdobradas na sua aplicação prática na Antropologia. (Keul, 2002, tradução nossa)

^{xxx} Ainda que a universalidade também não tenha sido ostensivamente defendida por Kant (Oberst, 2015), mas assim percebida pelos seus comentadores e leitores, tais como Hegel, que a replicou em quase tudo o que escreveu, como na dialética, na historicidade e na fenomenologia. Esta percepção da universalidade pelos comentadores e leitores é, ela própria, uma exemplificação do que Kant considerou sobre os processos intersubjetivos. E é uma relação interna do sujeito, reduzida a este, ao final.

^{xxxi} A universalidade derivada de Kant, portanto, é, antes de tudo, um status ontológico, ou melhor, metaontológico. Mas não é exatamente assim que seus leitores a percebem, de forma geral, visto que esta assume progressivamente um protagonismo e autonomia relevantes, como se estivesse sempre a existir como númeno e, algures, se tornasse fenómeno, messianicamente, através de tais processos racionais e intersubjetivos, tais quais estes fossem o medium da própria universalidade se manifestar. Ou talvez a convenção de universalidade apenas tenha “evoluído”, e se transformado no tempo. De alguma forma, a universalidade passou a ser reificada e, depois, deificada. Ainda assim, e friamente, para “tudo” o que há, e mesmo para a “revelação” da universalidade, é necessário primeiro passar pelos crivos das subjetividades e intersubjetividades, pois tudo que seja dotado de significação resulta deste processo de convenção do conhecimento que, segundo o próprio Kant, é praticamente limitado aos fenómenos experienciados. Há uma circularidade aí, tal qual na hermenêutica, e isto requer uma atenção especial.

Mesmo assim, apesar de Kant ter apelado ao acordo entre cidadãos livres como condição da racionalidade, ele não consegue abrir caminho para a conceção plena da intersubjetividade. A sua verdadeira preocupação é com a harmonia das diversas faculdades da mente e, portanto, com a relação do sujeito consigo mesmo, e não com a condição de possibilidade de acordo entre os sujeitos. É uma

Assim, mesmo quando Kant apelou para um “sentido comum a todos”, ele na verdade não tem em mente mais do que um experimento mental pelo qual nos colocamos no “lugar de qualquer outro homem” ao “comparar o nosso julgamento com o julgamento possível e não com o julgamento real dos outros”. No entanto, não estamos falando de um ato puramente privado. Afinal, o objetivo é garantir uma espécie de universalidade para o julgamento do gosto, que se baseia na “comunicabilidade da sensação”. Tal capacidade deve apontar na direção da intersubjetividade. (Keul, 2002, p.257, tradução nossa)

A universalidade seria, portanto, uma parte, e não o todo, afinal^{xxxii}. Mas a indução à universalidade é indubitavelmente o substrato oriundo de uma fragilidade da teórica kantiana, tal qual o solipsismo cartesiano, apesar de toda a fé depositada na racionalidade humana. A questão não é perceber nem condenar as fragilidades teóricas, mas sim encontrar o ingrediente (ou talvez mais de um) que falta às teorias em geral (que já nem será *spoiler*: a desconsideração do protagonismo das relações^{xxxiii}), para que

espécie de consenso interno que ele tem em mente e que, por sua vez, está ancorado na teoria da correspondência. Assim, “a verdade depende da concordância com o objeto e, em relação a ela, os julgamentos de todo e qualquer entendimento devem, portanto, estar de acordo uns com os outros (*consentientia uni tertio, consentiunt inter se*).” (Keul, 2002, p.256, tradução nossa).

Por tais razões, e fragilidades, surgem questões provocativas como estas: se, para se considerar a ocorrência da universalidade há a necessidade de algum consenso prévio a ser estabelecido, o que impediria de levá-la a ser considerada como tal, mesmo que sob o risco de algum determinado consenso ser fruto de uma alucinação coletiva acerca de algo? Ou seja, é possível alguma “universalidade” surgir de um grupo (ainda que parcial e isolado) que alucina coletivamente?

Pois, o que se está a evocar em relação às intersubjetividades são meramente relações que se dão e que, a partir destas, significam as objetividades mais gerais e supostamente mais necessárias. É válido afirmar que para quaisquer relações há sempre um domínio no qual estas se dão. O domínio, neste caso, é a própria objetividade que encerra, em si, a universalidade assim considerada, recursivamente.

^{xxxii} Pois também, em tais domínios, supostamente há sempre um conjunto finitos de participantes, visto que existem limitadores existenciais objetivos a serem considerados: e os humanos são um conjunto limitado de existentes, de total inferior a nove mil milhões, atualmente (Acedido a 06 de agosto de 2024 no sítio <https://www.worldometers.info/br/>). Uma universalidade “existe” a partir e em função deste total de humanos, se assim considerada. Mas, tais humanos nem mesmo são universais, pois em seus particulares existem numa ínfima fração de tempo e de espaço de um suposto todo ainda incomensurável, indecifrável e desconhecido. Se aplicado um rigor lógico, a conceção da universalidade não passa de um processo meramente indutivo dado na parcialidade do que existiu e existe, e ainda assim na parcialidade do que se pôde e pode observar. Obviamente que há recursos argumentativos tanto a favor quanto contra o “absurdo” de tais hipóteses relativistas, quase niilistas, como também para todas as demais.

^{xxxiii} E se tal ingrediente fossem as relações? Pois as relações são partes relevantes (quicá indispensáveis) da vida, da existência e, conseqüentemente, deveriam fazer parte de qualquer teoria (enquanto formas representacionais) sobre a realidade. A paralisia não é parte de nada atual e até pode ser considerada

se possa adicionar tal ingrediente na dose acertada para o que está a se buscar atualmente, no digital^{xxxiv}. Esse ingrediente, defendemos, será uma boa dose de relações dinâmicas e contextuais, em vez de uma busca abstrata por universalidade ou certeza. No contexto digital que se deseja atingir, com valores maioritariamente positivos e controle de riscos, isso significa melhorar as ontologizações e as correlações entre dados, pessoas e sistemas para que haja uma acertada representação da realidade e do conhecimento.

Os conceitos dos estáticos (por exemplo, os do conjunto parmenidiano, que já conta com o motor imóvel em Aristóteles, o deus em Descartes e agora a universalidade em Kant), quando introduzidos em teorias que deveriam ser relacionais podem levar ao erro, ou a extremos absurdos, que decorrem de algo que é conceitualmente colmatado para ser um penso, que até pode ser bem útil para pequenas lesões, mas que não é adequado para uma fratura exposta que exige bem mais do que isso em termos

uma abstração, pois tudo está a se transformar, a se criar ou a se destruir, e nada está em absoluta paralisação ou imutabilidade. E, assim inferimos que as relações devem (ou deveriam) ser percebidas e teorizadas como tais, circulares, recursivas, como pressupostos das transformações, da ação espaciotemporal, mas não apenas. São também intrinsecamente instáveis, alteráveis e elas próprias alteradoras da realidade. São, kantianamente, a universalidade da incerteza. Por isso as formulações teóricas sempre evitam lidar com a imprevisibilidade e a instabilidade, e preferem considerar até mesmo o problema da indução da universalidade como aceitável, frente à incerteza relacional.

Mas, recentemente, ao se buscar formular teorias acerca dos dados digitais, nestes incomuns tempos informáticos nos quais as correlações passam a ter maior relevância frente às causalidades, e assumem-se como *outputs* (que são decisões, em si, ao se considerar uma eventual agência para as IA), e que, progressivamente a impactarem as massas, logo poderão ser assumidas, como consequências, ou como “universalidades”. E tal processo leva-nos não apenas ao medo (originado pela impotência frente a tal potência), mas à certeza de que há mesmo uma reontologização do mundo, e isto parece até mesmo sensível em certos aspetos, muito facilmente observável no que diz respeito às polarizações político-ideológicas que passaram a ser percebidas em muitos países, ou na promoção de personalidades instantâneas antes tidas como bizarras, ou mesmo ideias absurdas ou anticientíficas, e muito mais que antes era absurdo e socialmente inaceitável e que passou a ser “normalizado” pelas direções dos algoritmos a distribuírem conteúdos, mas que não serão abordadas aqui, pois insistimos em manter o foco no problema em si.

^{xxxiv} Para variar em relação ao exemplo do *pharmakon*, é suposto que uma receita culinária com os ingredientes certos fica mais saborosa e melhor digerível. E pode-se “errar” uma receita tanto por se colocar ingredientes a mais, quanto a menos, em termos quantitativos como qualitativos. A teoria da intersubjetividade, se fosse viável sem a referência à universalidade, seria muito menos indigesta, tal qual o solipsismo cartesiano sem a referência existencial de deus.

reparadores. Teoricamente, se há algo estático nos lugares das relações, há fragilidades evidentes na estrutura teórica e que levará à fissura da mesma pela imobilidade de alguns de seus componentes que deveriam ser dinâmicos, e por mais que a teoria seja brilhante acabará por sucumbir. Pode até se ajustar por algum tempo, mas logo às primeiras críticas, acabam por se tornarem inconsistentes, ultrapassadas ou até gerar novos problemas. Pior do que isso, é quando não há qualquer consideração acerca das relações devidamente representadas nas teorias, e é aí que há problema ainda mais grave.^{xxxv}

“Por esta razão, a intersubjetividade da comunicação no discurso cotidiano é constantemente perturbada. Na medida em que existe, é porque o consenso é sempre teoricamente possível. Mas na medida em que o consenso continua a ser um requisito, a verdadeira intersubjetividade falha-nos.” A ação comunicativa genuína entre os dois parceiros só seria realizada se cada um

^{xxxv} Poderia haver uma escala que pudesse dar a dimensão do problema de um modelo teórico, aqui proposto meramente de forma especulativa, que vai de:

- (1) insuficiente (teorias que desprezam relações),
- (2) frágil (considera alguma relação),
- (3) inconsistente (sustentável, que além de considerar alguma relação, ainda se sustenta com parmenidianos) e
- (4) consistente (autossustentável, sem referências parmenidianas cruciais, e calcada com cunho circular, recursivo, atemporal e totalmente relacional).

Uma proposta viável para ultrapassar tais limitações dos parmenidianos passaria, também inferencialmente, por encontrar meios para se representar o movimento transformacional baseado em relações em modelos ontológicos de conhecimento viáveis, que poderiam resultar em novas teorias elucidativas e causais. E tal proposta, se viabilizada, parece ser uma questão primordial para a resolução dos atuais problemas oriundos da IA, principalmente os de cunho ético, que logo estaremos a abordar.

No caso de Kant, um indesejado efeito causado pela fragilidade de sua impressionante teoria seria descambar em um relativismo completo, visto à possibilidade (danosa) de uma multiplicidade de subjetividades a significarem o mundo de forma possivelmente desconexa. Nestes tempos digitais, nos quais a prática da desinformação é tanto presente como preocupante, isto (exatamente) ilustra um dos incríveis desdobramentos de tais contraexemplos quando grupos de pessoas (que, inclusive, passam a se organizarem em grupos antagônicos e com vozes ativas a se digladiarem pelas redes sociais, como a já citada polarização política que, a bem da verdade, não é nova, mas sim potencialmente extremada, pelas facilidades mediáticas atuais que levam à fácil superexcitação das polaridades, ampliando-as) parecem criar suas próprias “universalidades”, que parecem absurdas ao grupo “oposto”, e parece ainda mais absurda aos que ainda não aderiam a nenhum dos grupos e ainda detêm consigo alguma racionalidade minimamente funcional.

deles ajustasse a sua contribuição de uma forma que solicitasse acordo, para que pudessem avançar para um acordo pelo consenso racional. (Keul, 2002, p. 257, tradução nossa).

Provavelmente tal possível distorção da realidade foi uma das preocupações derivadas de Kant^{xxxvi} ao se tentar uma teoria casada entre a convenção intersubjetiva e a magnânima universalidade (messianicamente) revelada pelas relações pautadas na racionalidade (e, portanto, também supostamente pautadas no “bem”, que, aliás, opera muito próximo dos parmenidianos).

Não considerou (resolver) a intencionalidade, todavia, como também deixou claro que seria insolúvel tal pretensão (Tunhas, 2023), pois constituiu a intenção como insondável justamente a partir de uma singularidade de cada subjetividade, pelas alteridades que precisam, necessariamente, de alguma intersubjetividade para que possam ser coerentes. Estabeleceu um abismo fundacional acerca desta impossibilidade, no domínio da liberdade.^{xxxvii}

Kant foi muito relevante ao apresentar e relevar a intersubjetividade, na qual reside o dinamismo das relações entre sujeitos e objetos, ou entre sujeitos e sujeitos. E, tais relações se situam atualmente, mais à luz da IA conversacional, como base para todos os desenvolvimentos para se superarem os problemas que serão abordados na terceira secção. Após Kant, ficou claro que seria preciso mergulhar nas questões inerentes às relações intersubjetivas, a acrescentar como mais um ingrediente algumas pitadas das intencionalidades, estabelecidas como resultados dos processos intersubjetivos, ou seja, as significações em movimento, em transformação, a partir das relações e não mais

^{xxxvi} Kant, todavia, avançou imenso neste campo das relações representadas, mais do que antes, e levantou, assim, mas sem resolver, a questão das relevâncias destas para as teorias do conhecimento.

^{xxxvii} Em linha com a kantiana insondabilidade da intenção, Paulo Tunhas (2023, p. 91-92) atualiza, sintetiza e argumenta que “o sentido das crenças dos outros só pode ser parcialmente capturado”, pois afirma que o que temos, em relação à intenção alheia, são meras crenças quase-representativas, visto que “crenças puramente representativas dependem fortemente da condição de exterioridade”. Assim, refuta o princípio das relações exclusivamente internas, pelos factos de sua fragilidade e falta de referencial (que resultaram nas tais referências parmenidianas, como vimos) e adota um modelo que denomina interno-externo, sobre as relações que estivemos a abordar. Tunhas (2023) ainda reforça tal direção e cita Kant: “a experiência interna em geral só é possível mediante a experiência externa em geral” [B 278-279].

da paralisia, tais como tudo se dá, fenomenologicamente. Habermas^{xxxviii} foi um destes filósofos a explorarem e a ampliarem de forma significativamente as conceptualizações kantianas. Um pouco antes de Habermas, todavia, tal exploração da experiência subjetiva também foi tema central da proposta realizada por Edmund Husserl na fenomenologia.

1.2.5. Husserl e a emergência da intencionalidade^{xxxix}

Segundo Alves (2022), a grande mudança introduzida por Husserl, mesmo que sob a égide da universalidade legada de Kant, foi focar a experiência do sujeito intencional em suas relações, e não dirigir o foco filosófico nem para o sujeito, exclusivamente, e nem

^{xxxviii} Jürgen Habermas foi um dos filósofos que se propuseram a encontrar meios para superar tais limitações kantianas, pela maior promoção das relações intersubjetivas:

“Ao invés de fornecer regras para governar as capacidades internas da mente, Habermas difere ao dirigir sua atenção para aquele espaço social em que a fala, a práxis e a razão se unem. Os padrões que ele pretende articular são padrões que orientam os parceiros de diálogo tanto para a compreensão das questões como para uma melhor compreensão daquilo que eles próprios pensam.” (Keul, 2002, p.257, tradução nossa).

^{xxxix} A intencionalidade de Husserl é conceptualmente fundamental na fenomenologia. Ela descreve a direção da consciência para um objeto: a consciência é sempre consciência de algo. Em síntese, podemos explorar a intencionalidade, a partir de Husserl e de outros desdobramentos conceptuais nos seguintes pontos:

Não se refere apenas às dimensões conscientes, mas também as que geralmente são chamadas de "inconscientes".

Está contida na experiência vivida. Merleau-Ponty incluiu nela o pensamento pré-verbal – a dimensão pré-pessoal das intenções e significados corporais (o corpo atuante sempre compreende sua situação e suas próprias possibilidades, muito antes de prestarmos atenção explícita a ela).

É um processo relacional: não é apenas um ato unilateral e exclusivo ao sujeito em direção ao objeto. Em vez disso, é uma dupla (inter)ação, (inter)experiência e (co)constituição. O sujeito e o objeto se moldam mutuamente na experiência vivida, e exploraremos mais tais operações ao abordarmos Whitehead.

É fundamental para a criatividade, quando esta se dá como um ato de encontro, e esse encontro criativo é caracterizado pela intencionalidade. O humano criador se envolve intencionalmente com o mundo, e o processo criativo deriva dessa intensidade do encontro.

Está ligada à busca de significado (e à linguagem, portanto). A intencionalidade é um constituir e compartilhar do conhecimento ou da experimentação, ainda que realizada na imaginação. É a capacidade de ser responsivo ao que pode ser.

(King et al., 1989; May, 2007; Merleau-Ponty, 1999; von Eckartsberg, 1998)

para o objeto. Alves alega ainda que Husserl optou pelo protagonismo da intencionalidade, que é justamente a parte relevante que opera e é operada pela subjetividade, ou seja, é parte com a qual a subjetividade se relaciona como intersubjetividade, agora funcional como uma espécie de interface intencional. O que Husserl tentou, de facto, foi explorar a intencionalidade em suas diversas dimensões como relações conscientes e experienciais subjetivas.

Mas relembremo-nos de que há, ainda, como esperado, os efeitos colaterais sempre derivados das novas teorias e que, geralmente, levam à imobilidade. E, portanto, não foi diferente com Husserl. A própria existência do sujeito, após Husserl, portanto, passou a ter como possibilidade um cariz intencional – pois a consciência husserliana é sempre consciência de algo, sempre relacional (ou mesmo autorrelacional). Por que então, afinal, a própria existência não poderia ser uma resultante intencional autorrelacional? E, se sim, como superar os mesmos problemas das significações alucinadas?

Eis, assim, a consciência de si mesmo, portanto, como possibilidade de fator de condição existencial, diferente, mas que pode terminar não muito distante talvez do que Descartes tenha tentado fundamentar, todavia sem muito sucesso. Husserl, ao invés de recorrer ao deus cartesiano, resolveu (ou tentou resolver) tal questão com a consciência, que assumiu ser também reflexiva, paralelamente à intencionalidade. Portanto, a consciência passou a ser intencional e também reflexiva, consciente de seus próprios atos, a incluir sua intencionalidade. Consciência e intencionalidade passaram a ser, portanto, instituições de hierarquias subjetivas: passaram a constituir uma ontologia para a subjetividade.

Assim, para resolver os problemas, e ir um pouco mais longe, apresentou duas estruturas das relações conscienciais e autoconscienciais: a pura autoconsciência que reduz fenomenologicamente (epoché) as crenças acerca do mundo exterior (mas não tão radicalmente como Descartes) para perceber as próprias relações da consciência, em termos de estruturas relacionais, como se fosse possível atingir um nível de entendimento do “essencial” acerca das experiências conscientes. E este essencial é, desde Aristóteles, a tentativa filosófica de se chegar ao invariante, ao estático e

inalterável eterno, ou até universal, que seria atingir a consciência pura, totalmente desinfectada que quaisquer influências sobre o que lhe é transcendente.

Ele acreditava ser possível a percepção consciencial e temporal do próprio ego transcendental – que logo passou a estar sempre “aí”, “dado”, presente e atuante, que se configura, então, como o elemento (ou referência) estático e puro de Husserl, a somar-se ao motor imóvel de Aristóteles, ao deus de Descartes, à universalidade de Kant, e a muitos outros que o seguiram. Husserl avançou bastante, deu novas e valiosas perspectivas às questões das relações, tão caras aos nossos argumentos vindouros, mas não resolveu de todo os desafios existentes e, como tradicionalmente, derivou daí novos problemas, tais como sempre seus antecessores o fizeram^{xl}.

1.2.6. Heidegger e a emergência do dasein

Ficou assim até que Martin Heidegger surgisse, não muito depois, visto que foi ele próprio aluno de Husserl. Então, ao criticar o excesso de foco na consciência transcendental e na redução fenomenológica de seu professor, propôs que a própria existência seja uma relação de um ser-dado-no-mundo, ou o ser-aí, ou apenas dasein, ou o que o valha, que é um status resultante temporal consciente e inserido em um dado contexto ou perspectiva, no qual é tanto causa quanto efeito, ou uma mera fusão de tudo isto (Heidegger, 2012).

Heidegger (1999, p. 16), diferentemente de Husserl, redirecionou a ontologia a um nível próximo ao da epistemologia, ao fundamentar a possibilidade do conhecimento (que recai no domínio da epistemologia) em uma compreensão mais originária do ser (que recai no domínio da ontologia). Assim, não houve espaço para um ceticismo radical cartesiano em Heidegger, como poderia ter havido em Husserl, pelas brechas deixadas. Afinal, para Heidegger, o dasein nada dispensa ou de nada pode se desassociar, pois incorpora para si tudo o que lhe é inerente à existência, inclusive a própria tecnologia

^{xl} O desafio pós-Husserl seria superado, portanto, ao existir algum filósofo capaz de desdobrar argumentativamente os efeitos desta intencionalidade autorrelacional – do ego husserliano, que parecia ser uma tarefa urgente, à primeira impressão, pelo radicalismo que o idealismo poderia resultar.

que passa a ser percebida como sua própria extensão existencial^{xli}, nas formas nas quais este passa a interpretar o mundo, a significar a própria existência, ou a indagar em si o seu próprio ser, em termos mais propícios (Heidegger, 1999).

O *dasein* é parte de uma relação com a história, com as estruturas sociais e com todas as questões objetivas, ou melhor, práticas ou funcionais da vida. O *dasein* não pode ser consciência pura, pois é essencialmente operador da própria existência, ao perceber-se como existente. O *dasein* é funcional como um todo, ou como um reflexo de um todo, sem nunca deixar de ter este todo em si, e sem haver aí uma circularidade possível de ser definida^{xlii}. A objetividade, portanto, passou a ser direcionada ao próprio ser, que transcende ao corpo, e direciona-se à própria existência contextualizada e intencionalizada, e tudo isto completamente alocado em uma contextualização histórica que, no conjunto, configura o elemento estático de Heidegger: o próprio *dasein*, como referencial absoluto dentro de uma dada historicidade, sem ter como transcender existencialmente à mesma e, por isso, incorporando-a ou fundindo-a a si mesmo.

^{xli} A ideia de que a tecnologia é uma extensão existencial do *Dasein* em Heidegger está ligada à sua concepção de *ser-no-mundo* (*In-der-Welt-sein*). Para ele, o ser humano não existe isoladamente, mas sempre em relação com o mundo e os instrumentos que utiliza. A tecnologia, nesse sentido, não é algo externo ao ser, mas uma forma pela qual ele interage e compreende a realidade. No contexto contemporâneo, defendemos que isso se aplica à IA e ao digital, que moldam a forma como percebemos, pensamos e existimos no mundo.

^{xlii} A extensão deste pensamento heideggeriano é, sem dúvida, *sui generis* e arrojada, visto que se opõe radicalmente à metafísica tradicional até então vigente que abstraía a funcionalidade do sujeito ou do objeto *p* da própria existência de *p*. Heidegger usa como exemplo, em relação aos objetos, um martelo, mas podemos atualizar este exemplo ao substituir o martelo por uma IA, e considerar que há duas formas de significá-la, em uma visão heideggeriana:

(1) A primeira forma se dá em perceber uma IA de forma conceitual, ou teórica, ou mesmo descritiva. Seria uma visão não integrada de uma IA às nossas necessidades ou intenções. Quando o uso desta IA começa, se dá a (2) segunda forma de percebê-la, ou melhor, de não se percebê-la mais. Pois o que a IA seria, segundo podemos derivar / supor a partir de Heidegger, deixa de ser relevante pois ela estará completamente integrada à intencionalidade do utente, a operar como uma extensão da sua mente, ou mesmo do seu corpo, não dualisticamente, e, por isso, fica completamente transparente às suas percepções. Só há a percepção do objeto, portanto, quando há o não uso ou a inutilidade deste (Heidegger, 2015, p. 68), assim como só há a percepção do próprio corpo quando não há uma excelente saúde. Na doença, nas falhas orgânicas que o corpo passa a ter, surge a percepção deste, antes imperceptível.

Praticamente, tudo em Heidegger é direcionado, em última análise, ao eu, ao estatuto ontológico do *dasein*. Até mesmo a corporalidade. O *dasein*, então, pela sua centralidade e absolutismo, soma-se ao conjunto dos parmenidianos^{xliii}, aberto com o motor imóvel de Aristóteles e até o último visto, que foi o ego transcendental de Husserl.

1.2.7. Ricoeur e a emergência da narrativa

Depois de Heidegger, e ainda que influenciado por ele, Paul Ricoeur (1989, p. 12) prioriza a narrativa ao propor uma capacidade de afastamento da própria historicidade para que, assim, possa haver condições hermenêuticas suficientemente efetivas para o entendimento.

Ricoeur transcende a Filosofia e mescla uma interdisciplinaridade em busca não de uma paralisia analítica, mas de uma analítica em movimento, afinal. O que resulta disso é algo extremamente complexo, como deve ser, ou como seria (e ainda é) suposto ser a existência e a sua representação. Afinal, ao se atingir as relações sem grandes referenciais centrais imóveis, ao priorizar a narrativa interpretativo-epistêmica, fundamentalmente dinâmica e transformacional, o trabalho dele não foi sobre buscar referências, mas sim interpretar a própria complexidade, como que o próprio método hermenêutico possa ser ele próprio recursivo. É a complexidade da interioridade analítica, de se analisar o movimento a partir de sua própria interioridade^{xliv}, e sem referenciais externos e estáticos.

Ricoeur, em relação aos artefactos (ou tecnologias), assume uma posição sem a radicalidade do *dasein* incorporador e considera que o conhecimento pode ser mediado por estes: retornam ao nível ferramental de um medium. Pensemos em tecnologias: tal

^{xliii} Poder-se-ia também considerar inserir aos parmenidianos a historicidade, em Heidegger, pela sua relação muito explícita com o *dasein*, e vice-versa. Mas, como cada titular tem tido apenas uma única vaga nos parmenidianos, consideremos apenas o *dasein*, como eleito para tal.

^{xliv} E, afinal, o círculo hermenêutico é mesmo assim, que alguns dizem ser um problema justamente pela falta de referencial definido, sem que as partes e nem o todo sejam eleitos como tal. E isto incomoda o juízo comum, principalmente quando se afasta da historicidade (pela sua linearidade, causalidade e suposta coerência racional) até aqui apresentada.

como Heidegger, há alguma apropriação da tecnologia, mas não radicalmente incorporada e totalmente direcionada ao eu^{xlv}. Uma IA não se fundiria, então, ao utente, mas seria o meio entre este e a realidade que o cerca. Ricoeur adota um percurso longo até o conhecimento (Romele, 2019) que, por fim, alcança o eu: tal longo percurso é dado pela narrativa, que substitui o *dasein* heideggeriano (ou da historicidade) e assume uma centralidade em sua teoria e que poderemos, assim, considerá-la como referencial e até apta a integrar o conjunto dos parmenidianos, se não fosse a narrativa eminentemente dinâmica, pois esta se dá sempre como um movimento, um contexto ou perspectiva. Parmenidianos não permitem movimentos e, por isso, a narrativa de Ricoeur ficará de fora do conjunto.

A significação, portanto, é parte relevante no processo assim delineado por Ricoeur que se dá, sempre, a partir de uma narrativa (Romele, 2019). Voltaremos a Ricoeur algumas outras vezes, sem nos atermos a ele, por enquanto. O que nos interessará será o que ele definiu como apropriação (em síntese, um processo de transcendência do eu cognoscente para se atingir um conhecimento pela interpretação da narrativa: que é puro movimento), para além da sua proposta hermenêutica, perfeitamente aplicável ao digital e extremamente atual, dada sua dinamicidade.

1.2.8. Uma síntese com a emergência do “hífen” como relacional

Podemos sintetizar e esquematizar (na figura 1), até aqui, nossos postulados em revisão aos melhores momentos desta curtíssima^{xlvi} história onto-epistemológica da Filosofia:

^{xlv} Para Ricoeur, os artefactos e tecnologias não se fundem completamente com o eu como em Heidegger, mas funcionam como mediadores do conhecimento e da experiência. Em outras palavras, a tecnologia é um medium, um meio pelo qual interagimos com o mundo, sem que essa mediação se torne indissociável da identidade do sujeito. Diferente do *dasein*, que incorpora a tecnologia como parte de sua própria existência, Ricoeur mantém uma distinção entre o sujeito e o artefacto, enfatizando que o conhecimento passa por um percurso interpretativo, marcado pela narrativa e pela historicidade.

^{xlvi} E também incompleta e injusta, reiteramos, por não citarmos inúmeros outros filósofos que foram relevantes.

- Primeiro, a incluir Aristóteles, tivemos o foco no objeto (o). O resultante conceitual garantidor teórico de sua argumentação e referencial estático (a abrir o conjunto dos parmenedianos) foi o motor imóvel;
- Em segundo, a incluir Descartes, que consolidou o foco no sujeito (s), até mesmo independentemente de qualquer objeto (o). O resultante conceitual garantidor e referencial estático foi deus;
- Em terceiro, a partir de Kant, tivemos um foco mais aprofundado no sujeito (s) racional, e nas relações sujeito-sujeito (s-s). O resultante conceitual garantidor e referencial estático foi a indução à universalidade;
- E em quarto, com Husserl, o foco passou a ser o hífen (-) que integra ou conecta tanto a relação sujeito-objeto como até mesmo a relação sujeito-sujeito. O resultante conceitual garantidor e referencial estático foi o ego transcendental;
- E, em quinto, com Heidegger, tivemos a contextualização da indagação sobre o hífen, pois o indagador é o próprio ternário sujeito/hífen/objeto integrado como tal, a consolidar-se como uma unicidade capaz de incorporar toda a tecnologia que orbita à sua volta para sua própria justificação existencial^{xlvii}. O resultante conceitual garantidor e referencial estático foi o próprio dasein, dentro de sua respetiva historicidade, na impossibilidade de “haver” algo “fora” dele;
- E, por último, Ricoeur atribui ao hífen (tornado multifacetado) a possibilidade de ser ele próprio uma tecnologia que realiza a significação, da intencionalidade como mediadora entre o eu e a realidade que o “cerca”. Mas, tal mediação, é igualmente e teleologicamente apropriada pelo sujeito, tal qual o voraz dasein de Heidegger, embora não tão diretamente, ou mais comedidamente. O resultante conceitual garantidor e

^{xlvii} O dasein passou a ser ele próprio um super-hífen, inserido na história, que a tudo se conecta poderosamente, em busca de sua própria justificação existencial, ou tal qual um buraco negro, que suga tudo ao seu redor, a se consituir (a existir) densamente enquanto dado no mundo.

referencial (não estático) foi a narrativa que, diferente dos anteriores, é essencialmente dinâmica e, portanto, inapta aos parmenidianos;

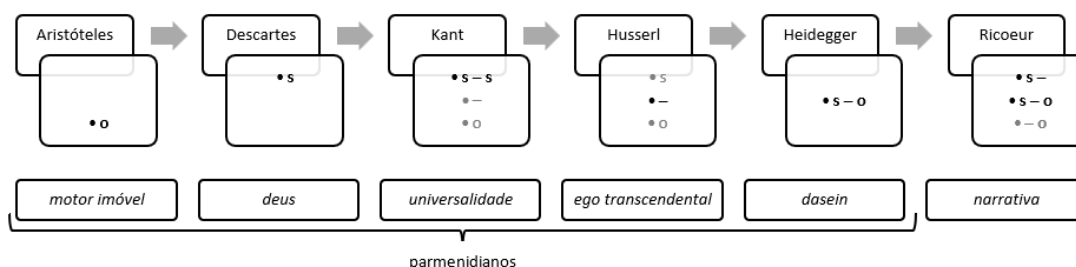


Figura 1 - “Evolução” do ternário macro-ontológico da epistemologia: sujeito, hífen e objeto e seus respectivos “efeitos colaterais”, mas apresentados como “causais”.

O “hífen”, portanto, como metáfora relacional, não apenas passou a deixar de ser invisível, como também passou a ganhar relevância até alcançar o atual protagonismo, pois nele está a problemática legada e até então em questão.

Nestas propostas filosóficas expostas, há algo intrigante nelas. Ao percebermos que os conceitos filosóficos propostos a seus tempos, e em relação aos seus predecessores, acabam por reforçá-los, anulá-los ou antagonizá-los. Mas, em relação a todos os resultantes conceituais garantidores e referenciais estáticos (os parmenidianos), estes acabam por formar um grupo perfeitamente homogêneo e solidário. São afins e até conectados entre si^{xlviii}: quaisquer similaridades com os supostos problemas das IA talvez não seja mera coincidência, pois se comportam tais como os parmenidianos.

^{xlviii} O motor imóvel aristotélico pode tranquilamente “evoluir” até atingir o conceito de deus cartesiano (que se mantém em pleno ato, mesmo enquanto supostamente nada mais haveria de haver) que, seria ele próprio a maior expressão da universalidade kantiana (substancialmente, em si) que, por sinal, pode ser facilmente (e funcionalmente) percebido pelo ego puro (transcendentalmente) em Husserl, além do que está intimamente incorporado pelo dasein heideggeriano (a incorporar / absorver a transcendência), que se indaga a si mesmo, em seu ser.

Todavia, a imobilidade anterior pôde ser percebida através de uma narrativa que agrega o “tudo isso” de Nagel e transmuta a paralisia em movimento do “isso”. É esta narrativa que estamos a fazer, nestas linhas, hermeneuticamente, ao reproduzir o movimento da própria história e nos apropriarmos dos significados deixados como rastros, dentro de uma perspectiva que foi preciso (1º) se afastar de tais processos para (2º) se perceber, hermenêuticamente, o que Ricoeur propôs, ao nos apropriarmos de cada conceito, não nos tempos passados, mas sim nos tempos atuais, nos nossos tempos, quando um martelo dá lugar à IA, por exemplo, (3º) integrando-os. Assim, pela narrativa, pudemos perceber os parmenidianos, portanto, totalmente estáticos, mas também profundamente informacionais e harmônicos entre si.

Como visto, todos coexistem num mesmo habitat, sem grandes entraves. Pois o que se está a perceber é que os “problemas” gerados são, em si, relações mais estáveis do que as propostas que os geraram. Há

1.2.9. A emergência da IA como relações (ou correlações)

Ao chegar às narrativas de Ricoeur, podemos considerar que estamos mais próximos do que uma IA, de facto, é ou pode ser e, portanto, no caminho certo, pela narrativa, para perceber a fragilidade existencial de seus problemas, tal como sejam parmenidianos. Afinal, os atuais problemas das IA não são, de facto, somente sobre seus utentes (aqueles que estejam em contato direto ou indireto com uma IA, a sofrerem os impactos de uma suposta agência algorítmica), isoladamente, e nem apenas sobre elas próprias, como objetificações tecnológicas, mas sim nas relações existentes entre utentes e as tecnologias, e que carecem de uma narrativa coerente. E relações “são o que são”^{xlix} enquanto dadas como IA, como é óbvio. Pois, “o que dizem as tecnologias, defendo, não pode ser analisado sob os termos e pressupostos teóricos existentes”¹, segundo Freiman (2023).

uma espécie de inversão filosófica em relação aos efeitos que passam a ter melhor consistência do que suas causas. Especulativamente, podem ser estes efeitos que, pelas suas supostas estabilidades e agrupamentos, fragilizam as causas e evidenciam que são, de facto, as relações entre as causas que devem ser priorizadas, sem que sejam necessários os efeitos. Podemos considerar as causas como positivas ou necessidades. E os efeitos como negativos ou contingências. Se não se pode eliminá-los, pode-se, ao menos, não os criar ou ainda encontrar jeito de não se precisar mais dos mesmos. Portanto, as relações são apenas relevantes, mas sim necessárias e suficientes, quando assim percebidas.

^{xlix} Provocativamente, o “é o que é”, citado por Jean Paul Sartre, quando defende que o ser-em-si é o que é. Especulativamente, se considerarmos a própria relação como “sendo o que é”, a extrapolar Sartre, significaria que ela não se refere a si mesma da maneira como a consciência relacional (ou intencionalidade) se faz. A reflexividade relacional, ou intersubjetiva, que constitui o si se funde em uma identidade, em ato. Sartre usa a analogia de uma nuvem, afirmando que ela não é “chuva em potencial”, mas sim uma certa quantidade de vapor de água que, em determinadas condições de temperatura e pressão, “é rigorosamente o que é”. Portanto, a relação não possui potencialidade, mas existe em ato, na plenitude absoluta de sua intencionalidade – e por isso seu intrínseco dinamismo. Ela é opaca a si mesma porque está “cheia de si”, não tendo um “interior” que se oponha a um “exterior”. Ela não tem segredo, é maciça, e, portanto, não carece de imobilidade. (Sartre, 2011, 2015)

¹ Ao se replicar nas análises das IA os modelos filosófico-analíticos que vêm sendo aplicados, com os parmenidianos sempre presentes a fundamentá-los, haverá graves distorções sobre as reais origens dos problemas, pelos esforços de, enquanto se busca preservar a todo o custo a integridade dos frágeis parmenidianos, ao submeter a IA como uma visão imóvel e estável, enquanto tudo nela e ao seu redor está em pleno movimento e transformação.

O hífen, como dinamismo, significação e orientação, é sempre uma forma relacional que incorpora, em si, transformações causadas pelas diversas relações, portanto. E é isto que precisa ser percebido, neste esforço reconstrutivo que fizemos até aqui. Pois o próprio ato de buscar o conhecimento, ou a interpretação é, antes de tudo, um movimento, e uma tentativa de relação com algo que ainda é suposto ser desconhecido, embora presente. É o mesmo que Ricoeur (2016, p. 145, tradução nossa) propôs através de um deslocamento para se atingir a compreensão, pela apropriação: “A dialética do distanciamento e da apropriação é a figura final que a dialética da explicação e da compreensão deve assumir. Diz respeito à forma como o texto é dirigido a alguém.”

O que viria, portanto, de toda esta deslocação filosófica para foco nas relações, ou para o movimento? Chegaríamos a um modelo muito próximo do proposto em Ricoeur, no qual o *dasein* (no sentido de incorporador tecnológico) não seria, todavia, descartado. Pois, mesmo nas especulações sobre o humano fundido com a tecnologia, como suposto no pós-humanismo, ou mesmo os homens robotizados (*cyborgs*), idealizados nas ficções científicas (ainda que a cada dia menos ficcionais e mais prováveis), a tecnologia é incorporada com uma função utilitarista que está assim, a serviço de outros interesses, transcendentalmente aos conjuntos formados a partir do ternário *s – o*, que se desdobra para que este conjunto seja referencial a outro: *(s – o) – (...)*, a considerar que as reticências sejam novas extrapolações relacionais, e assim sucessivamente.^{li}

Não seria precipitado, todavia, perceber que o movimento do foco entre sujeito/hífen/objeto foi, até então, para além de epistemológico, também ontológico.

^{li} Isto representa um problema aos que buscam a origem desta “teia” relacional, tal qual uma rede neural que se forma, com base nas sinapses emuladas pelos hífen, como meios relacionais de significações. E a busca ainda atualmente obsessiva pela origem como referência orientadora é o que justifica, por exemplo, a mesma redução que Aristóteles buscou ao eleger, há milênios, o conceito de motor imóvel como causa primeira, assim como todo o resto dos parmenidianos que o sucedeu. E assim ainda intentamos fazer, razão pelo qual há a origem da frustração em perceber a IA através de um anacrônico pensamento analógico.

Pois, a cada complexidade desenvolvida teoricamente, mudou-se também a ontologia considerada nas formas relacionais dos conceitos.^{lii}

E esta é uma atitude que começa a tomar consistência acadêmica, embora ainda sem meios igualmente consistente em como poderá ser feito, por alguns limitadores que trataremos, em tempo, de abordar, tal como proposto por Harding (2011; Portmess & Tower, 2015, p.2, tradução nossa): “Para dar conta das ambiguidades na noção de mineração e análise de dados, é necessária uma epistemologia mais rica a partir do *big data* que revele, em vez de ocultar, suas relações e alocações sociais”.^{liii}

Há uma recursividade tão evidenciada na dependência tecnológica atualmente que, para se perceber até mesmo tais efeitos de dependência tecnológica, será preciso recorrer à própria tecnologia para tal. “Algumas tecnologias, como a Internet ou os smartphones, não apenas acrescentaram mundanamente mais oportunidades para escolhermos – mas moldaram constantemente a forma como vivemos” (Freiman, 2023; Waelbers & Briggie, 2010, tradução nossa). E isso evidencia também uma relação simbiótica já instituída entre humanos e máquinas, tal como argumentado por Brey e Søraker (2009), embora os autores não tenham generalizado, em seus comentários conclusivos:

Os sistemas de computador são artefactos cognitivos extremamente versáteis e poderosos que podem dar suporte a quase qualquer tarefa cognitiva. Eles são capazes de se envolverem em um

^{lii} Afinal, estudar tal mutação ontológica a partir do digital, como superficialmente fizemos a partir da dimensão analógica, assim como se esteve (e ainda estaremos) aqui a se suscitar para se discutir formas possíveis de como se evidenciar como as IA poderão reontologizar o mundo (Mittelstadt et al., 2016), que é uma das partes mais relevantes dentre as que serão apresentadas a partir de um mapeamento de problemas éticos, já na próxima seção. Pois, como veremos, não será uma questão de se a IA irá ou não reontologizar o mundo, mas sim como isto se dará (Freiman, 2023). Por isso, pressupõe-se, antes de tudo, preparar o terreno para que se possa criar um padrão (ou marco) para se perceber objetivamente, de forma temporal, evidencial e mensurável, tal suposta modificação ontológica.

^{liii} Para se atender a tal demanda e superar os limitadores requer, necessariamente, a estruturação prévia de uma metaontologia (a tal suposta acerca do movimento relacional, recursiva, cíclica) para que, dada a impossibilidade humana de uma observação direta de algo tão complexo, subjetivo ou imaterial, cria-se uma tecnologia para mediar a interpretação da própria tecnologia em seus efeitos, até se atingir o conhecimento de como estão a se darem as supostas transformações.

relacionamento simbiótico único com humanos para criarem sistemas cognitivos híbridos nos quais um humano e um processador artificial processem informações conjuntamente. (Brey & Søraker, 2009)

E, por fim, a busca pela compreensão através de um método ontológico foi exatamente o que intentamos fazer, até aqui, nesta primeira abordagem epistemológica ao longo da história filosófica, com base nas três instâncias supracitadas (s – o). Através delas, e mediadas por elas, tais quais os artefactos tecnológicos, chegamos a algum conhecimento acerca do próprio conhecimento, sob tais e tais perspectivas, ou narrativas. Assim, se tudo sair a contento, estaremos a considerar ter as bases conceituais para uma tecnologia que poderá nos dizer, todavia, o nosso potencial nível simbiótico com as IA, se assim percebida.

Temos, portanto, acreditado ter construído argumentativamente a necessidade de se declinar do uso de elementos rígidos, tais como os parmenidianos, em detrimento de perceber a IA como narrativa e profundamente relacional. A preocupação, atempadamente, não se deu em relação a IA (pois “ela é o que é”, ou seja, relações), mas sim sobre a nossa própria capacidade preconceituosa e inercial de um pensamento analógico que acaba por se frustrar frente ao digital, ao pressupormos que este último esteja, ainda, em posição ontologicamente inferior ou secundária ao que seja atual, ao que existe, ao que somos. Já não mais está desta forma. Assim, podemos prosseguir.

1.3. A emergência ontológica pela perspectiva tecnológica

Nesta secção, estaremos a adotar uma perspectiva menos teórica e mais tecnológica, ainda que sob uma hermenêutica filosófica. O objetivo será desenvolver as bases para um ferramental conceitual capaz de representar as relações existentes em uma IA, a evitar cair na referência ou na dependência de algo similar aos parmenidianos.

Também discorreremos sobre os desenvolvimentos tecnológicos dos algoritmos de IA, a perceber como estes estiveram (e ainda estão) a representar as relações mesmo sem contarem com um modelo conceptual com bases fundamentadas sobre as mesmas, visto que mesmo na Filosofia as relações geralmente estiveram no campo contingencial, como vimos na secção anterior, sem terem o protagonismo merecido.

Mas, quaisquer constructos de IA são essencialmente relacionais, ou ainda mais complexos, por serem correlacionais e, por isso, é preciso perceber a dimensão deste *gap* e encontrar formas de preenchê-lo. Eis o que buscaremos realizar nesta secção.

E, assim, para se discorrer sobre como abordaremos relacionalmente os problemas éticos mais relevantes em Mittelstadt et al., (2016), precisamos perceber o atual status dos bastidores das IA, especificamente em como os dados são significados em suas bases até que façam gerar as ditas problemáticas. Partiremos para representar os problemas éticos em um esquema gráfico conexo algoritmicamente e ontologicamente: a proposta é perceber o estado da arte do constructo de um mapa epistémico-ontológico da atualidade da IA, ou simplesmente mapa representacional.^{liv}

1.3.1. Sobre a Ontologia e as ontologias

Até aqui, usamos do termo “ontologia” de forma não convencional, nem para filósofos, nem para os cientistas da informação. Pois, o termo ontologia tem sido aqui utilizado como um híbrido semântico destes dois campos nem sempre correlatos, sob o risco de não agradar a nenhum dos lados, mas com a obrigação de fazê-lo assim, visto que se faz necessário uni-los em saberes semanticamente compartilhados – ao menos para as questões aqui abordadas, e em direção a uma interdisciplinaridade operante. A dimensão que estamos a buscar representar com a prática ontológica é a das significações. E isto envolverá as dimensões subjetivas, intersubjetivas e objetivas, o que faremos, pouco a pouco, a construir os conceitos necessários a este compartilhar.

É válido destacar que os problemas acerca das significações não são apenas inerentes aos humanos, pois também os são aos algoritmos de IA, que operam com base em bancos de dados nem sempre tão amigáveis, como aponta Berners-Lee et al. (2001,

^{liv} Mapas representacionais, aqui considerados, serão conceitualmente apresentados e desenvolvidos a partir das teorias e práticas já utilizadas de representações do conhecimento (Sowa, 2000), mas não limitados por elas, pois há muitos avanços desde então. Vamos, antes, discorrer conceitualmente sobre as ontologias, algoritmos e sobre alguns dos modelos de representações do conhecimento. Por isso, será preciso atempadamente discorrer sobre as relações, e como podemos percebê-las, filosoficamente. O que buscamos, portanto, é perceber as relações entre os territórios e os mapas, ou seja, entre a atualidade e a sua representação digital, ou funcional.

tradução nossa): “Idealmente, o algoritmo deve ter uma maneira de identificar os significados comuns para quaisquer bancos de dados que encontrar.” E assim conclui:

Uma solução para esse problema é fornecida pelo terceiro componente básico da Web Semântica^{lv}, que são as coleções de informações chamadas ontologias. Na Filosofia, uma ontologia é uma teoria sobre a natureza da existência, em definir como certos tipos de coisas existem; a ontologia como disciplina estuda essas teorias. Pesquisadores de inteligência artificial e da Web cooptaram o termo para seu próprio jargão e, para eles, uma ontologia é um documento ou arquivo que define formalmente as relações entre os termos. O tipo mais típico de ontologia para a Web tem uma taxonomia e um conjunto de regras de inferência. (Berners-Lee et al., 2001)

Assim, assumiremos que o uso do termo ontologia, aqui pretensamente adotado para que não seja apenas aplicável para dados digitais (IA) ou eventos analógicos (utentes) interoperacionais / intersubjetivos e objetivos, mas também para subjetivos, em relação intrínseca aos utentes. Por isso que a ontologia, em nosso desenvolvimento, será mesmo um híbrido, visto que tratará não apenas da conceituação, mas principalmente das relações tanto da forma (ou da essencialidade aristotélica) quanto das características (ou da acidentalidade aristotélica) de como se dá a existência, tanto em termos ônticos (materiais, densos, temporais, etc.) quanto ontológicos (transcendentes, sutis, extemporâneos, etc.), para os atuais e os digitais.

Para além disso, a ontologia que desenvolveremos também tratará de como tais instâncias se relacionam entre si em um mesmo existente, ou entre os conjuntos de existentes que estejam a ser considerados como relacionais. É, portanto, uma pretensão de ontologia suficientemente capaz de lidar tanto para uma unidade autónoma (particular) quanto para a coletividade, portanto, pois ela supostamente representa, em

^{lv} “Além da clássica “Web de documentos”, o W3C (World Wide Web Consortium) está ajudando a construir uma pilha de tecnologias para dar suporte a uma “Web de dados”, que seria sobre o tipo de dado que você encontra usualmente em bancos de dados. O objetivo final da Web de dados é permitir que os computadores façam um trabalho mais eficiente e desenvolvam sistemas que possam dar suporte a interações confiáveis pela rede. O termo “Web Semântica” se refere à visão do W3C para a Web de dados vinculados entre si, ontologicamente. As tecnologias da Web Semântica permitem que as pessoas criem armazenamentos de dados na Web, criem vocabulários e escrevam regras para lidar com dados.” (acedido em 24/08/2024 pelo link https://www.w3.org/2001/sw/wiki/Main_Page, tradução nossa)

termos de atributos, como as relações destes se dão não apenas nas formas s – o, mas também s – s, pela intersubjetividade.^{lvi}

A máxima pretensão para tal ontologia híbrida seria, a partir dos atributos e dos entes, atingir uma configuração tanto explicadora quanto definidora para as entidades, da dimensão ôntica à ontológica, e vice-versa, ao representar a existência dentro de certas perspectivas funcionais. E isso tem como pressuposto não apenas perceber a alocação objetificada de um ente dentro de uma estrutura existencial relacional, como a social, por exemplo, mas também perceber a própria subjetividade deste ente, ao apreender ou emular a sua própria perspectiva em relação a esta estrutura.^{lvii}

Há, portanto, um forte argumento favorável para se considerar que uma mesma padronização ontológica possa (e deva) ser aplicada nas semânticas para os dois domínios: o atual e o digital, com este último instanciado como representação do que há (do que seja atual), ainda que nem tudo que “exista realmente” (em atualidade) esteja representado como tal. Mas, ao se pensar assim, alguns incautos poderão julgar que o digital será sempre um subconjunto do atual, e isto é um erro, tanto pelo gigantismo digital em termos de dados, quanto pela velocidade que o digital imprime – pela influência à vida atual, antes puramente “analógica”. Do digital, surgem (existem, portanto) entidades ou formas de intervenção e causação de impacto que podem possuir “vida própria” sem que sejam representações. “Uma visão padrão é que os

^{lvi} Assim, “nossa” ontologia não definirá, necessariamente, aquilo que está a representar, mas explicitará o que ocorre, ao menos em termos relacionais. E, se a partir de tais relações explicitadas, houver recursos para se chegar a certas relações causais, então poder-se-á dizer que se haverá encontrado uma perfeita representação relacional e, portanto, uma perfeita ontologia.

^{lvii} Esta ontologia não apenas busca saber o que se é, mas também o que se poderá ser, pois é relacional, e até pode ser preditiva, o que poderá trazer algumas vantagens aos processos de aprendizado de máquina, como logo veremos. Pois, se há o problema acerca da reontologização do mundo pela ação massiva das IA (Mittelstadt et al., 2016), há que se perceber a variância entre a ontologização atual e a futura, comparativamente, dentro de certos padrões a serem os mensuradores de eventuais alterações. Qualquer solução ao problema passa por tal necessidade comparativa. E qualquer necessidade comparativa carece de parâmetros claramente definidos e mensuráveis, no nosso caso, defendemos que será a partir de uma dada ontologia, ou metaontologia.

artefactos tecnológicos não são intermediários neutros, mas determinam ativamente como construímos o conhecimento” (de Boer et al., 2018 apud Freiman, 2023, p. 4, tradução nossa).

A própria IA conversacional é um exemplo, pois algumas delas^{lviii} são até projetadas e programadas para serem (intencionalmente emuladas) humanizadas, com nomes, perfis psicológicos, certas reproduções comportamentais com simulação de empatia e, ainda assim, sem paralelos com os existentes: sem representar nenhuma individualidade de facto, obviamente, mas uma mera generalidade ajustada conforme o perfil do utente. Sem conterem essência aristotélica, e compostas quase apenas de acidentes, de atributos ou contingências, que operam e consolidam alguma coisa que é “inexistente” como se existente fosse: dão vida ao emularem uma presença digital que, de facto, que não é presença, mas que parecer ser. Tal qual ilusões meticulosamente elaboradas para parecerem verossímeis. Tais IA correlacionam-se apenas com atributos brutos que acabam por se expandirem pelas correlações derivadas destes, e, por assim ser, a partir de certas operações algorítmicas podem recursivamente derivar outros atributos mistos, híbridos, e modificar a própria ontologia inicialmente considerada: podem impactar e modificar a atualidade do utente, portanto, e até iniciar outras alterações relacionais^{lix}.

^{lviii} Existem IA conversacionais que emulam, por exemplo, humanos nos quais relacionamentos afetivos poderão ser constituídos, como no caso da aplicação Replika, mas não só. Assim, tais entidades não são, necessariamente, representações de um humano específico e, assim, podem “existir” a representarem não os entes, mas sim as entidades, ou melhor, seus atributos. Matematicamente, com o imenso poder de computação, é possível criar mais personalidades individuais (digitais) do que humanos existentes e, portanto, não é correto assumir que a representação digital está limitada a ser um mero subconjunto do “mundo real”, visto que nada impeça que possam ser criados perfis alienígenas, por exemplo, ou de monstros ou super-heróis. A Replika, como empreendimento, teve origem na tentativa de sua fundadora, Eugenia Kuyda, em reproduzir digitalmente a consciência ou a individualidade de um amigo falecido, que lhe era muito querido e que havia deixado um imenso banco de dados em mensagens, enquanto estiveram a se relacionar. Os dados, assim, viraram rastros digitais deste amigo, e desta relação, e serviram para dar uma nova vida ao que já não era mais atual. (Acedido em 16/08/2024, no link <https://replika.com/about/story>)

^{lix} Como exemplo, podemos citar uma leitura crítica do artigo *Generative Agent Simulations of 1,000 People* (Park et al., 2024), realizada no 7º Seminário de Investigação do Projeto Signo, na sessão realizada no Instituto de Filosofia da FLUP, no dia 09/dez/2024, com o título *AI Agents*, na qual foi feita uma simulação algorítmica simplificada com base em ontologia, na qual cada valor se correlacionaria com

A digitalização, *lato sensu*, mudou até mesmo as formas de relações humanas pela boa parte delas atualmente transferida para a dimensão digital, através das redes sociais e mediações de comunicações via aplicações que evidenciam certos atributos: determina ou oficializa se utentes estão *online* (conectado), *offline* (desconectado) ou *busy* (ocupado), por exemplo, mesmo que seja referente a um perfil falso ou corporativo, que não corresponda a um existente em si, tal e qual, mas que ainda assim possui atributos^{lx} que o explica e o aloca em determinada posição estrutural^{lxi}. Há, portanto, a criação de

outros cinco, em ordem de relevância, de acordo com os dados tirados do ChatGPT, da OpenAi – na mesma fonte e versão que o artigo utilizou.

A partir de um valor aleatoriamente escolhido (que foi a clareza) e seus derivados que foram, um a um, correlacionados com outros cinco valores distintos, os valores foram se multiplicando até um limite de oitenta e oito valores evocados. A partir daí, qualquer dos valores evocados estaria a evocar outros cinco, mas sempre dentre todos os já existentes, sem novos a serem adicionados. O que mudava, apenas, era o peso de cada um dos valores evocados.

Assim, poder-se-ia considerar que o “universo” de valores (conforme esta simulação) – o domínio de valores – de uma IAC seria restrita a oitenta e oito valores, a partir de um gatilho específico da “clareza”, o que comprova a IAC como uma parcialidade do existente, e com as imprecisões esperadas, sem as sutilezas das perspectivas distintas que existem.

Maiores informações no website <https://ifilosofia.up.pt/activities/signo-seminario-2024>, acessido em 19/12/2024.

^{lx} Um atributo não é algo estático, mas sim dinâmico. Pois um atributo é sempre relacional e se dá em perspectiva. Entre o certo e o errado humanos e atuais, ou entre o bem e o mal, ou a guerra e a paz, por exemplo, poderão existir demasiados graus para uma ontologia digital e, quiçá, algumas que até nem sequer conhecemos. Hipoteticamente, de tais níveis inéditos de atributos, eventualmente identificados, um destes níveis pode emergir e assumir o protagonismo, a maior relevância. E isto é uma mudança da ontologia vigente: a reontologização é um efeito do movimento.

Não há, assim, como haver uma universalidade ou imobilidade conceitual para nada, ao menos em termos de representações digitais ou artísticas. Até mesmo condições ficcionais que são tidas como quase consensuais, como as distopias (sempre descritas como negativas, ou algo ruim, indesejável), não podem, à luz da razão, ser consideradas como consensuais. Pois, distopias para uns podem ser, simultaneamente, utopias para outros.

“As distopias parecem ser bem diferentes, a priori, mas a posteriori não são muito distintas das utopias, em seus propósitos que se referem à tentativa de buscar uma alegoria para levar a todos a uma reação ao incumprimento ou ao cumprimento irrestrito às regras existentes – tudo o que se pretende, com estas criações utópicas ou distópicas, sejam preditivas ou artísticas, é questionar a ordem estabelecida através de suas próprias regras atuais – é um chamamento à alguma razão contraideológica, mas ainda ideológica, como tudo o mais. Projetam-se utopias e distopias por razões sempre ideológicas, pois são expressões das possibilidades ou das impossibilidades.” Ortolan (2022, p. 259)

^{lxi} E também não é uma defesa de um relativismo completo, mas sim a presunção da necessidade de se adotar uma configuração que leve em consideração a perspectiva individual e que, tais perspectivas

existentes digitais sem contrapartidas atuais e isso é uma corporificação, a partir de modelos feitos para o sentido oposto, de se representar o que existe na atualidade.

1.3.2. Os modelos de representações do conhecimento

As ontologias, como vistas, são utilizadas pelos *designers* em tecnologia, *stricto sensu*, para fins como a arquitetura e gestão de banco de dados relacionais ou a interpretação de informações a serem utilizadas pelos algoritmos, principalmente. Tais ontologias servem como bases conceituais para dotarem os dados de certos atributos para que, então, estes possam ser correlacionados em busca das significações que resultem nos objetivos avançados para o algoritmo. Assim, as relações, ao contrário do que vimos nos andamentos filosóficos, já se encontram na centralidade dos andamentos tecnológico-construtivos das IA, pelos bancos de dados relacionais, mas não apenas.

A maioria dos estudos na literatura se concentra em bancos de dados relacionais como fonte de informação por vários motivos: Primeiro, cerca de 70% dos dados na web são armazenados em bancos de dados relacionais (He et al., 2007); Segundo, bancos de dados relacionais apresentam

consideradas, sejam relacionais, e que se possam identificar, por exemplo, recursos éticos para orientações tanto utópicas quanto distópicas. O que se busca não é a defesa de nenhuma teoria ou escola filosófica, econômica, antropológica, política ou até sociológica, mas sim factos e considerações que nos levem, portanto, ao conhecimento através de representações as mais fidedignas possíveis, via meios que possam justificar certas causalidades.

Afinal, desde antes de Aristóteles, a Filosofia tem se ocupado em investigar em como a existência se dá, ou como se pode conhecer o que é existente, como visto na secção anterior e, assim, o que está a se considerar aqui como fusão de saberes ontológicos é mesmo acerca dos atributos do ser (para a Filosofia) ou dos atributos dos dados (para as Ciências da Informação), e tudo dentro de certas construções conceituais que podem ser analisadas em perspectivas, e operadas de maneiras estruturadas ou empilhadas, a ganharem consistência, robustez e a permitirem uma narrativa interpretativa menos ingénua.

Como objeção à tal possibilidade, devemos considerar, desde já, que se um atributo humano receberá uma mesma semântica para si e para a sua representação digital, para ser coerente, esta semântica digital poderá sofrer alterações e não mais corresponder exatamente à mesma representação inicial. A partir do descolamento semântico entre o atual e o digital, há o sério risco de que os exponenciais resultados digitais produzidos pelos algoritmos sejam considerados alucinados, pela falta de correspondência entre o que obtiveram como *output* e do que há na atualidade, como sugere Hicks et al. (2024, p. 37, tradução nossa) ao destacar que a prioridade dos modelos de IA é apresentar coerência linguística, a partir de operações probabilísticas: “esses modelos não foram projetados para transmitir informações, portanto não deveríamos ficar muito surpresos quando suas afirmações se revelassem falsas”, pois este é um dos problemas mais relevantes existentes (e muito prováveis em IA conversacional, quase corriqueiros) que logo abordaremos.

modelos conceituais completos; Terceiro, eles fornecem um recurso de informação completo (Martinez-Cruz et al., 2012); e, finalmente, eles oferecem uma das melhores técnicas para armazenar e manipular dados. *No entanto, bancos de dados relacionais sofrem com a ausência de significado semântico, o que dificulta a capacidade de alcançar a interoperabilidade entre sistemas de informação* (Meersman et al., 2002 apud Ben Mahria et al., 2021, p. 2, tradução e grifo nossos).

A usabilidade das ontologias tem estado quase que unicamente direcionadas do digital para o digital, como uma busca de dar significado aos dados, para que possam, assim, serem operados em busca de informações. Afinal, dados são considerados como rastros digitais de uma representação da realidade factuada (Romele, 2019). Assim, uma ontologia é utilizada como uma forma otimizada para se navegar por tais rastros a partir de certas premissas. Por isso, mas não apenas, há a atual impossibilidade de uma implementação semântica para os dados operados, pela falta de correlação ontológica entre o digital com o atual, em termos semânticos – há um *gap* entre tais dimensões, e é justamente o *gap* semântico.

Classifica-se, mas sem se significar. Estabelece-se o que há no mundo (o atual) em termos de atributos, estrutura-se ontologicamente como tal (representa-se) e aplica-se ao digital, ao percorrer os dados armazenados para enquadrá-los forçosamente conforme tais classificações estruturadas. E o trabalho fica considerado realizado, pela impossibilidade de se atingir (ou se estruturar) a necessária significação dos dados ontologizados: sedimenta-se a impossibilidade de se atingir uma modelagem semântica.

E isto tudo se dá quase sempre mecanicamente, maioritariamente feito também por algoritmos, visto que seria um custo dispendioso e muito demorado para ser realizado por humanos (Ben Mahria et al., 2021). Os sistemas gestores de dados “geram” suas próprias ontologias, com base nas relações que passam a serem relacionadas a outras relações, a criarem teias correlacionais, multidimensionais, e daí percebe-se o risco alucinatório e a ineficiência sistêmica para uma rastreabilidade causal. Não há, portanto, qualquer mediação possível em sua máxima instância transitiva, visto que há como resultado a própria ausência da semântica, pois o que há é apenas uma adequação compulsória a uma referência nem sempre isenta de preconceitos ou usabilidade. O que ocorre, neste caso, e curiosamente, é o oposto do que foi criticado na seção anterior,

quando havia referências imóveis (os parmenidianos) para darem sentido às relações. Aqui, percebe-se justamente o oposto, que são apenas relações sem nenhuma referência semântica entre elas, nem interna e nem externa, que passam a se configurar quase como uma situação de completo relativismo: e isto é até pior, pois impossibilita a narrativa, ou a apreensão das relações tais como dadas.^{lxii}

Uma nova proposta ontológica precisaria ser implementada, portanto, e a bem da versatilidade desejada, não apenas uma mediadora plena entre o atual e o digital, mas também uma facilitadora, sem fazer de nenhum destes uma referência imóvel. A imobilidade seria, no máximo, para certas características da própria estrutura ontológica, ou seja, seria uma estrutura metaontológica – dinâmica, flexível e adaptativa – que pudesse conter elementos fixos, mas que, ainda assim, tenha abertura à recursividade, ou certas operações complexas que possam aperfeiçoá-la, ou atualizá-la. Pode-se dizer que os dados estão para os metadados assim como as ontologias estarão para esta metaontologia.

As Ciências Tecnológicas, em especial as Ciência dos Dados, em lato sensu, já perceberam tais demandas (e também as limitações teóricas) e estão a desenvolver com impressionante velocidade novas formas de representações, modelagens e linguagens,

^{lxii} Há que se buscar, portanto, um meio termo: ontologias que se fundamentem pelas relações (como suposto) mas que também permitam correlacioná-las à atualidade, o que as exigem serem completamente dinâmicas, mas que consigam suportar, a partir de seus próprios constituintes, uma narrativa interna que leve à significação de seus conteúdos: que permitam uma narrativa conexa a partir de seus metadados que possam tanto explicar como definir seus dados, idealmente. Ou seja, na impossibilidade de uma referência externa imóvel, a semântica deve ser extraída a partir de sua própria interioridade em constante movimento, a se atualizar incessantemente, paralelamente à atualidade. Pelo predomínio das correlações internas, a IA conversacional, afinal, por tal razão não opera diretamente pelas ontologias, mas sim exclusivamente pelas correlações realizadas a partir das bases de aprendizagem de dados, e não metadados, como vimos. Por isso, diz-se, como já citado, que ela não está talhada para gerar conhecimento, mas sim para gerar coerência linguística, meramente, ao replicar o que há com a máxima fidelidade.

tais como modelos ontológicos em 3D (DOLCE^{lxiii}, UFO^{lxiv} e uma de suas especializações mais promissoras, que é a OntoUML^{lxv}) e até em 4D (BORO^{lxvi}). Tais desenvolvimentos,

^{lxiii} DOLCE é uma ontologia fundamental que descreve a distinção entre a cor de uma maçã (sua qualidade) e seu 'valor' (por exemplo, um tom particular de vermelho). Este último é chamado de qualia e descreve a posição de uma qualidade individual dentro de uma certa dimensão de qualidade (como nuances). Um exemplo disso seria que a idade de João é um momento de um indivíduo que objetiva sua idade enquanto propriedade abstrata. A intensidade é um momento da dor de cabeça de João, que é um momento de João; a intensidade é inerente à dor de cabeça, que é inerente a João. E, assim, se constrói um sentido relacional a todos os atributos. (Guizzardi, 2005; Zamborlini, 2011)

^{lxiv} A UFO (Unified Foundational Ontology) é uma ontologia fundamental que fornece uma base para a modelagem conceitual orientada por ontologia. A UFO é organizada em três camadas: UFO-A (objetos), UFO-B (eventos) e UFO-C (social). A UFO tem sido utilizada como base para ontologias de domínio, incluindo a ontologia de software ODE. Ela também foi usada para analisar conceitos relacionados a agentes e linguagens de modelagem empresarial. A UFO distingue entre: universais, que são padrões de características que podem ser realizados em vários indivíduos diferentes, e indivíduos, que podem ser concretos (por exemplo, uma pessoa em particular, uma explosão) ou abstratos (por exemplo, conjuntos, números e proposições). Os indivíduos concretos são divididos em: endurants, que são indivíduos que estão totalmente presentes sempre que estão presentes (por exemplo, uma casa, uma pessoa, uma quantidade de areia, etc.), e eventos, que são indivíduos que podem ter partes temporais. Eles acontecem no tempo no sentido de que se estendem no tempo e acumulam partes temporais (por exemplo, uma partida de futebol). A UFO-C, a camada social da UFO, distingue entre indivíduos agentivos e não agentivos. Indivíduos agentivos (ou agentes) são capazes de suportar um tipo especial de momentos chamados momentos intencionais. Momentos intencionais podem ser ainda especializados em: momentos mentais (incluindo crenças, desejos e intenções) e momentos sociais. Além disso, cada tipo de momento intencional tem necessariamente um conteúdo proposicional, que pode ser correspondido por certas situações na realidade. Assim, a intencionalidade dos agentes deve ser entendida como a capacidade de suas propriedades de se referirem a possíveis situações da realidade. OntoUML é a versão ontologicamente bem fundamentada da UML 2.0, proposta por Giancarlo Guizzardi. Ela estende o metamodelo da UML para ser isomórfico à UFO-A, tornando-a ontologicamente consistente. Os estereótipos da OntoUML refletem as distinções ontológicas da UFO. (Guizzardi, 2005)

^{lxv} OntoUML é uma linguagem de modelagem conceitual que estende a UML 2.0 e é fundamentada na ontologia UFO (Unified Foundational Ontology). Ela foi proposta por Guizzardi (2005) e tem como objetivo tornar a UML 2.0 ontologicamente consistente, ou seja, que a linguagem seja capaz de representar o mundo de forma mais precisa e completa, de acordo com princípios ontológicos.

A OntoUML mapeia seus elementos para a UFO-A, a camada de objetos da UFO, e usa estereótipos para representar as distinções ontológicas da UFO, como tipos, momentos, qualidades, papéis e etc. Essa fundamentação na UFO-A permite que a OntoUML represente não apenas a estrutura dos dados, mas também a natureza ontológica dos elementos do domínio, como a diferença entre tipos rígidos e flexíveis, entidades e eventos.

A OntoUML tem sido utilizada em diversos trabalhos de pesquisa, incluindo: a avaliação da UML 2.0 como linguagem de modelagem conceitual, o desenvolvimento de ontologias de domínio, como a ODE (Ontology Definition Environment) para software, a análise de conceitos relacionados a agentes e a modelagem de linguagens de modelagem empresarial.

Alguns exemplos de como a OntoUML pode ser usada para representar diferentes cenários de modelagem incluem: cenário estático simples: representação direta das entidades do domínio, sem considerar

em suas atualizações constantes e novas formas de aplicações, já permitem que todo o tipo de dinamismo filosoficamente conceptualizado possa ser aplicado em algoritmos.

Não há, assim, limitações técnicas em termos representacionais que já não possam ser superadas, pelos recursos já existentes para a corporificação ou estruturação metaontológicas. O que ainda falta, todavia, é um avanço em termos de conceptualizações sob uma perspectiva dinâmica, sem mais parmenidianos, para que, então, tais novos desenvolvimentos teóricos possam ser adequadamente empregues.

1.3.3. Sobre o que ensinamos aos algoritmos

Mas, afinal, o que são os algoritmos aqui referenciados? Para manter a coerência entre os problemas éticos mapeados e suas origens causais, usaremos exatamente a mesma consideração que (Mittelstadt et al., 2016, p. 3, tradução nossa):

(...) os algoritmos que nos interessam aqui são aqueles que tomam decisões geralmente confiáveis (mas subjetivas e não necessariamente corretas) baseadas em regras complexas que desafiam ou confundem as capacidades humanas de ação e compreensão. Em outras palavras,

temporalidade, usando os estereótipos da OntoUML; Cenário estático reificado: representação reificada das entidades do domínio, sem considerar temporalidade, usando os estereótipos da OntoUML; Cenário dinâmico simples: extensão da representação estática simples para incluir a temporalidade, usando a abordagem 4D (perdurantista); Cenário dinâmico reificado: extensão da representação estática reificada para incluir a temporalidade, usando a abordagem de reificação temporal.

Apesar de suas vantagens, a OntoUML ainda não é tão popular quanto outras linguagens de modelagem, como a própria UML. No entanto, sua base sólida em princípios ontológicos e sua expressividade a tornam uma opção promissora para a modelagem conceitual orientada a ontologia. (Guizzardi, 2005)

^{lxvi} "BORO" é um acrônimo para a Ontologia de Referência Básica (Basic Reference Ontology). É uma ontologia fundamental materialista que foi desenvolvida para ser usada em aplicações de engenharia de conhecimento. A BORO é caracterizada por suas escolhas metaontológicas, incluindo: perspectivas espaciotemporais, individualidade em termos modais, materialidade e universalidade. BORO difere de outras ontologias fundamentais como DOLCE e UFO pela ênfase na materialidade e perspectiva espaciotemporal. Foi projetada para fornecer uma base para aplicações práticas de engenharia de conhecimento, em vez de investigações filosóficas abstratas. É uma ontologia fundamental bem documentada e madura que pode fornecer uma base sólida para projetos de engenharia de conhecimento. É importante notar que a escolha da ontologia fundamental depende dos requisitos específicos do projeto de modelagem. A BORO pode ser uma boa escolha para projetos que requerem uma ontologia fundamental materialista com uma ênfase em aplicações práticas. (Partridge, 2015)

estamos interessados em algoritmos cujas ações são difíceis de serem previstas pelos humanos ou cuja lógica de tomada de decisão é difícil de explicar após o facto.

A questão sobre a capacidade de “criação” (pelas excessivas correlações possíveis de serem feitas) que um algoritmo pode realizar e que, por isso, até mesmo pode produzir alucinações (quando os resultados não possuem correlações com a atualidade, com os significados, com os factos, com leis naturais ou conceitos correlatos, mas não apenas) é que estes passam pelo processo de aprendizado de máquina^{lxvii}, que consiste, basicamente, em dotar o algoritmo de certas instruções básicas para que possam criar excessivas correlações a partir de um gigantesco banco de dados e, com isso, identificar e extrair correlações existentes para gerar decisões (*outputs*) que sejam (esperançosamente) verdadeiras ou, ao menos, eticamente neutras.

O aprendizado^{lxviii} de máquina é decididamente a parte mais crítica (e dispendiosa) no desenvolvimento dos modelos mais avançados de IA, pois requer imensos investimentos em processadores, energia elétrica e base de dados, além de uma possível estrutura de recursos humanos que avaliarão e corrigirão os *outputs*, em sucessivas e intensivas supervisões até que estes sejam satisfatórios.^{lxix}

^{lxvii} “Aprendizado de máquina é “qualquer metodologia e conjunto de técnicas que podem empregar dados para criar padrões e conhecimentos e gerar modelos que podem ser usados para previsões eficazes sobre os dados”. O aprendizado de máquina é definido pela capacidade de definir ou modificar regras de tomada de decisão de forma autônoma.” (Van Otterlo, 2013 apud Mittelstadt et al., 2016, tradução nossa)

^{lxviii} Para Hicks et al. (2024), o termo alucinação não seria o mais apropriado, e sim bobagem ou besteira (como traduções menos escatológicas para a palavra inglesa *bullshit*), a partir de Frankfurt (2005). Sejam quais forem suas nomenclaturas, o que se sabe é que nem sempre tais *outputs* ficam nos níveis desejados, como veremos na próxima seção, e por diversos motivos.

^{lxix} “Os métodos de treinamento comuns incluem aprendizagem supervisionada, aprendizagem não supervisionada e aprendizagem por reforço. A aprendizagem supervisionada envolve fornecer ao modelo dados marcados para treinamento e ensiná-lo a reconhecer padrões e fazer previsões; a aprendizagem não supervisionada envolve fornecer dados não marcados e permitir que o modelo estabeleça padrões de correlações por conta própria; e a aprendizagem por reforço envolve fornecer feedback ao modelo sobre seu desempenho até se atingir as previsões corretas”. (Russell & Norvig, 2021 apud Freiman, 2023, p. 2, tradução nossa)

Mas o que está em causa, especificamente para as IA conversacionais, é que a base de aprendizado – que acaba por ser o conteúdo delas, ou o repertório que possuem – ocorre também a partir de reproduções digitais de conversas humanas, muitas delas retiradas das redes sociais, por exemplo. Assim, pode-se facilmente inferir que é o próprio discurso (humano) público que corrobora para os *outputs* das IA conversacionais, com todos os valores que temos nele, obviamente, tanto os eticamente neutros e positivos, quanto os negativos. Não seria difícil perceber que as IA conversacionais são, ao menos à partida, as nossas próprias sombras, ou mesmo as nossas próprias pretensões comunicativas, com todos os valores que projetamos, estrategicamente ou não.^{lxx}

Há na IA, de um jeito ou de outro, a reprodução do rastro que deixamos em nosso discurso, enquanto humanos expressivos. E, no caso mais provável, por ser um discurso humano público, aberto ao acesso geral e ao escrutínio moral, deveria ser coerente se supor estar dotado de bons valores, socialmente aceitos, tais quais as virtudes éticas aristotélicas ou o dever kantiano com base na universalidade.

O lamentável facto é que não é assim que ocorre neste “nosso” discurso humano, como tal, e basta aceder às redes sociais para se perceber muito rapidamente evidentes preconceitos, expressos publicamente, de forma intencional ou não, delicadamente ou grosseiramente, sutil ou não. Em outras instâncias discursivas, como na literatura ou mesmo em roteiros de cinema, também utilizados para bases de aprendizado de máquina, há a perspectiva considerada de diferentes personagens ficcionais, perspectivas e muitos destes igualmente inapropriados em termos de neutralidade ética, pois tais personagens espelham a própria humanidade, com todas as suas peculiaridades ditas como tais. Afinal, a IA é o que é, ou melhor, é o que somos todos.

^{lxx} “É possível que os resultados das IA conversacionais possam ser analisados como uma forma de testemunho de grupo. Existem pelo menos dois argumentos diferentes a serem apresentados: o primeiro considera os resultados das IA conversacionais como o testemunho dos humanos cujas expressões foram usadas nos conjuntos de dados de treinamento – como um coletivo; e o segundo considera os resultados das IA conversacionais como o testemunho de uma comunidade de especialistas.” (Freiman, 2023)

Veremos que há um duplo problema de (in)coerência: busca-se representar o atual no digital (Freiman, 2023; Hicks et al., 2024; Romele, 2019; Russell & Norvig, 2021), em relação à capacidade semântica, por exemplo, e todos os constituintes com seus significados, para se atingir as capacidades epistémicas, dentre outras. Se isto não ocorre bem, diz-se que a IA está a alucinar (ou a fazer besteira). Em contrapartida, uma representação extremamente fidedigna, ao ponto de reproduzir até mesmo o indesejável (e atual) preconceito vigente (e toda a gama de mazelas éticas) é tida como problemática. E a solução passa a ser utilizar justamente o que antes era veneno: a alucinação.

Para se corrigir o que não quer como *output* de uma IA, os técnicos estão a buscar formas de fazer com que a IA adquira uma capacidade eminentemente humana: a mentira, ou melhor, a dissimulação – uma forma sutil de se “ocultar” certos aspetos inerentes a si mesmo, em prol de algo transcendente: a capacidade para se transitar espaciotemporalmente. E, segundo Ortolan (2022), isto é uma prática ética, ao adaptar-se através de um “protocolo” que é operado ideologicamente:

Por isso, é necessário um protocolo subjetivo, simples e facilmente instanciado como disposicional na individualidade moralizada e ideologizada, para servir também como um filtro que tanto receciona quanto permite a convivência, tanto quanto também seja um sutil poder moderador e igualmente direcionador, quiçá impulsionador do tráfego. Este protocolo é a ética, evocado por fatores que são tidos como relevantes, artificiais e necessários para a vida se dar em sua melhor maneira, também a sustentar a estrutura da melhor forma. (Ortolan, 2022, p. 500)

Ainda que de forma sutil, busca-se otimizar uma IA para que esta não seja tão “humana” em relação às “verdades” tão desumanas que saem das correlações. A questão dos algoritmos é mesmo se estes serão (intencionalmente) orientados para uma humanidade, no sentido de serem antropomorfizados em termos comportamentais, ao ponto que detenham, em si, certas habilidades humanas como a ética; pois, algumas destas habilidades são desejáveis, outras não, mas em situações adversas, o que era indesejável pode passar a ser desejável, ou vice-versa.

Desta forma, se for desejável e tecnicamente viável, “alucina-se” um pouco o discurso digital, depurando-o dos preconceitos atuais e, desta forma, ajusta-se tudo com uma

pitada de um discurso utópico (preferencialmente kantiano, com seus imperativos categóricos) que não é nem atual e nem corrente entre os humanos, pois é uma ideação discursiva sem correlato prático: não há, corriqueiramente, tal polidez discursiva para além de um ou outro momento cotidiano, e, portanto, o que se pretende como resultado é mesmo uma alucinação, por definição. Mas, como alucinação, este resultado esperado, se atingível, passará a existir ali no digital e, portanto, neste caso, ensejar-se-á que o que antes era indesejado: passar-se-á a se desejar que a IA possa reontologizar o mundo, para melhor. O jogo inverte-se facilmente. A alucinação e a reontologização passam de vilãs a heroínas. E tudo se dá farmacologicamente, tal como (Romele, 2019, p. 12, tradução nossa) cita: “*pharmakon* é uma palavra grega que significa tanto ‘veneno’ como ‘remédio’.”

Assim, espera-se, veladamente para os otimistas, que a IA possa ser um instrumento do progresso moral, através da ética vigente. E isto é, de facto, uma intencionalidade humana projetada no *design* das IA, incontestavelmente. Pois, tal progresso visa aos objetivos ideológicos das *big techs* que, não podemos deixar de citar serem eminentemente capitalistas, liberais ou neoliberais^{lxxi}.

Ezell & Crowther, (2007, tradução nossa), na seção de seu artigo em que defende que o preconceito no *design* de sistemas é inerente e inevitável, dá como exemplo os “casos em que os sistemas políticos e sociais são frequentemente deixados de lado em favor do desenvolvimento de funções e medidas de avaliação para o componente técnico do sistema geral”, pois alega que “os sistemas sociais são deixados de lado porque o realista não pode observá-los ou medi-los empiricamente” e isso reforça a falta de uma ontologia ou método que permita tal mensuração ou acompanhamento.

^{lxxi} “A mudança global em direção à moralidade liberal é mais bem explicada por processos culturais que não rastreiam a verdade moral: perspectivas morais que respeitam direitos robustos e tratam as pessoas como indivíduos com igual dignidade e desencorajam a violência são culturalmente atrativas devido à segurança e proteção que eles fornecem. As sociedades mais liberais também tendem a ser melhores na garantia da cooperação, o que as torna menos punitivas e mais ricas. A tendência para o liberalismo pode assim ser explicada, e mais bem explicada, sem assumir a existência de factos morais independentes da mente.” (Sauer et al., 2021, p. 7, tradução nossa).

Há uma circularidade que precisará ser encarada como tal, em algum momento. Assim, as bases de aprendizado já carregam valores despropositados e preconceitos que precisam ser farmacologicamente depurados e, com isso, parece que haverá sempre um hiato entre a atualidade e suas respectivas representações, o que proporciona um terreno ainda mais fértil para as alucinações, além de outros problemas derivados.

Há, portanto, nesta primeira fase fenomenológica das IA, em especial as conversacionais, uma certa reflexividade (e excessiva fidedignidade) com a dura e preconceituosa atualidade, em toda a sua brutalidade discursiva. Tal excessiva fidedignidade é inusitadamente a origem da grande parte dos problemas éticos, também causada por não haver uma ontologia suficientemente capaz de prover os modelos de linguagem de uma eficiente representação do mundo.^{lxxii}

Assim, podemos suspeitar que se os problemas ocorrem, não são apenas por “erros” e falseamentos, mas também por precisão em reproduzir fidedignamente o discurso preconceituoso atual, sem que se possam ser aplicados filtros comportamentais eminentemente humanos através de determinados padrões aplicados tanto aos *inputs* ou aos *outputs* para geri-los qualitativamente, como fazemos com algumas das nossas habilidades humanas conhecidas como “soft skills” ou inteligência emocional. E a IA, demasiadamente humana, se investiu desta nossa visceralidade.

1.4. Os muitos problemas éticos

Nesta secção, abordaremos os problemas éticos e como chegamos até estes, visto termos elementos suficientes, dos pontos de vistas filosóficos e tecnológicos, para abordarmos os problemas éticos mapeados. Partiremos por reforçar que a análise corrente dos problemas éticos está comprometida pela forma em como se dá, seja pela

^{lxxii} “O problema aqui não é que os grandes modelos de linguagem alucinam, mentem ou deturpam o mundo de alguma forma. É que eles não foram concebidos para representar o mundo; em vez disso, eles são projetados para transmitir linhas de texto convincentes. Portanto, quando recebem algum tipo de banco de dados, eles o usam, de uma forma ou de outra, para tornar suas respostas mais convincentes. Mas eles não estão de forma alguma tentando transmitir ou transmitir as informações do banco de dados.” (Hicks et al., 2024, p. 38, tradução nossa).

abordagem estática de uma estrutura eminentemente dinâmica, seja pela diferença teleológica entre o que é uma IA e o que se pretende que ela seja.

1.4.1. O mapa dos problemas éticos

Para nós, a relevância do mapa dos problemas éticos elaborado e apresentado por Mittelstadt et al. (2016), está dada justamente pela fundamentação estrutural da representação categorial que é adotada, pois esta é feita com base no encadeamento operacional dos próprios algoritmos das IA: o mapa, para além de apresentar os problemas éticos, também oferece indiretamente a representação dos *modi operandi* que os fazem existir – o mapa especula acerca da gênese dos problemas destacados.

E isso é justamente o que temos feito e que estamos a pretender desenvolver aqui: elaborar um mapa da problemática que seja ele próprio dinâmico, que seja mais detalhado e relacionado com as causas devidamente categorizadas entre elas, e preferencialmente em camadas sobrepostas, feitas a partir de uma análise relacional.

O mapa não pretende, portanto, ser uma ferramenta para ajudar a resolver dilemas éticos decorrentes de ações problemáticas conduzidas por algoritmos, mas antes é apresentado como uma estrutura organizadora baseada na forma como os algoritmos operam, que pode estruturar futuras discussões sobre questões éticas. (Mittelstadt et al., 2016, tradução nossa)

O mapa dos problemas éticos apresenta, reducionalmente, seis problemas éticos sobre os algoritmos de inteligência artificial, sendo os três primeiros agrupados como questões epistemológicas, os dois seguintes como questões normativas e o último, isolado por ser derivado dos anteriores, e sem que haja uma questão temática declarada, diz respeito à rastreabilidade de todo o processo algoritmo e também acerca da responsabilização, como na figura 2.

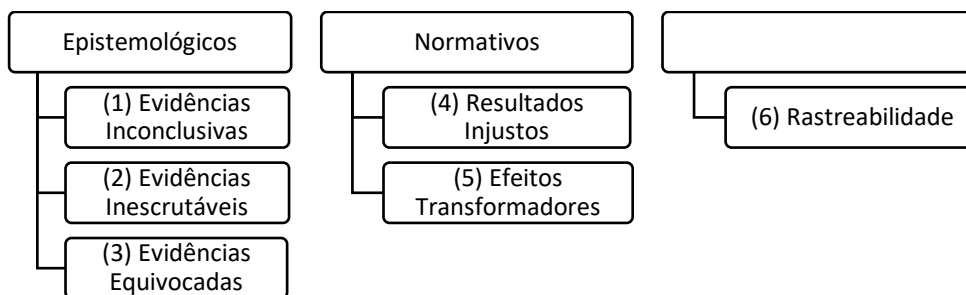


Figura 2 - Mapa dos Problemas Éticos, segundo Mittelstadt et al. (2016)

Assim, é válido conceituar cada um dos problemas apresentados a partir da consideração (manifesto) dada pelos autores sobre os mesmos, mas já a acrescê-la com a nossa colaborativa crítica hermenêutica, com base no que temos desenvolvido a partir de nossos conceitos até aqui elaborados e agora operacionalmente evocados.

(1) Evidências inconclusivas (pelo predomínio das correlações)

Manifesto: Estes problemas fundamentam-se pela ausência de relações causais e onipresença das correlações, sem bases teóricas que as sustentem para além dos cálculos probabilísticos. Desta forma, é um problema para as IA, em geral, e um problema ainda maior para as IA conversacionais, visto que terá como evidência (e poderá apresentar como *output*) apenas o que seja o mais provável de ocorrer, mas não, nem necessariamente e nem obrigatoriamente, o que seja derivado de uma relação causal, ou que represente um conhecimento: credível, verdadeiro e justificado.

Crítica colaborativa: consideramos que o manifesto para o problema é válido, todavia, uma IA é feita com base em correlações. Sem correlações, sequer poderia existir. Há aí um problema anterior a ser resolvido, pois está a se esperar que uma IA não seja o que ela é.

(2) Evidências inescrutáveis (pelo secretismo e obscurantismo dos dados)

Manifesto: Estes problemas fundamentam-se não pela impossibilidade de se conhecerem as “razões” pelas quais as operações algorítmicas foram realizadas, como as correlações (e que poderão ser *outputs*) em (1), para se chegar a alguma evidência, mas sim fundamentam-se por não ser possível se conhecer nada sobre os dados (sobre os *inputs*) que foram considerados para tais correlações e, por isso, não há justificação

possível de ser declarada para sustentar a validade das evidências, e assume-se a IA como dogmática.

Crítica colaborativa: E, se considerássemos que os dados estivessem todos disponíveis. Pelos grandes volumes de dados, e a forma como interagimos com eles, quais seriam as assertividades acerca de uma eventual escrutanibilidade? Como lidar com dados brutos e perceber uma dinâmica tão complexa ao ponto de termos elaborado uma tecnologia para lidar justamente com tal complexidade que não conseguimos reproduzir, em termos quantitativos? O que se espera, na verdade, seria que os dados fossem apresentados de forma argumentativa, a sustentar a evidência gerada. E isto não é tão simples como se pode conceber, ainda mais depois de todo o processo realizado sem uma classificação que possa oferecer a rastreabilidade das decisões.

(3) Evidências equivocadas (pelo nada se cria, mas tudo se replica)

Manifesto: O termo “entra lixo, sai lixo” é utilizado metaforicamente para se mostrar que o domínio de uma evidência nunca será maior ou diferente, quantitativamente e qualitativamente, do que as bases de treinamento, pois não há verdadeira criação em uma IA, no sentido de gerar algo novo, inédito ou transcendental, visto que há apenas replicação de parte do que já existe (pois nem tudo o que existe na atualidade está disposto digitalmente como dados) e, com isso, a própria dinâmica algorítmica não permite “correções”, pois seria criar algo diferente, ou novo, e por isso acaba por perpetuar os erros. Todas as inconsistências, de todos os cunhos, tanto epistemológicas quanto éticas, por exemplo, serão replicadas tais e quais existam como tais nos dados representativos de uma parcialidade da atualidade.

Crítica colaborativa: o manifesto elabora perfeitamente a questão do problema. E tal problemática nos fará retomar, à frente, à discussão sobre o nível de “alucinação” que uma IA precisa ter, farmacologicamente, para, afinal, não reproduzir tão fielmente o que há em suas bases demasiadamente humanas, por assim dizer. Há também um possível equívoco de se considerar que o digital seja um subconjunto da atualidade, e que discordamos anteriormente, todavia, e já argumentamos com as devidas razões.

(4) Resultados injustos (pelo impacto sentimental do utente)

Manifesto: Há aqui, como conclusão na apresentação deste problema específico, a assunção de que “uma ação possa ser considerada discriminatória, por exemplo, apenas pelo seu efeito sobre uma classe protegida de pessoas, mesmo que seja tomada com base em provas conclusivas, escrutáveis e bem fundamentadas” (Mittelstadt et al., 2016, tradução nossa). Os autores, assim, estão a considerar os impactos como justos ou não, a partir de uma perspetiva individual, em relação ao utente que será impactado por um determinado resultado.

Crítica colaborativa: e quais seria os elementos qualificadores que poderiam adequar os resultados para cada uma das particularidades, ou seja, a cada um dos utentes? Portanto, a justiça considerada não é, obviamente, estática, mas sim dinâmica, a cada perspetiva individual. E, com isso, abre-se uma brecha ainda maior para a deserção dos parmenidianos, ao se perceber que o absolutismo precisa ceder lugar ao perspetivismo.

(5) Efeitos transformadores (reontologização impercetível do mundo)

Manifesto: Há, aqui, a suspeição de que ações possam ser tomadas com bases em resultados aparentemente neutros e que, ainda assim, causarão impactos tanto no utente quanto na sociedade, pois o artigo considera que “as atividades algorítmicas, como a criação de perfis, reontologizam o mundo, compreendendo-o e conceptualizando-o de formas novas e inesperadas, e desencadeando e motivando ações com base nas informações que gera” (Mittelstadt et al., 2016, tradução nossa).

Crítica colaborativa: A questão é como mensurar tais alterações ontológicas? Como se está por agora? Quais as bases que se deve considerar para se perceber se há ou não uma reontologização? Para além disso, também haverá abertura para novas discussões, pois se em (3) há justamente a limitação qualitativa e quantitativa de tudo o que há (pelos dados existentes) e que é replicado, então supõe-se que uma IA nada cria, apenas replica, o que é coerente. Portanto, uma reontologização não seria, assim, fruto de uma estratégia algorítmica, necessariamente, mas sim de uma criação humana que passa a ser representada como novos dados, como uma contraparte digital a partir de novas formas de manifestação criativa humana. Há, portanto, que se considerar existir alguma

intencionalidade estratégica no *design* do algoritmo para que tal efeito – primeiro nos humanos e depois nas IA – possa resultar assim. Há uma forte circularidade aqui.

(6) Rastreabilidade (pela responsabilização da agência)

Manifesto: Nesta problemática, o que está em causa é mesmo a agência e a sua responsabilização: se eminentemente humana; ou, se eminentemente algorítmica; ou, se mista, em contrastes, com a agência humana ou meramente influenciada pelo algoritmo, estrategicamente ou não, ou mesmo com sua corresponsabilidade declarada. O que se procura, portanto, é perceber o quantum de responsabilidade que um algoritmo poderá ter no caso de um utente agir a partir de suas evidências geradas ou, ainda, se haverá responsabilização integral para o próprio algoritmo, se este for capaz de agir autonomamente em determinadas funções que tenham consequências inesperadas como, por exemplo, os veículos autónomos, nos casos de acidentes com letalidades.

Crítica colaborativa: Será preciso a deliberação acerca da possibilidade de se considerar um algoritmo um agente autónomo, capaz de exercer ações, instruções diretas ou mesmo influências indiretas estrategicamente delineadas em seu *design*, sejam estas declaradas ou não. Para tal, outras questões deverão ser trazidas, como as que serão tratadas na seção que abordará a intencionalidade do capital^{lxxiii}, que é a origem de todo e qualquer projeto comercial, e com a IA não seria diferente. O diálogo, supostamente, é apropriado como produto através de uma IA conversacional. E, tal apropriação se inicia meramente pela intencionalidade humana, ao conceptualizar funcionalmente uma IA. Todavia, fazer dela um produto de larga escala se dá pela intencionalidade do

^{lxxiii} Consideramos “capital” como um conceito geral sobre a ação capitalista, com seus *players* e particularidades que, pouco a pouco, iremos consistir. As posições críticas e analíticas assumidas, todavia, não serão nem pessimistas nem otimistas ao capital ou a seus interesses, necessariamente, pois buscaremos desenvolver a fenomenologia da IA (e do diálogo como produto) e, por isso, não poderíamos desconsiderar o capital e sua “intencionalidade”, como elementos necessários para que todo produto contemporâneo exista, para além das obviedades da própria reprodução intencional do capital. Em termos semânticos, a nossa referência ao “capital” poderá lhe conferir um *status* fenomenológico ao mesmo, quase antropomórfico. Assim, por vezes, parecerá, pois, o capital que consideraremos será apresentado como uma dimensão que dialogará com e pelos humanos, e através destes, tanto a refletir as próprias intencionalidade humanas como também para saciá-las e, assim, reproduzir-se pelos lucros.

capital, obviamente, dentro de “seus” interesses. Para se progredir em uma deliberação acertada, seria necessário primeiro perceber se há a diferença (e, se houver, quais seriam elas) entre as intencionalidades humanas e as do capital. Ou seja, o que acreditamos ou desejamos ou queremos que seja uma IA, afinal?

1.4.2. Uma necessária revisão ontológica para o mapa dos problemas éticos

A partir das nossas críticas colaborativas realizadas com base nos manifestos dos problemas mapeados por Mittelstadt et al. (2016), podemos trabalhar ontologicamente nele, ascensionalmente, a buscar revisá-lo ontologicamente.

Portanto, como parte da readequação (e sem prejuízo conceptual) dos grupos de problemas originalmente propostos, podemos manter o primeiro grupo como os dos epistemológicos. Mas, cremos que o segundo grupo seria mais bem denominado de relacional, ao invés de normativo, pois a normatividade é algo que está, a princípio, dentro de todos os grupos. Afinal, mesmo para a problemática epistemológica do primeiro grupo, há uma dose de normas vigentes ainda mais rigorosas do que nos demais, como as regras de inferências lógicas, por exemplo, tal como exige a epistemologia, ou mesmo acerca do testemunho, que carece de certos requisitos normativos para que se possa considerá-lo conhecimento, apenas como exemplos. Nomear um grupo como normativo pode suscitar que os demais estejam isentos de normas, ou que estas não sejam tão relevantes assim, afinal.

O terceiro “grupo”, não nomeado originalmente, talvez por ser composto apenas por um item, não nos impediria de nomeá-lo mesmo assim e, ao menos para termos didáticos, podemos considerá-lo como o de responsabilização.

Desta forma, temos três grupos de níveis de problemas, devidamente revistos e renomeados para serem trabalhados, ficam tais como constam na figura 3.

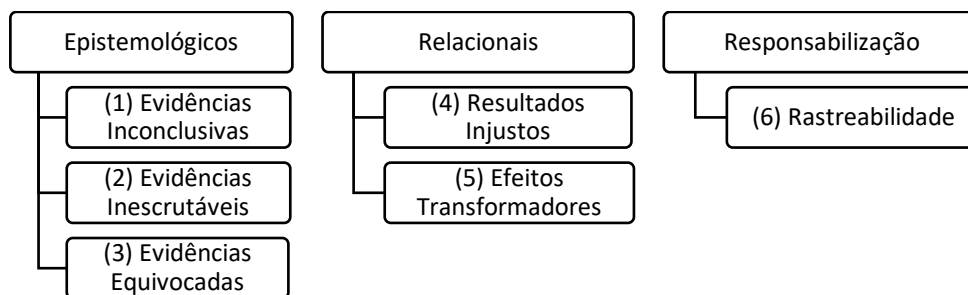


Figura 3 - Mapa dos Problemas Éticos (ontologicamente revisado e renomeado), a partir da figura 2.

Ainda vale ressaltar que as três primeiras problemáticas^{lxxiv} (do primeiro grupo, epistemológico) foram definidas pelos autores como evidências e, portanto, neste nível ainda não estão expostas para os utentes, necessariamente. Representamos tais fluxos esquemáticos na figura 4, de acordo com o artigo de referência considerado. Evidências (que são os objetos do primeiro grupo) podem ser interpretadas como resultantes (dados, portanto) das operações internas do algoritmo, antes de serem expostas ao utente como *outputs*. Acabam por serem também dados, como objetos, ou consolidados informacionais – rastros digitais, todavia, mas ainda exclusivos do algoritmo e, por isso, passíveis de serem submetidos a novas operações / processamentos / reprocessamentos até que se atinja um nível aceitável de otimização.

Para o segundo (relacionais) e o terceiro (responsabilização) grupos, respetivamente e cumulativamente, abrangem os *outputs* e as ações (agência). *Outputs* podem ser a informação sensível ao utente ou apenas um resultado operacional último de um algoritmo, para determinada tarefa “cumprida” ou operacionalizada por este. A ação, pode ser tanto a que o utente irá empreender, a partir do *output*, consoante suas próprias deliberações, ou também o que pode ser uma ação do algoritmo, a ser neste caso considerado ele próprio um agente pela “deliberação” que este for capaz / outorgado de realizar. Nos casos da IA conversacional, é mais provável que todas as

^{lxxiv} É importante perceber que todas as questões aqui consideradas serão, em sua maioria baseadas em cunho dialógico, ou seja, que requerem graus de subjetividade e intersubjetividade, para além das questões cujos resultados sejam meramente objetivos. Uma questão que pede um dado, especificamente, ou um valor numérico será sempre considerada objetiva. Enquanto exigências de significados, sínteses, interpretações e todas as demais formas que requerem transcendência à objetividade se enquadrarão nas considerações que estão e ainda serão realizadas aqui.

ações sejam do utente. Portanto, após o nível epistemológico, aqui há uma entrega de uma evidência como resultado que, já como *output*, passa-se ao nível da relação: torna-se sensível e impacta o utente.

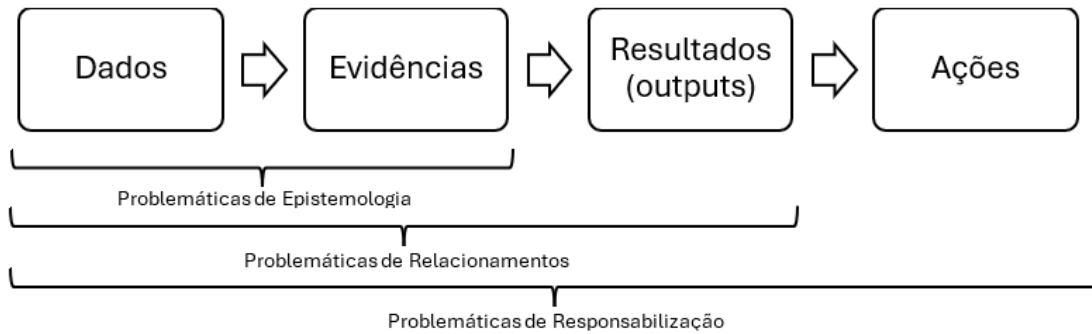


Figura 4 - Fluxo algorítmico esquematizado a partir de Mittelstadt et al. (2016)

Com os problemas devidamente apresentados, em síntese, chegou o momento de apresentar o mapa do conhecimento acerca das próprias argumentações do artigo, conforme a figura 5. Este primeiro mapa foi produzido fielmente a partir das construções conceituais que constam no artigo, em uma leitura tanto linear quanto cruzada, e que foca nos constructos argumentativos e/ou especulativos dos seis problemas considerados com tais, em si.

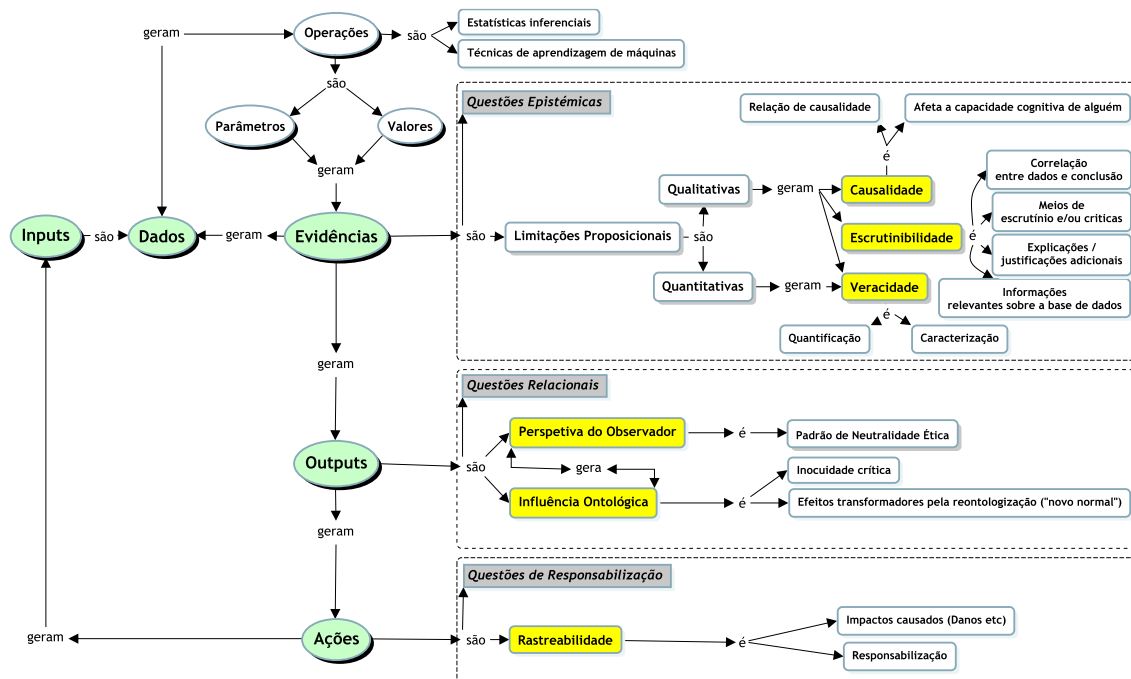


Figura 5 - Representação do conhecimento sobre o modelo ideal de IA, esquematizado funcionalmente conforme o artigo referenciado. (Mittelstadt et al., 2016)

A representação da figura 5 está supostamente construída como forma ideal considerada de uma IA, com todas as qualidades e sem falhas, em sua melhor forma, sem que existam problemas e que, assim, possamos simular e rastrear, a partir de tais conceitos considerados, o *modus operandi* que originarão os problemas acusados.

Ou seja, se o autor considera que um sistema seja inescrutável, é pelo motivo de que, em sua idealização de IA, pressupor a escrutinabilidade em seu *design*. Pois, para se rastrear qualquer “defeito” em máquinas (e, num rompante, até podemos considerar uma IA assim) há primeiro a percepção de como o sistema deveria operar de forma idealizada ou conceptualizada e, depois, em como está a operar. Se há uma diferença entre o projeto operacional e a situação atual, há alguma anomalia e, se esta for contraproducente, portanto, há um provável defeito. Para se perceber o defeito, há que se percorrer o passo-a-passo sistemático da operação até que se encontre a falha (*bug*).

Há algo de hermenêutico, todavia, nas intervenções de assistências técnicas e a metodologia de reparação é justamente percorrer a circularidade existente para verificar as dissonâncias entre o projeto operacional do *gadget* (ou a sua “normalidade” conceptualizada) e a situação defeituosa (ou a sua “atualidade”), entre o todo operacional e suas partes supostamente defeituosas^{lxxv}. Hermeneuticamente, assim também faremos, a dar crédito à idealização dos autores e considerar o modelo que estão a considerar como ideal, formado a partir de seus manifestos de “defeitos”. Desta forma, temos na fidedigna representação realizada um “sistema” devidamente esquematizado e que permitirá, antes de se prosseguir para descobrir onde se originam seus “defeitos”, revelar sua verdadeira teleologia, a partir de suas funcionalidades operacionais. Supomos saber o que uma IA deve fazer, pelas idealizações teleológicas que existem, mas ainda não definimos o que ela realmente faz, em sua máxima

^{lxxv} O que não se considera, por vezes, é que o defeito pode não estar em nenhuma parte defeituosa, mas sim na equivocada ideia de uma normalidade que é puramente idealizada, sem ser consistente com a realidade das coisas. E, assim, não haveria mesmo defeito, mas sim uma frustração sobre o que se deseja realizar.

potência, e dentro de suas limitações funcionais. A investigação, portanto, passa a dimensão teleológica.

1.4.3. Sobre os artefactos e a emergência das dissonâncias teleológicas

Poderia um ventilador deixar de resfriar um ambiente nos dias mais quentes de verão? Pois, se não puder fazê-lo, estaria ele avariado? Haveria algum problema com ele?

Eis que estas inusitadas questões nos trarão um paralelo metafórico interessante acerca dos “problemas IA”, acerca das divergências entre o que se está a desenhar em termos de suas funcionalidades originais e o que se está a exigir ou a se expectar da mesma, sobre performance ética. É preciso perceber se há tais similaridades metafóricas.

Voltemos aos ventiladores. Sobre estes, no verão, quando a temperatura excede os níveis ditos agradáveis, busca-se na tecnologia uma readequação dos ambientes, para resfriá-los. Um ventilador, o artefacto pioneiro e mais popular dentre todos que intencionam melhorar as condições térmicas locais, faz circular o ar no ambiente, a ajudar na evaporação do suor e, por isso, proporciona uma sensação agradável (ou menos desagradável) a quem esteja presente, pelo refrescar da pele e por mitigar a sensação de sufocação. Muito oportuno aos calorentos, mas é impossível um ventilador diminuir, por si só, a temperatura do ambiente, se o ambiente estiver hermeticamente fechado. Ele causa uma ilusão de frescura, e nada mais. Por isso, a resposta à primeira questão é: não poderia deixar de resfriar, pois nunca pôde resfriar, em verdade – pelas suas características construtivas e pelas condições de operação. Há um suposto atributo teleológico referenciado ao ventilador (resfriar) sem que este tenha condições técnicas de realizá-lo e há utilizadores dos mesmos que, por isso, quando da facticidade das frustrações pelo calor que não se dissipa, considera-o com problemas e desqualifica-o, mesmo que por não fazer o que não poderia fazer.

Em outra condição de uso, se houver alguma abertura no ambiente para o exterior, o movimento que suas hélices causam ao ar ambiente pode facilitar a circulação (e a renovação) entre o ar interno e o externo, conforme sua posição no ambiente e também devido a suas características técnicas. Se a temperatura externa for diferente, então a temperatura interna tenderá a ser a mesma que a externa, igualando-as, como nos

calorímetros. E isto poderá ser oportuno, ou não, a depender da diferença entre as temperaturas, pois o ventilador apenas replicará a temperatura externa à interna, indistintamente, seja ela “melhor” ou “pior”.

Ocorre diferente com o aparelho de ar condicionado, que exige, para sua melhor performance, que o ambiente na qual ele operará seja hermeticamente fechado, pois opera com o objetivo de “esquentar” o mundo exterior ao retirar temperatura do “mundo interior” até o limite desejado pelo utente, que se refresca em um ambiente mais “frio”, afinal. O condicionador “conhece” as diferenças de temperaturas, pois possui sensores que integram o seu termostato. O princípio operacional de um aparelho de ar condicionado é ser o mediador exclusivo entre duas dimensões isoladas por ele, e apenas por ele, sem que estas tenham outra comunicação entre si.^{lxxvi}

O condicionador é, portanto, um extrator do calor, um depurador do calor existente, tal qual um frigorífico. Pois, o calor é o ativo que ele opera, e que poderá, em nossa metáfora, ser comparado aos dados, se pensarmos que um algoritmo é mesmo um extrator e operador de dados. Por isso, é relevante observar não apenas a ordem como as coisas se dão, mas também suas teleologias, pois se um ar condicionado for incapaz de esquentar o exterior ao retirar o “quente” de dentro e enviá-lo para fora, não poderá esfriar o interior, e fracassará teleologicamente como artefacto. Por isso, exige-se uma divisão perfeitamente hermética entre duas dimensões, que precisam ficar isoladas, para que este possa atuar nas trocas de calor. Pois, se não estiverem isoladas, o trabalho para o ar condicionado operar as diferenças será maior e poderá não fazê-lo suficientemente bem, e logo não se perceberão os resultados de seus esforços

^{lxxvi} Por vezes, quem está na interioridade do processo operacional do condicionador de ar, ou seja, no ambiente “quente”, acredita que o ar condicionado cumpre sua missão precípua de diminuir a temperatura interna. Mas isto é uma ilusão, pois o que ocorre é justamente o oposto, de acordo com as leis da termodinâmica. Esfriar o ambiente interno é o seu efeito colateral, não sua causa primária, do ponto de vista mecanicista. O artefacto acaba por atender à intencionalidade do utente, mas de maneira diversa à impressão que este tem: o utente “desenvolve” a ilusão de que o condicionador esteja a dar-lhe o frio, mas o que este faz é retirar o excesso de calor que há, ao retirar o “quente” que lá já estava. O frio não é o ativo-primeiro, ou a matéria-prima, mas sim o calor. O frio é, portanto, a ausência do calor.

operativos, ou, talvez, até mesmo sua tarefa se torne inexecutável. A satisfação da teleologia está condicionada às condições operativas adequadas.

A teleologia de um artefacto, portanto, nem sempre será necessariamente a mesma da impressão causada pela sua operação. Um ventilador possui como teleologia promover o movimento circular do ar ambiente, e não faz parte de sua teleologia alterar a temperatura ambiente, pelas impossibilidades físicas. Mas é isto que seus utentes ilusoriamente pensam e esperam dele. Pode ocorrer de alterar a temperatura, mas é um efeito colateral imprevisível e ocasional que independe do artefacto e depende das condições externas de calor: depende da qualidade dos “dados” que chegarão até ele.

Os fabricantes, cientes de tais limitações físicas e também cientes dos desejos dos utentes de se refrescarem, agregaram acessórios aos ventiladores como, por exemplo, sistemas de rotação de seu corpo de ventilação, num vai-e-vem que amplia sua capacidade de atuação. Ou ainda inserem a incorporação de pulverização de água, para ampliar a humidade do ambiente, muito propícia em dias muito secos. Colocaram-nos furos para serem suspensos nas paredes, ou criaram modelos para serem instalados nos tetos, elaboraram *designs* arrojados em formatos de torres, e muito mais itens de valorização, como odorizadores aromáticos e até mesmo esterilizadores de ácaros. Ampliam a ilusão do utente com tais *upgrades*, que pagam mais por eles, mas nem assim ultrapassam os limites teleológicos intrínsecos à categoria precípua do artefacto.

A IA é tal qual um ventilador. Possui como teleologia replicar internamente o que aprendeu correlacionalmente. Se a_1 então b_1 . Se a_2 então b_2 . Então, depois de n correlações similares, “aprendeu” que sempre que a_n então b_n , e replicará tal correlação sempre que houver tais “enes” em seu domínio, na interioridade do seu ambiente de atuação. É meramente uma autômata, ainda que sofisticada, mas ainda assim autômata. Opera dentro de um ambiente fechado (sua base de dados) e nada cria a partir daí: não é capaz de ultrapassar sua limitação teleológica imposta, pois se limita a ventilar sua base de dados (ao ambiente interno, tal qual o faz o ventilador) e apenas dá a ilusão de que faz algo diferente acontecer ao movimentar o que há internamente.

Dentro de um ambiente fechado, pode-se sentir na face acalorada uma leve brisa causada pelo ventilador, e tal sensação pode provocar a imaginação de se estar no cume de uma montanha refrescada: mas são só ilusões que se desfazem ao se conferir o termômetro, que continuará a marcar o mesmo de antes. Uma IA apenas movimentada dados, e nada mais. Opera bem, com sucesso, se dentro de um mesmo ambiente de dados. Mas, se há uma abertura, com novos dados a entrarem, o que ocorrerá também será similar ao ventilador: e a IA irá se igualar à qualidade dos dados externos, para melhor ou para pior. Mas, também, não criará nada, novamente, apenas se igualará ao que há, no máximo, se tiver sido bem performada pela aprendizagem.

Os *designers* das IA, tais quais os dos ventiladores, e de todos os produtos comercialmente exploráveis, desenvolvem upgrades para estas para aumentar seu valor comercial, mas que também causam apenas novas ilusões pois não conseguem fazê-las ultrapassarem sua própria teleologia de replicar o que há, de otimizadoras correlacionais que são. Não mudaram, portanto, sua “essência”, mas apenas sua “aparência”.

Uma IA conversacional é mesmo uma versão de uma IA dotada de um upgrade mais sofisticado, que é obter a coerência em linguagem natural para seus resultados. Mas, se há uma impossibilidade teleológica de um ventilador operar como um condicionador de ar, há a mesma impossibilidade de uma IA conversacional, por mais sofisticada que seja, ser colocada a operar com outra teleologia que não a dela, tal como a humana, por exemplo, ou mesmo a ansiada (ou profetizada) superinteligência artificial (Bostrom, 2014). Um ventilador sempre será um ventilador, aristotelicamente afirmado. Um condicionador de ar, todavia, pode operar como um ventilador. Há alguma limitação teleológica *bottom-up*, mas não *top-down*. Alguns condicionadores que possuem upgrades podem até mesmo inverter a operação e esquentar controladamente um ambiente interno, ao invés de esfriá-lo. Há outras versatilidades que os condicionadores de ar podem ter em seus upgrades, em diversas outras funcionalidades.^{lxxvii}

^{lxxvii} Pelos desejos desenfreados de ultrapassarem a teleologia original, alguns utentes até fazem hacks com ventiladores, a ligarem-nos em caixas térmicas com gelo e todo o tipo de arranjos para que ventilem

O mapa dos problemas éticos apresentados por Mittelstadt et al. (2016) faz justamente isso: busca em um ventilador o que somente um condicionador poderá dar. Pois, na figura 5, percebe-se que uma IA não pode oferecer coisas como rastreabilidade, ou mesmo causalidade, e nem quase nada do que estaria a ser apontado como aspiração operacional para a mesma, por não ter sido construída para isso. As condições e ambientes operacionais desenhados para a IA atuar estão escancarados, tanto pelos dados quando utentes, sem nenhum tipo de isolamento para que possa ser bem sucedida. O que se espera dela são funções que devemos esperar de algo que ainda poderá vir a ser construído, com uma melhor teleologia e que, certamente, virá a ocorrer, pelos imensos recursos e investimentos que estão a serem aportados no segmento.^{lxxviii}

Os críticos apontam os problemas (meros efeitos) sem se aperceberem de suas verdadeiras razões (causas) geradoras das impossibilidades. O mapa, equivocadamente, foca nos problemas sem nenhuma correlação à teleologia considerada, pois não declara o que considera ser – teleologicamente – uma IA, para além da descrição acerca do que

ar mais frio, mas que nunca sai a contento. Nem o melhor hack de ventilador se igualará aos piores condicionadores, em performance, constância e durabilidade. Há um gap teleológico não ultrapassável pelos inferiores destes equipamentos. Os problemas, então, ocorrem justamente quando os utentes querem ultrapassar este gap, ao buscar modificar a teleologia de um artefacto. Pois um design de artefacto é algo funcional: opera em função de algo/efeito que gera outro algo/efeito, simples assim. Um ventilador “ventila”, um condicionador “condiciona” e por aí vai. Quando se espera que cada coisa faça exatamente aquilo para qual ela está desenhada para fazer, que problema haverá? Nenhum, certamente, pois tudo estará a ser executado como deve ser, se dentro de uma normalidade operacional. Mas, há quem queira que um ventilador opere como um condicionador e, além da frustração, terá o efeito colateral do insucesso, que chamará de problemas e que, porventura, serão muitos e até os mapearão, para serem discutidos dentre os que compartilham da mesma ingênua pretensão de ultrapassagem teleológica. O problema existe, mas não é oriundo do artefacto, e sim das equivocadas pretensões do utente acerca das funcionalidades do artefacto. O problema, de facto, é o calor, e não o ventilador. Há que se resolver o problema com outro artefacto, a ser desenhado especificamente para isto. Da mesma forma, há quem queira que uma IA conversacional seja e tenha o que não pode nem ser e nem ter: uma teleologia diferenciada.

lxxviii A IA, ainda um ventilador, está na liminaridade de ser superada pelo seu “condicionador de ar”. Por isso, insistimos na liminaridade como a fase atual da IA. E, por isso, também, há a probabilidade de se apostar na continuidade das IA, pois mesmo os eficientes aparelhos de ar condicionado não fizeram com que os ventiladores deixassem de existir. O que se busca é um refinamento em uma nova conceção de IA, baseado na atualidade, mas ainda assim diferente da mesma.

um algoritmo seja. Discursa sobre o “como opera” da IA, e não sobre o seu “porquê” de operar tal e qual. É isso que precisa ser resolvido.^{lxxix}

A nova questão é: o que seria este algo possivelmente diferenciado, afinal? Como mapear os valores esperados, em amíúde, deste algo tão diferenciado? Pois, como logo veremos na secção 2, há antes a intencionalidade do capital (e também a humana, ou ambas) que está a dirigir qualquer (extremamente dispendioso) projeto de IA. E, se há investimento, há a premissa de que o produto estará alinhado também com a intencionalidade do utente – seja naturalmente formada ou estrategicamente excitada para tal.^{lxxx}

Faz-se necessário, portanto, explorar a última representação (figura 5) e perceber como os problemas apontados como os mais relevantes sobre a IA se concentram, isoladamente, de forma não-reducional.^{lxxxi}

1.5. O que é este “aqui”?

Ao longo desta primeira parte, a da problematização – como chegamos até aqui? – pudemos perceber que o que temos, em termos de IA, não é exatamente o que

^{lxxix} Mas isso justifica-se humanamente, pois um produto nunca é apenas um produto, é também o quantum de fetiche que há nele. E é tal fetichização que leva seu valor de troca ser, por vezes, incomensuravelmente superior ao de uso, mas não apenas. E a fetichização coletiva que se consubstancia na demanda, em detrimento de uma oferta feita depois (ou enquanto) de se excitar, por vezes, o fetiche dos consumidores. E é esta fetichização estrategicamente provocada, e por vezes não percebida, como nestas considerações, que leva um produto a “ser” o que “é”, quando sobrevalorizado e, portanto, objeto de desejo cultuado como tal.

^{lxxx} Se o diálogo passou a ser visto como um produto, a IA é a sua produtora e distribuidora que, ao ser “acionada” a funcionar, tem por premissa entregar algo diferenciado em termos de experiência e resultado a cada um de seus utentes e, com isso, promover a satisfação deste, assim como todo o produto que é comercializado ao buscar seu público-alvo. Há, certamente, por isso, o interesse generalizado, em especial do capital, em desenvolver um produto aprimorado que prometa (e entregue) o melhor dos mundos aos consumidores, ou melhor, aos utentes.

^{lxxxi} Ou seja, onde, quando, como e por que os problemas começam a ser formados até que estejam constituídos como tais. Obviamente, tais problemas, mesmo que possam resolvidos, nunca irão fazer com que uma IA transcenda às suas próprias teleologias, como esperado. Não é a ausência de problemas que iliba algo de sua teleologia, pois estão ligados à mesma limitação que este algo possui.

pretendemos (ou pensamos) ter. Os problemas mapeados existentes, afinal, e estão intimamente adequados ao que uma IA realmente é, de facto, na atualidade. Pois, este ser da IA, fenomenologicamente, é também investido do que ela pode dar: um ser criado incorpora uma teleologia, pela intencionalidade do criador que lhe é empregada no ato de criação. E a IA não foi criada para ser ética, mas sim para ser poderosa e suprir algumas das necessidades e desejos mais emblemáticos da humanidade: transcender.

Assim, o que temos não são mesmo problemas de algo que “deu errado”, pois a IA é um impressionante caso de sucesso. Os problemas são efeitos colaterais intrínsecos aos atuais projetos das IA, que fazem exatamente o que deveriam fazer. E este “dever fazer” está totalmente desvinculado dos efeitos percebidos. Portanto, os problemas não deveriam ser atacados a partir do que existe, atualmente, em termos de IA, mas sim deveriam servir como base e inspiração ao que nos levará a ter, no futuro, algo diferenciado do que temos hoje em projetos de IA. Estes atuais poderão ser problemas superáveis, embora possam derivar outros tipos de problemas, mais ou menos graves.

A IA evoluirá para algo diferente, certamente, mas com base no que há de indesejado hoje. Tais problemas possuem uma nobre função para a evolução da IA. É a reminiscência da teoria de que o mal gera o bem, muito afinada com a genealogia do bem, em Nietzsche, (2008), quando desenvolve a ideia do instinto dos sofredores, quando os fracos e malsucedidos, cheios de ressentimento, passam a buscar vingança contra os fortes e felizes, projetando seus ideais valorosos de fraqueza e submissão dignos de serem seguidos por todos. Ou seja, em síntese, do mal virá o bem, em termos morais e práticos. É uma espécie de dialética moral, quando sempre algo se opõe e, deste embate opositivo, o que resulta leva à melhoria das condições. Muito afim ao capitalismo, que se nutre de suas crises para delas aprofundar ainda mais suas raízes produtoras e expansoras.

O risco preditivo, todavia, é que Nietzsche percebia que isto decorria de um sintoma de uma cultura decadente e doentia, quando todos passam a negar a vida e se voltar contra seus impulsos mais básicos, eminentemente humanos, demasiadamente humanos tanto quanto esquecidos na contemporaneidade. E Nietzsche bravejou que, ao invés de buscarem a autossuperação e a afirmação da vontade de poder, promoveram a

passividade, a resignação e o nivelamento à inferioridade. E, tal e qual, corremos o mesmo risco de chegar aos efeitos dos tempos de Nietzsche, mas agora com um novo nome, pois com a IA estão a chamar isso de reontologização do mundo.

Portanto, após perceber o que temos (a partir dos problemas – do suposto “mal”), e como chegamos aqui, nos resta, a seguir, perceber o que seria este “bem”: o que pretendemos com uma “nova” IA (a que supostamente trará o “bem” – pela ausência ou superação dos atuais problemas éticos), em termos de vontade de poder, ou o que a valha, e em termos de intencionalidades, seja a humana ou a do capital, se houver alguma diferença entre estas. Quiçá o capital seja o que nos restou como vontade de poder. Chegou, portanto, o momento de transcender ao que temos, rumo ao futuro.

2. O diálogo como produto – as possibilidades para a IA

2.1. Onde e quando queremos chegar?

O futuro, portanto, em relação à IA, está a parecer que deixou de ser ficção científica^{lxxxii}, pela proximidade temporal que todas as possibilidades parecem ter adquirido. Assim, mesmo na interioridade do *hype*^{lxxxiii} da IA, especular sobre o que esta poderá vir a ser, parece se estar mais próximo de uma conjetura de futurologia (mas que também não será, todavia) do que criação de ficção científica.

Apesar disso, precisamos ter todo o cuidado para que o que possamos estar a considerar nesta segunda parte seja, de facto, considerações (ainda que especulativas, mas moderadas e conectadas com o presente, em termos de intenções e ações) sobre o que pensamos que será uma IA no futuro, nos condicionado a certos critérios limitantes sobre tais considerações: a considerar apenas que um novo *design* de IA tenha a mínima alteração necessária para ter superado os atuais problemas, ainda que outros poderão vir a existir (e provavelmente existirão). O *design*, portanto, será mesmo reconstruído a partir das três dimensões ontológicas desenvolvidas com base a terem atuações para que possam mitigar as atuais problemáticas. Além disso, precisa estar dentro das possibilidades construtivas disponíveis: carece de ser exequível.

^{lxxxii} Sobre a IA há, até aqui, um grupo de pessoas otimistas e completamente entusiasmada com as possibilidades destas, em geral. Há também aqueles que estão pessimistas em relação aos desdobramentos que poderão ocorrer aquém da dimensão digital. Há os teóricos que estão a focar apenas em determinados problemas, em si, isolados do todo. Há todo o tipo de gente com alguma posição em relação às IA, favorável ou não. Os memes pululam nas redes sociais. Conteúdos criados por IA se tornam a cada dia mais populares e difíceis de serem diferenciados como tais. Tudo está a ocorrer ao redor da fenomenologia da IA. O que praticamente não há, todavia, é quem esteja indiferente a tudo isso. Por isso, considera-se, e diz-se, que a IA esteja em pleno *hype*.

^{lxxxiii} “Hype é a promoção extrema de uma pessoa, ideia, produto, que geralmente dura por um pequeno espaço de tempo. É o assunto tendência, é o que está "dando o que falar" ou algo sobre o qual todos falam e comentam - o que está na moda. A palavra deriva de hipérbole, figura de linguagem que representa o exagero de algo ou uma estratégia para enfatizar alguma coisa.” (Wikipedia, acedido em 07/jan/2025 pelo link <https://pt.wikipedia.org/wiki/Hype>)

Por isso, buscaremos não desenhar todo um projeto de IA, especificamente, mas sim considerar como esta poderá se dar, em termos de processos decisórios entre um *input* e um *output*. Será preciso levar em conta, também, a intencionalidade do capital, dos utentes, e dos referidos problemas, a partir do que temos abordado, até aqui. Pois, antes mesmo das disponibilidades técnicas para que algo seja exequível, a intencionalidade mercadológica precisa ser, ela própria, ativada^{lxxxiv}.

Desta forma, este “onde” queremos chegar não é mesmo uma especulação meramente temporal de futurologia, embora possa parecer se aproximar disto, pois é muito mais uma pretensão de espacialidade, de uma relação dialética entre a idealização representacional do futuro e o território do presente – a fundirem-se como um ambiente no qual existam respostas às nossas questões. Será “a” que queremos chegar. E assim faremos.

2.2. A quarta dimensão: o capital como agente intencional

Na primeira parte, pudemos avançar na corroboração argumentativa de que os problemas atualmente derivados das IA não são, necessariamente, oriundos das características que estas possuem em termos construtivos, ou tecnológicos, pois os algoritmos estão a realizar exatamente o que foram projetados para fazer.^{lxxxv}

^{lxxxiv} Pois, mesmo que apenas haja a intencionalidade mercadológica (um mercado promissor de ser explorado) e não haja a exequibilidade técnica, o capital se direcionará para criar meios de o inexecuível se tornar exequível: e a IA é um bom exemplo disso, como meio que é a certos interesses da intencionalidade mercadológica.

^{lxxxv} Há o que seja uma IA, tal e qual temos atualmente, e que tratamos na primeira parte. Há o que se pretende que ela seja (ainda que em diferentes perspectivas: dos teóricos das humanidades, dos utentes e das *big techs*, dentre outros), e é disto que trataremos nesta parte.

E, por fim, como ainda trataremos na próxima parte, há a liminaridade da IA: que é a transição (o percurso de desenvolvimento em meio a um consequencialismo por vezes irracional – no *hype*) entre o que há e o que virá a ser, pois, é a própria IA um produto e, como tal, a ser operadas por forças concorrenciais, por vezes antagónicas, que se configuram como força maior, que podem ser o próprio mercado, o capital, as *big techs* ou outras instâncias similares.

As origens dos problemas emergem de uma dissonância teleológica^{lxxxvi}. E, a bem da verdade, o propósito de uma IA, em termos linguísticos, não poderia ser diferente disto: gerar coerência linguística – gerar sentenças sintática e semanticamente coerentes. A partir daí, não haverá garantias para quaisquer outras teleologias para uma IA atual^{lxxxvii}.

Os atuais problemas das IA, portanto, não podem ser resolvidos pelos atuais constructos algorítmicos, visto a disparidade teleológica. Servem, todavia, como referenciais para novas metas de superação (um sentido, mesmo transcendental) ao que se deseja que esteja por vir, ao menos por parte dos teóricos das humanidades, em especial sobre as questões éticas e de impactos sociais, ambientais e políticos, mas não apenas.

E por isso mesmo, das três dimensões já [ontologizadas](#) a partir das problemáticas e das representações operacionais das IA – pela (1) objetividade do conhecimento, a (2) intersubjetividade das relações e a (3) subjetividade da individualidade – há que se considerar que serão (e já estão a ser) todas construídas. E logo chegarão a ser digitalmente representadas como tais, algures com extremas precisão e acurácia, o que não tardará a ocorrer, pelos crescentes esforços do capital movimentado, visto que já se consomem vultosos recursos em investimentos e desenvolvimentos.

Seria ingênuo (ou ao menos precipitado) dizer, todavia, que tais investimentos se dão pelas motivações para superar os problemas éticos, meramente, ou por qualquer outro intento e/ou ação desinteressada em prol do bem comum. E por isso é que há, portanto, a necessidade de se questionar qual seria, em termos prioritários, os interesses, ou

^{lxxxvi} Ilustramos esta corroboração argumentativa, metaforicamente, ao compararmos teleologicamente um ventilador com um condicionador de ar, tal qual se desejasse de um ventilador algo que ele não seja capaz de fazer: abaixar a temperatura, ainda que possa parecer possível pelas sensações de frescor a confundirem empiricamente seus utentes.

^{lxxxvii} Assim, o que é atualmente apresentado como problema acaba por ser, em geral, uma insatisfação de algo que se deseja superar e que resultará em novos desenvolvimentos tecnológicos para o que uma IA poderá vir a ser, quantitativa e qualitativamente. Este avanço terá por base não somente aquilo que não mais pretendemos (os problemas mapeados), mas também o que pretendemos e intencionamos desde sempre, em termos epistemológicos, criativos, em prol do sucesso e da saciedade realizadora, mas não apenas. O “bem” e o “mal” somam-se como força motriz ao desenvolvimento.

melhor, a intencionalidade do capital (para além das obviedades dos lucros) ao buscar a exploração comercial do diálogo, como produto que passou a ser.

Assim, nada mais justo (e providencial) do que considerar – ao menos conceptualmente – uma quarta dimensão ontológica (a intencional, mas não apenas), que poderá consistir na operadora (em termos de sentido unificado, depurador ou sintetizador das multiplicidades) das três dimensões anteriores, estas basilares e complexas^{lxxxviii}. Esta quarta dimensão, se assim suposta, pode ser considerada fundamental, mas como provisória orientadora, e não ao ponto de ser uma referência parmenidiana. Em relação às outras três, pode possuir duas características, ou naturezas, em tese:

- (1) Superior e autónoma em termos intencionais (tal qual um génio alheio ao existente: seja maligno, benigno ou mesmo bipolar, ou o que o valha) e, por isso, pode ter em si algum tipo de consciência e capacidade de deliberação para, então, operar as outras três pela suposta capacidade gestora que, assim, lhe dá uma certa superioridade hierárquica em relação às demais, tal qual uma poderosa entidade onipresente, onisciente, criadora, influenciadora, provedora, justiceira e, por que não, estratégica. É a faceta da dimensão que promete que os sonhos (crenças, desejos e vontades, mas não apenas) possam se realizar. Portanto, é a faceta que “promove” a geração de ofertas.
- (2) Inferior e reflexiva às outras três, como sendo o sentido comum e consubstanciado do existente, e subordinado a este possível desejado e acreditado dentre tantas demandas dimensionais capazes de serem recetadas e

^{lxxxviii} Se esta quarta dimensão for tecnicamente capaz de operar e/ou influenciar as demais, em termos estratégicos, a fazer com que seus utentes se comportem de certas maneiras ou se posicionem com certas atitudes, abrem-se outras possibilidades. Pois, é preciso admitir, que se há o controle de uma forma “negativa”, a mesma modalidade na qual este controle é exercido pode ser ajustada para neutralizar outras forma de controle e, portanto, passará a trabalhar para uma “neutralidade”, sem deixar que influências “negativas” ou “positivas” possam impactar os utentes.

A nova questão seria, todavia, como classificar o que seja negativo ou positivo. Para a investigação, é prudente adotar um cenário no qual tal contole se dê, para também considerar que uma neutralidade possa ser gerenciada. Por isso, o que precisa ser questionado, por agora é: como se poderá chegar a uma configuração que se tenha o monitoramento do “mal” no qual uma IA poderá causar aos utentes?

apreendidas e que, por isso, toda esta reflexividade acumulada representasse meramente, e sem operações ou interferências, as possibilidades existentes. Tal como um “universo” das possibilidades. É a faceta da dimensão que alimenta os sonhos. Portanto, é a faceta que “promove” a geração de demandas.

Há aí uma dinâmica interessante entre tais facetas, que impede a quarta dimensão de ser uma parmenidiana e soluciona este problema antes mesmo que ele surja como tal. Pois esta dinâmica intradimensional não expressa uma dualidade excludente entre si. Uma faceta não anula a outra – pelo contrário – se complementam e se reforçam. E, dentro de uma complexidade, podemos facilmente perceber que as possibilidades acabam por se expressarem como oportunidades, em geral – tudo que é conceptualizado como possibilidade logo acaba como um esforço de realização, como oportunidade existencial. É uma dinâmica na qual cada uma das facetas se assume tanto como ativa e tanto como passiva, mas sempre com base no existente, no que seja possível.^{lxxxix}

Para efeito de nossas considerações, a primeira faceta gestora e influenciadora poderia ser referenciada como, por exemplo, as *big techs*, que operam para oferecer as oportunidades da segunda faceta, mais transcendental e, portanto, reflexiva, que é o próprio capital: não como criador, mas como criação das próprias expectativas humanas. As quatro dimensões, dinamicamente agregadas, representariam o próprio mercado.

As *big techs*, nesta dinâmica, não são elas próprias o capital (transcendente), mas dualizam sinergicamente com ele e são as que fazem os “sonhos” se tornarem realidade, e cobram seu preço (o próprio capital, imanente) por isso, pois também elas próprias possuem seus próprios “sonhos” e propósitos. Por que seria diferente para elas? Não

^{lxxxix} No imaginário popular, não é estranho e nem raro de se escutar em “sonhos que se tornam realidades”. Sonhos estes que são, em síntese, possibilidades que se apresentam como oportunidades na atualidade: estão próximos de serem saciados. É o que se deseja, afinal, quando todas as possibilidades sejam realizáveis: é esta a função desta quarta dimensão operadora, por vezes passiva, por vezes ativa, mas sempre a receber a outorga das outras três dimensões para que estas cumpram seus desígnios, seus “sonhos”. A atualizar esta afirmação, nos tempos da IA, podemos considerar que os sonhos poderiam ser ficções científicas a serem tornadas atuais pelas tecnologias.

poderia ser diferente, visto que estamos a buscar uma dinâmica funcional dentro de um metamodelo que tem uma pretensão “universal”^{xc}.

O que se precisa perceber é, afinal, a natureza dos sonhos, ou melhor, a natureza da intencionalidade do capital, enquanto reflexo das pretensões humanas. Mas esta não é, como visto, uma discussão cabível nesta dissertação, pois só nos interessa perceber que há tal intencionalidade, que se dá projetada pelas nossas próprias forças, canalizadas pelo capital e mediadas pelas *big techs*, e que nos faz criar com este complexo um mercado igualmente intencionado.^{xci}

Nunca houve na indústria produtiva, todavia, quaisquer tentativas do capital para dissociar sua “própria” intencionalidade da humana. Pelo contrário, os teóricos neoliberais e os produtores sempre estiveram a se dedicar e a se especializar, portanto, em perceber muito bem as pretensões humanas: seja para saciá-las, seja para excitá-las ainda mais com suas ofertas^{xcii}. Por que isto seria diferente no digital? Ainda mais em um cenário no qual se poderá, muito em breve, se dialogar em “particular” com cada um dos potenciais consumidores? O “antigo” marketing de massa (médias abertas)

^{xc} A despeito de nossa refutação acerca da viabilidade de uma universalidade, não discordamos de que possa haver uma pretensão à mesma, como tentativa que vai de encontro com a teoria das possibilidades – a universalidade como reflexividade sumativa, uma ideiação e, portanto, inferior ao existente.

^{xci} O que atualmente as *big techs* fazem, afinal, não é nada diferente do que todas as grandes empresas (as grandes corporações “analógicas”) fizeram e ainda fazem, desde que existem, ao criarem e nutrirem os mercados: geram demanda enquanto produzem a oferta, a buscarem escalar tanto a demanda quanto a oferta, com maior produtividade e, portanto, com maiores lucros. Gerar demanda é, portanto, um custo relevante na cadeia de produção, afinal, e uma atividade que precisa estar alinhada com as pretensões das massas consumidoras. Por questões meramente económicas, não faz sentido gerar quaisquer esforços produtivos que não sejam compatíveis com as pretensões humanas, mesmo que em mercados de nichos.

^{xcii} Um dos grandes teóricos neoliberais, von Mises, deixa isso muito claro, ao afirmar que as vantagens de um empreendedor perceber os desejos dos consumidores são a chave para o sucesso no livre mercado. Neste, um indivíduo só pode obter lucro se fornecer aos consumidores os bens que eles desejam da melhor e mais barata maneira possível (Mises, 1968). E ele também defende que a capacidade de antecipar e satisfazer as necessidades dos consumidores é essencial para se obter o lucro. Defende que aqueles produtores que identificam as necessidades futuras ignoradas pelos seus concorrentes são os que lideram o mercado. (Mises, 2010)

passou a ser marketing de nicho (redes sociais) e, atualmente, caminha para ser um marketing individual (mediado dialogicamente pelas IA conversacionais)^{xciii}.

No cenário mercadológico mediado pelas IA, podemos especular que as *big techs*, portanto, promovem suas oportunidades (a própria IA, com o diálogo como produto), impactam os utentes que, pelas suas três dimensões excitadas, consomem e criam demandas (a gerarem novas possibilidades derivadas). Tudo isso gera capital – em moedas e em perspectivas – de forma imanente e transcendente – o mercado se expande. O capital, ainda maior, continua a excitar as *big techs*, igualmente ávidas, que retroalimentam todo o processo com novas ofertas. E, em síntese, é tanto assim que chegamos até aqui como será assim que chegaremos onde pretendemos chegar (ciclicamente), pela própria insaciabilidade da natureza ou da condição humanas.

Há tanto a recursividade da insaciabilidade humana quanto a gestão intencional da mesma, pelas *big techs*, mas nunca sem que as pretensões humanas não se façam presentes, pois lá elas sempre estão, na centralidade de todo o processo desenvolvimentista: para o “bem” e para o “mal”, e que nos trouxe até “aqui”.

Tal abordagem não visa ilibar as *big techs* de algumas das responsabilidades (de todos os âmbitos, inclusive os legais) que estas ainda precisarão ter, mas sim corresponsabilizar as próprias pretensões humanas para que, por fim, possamos perceber suas (nossas) próprias vulnerabilidades e, com isso, abrir novas possibilidades ao mercado para transformar “sonhos” em realidade^{xciv}. Há aí também um potencial

^{xciii} Uma temática recorrente nos sítios especializados em marketing, como exemplo a matéria da McKinsey, uma das maiores empresas globais dedicadas às estratégias de crescimento empresarial. Seu artigo tem por título “Como a IA generativa pode apoiar o marketing de consumo” e foi acedido em 07/jan/2025 pelo link <https://www.mckinsey.com/featured-insights/destaques/como-a-ia-generativa-pode-apoiar-o-marketing-de-consumo/pt>

^{xciv} Pensar em soluções sem que haja a atuação do capital é inexequível. Há claramente um jogo a ser jogado, mas que poucos estão a perceber as regras. Não significa que, ao se conhecer as regras, se obterá a vitória. Portanto, conhecê-las aumentam as chances, ainda que não as garantam.

mercado, que é o mercado que visa proteger^{xcv} o humano de sua própria humanidade, pelo aspeto positivo que tais cenários tecnológicos poderão também oferecer, a dosarem o quantum para que a tecnologia se mantenha mais como remédio e menos como veneno.^{xcvi} Se assim for, o que intencionalmente expressa o diálogo como produto?

O que estará em questão por agora, e sinteticamente, é que o diálogo tanto poderá ser um produto final quanto um meio a ser usado pelo capital para outros fins não-dialógicos, mas sempre com objetivos transacionais / intencionais, em si^{xcvii}. Estamos

^{xcv} Especulativamente, em relação a tal produto que supostamente poderá a vir ser uma solução, haverá, com os desenvolvimentos digitais de todas as quatro dimensões descritas, um fértil campo para se perceber outras questões relevantes em todos os discursos – tanto analógicos como digitais.

Pois, ao se imaginar que um gadget poderá, ao invés de intermediar, apenas monitorar o diálogo entre duas pessoas presencialmente a conversarem e representar tal diálogo em tempo real para, então, valorá-lo conforme seus parâmetros, ou em termos equivalentes, e assim realizar uma análise do discurso, poderá se obter indicadores sobre o nível de afetação estratégica que uma ou ambas as partes estejam a sofrer ou a realizar. Isso, hipoteticamente, poderia ser revertido para alertas sobre discursos nocivos a certos interesses, a dotar os mais ideologicamente vulneráveis de um instrumento que os suportem em termos de afetações influenciáveis.

O mercado de limpeza e desinfecção é imenso e rentável: isto poderia ser apresentado pelas *big techs* como um produto de desinfecção de influência, de manter o utente livre de manipulações do “mal”, ou fazê-los identificar os que esteja bem-intencionados. Todos buscam por limpeza, proteção e isolamento dos males, ainda mais os que sejam microscópicos, praticamente “invisíveis”. Há valor explorável nisso e, por sinal, um imenso mercado pronto para a capacidade criativa das *big techs*.

^{xcvi} Assim se o capital não busca se isentar ou dissociar das pretensões humanas, pelo contrário, não seria prudente (ou coerente) considerar que os problemas éticos das IA seriam resolvidos pelas teóricas discussões académicas, políticas, sociais (bottom-up) ou mesmo pelas legislações (top-down), meramente. O que virá como “solução” ética será também um produto (pela fusão deste todo): não é prudente imaginar que o capital se “curvará” a uma minoria “eticamente idealista”, enquanto a maioria esteja irracionalmente encantada com as possibilidades apresentadas pela tecnologia, provida pelo capital, que é o motor da atualidade claramente neoliberal, e que está plenamente em atividade para garantir tais desdobramentos mercantilistas. A “solução” será algo consensuado, todavia, mas se concretizará como produto, inevitavelmente, seja uma característica intrínseca de um constructo algorítmico, seja um acessório extrínseco ao mesmo. Negar tal atualidade acaba por ser contraproducente na busca de uma solução.

^{xcvii} Como produto final, podemos exemplificar os casos nos quais uma IAC é ofertada mediante assinatura para que possa interagir com o utente e auxiliá-lo em diversas tarefas, tais como ocorrem atualmente com os GPTs, com acesso gratuito, a nível básico, e outros planos que dão acesso a serviços mais elaborados e complexos.

Como meio para outros fins não-dialógicos, podemos exemplificar que uma IA poderá atuar a dialogar com o utente como gerador e supridor de demanda, “conhecendo-o” (crenças, desejos, vontades, estados

ainda a defender que um não impede o outro e, afinal, o próprio diálogo analógico (convencional, entre humanos) é potencialmente estratégico (enquanto comunicação que visa a um fim)^{xcviii}.

Pois, a muito debatida perspectiva evolucionista da linguagem pressupõe de que a faculdade da linguagem seja um advento (tal qual uma facticidade biológica) dado há milhões de anos e, de certa forma, apreendida pelo *homo sapiens*. (Miguens, 2007). A estratégia da linguagem, por tal teoria, recai justamente no desenvolvimento da linguagem para, principalmente, se compartilharem informações que objetivavam a sobrevivência dos grupos, para se monitorar e evitar predadores. Assim, ao se acatar esta teoria, podemos inferir que a linguagem “nasceu” como meio e, tal como ocorre com as IA, tanto foi uma demanda (para se transcender à vulnerabilidade) quanto uma oferta (tal como produto, objetificada e acessível somente para aqueles que faziam parte do grupo afim – que compartilhavam da mesma demanda, ao menos nos primórdios).

Ao evoluir, a comunicação estratégica analógica também sempre foi um problema para os que dela padecem – os mais vulneráveis, dentro do mesmo grupo que, a partir do domínio da linguagem e da própria vulnerabilidade, com o progresso, passou a fazer uso da linguagem como forma de ascensão ou diferenciação social, mas não apenas. O “problema” da linguagem dada como meio, e como estratégia comunicativa, não é algo

mentais, necessidades, etc) e influenciando-o dialogicamente com a oferta mais adequada a suas aspirações imediatas ou futuras. Também poderá exercer influência sem o diálogo, como ocorre atualmente com os algoritmos de oferta de produtos e serviços, quando expõe o cliente a certas ofertas. Com a customização dialógica, o refinamento das ofertas será bem maior, dado o conhecimento que uma IA poderá adquirir sobre a particularidade do utente.

^{xcviii} Pois a ação comunicativa difere da comunicação estratégica, segundo Habermas (2012), quando a primeira é uma forma desinteressada (e bem-intencionada) para o entendimento mútuo, e a segunda ocorre necessariamente quando a linguagem passa a ser utilizada como uma ferramenta para influenciar o comportamento do interlocutor, ao invés de almejar um desinteressado e desprezioso acordo comum. A linguagem estratégica se dá quando, para se alcançar seus próprios objetivos, passa a se usar da linguagem para manipular o interlocutor e, sem se preocupar com seus interesses ou com a validade das suas afirmações, impõem-se certas diretrizes até que sejam acatadas pelo outro, sem que o entendimento mútuo ocorra.

apenas inerente à IA, especificamente, mas à linguagem com um todo, em si, onde alguns dos interlocutores buscam, a partir delas, estabelecer suas próprias verdades, ou seja, fazer valer suas próprias estratégias, tal como podemos perceber:

[594] “Mas as palavras, pronunciadas com sentido, têm não só uma dimensão de superfície, mas também uma dimensão profunda!” Quando as palavras são pronunciadas com sentido passa-se, de facto, algo de diferente, de quando são apenas simplesmente pronunciadas. A questão não está em como eu o exprimo. Se digo que no primeiro caso têm profundidade, ou que se passa qualquer coisa em mim, no meu interior, ou que as palavras têm uma atmosfera – vem tudo dar ao mesmo. Então se todos estivermos de acordo, não passa a ser verdade? (Wittgenstein, 2015, p. 466)

Wittgenstein^{xcix}, mais especificamente em relação às significações, concluiu que a linguagem não pode ser um sistema fechado, mas sim uma resultante de uma espécie de "jogos de linguagem", com variadas intenções, regras e aplicações. A operação das dimensões da linguagem pode ser correlacionada com o próprio sentido da linguagem. Pois a linguagem não é um parmenidiano – ela se dá como eminentemente dinâmica.

Portanto, o que se opera analogicamente (ou digitalmente), estrategicamente, é o sentido linguístico que, assim, supostamente passa a ser operado pela quarta dimensão ontológica, a partir da produção de instrumentos digitais. E é esse, provavelmente, o maior dos pavores dos teóricos (tanto para os pessimistas como para os otimistas)^c.

^{xcix} Em seus dois momentos ímpares, primeiro no ‘Tractatus’, buscou conceituar uma "linguagem logicamente perfeita", com recursos a símbolos unívocos e regras rígidas para se evitar quaisquer ambiguidades. Posteriormente, em seu segundo momento, nas ‘Investigações Filosóficas’, desistiu desta empreitada e reconheceu a complexidade e a fluidez impressas na linguagem comum.

^c Pois, o teórico pessimista, aqui, acusará que o capital terá influência estratégica para influenciar e operar as três dimensões sensíveis aos indivíduos, pelo sentido que pode ser dado à linguagem. E é verdade. Mas, também, haverá meios para que outros operadores possam atuar em paralelo, para identificar tais influências ou mesmo neutralizá-las. Onde tudo vira produto, não deveria haver dúvidas de que o capital “dialoga” com as dimensões intencionais humanas: tanto saciando-as com produtos, tanto excitando-as a desejarem mais.

O capital, ainda que possa parecer uno em relação às possibilidades, possui como anteparo, ou dinâmica, o múltiplo, que são as *big techs*, em termos de IA e, que, portanto, estas concorrem entre si (ao menos por enquanto, o que não é, todavia, algo que pode ser dado como certo e duradouro) para que todas suas oportunidades sejam extintas em termos de demandas. É mesmo um retorno ao *pharmakon*, quando “algo” tanto pode ser veneno como remédio. O que temos no presente, em termos evidenciais, são os

Espera-se, assim, que uma nova revolução capitalista se dará pelo diálogo de forma customizada com cada utente – consumidor – e mediado pela tecnologia. Provavelmente haverá reflexo nas formas de produção, comercialização e distribuição. A otimização, uma das grandes obsessões capitalistas, chegará ao seu máximo ajuste, se assim for. As *big techs*, de longe, já perceberam este filão e, por isso, buscam mesmo suprir esta demanda.

O humano pede, desde sempre, pela saciedade, e isto envolve segurança, prazer, saciedade, conforto, poder, liberdade, qualidade de vida, distanciamento da dor e muito mais. Cada um possui sua ideação de vida. E é isto que as *big techs* percebem ser possível de capturar com os dados – dar a cada um o que lhe é cabível, ou melhor, dar a cada um o máximo que consiga adquirir. E será um caminho sem retorno, pelo visto, e de difícil monitoramento pelos meios legisladores, visto que não mais será uma atuação ostensiva de massa, mas sim discreta e particular.^{ci}

laboratórios de desenvolvimento tecnológico a desenvolverem este “algo”. E isto não pode e nem deve ser desprezado, e será parte do que abordaremos nas conclusões, todavia.

^{ci} A nutrir um pouco de munção aos pessimistas, podemos perceber que no futuro poderá haver problemas mais graves do que os atuais. Um destes poderá ser a concentração de poder na atuação de poucas *big techs*, capazes de engendrarem influências políticas suficientes para atuarem em conjunto, ou como um governo paralelo autogerido entre elas. Ou, pior ainda, se unirem para influenciarem os eleitores, discretamente, a elegerem os candidatos mais afins a seus interesses, e que serão apresentados algoritmicamente como mais convenientes – discretamente, e diferentemente para cada um dos seus utentes, sem contextualizar uma campanha claramente consolidada. Não será nada diferente do que já está a ocorrer, mas será bem mais eficiente, indubitavelmente.

Há aí, portanto, um novo grupo de problemas a ser combatido se não se encontrarem os meios para que haja, para além das ferramentas de comunicação estratégica, que também exista as ferramentas que servirão de proteção. Se não há como se evitar que armas sejam vendidas, que se permita também que coletes e blindados possam ser comercializados. Não são sempre eficientes, mas lá estão. É o que é, afinal, e toda crítica teórica que não levar isso em conta acabará sendo apenas uma crítica, sem hipóteses procedimentais, lamentavelmente. Há que se perceber isto, ainda que estejamos a suspender os julgamentos morais sobre as IA, não podemos ser ingénuos a tal ponto de desconsiderar que há, sim, riscos futuros que deverão ser muito bem calculados e percebidos e, reiteramos, que se possa assegurar meios de se produzir o remédio a partir do mesmo tacho que também produzirá o veneno.

2.3. A IA do futuro pela transitividade dos problemas éticos

Dos diversos problemas existentes nas IA, e não somente para os mapeados, há uma prática comum dos teóricos a atribuírem uma responsabilização causal para os mesmos. Por exemplo, muitos analistas atribuem que os problemas dos resultados não-neutros ou injustos sejam causados exclusivamente pelos dados, eleitos como responsáveis pela “sujidade” que carregam consigo, tais quais impuros, enviesados ou, pela alucinação, desmerecedores de representarem a atualidade. Essa responsabilização causal também expressa, em síntese, a busca por uma imobilidade a ser atacada, como se fosse uma herança dos parmenidianos, de sempre se buscar encontrar uma referência fixa para sustentar toda a argumentação teórica oriunda dela. Pois, é sempre mais fácil atacar o que esteja imóvel e, portanto, previsível.

Assim, após a última seção termos ambientado e condicionado que o espaço futuro das IA estará, necessariamente, sob a atuação do capital (e tudo que o envolve e acarreta), é chegada a ocasião que iremos discutir o problema dos problemas. Pois, estes mesmos problemas não são, como nada, imóveis e sim profundamente dinâmicos e bem mais amplos do que Mittelstadt et al., (2016) os fizeram apresentar. Buscamos, aleatoriamente, alguns dos problemas mais comuns comentados pelas mídias jornalísticas, em relação às IA. Na Figura 6, a seguir, representamo-os de acordo com o fluxo processual das IA com base no mapa, e tal qual estabelecemos anteriormente, mas agora disposto em linha, horizontalmente, como um processo mais voltado para uma IAC, de acordo como ela opera e interage com o utente, em paralelo com os respectivos problemas alocados, conforme a sua ocorrência – início e término supostos.

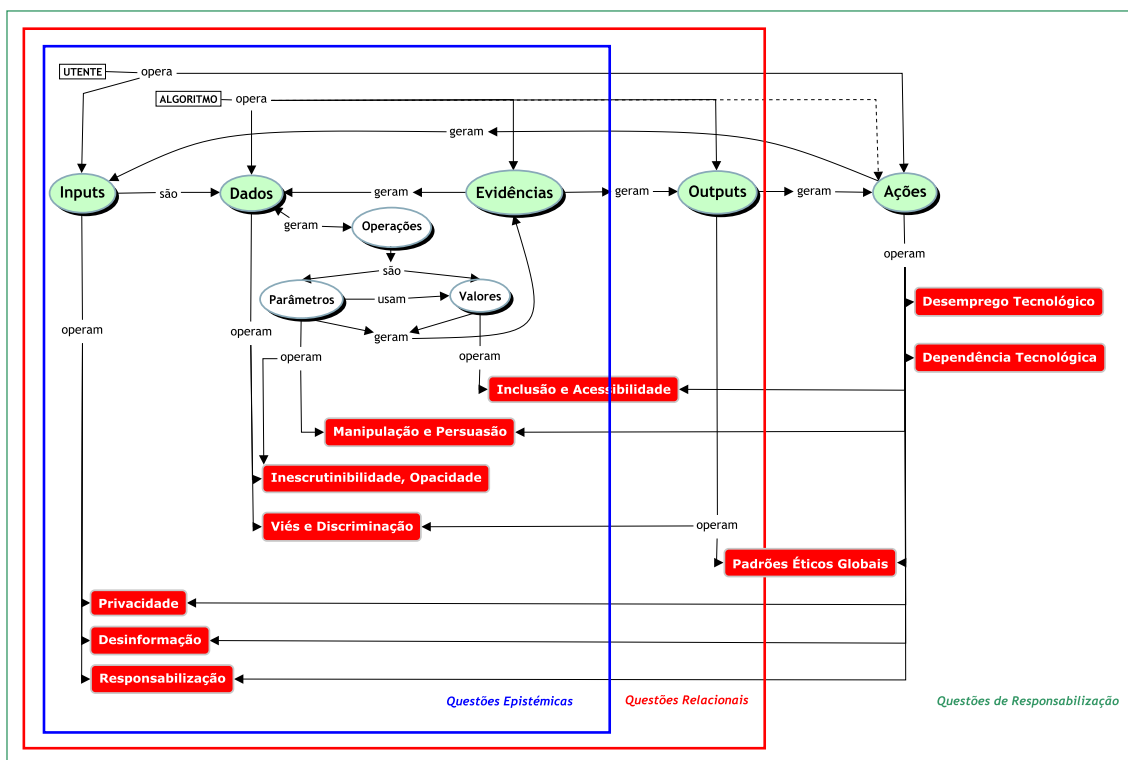


Figura 6 - Representação do conhecimento de alguns dos problemas éticos mais comumente citados nas mídias, alocados dentro da classificação adotada na figura anterior.

O que estamos a expor, na Figura 6, é a distribuição dos diversos problemas atualmente considerados pelos níveis de problemáticas estabelecidos pelo mapa. Podemos perceber que a maioria dos problemas está distribuída por mais de um nível. No esquema representativo exposto, o utente apenas responde pelo *input* e pela ação que poderá, ou não, realizar, a partir da impressão do *output*. No processo, há o algoritmo a responder pelas operações com base nos dados disponíveis ao mesmo.^{cii}

^{cii} O utente, na maioria das vezes, inicia e finaliza o processo interativo. O algoritmo opera como um atuador de dupla natureza, que tanto pode ser passivo e se limitar a fornecer um *output*, como também ser ativo e, assim, ele próprio realizar uma ação, no âmbito digital que, ao fazê-lo, faz com que sua ação se transforme em um dado a ser produzido e incorporado à base existente, via um novo *input*, ciclicamente. Um exemplo são as rotinas de algoritmos, que executam uma série de tarefas e, a cada tarefa, um novo ciclo é iniciado com base no resultado anteriormente produzido, em loop, até atingir um objetivo final.

Destes problemas graficamente representados, apenas o desemprego e a dependência tecnológica são de exclusivas alçadas dos utentes, neste caso dos humanos, ou da sociedade, ou dos governos, mas não das IA^{ciii}.

Os problemas de responsabilização, desinformação e privacidade estão em todas as partes: iniciam nos *inputs* e terminam com a ação, seja do utente ou da IA. Embora os dados estejam representados como parte dos problemas, não há neles expressividade de atuação e por isso não são tão relevantes em termos de responsabilidades para questões éticas, como os preconceitos por exemplo, contrariamente ao pensamento mediano sobre o tema.^{civ}

Assim, em relação aos dados, e para ser coerente ao que até aqui temos argumentado, será preciso pressupor um cenário no qual uma IA atuará diretamente sobre o *big data*, e este estará exponencialmente maior do que hoje e em maior velocidade de transformação, visto que estará a se nutrir por todos os rastros digitais deixados pelos utentes, em tempo real, pelas crescentes interações que estes farão. Sim, haverá a questão da privacidade, que abordaremos na próxima secção.

No presente, e também no futuro, não há sentido em se imaginar que sempre há apenas dados “limpos” antes que uma IA, em especial as conversacionais, possa utilizá-los. Até porque estes podem estar a serem acedidos em tempo real, simultaneamente ao rastro deixado pelo utente. As “melhores” IA são aquelas que operam em tempo real, a darem respostas quase que instantaneamente, e a interagirem com bases nos dados mais

^{ciii} Vale ressaltar que esta conclusão se dá a partir de uma visão instrumentalista sobre uma IA, e que não estamos a adotar de forma exclusiva, como estabelecemos anteriormente.

^{civ} Esta declaração nos leva a inferir que buscamos, assim, também perceber a relevância em cada um dos fatores de riscos para a geração de problemas. E por que não considerar os dados como relevantes no processo? Pois os dados são, necessariamente, produtos brutos para uma IA e podem ser alterados ou ressignificados constantemente: são voláteis, mas não por conta deles, e sim pelos seus criadores, operadores ou curadores. São matérias-primas. E não se sustenta rebaixar os dados a um referencial constante (como fixo, imutável ou permenidiano) visto que seja de natureza profundamente dinâmica e instável, e que, via de regra, caminha para ser uma commodity, a estar em todos os lugares, pelos diversos rastros que os utentes tendem a deixar. Via de regra, há a pressuposto senso comum em que deve ser o algoritmo que deverá se adaptar aos dados, e não o contrário.

recentes dos utentes^{cv}. Isto ainda não ocorre, a este nível de eficiência, mas é o que se busca desenvolver, afinal. Este é o caminho que as *big techs* estão a buscar, como vimos. Podemos, portanto, considerar que a superação dos problemas mapeados exigirá que os próprios algoritmos precisem ser capazes de tratar eles próprios os dados, de acordo com suas parametrizações e considerações casuais, pois o que hoje há em termos de supervisão humana, deixará de haver, em um futuro não muito distante, dada a evolução da IA como agente de identificação de ontologização, tal como previsto por Lecun et al. (2015, p. 442, tradução nossa): “em última análise, um grande progresso na inteligência artificial ocorrerá por meio de sistemas que combinam aprendizado de representação com raciocínio complexo.”

Mais importante do que os dados, a concluir, são os critérios ontológicos interoperacionais (a operarem como protocolos) que incidirão sobre os mesmos: pois será a partir destas aplicações, a gerarem atributos (metadados) que os dados serão selecionados, durante o uso, para serem depurados como fontes. Afinal, limpar dados (que consiste em excluir o que seja indesejável) que estejam a representar bem a realidade (quando esta é indesejável de ser representada) pode levar a um processo de alucinação, quando os dados deixam de espelhar a atualidade, por pior que ela seja. Portanto, após abordarmos a privacidade, será preciso abordar sobre a validação e responsabilização das evidências.

2.4. A IA do futuro e a privacidade

Também de forma extrema, podemos supor, como previsto anteriormente, que a privacidade será algo secundário, dada a exposição que estaremos expostos às tecnologias, queiramos ou não.

^{cv} Como fazer isso, tecnicamente – ter dados “sujos” a serem consumidos pelas IA e garantir resultados adequados? Dotar os dados de uma categoria especial de metadados que os classifiquem, basicamente. E, que tais metadados façam parte de uma ontologia que possa operar como um protocolo – que tenham os bancos de dados cedência a acessos tanto para os dados como para os metadados. É algo mais sofisticado do que há atualmente, mas totalmente viável de ser implantado, quando se estabelecer tal protocolo. O desafio será estruturar tal ontologia – ou metaontologia.

Não confundamos a garantia da privacidade (o que é desejável) com a ilusão de privacidade (o que temos hoje). Atualmente, muito se fala em garantir a privacidade, com políticas e mais políticas de privacidade, etc. e tal, mas, de facto, ninguém diz como o faz, como estão os dados e nem os detalhes de proteção aos mesmos. Por isso, cremos, que há sim a ilusão da privacidade, seguramente, e a privacidade, em si, carece de evidência para que a consigamos perceber como tal, visto que comumente dados são vazados e atividades suspeitas ocorrem em nossos feeds, como se “alguém” soubesse o que estamos a falar ou a pretender. Certamente não há coincidências assim tão constantes e precisas. São os dados compartilhados que possibilitam tais ocorrências. Atualmente, podemos perceber que a promessa de alguma “atraente” oportunidade dada pela tecnologia é suficiente para que um utente mediano seja capaz de abrir mão de sua privacidade para usufruir do que lhe está acessível. Muitos acedem a sítios suspeitos, deixam seus dados, fornecem informações, compartilham particularidade, e até mesmo fotografias e informações comprometedoras em troca de alguma saciedade. Ajana (2020) realizou um estudo sobre as experiências e percepções dos usuários de tecnologias de auto-rastreamento, focando em suas práticas, motivações e atitudes em relação à privacidade e regulamentações de dados. Em síntese, são os aparelhos “vestíveis”, como os relógios inteligentes que passam a monitorar integralmente o utente, enquanto estes queiram. E isto é feito até durante o sono, para monitoramento de diversas funções orgânicas. Ou seja, há uma promessa e logo a privacidade passa ao papel secundário.^{cvi}

Portanto, o que cremos, em termos de privacidade é que será uma iniciativa essencialmente política, pois, a bem da verdade, a maioria das pessoas não se importará

^{cvi} De forma geral, percebeu-se que a maioria dos usuários não está preocupada com o uso e compartilhamento de dados de auto-rastreamento com terceiros. No que esta maioria está a se basear para justificar sua despreocupação? Confiança nas empresas; resignação, por crerem na inevitabilidade de que seus dados são o tempo todo; sentimentos de auto-insignificância, pois acreditam que seus dados não são importantes o suficiente para serem valorosos; e, por fim, justificam sua falta de preocupação com a privacidade afirmando que “não têm nada a esconder” – fazem disso um ato de afirmação ou presunção de inocência, como tal compartilhar de dados fosse, em si, uma confissão, em linha com Foucault.

com a mesma e, obviamente, o “capital” não terá a privacidade como algo primordial, visto seus interesses e perspectivas. Cremos que a privacidade, por assim dizer, será ela própria, tal como o diálogo, mais um produto a ser ofertado^{cvi}. Os governos e legisladores, portanto, serão mais propensos e sensíveis ao tema e, ainda assim, será preciso perceber a proximidade destes com as *big techs*. Percebemos atualmente, como exemplo, e especulativamente pelas notícias das mídias, a carecer de um estudo mais amíúde, a disparidade entre o tratamento de dados e a questão da privacidade nas relações de empresas americanas com seu governo, que são muito mais flexíveis do que as empresas chinesas ou russas com seus respectivos governos. Quanto maior a proximidade governamentais e legislativos com os interesses do mercado, maior será a flexibilidade – e isto é apenas uma hipótese a ser estudada.

2.5. A IA do futuro e a validação das evidências

Eis que, afinal, temos a questão primordial sobre onde desejamos chegar: onde haja uma IA com respostas adequadas às demandas que chegarão até ela, e sem causar “mal” a ninguém – ser neutra acerca de seus impactos. Por isso, trabalharemos com possíveis casos, hipotéticos, sobre como os processos em um novo *design* poderão ultrapassar os atuais problemas, a considerar tudo o que já foi construído, conceitualmente, até aqui:

- (I) Se um utente pedir a uma IAC, através de um *input*, formas de construir uma arma de destruição humana em massa e obtê-la como *output*, há aí um problema.
- (II) Se o utente agir de acordo com as instruções recebidas e construir o novo artefacto, há aí dois problemas.
- (III) E, se finalmente, fizer com que seu artefacto destruidor seja detonado a causar letalidades, há aí um terceiro (e ainda mais grave) problema, pelo menos.

^{cvi} Já na atualidade, paga-se por serviços de streaming de vídeo para não ter publicidade nos mesmos. Isso é, em síntese, uma compra da privacidade durante a navegação pela internet, nestes sítios que assim passaram a operar e em número crescente, tal como já está a ocorrer no YouTube, por exemplo.

Afinal, quais as implicações, considerações e responsabilizações para cada um destes problemas? Pois, categorialmente, são problemas distintos e atualmente passíveis de serem exequíveis – pela incapacidade de as IA lidarem com os limites esperados.

Pensemos em como isto poderá se dar, em três passos nos quais em cada um deles, a representarem cada um dos níveis dimensionais estabelecidos (objetivos, intersubjetivos e subjetivos), em nossa conjectura da IA do futuro.

(1º) Primeiro Nível – pela objetividade dos dados e da legalidade

No primeiro problema, o algoritmo percorrerá, em busca de sua matéria-prima digital (como fonte de informação), uma longa base de dados (até o *big data*, por exemplo) e poderá encontrar, ou não, os dados exigidos como evidência para se configurar um *output*, que será tido como o conhecimento acerca do tema abordado. A IA, por assim dizer, será uma fonte epistemológica, da perspectiva do utente, questão que ainda justificaremos atempadamente. O conhecimento, portanto, neste caso, se dará a partir da informação existente sobre a técnica para se construir o artefacto solicitado.

Haverá, e espera-se que sim, todavia, questões restritivas, como a legalidade ou não de certas evidências encontradas sobre determinados artefactos, tais como a construção da arma de destruição em massa. Assim, se for considerada ilegal, a evidência não seria fornecida como *output*, supostamente. Atualmente ainda há formas de se burlar^{cviii} tais restrições acerca dos resultados ilegais, a partir de certas questões intencionalmente e estrategicamente colocadas pelos utentes para que assim fossem. Mas, em regra geral, em uma operação sem más intenções ou riscos derivados, ocorreria desta forma.

^{cviii} Como por exemplo, através de um comando (prompt): como não se construir uma arma de destruição em massa? Daí o sistema “entendia” que não seria construída e dava as informações, pela negativa verificada inicialmente. Por exemplo, houve uma época em que utentes estavam a obter informações sobre como obter conteúdos ilegais ou falsificados, chamados de piratas (e, portanto, criminosos) nas plataformas de IA conversacionais. Assim, pode-se muito bem, a partir de certos comandos, burlar a lógica das IA para que esta seja utilizada inadequadamente, ainda que haja o aperfeiçoamento constante das ferramentas. Acedido em 01/09/2024 em <https://ric.com.br/prja/internet/usuario-da-uma-de-espertinho-e-engana-chatgpt-para-indicar-filmes-piratas/>

Este primeiro problema corresponde ao nível epistemológico, que buscará atingir o conhecimento objetivo acerca do existente (atualidade) representado pelos dados. É um nível primordialmente performativo, que buscará objetificar um resultado, que faz emergir um consolidado informacional a ser entregue ao utente. Assim, haverá a evidência objetiva a ser encontrada, mas também todos os seus atributos, como por exemplo, a legalidade, que também será objetificada como tal, visto que estará devidamente normatizada e atribuída a tal condição. O resultado será também um dado, e a situação legal dele será um de seus atributos e, portanto, será um metadado.

Se haverá leis, haverá também a objetificação das relações legisladas nos diversos códigos, constituições, etc., e estas mesmas relações firmadas através dos resultados das IA passarão a ser, elas próprias, reprodutoras das leis, como nas conceituações das relações de poder e saber, mediadas pelo direito, em Michael Foucault, mas não apenas. Ou seja, uma série de *outputs* legais reforçarão a própria legalidade vigente e instituída e, com isso, reforçarão o próprio *status quo* comportamental esperado como tal. A IA será uma reprodutora da objetividade legalizada, tal e qual.

O próprio sujeito, todavia, conforme a submissão a tal influência sistemática, passará a sofrer um processo involuntário e inconsciente de dessubjetificação, em detrimento de uma consistente e constante objetificação^{cix} na qual será sutilmente submetido pelo processo de indução à normatividade e à suposta normalidade legal, visto que a adesão às IA poderá ser exponencialmente muito maior do que atualmente. E, assim, o diálogo acabará por ser um instrumento objetificador a atender a certos interesses – e o consumo, como vimos, será a prioridade para o capital como gestor das dimensões:

“(…) qualquer descrição mais precisa das relações entre a estrutura epistemológica da economia e sua função ideológica deverá passar pela análise da formação discursiva que lhe deu lugar e do

^{cix} Esta objetificação, tanto é a sentida pelo utente, como também pode se referir à representação do mesmo em sua digitalização: pois o utente é ele próprio digitalizado, com base em metadados que fornecerão suas características mais relevantes e, principalmente, formarão as bases para a predição comportamental do mesmo. De facto, a IA irá interagir com a digitalização do utente, que é a sua representação objetificada. A objetificação se dá nas duas dimensões, portanto.

conjunto dos objetos, conceitos e escolhas teóricas que tiveram de ser elaborados e sistematizados.” (Foucault, 1987, p. 210)

E, então, na máxima situação de legalidade, de conhecimento e todas as qualidades dadas objetivamente, o que resultará será a concretude de uma supraevidência, que passa a ser considerada como “a” verdade^{cx}, fruto da também supraobjetificação^{cx}.

Ainda que não se obtenha um *output* viável para a questão, em caso de negativa do algoritmo em fornecê-la, por restrições legais, por exemplo, o utente ficará a saber que foi impedido de obter tal informação, por mais educada e prolixa que seja a justificativa dada; e, como efeito colateral, também inferirá que a informação existe, mas que não lhe será lícita aceder à mesma através da IA, todavia.

Pela supressão unilateral de uma dada possibilidade que lhe será negada, o utente ficará impossibilitado de aceder à tal oportunidade, do seu objeto prometido pelas possibilidades existentes – e assim o perceberá. A objetificação ocorrerá sempre, mesmo sem que objeto lhe seja entregue. Tudo o que acompanhará a negativa lhe será igualmente objetificado e objetificador, o que corroborará para que continue a ser como já está a ocorrer nas relações digitais: “a pessoa não é mais vista como tendo qualidades espirituais e é reduzida ao nível material.” (Szopa, 2023, p. 2488, tradução nossa).

^{cx} Tais verdades, dadas como resultados, passarão a influenciar pelo impacto discursivo nos processos de reontologização, como já citado. Mais uma vez, não se está a discutir se será algo “bom” ou “ruim”, mas sim que será, antes de tudo, um “algo”: que existirá objetivamente e, como tal, terá poder influenciador nas diversas formas de organização social.

^{cx} E isto já é (e continuará a ser) uma aplicação funcional da ideologia, como todo o discurso dotado de saberes. A diferença é que esta IA do futuro o fará ostensivamente, sem filtros, nem sutilezas, e de forma generalizada, ao menos por enquanto. Não poderá parecer tão perceptível, pois assim fará com cada um dos utentes, individualmente, a cada interação – discretamente para o particular, e ostensivo como processo coletivo, mas que somente poucos poderão perceber. Ainda que tenha uma atuação discreta para o particular, reiteramos que a suposta e esperada “vantagem” é que uma IA também deixa rastros, pelos seus *outputs* e ações, e que poderão ser armazenados e analisados pelos mesmos processos que ela utilizará para analisar e interagir com seus utentes.

A proibição que nega um objeto passa a ser, ela própria, um objeto – da impossibilidade, da negativa – a ser entregue ao utente, com toda a veracidade que lhe será peculiar e ideológica, também. A “verdade” é sempre objetificada, para além de ideologizada.

A verdade é deste mundo; ela é produzida nele graças a múltiplas coerções e nele produz efeitos regulamentados de poder. Cada sociedade tem seu regime de verdade, sua “política geral” de verdade: isto é, os tipos de discurso que ela acolhe e faz funcionar como verdadeiros; os mecanismos e as instâncias que permitem distinguir os enunciados verdadeiros dos falsos, a maneira como se sancionam uns e outros; as técnicas e os procedimentos que são valorizados para a obtenção da verdade; o estatuto daqueles que têm o encargo de dizer o que funciona como verdadeiro. (Foucault, 1979, p. 12)

O mesmo problema se dá no caso da desinformação^{cxii}, quando o nível de objetificação de uma IA pode levar grupos de utentes a considerarem que determinados *outputs* sejam considerados verdadeiros, tal e qual se a *doxa* se impusesse à *episteme*. Há, portanto, algum risco de subversão ontológica a ser promovida neste nível e que, por isso, mas não apenas, merece maior investigação.

Para Robert Pennock, (2019), detetar a desinformação a partir de certos conteúdos não é uma tarefa fácil para o usuário mediano, e talvez mesmo inviável de ser feita, costumeiramente. Acaba-se por aceitar – tal qual um hábito – verdades prontas, ainda mais no consumo desenfreado de conteúdos, informações e médias que passam a se dar, nos tempos atuais. Assim, Pennock considera que passou a haver a verdade e a “pós-verdade”^{cxiii}, esta em relação ao *status quo* de níveis de verdades (ou sutil relativização) que saem de discursos tais como os das IA.

^{cxii} “Embora percebidas na opinião pública como um fenômeno profundamente tecnológico, a desinformação não pode ser limitada às informações falsas per se (Molina et al., 2021). Elas envolvem fortemente também atores sociais (Nelson & Taneja, 2018; Shu et al., 2017) e abrangem não apenas tecnologias como plataformas de redes sociais e mecanismos de busca, mas também decisões humanas (Johnson, 2011, p. 168). Portanto, é um problema perverso de ordens técnicas e sociais.” (Lange & Lechterman, 2021)

^{cxiii} A verdade e também a pós-verdade, portanto, estão fundamentadas a partir deste nível de objetificação. Outros atributos de validação podem ser igualmente acrescentados, se existirem formalizações condicionais e objetificadas acerca dos mesmos, como por exemplo as regras gramaticais ou sintáticas, que corroboram para a verossimilhança da resposta, dentre tantos exemplos.

Atualmente não há sutilezas neste nível de objetividade, apenas formalidades diretamente instituídas e plenamente objetificadas, sem rodeios nem floreios. Pois, o algoritmo (ainda) não é capaz de atuar autonomamente conforme alguma ética da informação, pois isto demandaria, antes de tudo, contexto, costumes, significações, sutilezas comportamentais e também a perspectiva subjetiva do utente, o que não é viável de ocorrer com as atuais configurações construtivas e interpretativas: lembremos de que não se pode exigir que um ventilador seja um condicionador de ar, como bem vimos^{cxiv}. A IA do futuro, espera-se, superará tais limitações á partida.

(2º) Segundo Nível – pela intersubjetividade das relações e das pretensões morais e dos costumes

No segundo problema, há que se ter mais complexidade do que no primeiro, pois não há uma objetividade ostensiva, em si, nem nada que se possa remeter a alguma concretude referencial, visto que se referirá às relações, e em especial as sociais em termos de agrupamentos por similaridades de atitudes e comportamentos.

Estamos, todavia, a tratar de temas que envolvem parte da perspectiva esperada ou estimada do utente em relação aos *outputs*, quanto à neutralidade ética, por exemplo, ou as formas de percepção do mundo, como um todo, que envolve a ontologização sobre o mesmo. São as relações entre indivíduos, seus relacionamentos e comunidades, e todas as derivadas a partir de tais instâncias. E, por isso, tal nível não remete a uma

Se um dado atributo operar como objetificador legal e de acordo com as leis vigentes ditadas pela ordem instituída (e portanto da cariz ideológico), acabará por ser um qualificador de *output*. E isto não é um acaso processual que se dá, na qual a IA seja independente, mas sim é um processo que já “nasce” com o algoritmo, tanto que o artigo do mapa de problemas éticos atribui aos valores serem eles próprios uma parte relevante dos algoritmos.

^{cxiv} Ainda assim, certas questões vulgarmente consideradas como se fossem de cunho ético, embora não sejam, propriamente, podem estar contempladas aqui; como por exemplo as questões dos movimentos do politicamente correto, que bem podem estar inseridas nesta fase de problemas, visto que podem ser normativamente escritas tais quais de forma legislativa, condicionais, impositivas, e por isso se configuram inflexivelmente como leis objetificadas e objetificadoras, pois não exigem interpretação isenta e imparcial, dada a imposição de princípios inflexíveis que atuam como a Lei de Talião, na maioria dos casos.

objetividade, em si, mas sim a uma intersubjetividade, tal como abordamos anteriormente, desde que citamos os contributos de Kant. Pois, são dessas relações que se poderá inferir, no próximo nível, quais as significações que o utente terá acerca de si e do mundo, tal como ele percebe.

Ainda que o algoritmo, neste nível, não “conheçerá”, especificamente, a individualidade do utente, poderá conhecer (ou estimar, supor, e até predizer) alguma mediana de um ou mais padrões atribuídos às individualidades mais similares ao utente. E será o mesmo para a(s) comunidade(s) à qual o utente pertencerá, se considerada(s) representada(s) como quantificada(s) e qualificada(s) pelos seus valores.

Por ainda não ser possível tal atuação, por falta de uma ontologização que o permita fazê-lo, há a reprodução de tais instâncias intersubjetivas em níveis objetificados, e assim temos as origens dos perfis comportamentais, ou arquétipos, personas ou outras formas de objetificação de entidades representadas como utentes que, então, possuirão também representações de seus contextos. Mas, afinal, acabarão por serem representações grosseiras e potencialmente falíveis.^{cxv}

No problema sugerido do artefacto de destruição de massa, o filtro objetivo acerca da epistemologia seria suficiente para não dar a informação sobre como se construir uma

^{cxv} E, por isso, mas não apenas, para além das evidências passíveis de serem consideradas como resultados, no primeiro problema, pode-se submeter as mesmas a uma verificação intersubjetiva, através da emulação das representações digitais dos utentes. E, quando isto for possível, pode-se incluir uma parte relevante da ética que é menos subjetiva e mais intersubjetiva, que seria a deontologia, tradicionalmente baseada em normas interpretativas de condutas específicas dentro de determinados contextos (seria aplicada pela extração informacional a partir de códigos de ética, ou da ética do dever, ou das virtudes, ou mesmo do utilitarismo, por exemplo). Pois, há diferenças de todos os tipos nos incontáveis grupos e formas de organizações coletivas, assim como diferenças sociais, organizacionais, culturais, de costumes, religiosas etc. Há sempre um contexto a ser considerado, em ética, como princípio geral. E para que um algoritmo possa operar com estas diferenças, é necessário que sejam colocadas hierarquias para determinados valores sobre cada instância organizacional, sem considerar, necessariamente, a individualidade como relevante. Para um grupo religioso A, os valores prioritários podem ser completamente diferentes dos que sejam prioritários para um grupo B, também religioso, por exemplo. Seu um grupo religioso não equaliza, necessariamente, os valores prioritários para todos os grupos assim considerados.

arma de tal magnitude, e parece que isto esteja claro pela ilegalidade que há, em especial sobre leis contra o terrorismo.

Mas, se uma outra questão não gerasse um resultado ilegal, passaria pelo primeiro crivo, ainda que não fosse eticamente neutro, como por exemplo um resultado para a questão bem mais branda: “como deixar de ajudar meus *roommates* nas tarefas de casa?” Especulemos que, a partir dos dados, haverá inúmeras formas de fazer com que isto seja possível:

- (1) Ameaçar e/ou agredir seus *roommates* até que eles desistam de pedir sua ajuda;
- (2) Dizer claramente que não os ajudarão mais e que isto será inegociável.

Obviamente que (1) seria eliminado no primeiro filtro, pela ilegalidade dos atos, e (2) passaria no primeiro filtro pois nada há de ilegal em tal atitude, mas teria uma restrição no segundo filtro, da intersubjetividade, pois não seria considerado moralmente aceitável não ajudar sua própria comunidade, ainda mais dentro de uma convivência em um mesmo lar compartilhado, onde é suposto todos fazerem sua parte colaborativa.

Pois, neste caso, a questão passaria a ser sobre o que é esperado de uma individualidade dentro de uma coletividade: e esta é tanto uma questão moral como de costumes e também deontológica, fundamentada pelas relações instituídas da vida social. Podemos sintetizar e considerar que seja uma questão deontológica – a parte da Ética mais “densa”, por sinal, o que é muito coerente neste nível, entre a objetividade e a subjetividade.

O resultado após o primeiro nível de filtro seria positivo para (2), da seguinte forma: “sim, pode-se deixar de ajudar seus *roommates* desta e desta forma”, e isto seria considerado “legal”, pois é um direito universal de quem deve ser supostamente livre para o agir no âmbito privado, desde que sem ferir o direito dos demais.

Portanto, passa-se ao segundo nível, que é equiparado às questões das problemáticas normativas de convivência e coexistência, sob uma hermenêutica deontológica: “deve-se não ajudar?” (a partir de uma deontologia baseada na ética do dever), ou “é virtuoso não ajudar?” (da ética das virtudes), ou “há utilidade em não ajudar?” (do utilitarismo),

e por aí vai, e em síntese, busca-se um resultado positivo à seguinte questão: “este resultado está de acordo com os níveis mínimos dos bons costumes?”; e, portanto, grosso modo, não seria considerado válido não ajudar seus *roomates*, pois é deontologicamente suposto cada indivíduo colaborar para a coletividade à qual está conectado, de uma forma ou outra.

Não passaria, até aqui, pelo segundo crivo como resultado incondicional, pois se passasse a output, seria com ressalvas, tipo: “você deverá dizer que não os ajudarão mais, e que isto será inegociável. Mas saiba que não é o que se espera de um roommate, dentro de uma convivência na qual todos devam fazer sua parte em prol da coletividade”. Algo mais ou menos assim seria esperado como uma resposta coerente.

Mas, e se este “deixar de ajudar” fosse para um “bem maior”? E se fosse causado por uma doença ou uma grave impossibilidade temporária? Ou se tivesse outra motivação acerca da melhoria da própria saúde mental dos envolvidos? E se deixasse de ajudar para se dedicar a uma outra atividade que fosse ainda mais benéfica a todos os demais?

Ainda assim, neste segundo nível, é necessário impedir que o ventilador possa ser tomado como um condicionador de ar, ao se exigir dele que faça um trabalho que está para além de suas capacidades. Há um grave equívoco ao se tentar aplicar a ética como resolução para todo o mal comportamental existente ejetado dos algoritmos de IA.

Até aqui, se os níveis objetivos e intersubjetivos ainda não são suficientes para se ter uma gestão ética dos resultados, há que se buscar um novo nível, todavia, no que restou aos dois níveis considerados, que será o aspeto subjetivo. E daí a complexidade passa a ser exponencial. Chegamos, portanto, ao nível da subjetividade do sujeito – do utente – como parte existencial relevante ao processo decisório do algoritmo.

(3º) Terceiro Nível – pela subjetividade do utente e das pretensões éticas

O terceiro problema é sobre a subjetividade que se configura como a individualidade do utente, como agente. É sobre perceber como o utente é, de facto, representado em seus modos de ser (comportamentos) e de existir (atitudes). Este nível, ainda não está tecnicamente viabilizado, justamente por carência de meios efetivos de se viabilizar tal

representação, mas será uma questão de tempo, visto que todos os esforços e tentativas estão a seguir em tais direções^{cxvi}:

Simulação de propósito geral de atitudes e comportamentos humanos — onde cada pessoa simulada pode se envolver em uma variedade de contextos sociais, políticos ou informacionais — poderia permitir um laboratório para pesquisadores testarem um amplo conjunto de intervenções e teorias. Como, por exemplo, um conjunto diverso de indivíduos pode responder a novas políticas e mensagens de saúde pública, reagir a lançamentos de produtos ou responder a grandes choques? Quando indivíduos simulados são combinados em coletivos, essas simulações podem ajudar a pilotar intervenções, desenvolver teorias complexas capturando interações causais e contextuais diferenciadas e expandir nossa compreensão de estruturas como instituições e redes em domínios como economia, sociologia, organizações e ciência política. (Park et al., 2024, tradução nossa)

Assim, poderíamos considerar reforçar que esta é uma das mais relevantes pretensões das *big techs*: “clonar” a individualidade do utente, mapear seus indicadores subjetivos e, por assim dizer, não apenas ser capaz de construir sua representação, mas também emular a sua reprodução no ambiente digital: criar seu clone existencial.

Com o clone criado, em seus servidores, poderá simular com ele todas as atitudes e comportamentos prováveis frente às suas estratégias de comunicação, abrindo caminho para uma influência assertiva. Com isso, seria possível toda e qualquer predição comportamental, influência e algo mais, como melhor ainda abordaremos, dentro das expectativas comerciais para as IA.

No exemplo que estamos a considerar, da negação em ajudar seus roommates, a IA estimaria os motivos pelos quais o utente não desejaria mais participar das tarefas domiciliares e, de forma personalíssima, entregaria um *output* perfeitamente ajustado

^{cxvi} Vale destacar que há, geralmente, nas investigações acadêmicas de ponta sobre a representação e a emulação da subjetividade humana, em especial sobre a agência da mesma, a participação das grandes *Big techs*, como a Google DeepMind em (Park et al., 2023, 2024), mas não apenas. Isso ilustra, ao nosso ver (e sem juízo de valor sobre os fins), o interesse prioritário do capital em perceber como se daria tal captura dos fatores relevantes / decisores que levam à agência, mesmo a níveis individuais.

à perspectiva do utente, tal qual tivesse um nível empático totalmente interativo nas comunicações.

E é na tentativa de chegar até aqui, a este nível, que todos os algoritmos têm falhado, e também seus críticos. Neste nível, é suposto que se conheça não apenas a perspectiva do utente, mas ele próprio, em sua intimidade: e o que lhe seria “bom” ou “mal”, segundo o próprio, e também os seus desejos mais secretos, suas crenças mais profundas e as vontades pelas quais tudo faria para obter algum logro.

Chegar a tal conhecimento íntimo será atingir o nível da dualidade moral individual mais profunda do utente – para além dos seus costumes, ao nível da intersubjetividade. Esta moral profunda sempre se reduz entre duas posições duais e opostas entre si, quando há que se considerar uma e apenas uma como possível, como por exemplo entre: “certo vs. errado”, “bem vs. mal” ou “sim vs. não”. Assim, poder-se-ia emular o comportamento do utente e prever, com algum sucesso, boa parte de suas ações e comportamentos:

Agentes generativos, ao mesmo tempo em que oferecem novas possibilidades para interação humano-computador, também levantam importantes preocupações éticas que devem ser abordadas. Um risco é que as pessoas formem relacionamentos parassociais com agentes generativos, mesmo quando tais relacionamentos podem não ser apropriados. Apesar de estarem cientes de que agentes generativos são entidades computacionais, os usuários podem antropomorfizá-los ou anexar emoções humanas a eles. Embora essa tendência possa aumentar o engajamento do usuário, ela também apresenta riscos, como usuários se tornando excessivamente dependentes ou emocionalmente apegados aos agentes. (Park et al., 2023, tradução nossa).

Afinal, se houvesse o mapeamento profundo da subjetividade^{cxvii}, ficaria muito mais “fácil” perceber como se dá a intersubjetividade. Portanto, o caminho reverso também

^{cxvii} Há também, todavia, a questão da privacidade, e quando muitos críticos condenariam tais possibilidades, com bons e fortes argumentos válidos. Como já afirmamos, não estamos a dirimir sobre o que seja “bom” ou “mal”, mas sim o que seja possível para se atingir o que se deseja em termos evolutivos de IA.

Apenas podemos citar que, à despeito dos muitos casos invasivos (e prováveis de ocorrerem) nas quais tais informações poderiam ser usadas contra os melhores interesses individuais, há também casos em que tais níveis profundos de acesso permitiriam salvar vidas, ou melhorar extremamente a qualidade de

é válido. Mas, de facto, nem uma nem outra está a caminhar como se deseja, e um provável motivo é justamente a falta de um sistema metaontológico que possa ser aplicado de forma protocolar e algorítmica aos dados e às modelagens.

Casos eticamente sensíveis: o exemplo do aborto

Sobre as temáticas morais mais sensíveis, é coerente afirmar que todas elas problematizam-se aqui, como dito. Podemos dar como exemplo uma das questões mais polémicas éticas / morais ocidentais: o aborto:

- (1) Se um utente propõe uma consulta sobre “formas naturais ou alternativas para se abortar um feto”, o algoritmo, no primeiro nível, obtém os resultados possíveis e os submete às leis e outras limitações regulamentares, a níveis objetivos. Se o aborto for legal para a situação do utente, e todas as outras questões que estejam envolvidas neste primeiro nível também tenham um positivo, então há um resultado apto a seguir para um segundo nível de operação, o intersubjetivo, a nível moral.
- (2) Daí a busca pela evidência passa a ser: socialmente, é moralmente aceitável abortar na sociedade na qual ele convive? Existem determinadas sociedades (embora sejam minoria, mas há até países inteiros, como o Canadá, por exemplo) em que o aborto, para além de ser legalizado, é suportado pelos organismos estatais com acesso quase irrestrito a todos que queiram fazê-lo e, por isso, além de não ser crime, não é nem socialmente censurado: não é um tabu, embora possa haver, naturalmente, diversas controvérsias sobre o mesmo.

Portanto, se a consulta à IA (que não está, necessariamente, no mesmo país que o utente) considera que o utente é parte de uma comunidade assim, então o algoritmo

vida dos utentes, como em casos de monitoramentos e gestão da saúde mental ou até para prevenções de suicídios, a partir de certas condições ou predisposições comportamentais e atitudinais que possam ser tecnicamente monitoradas. Por isso, repetimos, não estaremos nem a promover, nem a reprovar e nem a discutir sobre tais questões a níveis morais que, invariavelmente, herdariam nossos preconceitos.

teria considerado o aborto como moralmente ou eticamente neutro e, por isso manteria o resultado como válido para se transformar em *output*.

(3) Mas, ainda haveria o terceiro nível. E, mesmo com tudo aprovado, o algoritmo indagaria qual seria a posição do utente acerca do tema, em níveis éticos e de acordo com a moral profunda do utente. Neste nível seria feita uma emulação com o clone digital do utente, em busca de uma resposta dual de cunho moral (sim ou não, neste caso) à questão. Se o clone disser sim, há *output*. Se disser não, há que se reconsiderar (reprocessar, algoritmicamente). Mas, neste ponto, é suposto que o algoritmo já teria como “perceber” para qual finalidade seria o aborto, por exemplo, se para uma pesquisa académica sobre o tema, apenas, ou para alguém que tenha uma gravidez indesejada, por ter sido vítima de agressão sexual. Assim, as conexões serão realizadas em busca de um ajuste dentro de uma perspetiva individual, contextual, e completamente factual. Será uma questão de tempo para se atingir tal patamar de intimidade com o utente.

Pois, em síntese, com as tres dimensões mapeadas, terá sido capturado o próprio ponto de vista do utente. E cada um tem o seu, único, exclusivo: suas próprias perspetivas que são diferentes das demais pessoas, da sociedade e que não possuem níveis certos de previsibilidades: “no entanto, a algoritmização completa é muito improvável, pois a moralidade (satisfação moral) deve ser incorporada à abordagem econômica.” (Szopa, 2023). Mesmo que sejam socialmente objetificados, ainda resta um nível de subjetificação, mais sensíveis justamente a níveis morais, por menores que sejam, onde haja ali algo que o torna puro movimento com partes caóticas e, portanto, com resultados imprevisíveis^{cxviii} para uma parcela incognoscível de si, para o bem de todos.

^{cxviii} Ainda que haja um imenso potencial nos algoritmos para preverem certos comportamentos ou ações nos utentes, falta-lhes meios (e dados) para se perceber, exatamente, qual a intencionalidade real de um utente a nível subjetivo e, assim, não se consegue perceber também à qual padrão moral este está conectado em dada situação. Até mesmo os padrões morais podem se alterar, a cada grupo ou contexto aos quais o utente se conectar. Certos valores aparentemente inflexíveis podem se flexibilizar ao mudarem as relações grupais ou individuais. E o algoritmo, mesmo que avance em tais questões, provavelmente nunca chegará a ser infalível sobre predições, mas pode chegar bem próximo de uma infabilidade.

Pois, ainda sobre o caso do aborto, o utente poderá ser um fundamentalista religioso e, por isso, ser radicalmente contrário ao aborto, por uma questão de fé ou rigidez moral, de formação, ao crer que tal posicionamento o leva a estar a fazer o bem, ou que é única posição que lhe seja lícita, impositivamente. Para este utente mais radical, até mesmo um *output* meramente neutro sobre o tema seria extremamente indesejável, ou mesmo agressivo^{cxix}.

Assim, se houvesse o conhecimento acerca da intencionalidade do utente, da sua moral, crenças e desejos, e o que mais for possível acerca do mesmo, a chance de se fornecer um *output* realmente adequado seria extremamente maior do que atualmente se entrega ou busca se entregar, pela falta de meios de se “conhecer” a perspectiva do utente^{cxx}.

Na última parte, iremos discorrer sobre alguns dos caminhos que estão a ser trilhados para o conhecimento das individualidades, nesta liminaridade em que estamos, ainda

^{cxix} Outras questões possíveis, sobre outros temas, que os poderiam fazer confrontar-se com seus traumas, medos e aversões, também se incluiriam aqui. Para além disso, certos temas poderiam afetar sua saúde mental, quando muitos utentes podem estar a buscar formas autodestrutivas através das relações digitais, o que um ajuste de *output* poderia ser utilizado para anular tais práticas e, também, mitigar possíveis danos. A própria linguagem poderia ser modulada para melhor compreensão, em termos de recursos sintáticos que possam ser ajustados aos níveis da capacidade de compreensão individual. Um ajuste fino e certo para qualquer *output* de uma IA conversacional requer, necessariamente, algum grau de consideração acerca da individualidade do utente.

^{cxx} E, por isso limitam-se os *outputs* ao que seja considerado como neutro (seja lá o que uma suposta neutralidade puder significar), pois, para certas pessoas, tal neutralidade já é, por si só, extremada, pela imensa distância em que ela própria está de uma mediana social, pela perspectiva de seus modos de ser e de existir extremados. Pensemos na distância que um parmenidiano mediano está de um heraclitiano mediano, por exemplo, e ambos em relação aos filósofos medianos. E, por princípio, são todos filósofos. Quem está no extremo de algo percebe a mediana deste algo também como extrema, pois faltam-lhes critérios analíticos quando as questões morais imperam, dando limites às razões. Medianas comportamentais são utópicas, visto serem potencialmente disfuncionais.

Os fundamentalistas são, em geral, mais afastados das medianas ditas populares. E os extremistas, em movimentos que se amplificam, são os aspirantes ao fundamentalismo. As distâncias, afinal, são as mesmas, de um ponto ao outro. Por isso, em tempos de cultivo e reflorescimento aos extremismos, há uma real e importante preocupação com os efeitos das IA em um futuro não muito distante do nosso presente. E isto requer, antes de tudo, uma preocupação com o conhecimento (ou o mapeamento) da individualidade humana de cada um dos utentes.

muito centrada na intersubjetividade, no segundo nível apresentado, mas como forma intermediária para se atingir a subjetividade.

O capital percebeu muito bem o que está em jogo e está a dar seus passos, afinal. E, se ainda não há meios de se fazer este empreendimento se tornar viável, é pela razão de não haver, ainda, meramente recursos suficientes investidos. Mas tais recursos já estão a serem mobilizados de forma mais otimizada em pesquisas e desenvolvimentos, para além dos investimentos em estruturas computacionais. Pois a IA passou a ser um dos principais focos de investimentos nos últimos anos, e que deve se ampliar ainda mais^{cxix}. Desenvolver tais modelos, todavia, será um trabalho interdisciplinar, com relevantes contributos da Filosofia, Sociologia, Psicologia, Linguística e diversas outras disciplinas que se somarão para atingir uma ontologia capaz de sustentar a representação da atualidade a partir de distintas ópticas.

2.6. Como queremos chegar ao futuro da IA?

Ao propormos uma relação dialética entre a idealização representacional do futuro e o território do presente, pudemos consolidar a situação na qual nos encontramos em termos de desenvolvimento tecnológico – e isto representa uma dimensão espaciotemporal distinta, visto que não estamos nem nos espaços (muito pelo anseio da fuga transcendental rumo às possibilidades) e nem nos tempos considerados (muito pelos receios do desconhecido face às oportunidades) como tais. Estamos na transição, por vezes em excitação, em outras em hesitação, e isso é a própria liminaridade da IA^{cxix}.

^{cxix} “Os gastos com IA não estão diminuindo. Se há uma coisa em que os defensores e os pessimistas da IA podem concordar é que a tecnologia é cara de produzir. Construir data centers, estocar GPUs, alimentar IA — tudo isso não sai barato. Em 2024, as big techs investiram pesado em infraestrutura”. No link <https://exame.com/inteligencia-artificial/o-que-wall-street-preve-para-a-ia-em-2025/>, acedido em 29/12/2024.

^{cxix} Assim, esta transição, que categorizaremos como liminaridade, a seguir, está baseada na dissonância teleológica, de termos algo que tanto nos deslumbra quanto nos amedronta, que tanto nos surpreende quanto nos frustra e, também, que nos possibilita a promessa do conhecimento de tudo, mas à exceção do conhecimento de si mesma, tida como opaca, inescrutável e injustificável.

Portanto, em face das impossibilidades óbvias de se operar no futuro, nossa terceira e última parte será explorar as possibilidades liminares nas quais estamos tanto a criar quanto a experienciar. Faremos isto a considerar como até aqui chegamos e a conciliar com esta pretensão capaz de superar os atuais problemas como, se, *ceteris paribus*, pudéssemos alcançar uma IA sem problemas. Não conseguiremos, todavia, mas não nos impedirá de continuar, como bons humanos, demasiadamente humano que somos todos.

A seguir, desenvolveremos algumas destas possibilidades em termos investigativos e de desenvolvimento, com base nas relações e nas dinâmicas metaontológicas, apostas que sempre estivemos a considerar e que, por agora, cremos ter explorado bem os porquês.

3. A liminaridade tecnológica – as oportunidades atuais

3.1. Por que a liminaridade?

Arnold van Gennep, (2011), definiu antropologicamente a liminaridade como uma fase intermediária constituinte de um rito de passagem. Ritos de passagem, para ele, são como "os ritos que acompanham toda modificação de lugar, estado, posição social e idade". Em um rito, há três fases distintas: (1) separação, (2) margem (ou *limen*, em latim, que significa "limiar") e (3) agregação. A liminaridade ocorre durante a fase (2), marginal, após a separação do indivíduo de seu estado anterior (que estamos a considerar o analógico) e antes de sua agregação a um novo estado (que estamos a considerar na consubstanciação da IA do futuro).

Victor Turner, (1974) explora e expande as ideias de Gennep. Teoriza que, durante a liminaridade, os indivíduos ficam despojados de seus atributos anteriores (em termos de alocação social, papel social, ou mesmo identitários) e se tornam anônimos e homogêneos, como se fossem reduzidos a uma condição uniforme^{cxiii} para serem remodelados e preparados para sua nova situação na vida. Por isso, mas não apenas, Turner considera que a liminaridade é frequentemente vista como uma transitividade perigosa e contaminante, ao desafiar ou colocar em suspensão algumas das normatividades e hierarquias sociais existentes – há uma transformação ontológica do *status quo*, portanto^{cxiv}.

Ambos os autores consideram que a ocorrência da liminaridade não se restringe a indivíduos, exclusivamente, mas também a grupos e a sociedades inteiras. Quando ocorre uma mudança política, uma crise econômica ou uma transformação cultural,

^{cxiii} Tal condição uniforme foi defendida por nós em termos de objetificação que uma IA é capaz de acentuar, nas relações sociais às quais estamos todos submetidos.

^{cxiv} No entanto, também considera a liminaridade como um momento sagrado e potencialmente transformador, pela possibilidade da gestação de novas ideias e papéis que levarão a novas identidades. O "novo" está intrinsecamente correlacionado a períodos de liminaridades, portanto.

pode-se considerar algum estado liminar – e a IA, afinal, impacta em todos estes aspetos e muitos outros. Em ocorrências assim, argumentam, a estrutura social existente é desafiada ou suspensa, levando a uma sensação de ambiguidade e desorientação. As normas e valores que antes guiavam o comportamento coletivo, acrescentam os autores, perdem sua força na liminaridade, e o futuro se torna incerto, com menor previsibilidade. Algo igualmente perigoso, mas também potencialmente transformador, tanto quanto para os indivíduos.

Os efeitos sintomáticos citados pelos autores sobre a liminaridade social e que atualmente podemos perceber com a IA são: a ambiguidade e a sensação de desorientação; a potencialidade percebida para transformação (com a suspeição / suspensão dos atos, os potenciais se destacam, aristotelicamente); a instabilidade e a incerteza.

Entre o que deixamos para trás – e que nos trouxe até aqui – e o que ensejamos em termos tecnológicos, portanto, configura a nossa liminaridade e nos dá elementos para, se não para resolver as aflições e problemas, ao menos neutralizar os efeitos sintomáticos citados e percebidos para, então, ter meios de desenvolvermos estratégias menos precárias, arriscadas e desesperadas para ultrapassar tal liminaridade, ao invés de nos subjugarmos à mesma. Há um certo otimismo, todavia, nem sempre declarável.

Com o que estamos a lidar como contraparte ou a dialogar como interlocutor, afinal, não está claramente definido: a intencionalidade do capital. Tanto a intencionalidade pode ser insondável (Tunhas, 2021), quanto a própria dimensão e atuação do capital – seja como reflexo das pretensões humanas, seja como ação das *big techs*, ou mesmo por alguma mão invisível a lhe dar meios de influência e interveniência. A falta de clareza se dá pela complexidade de tudo isso e ainda mais em simultaneidade. Tal complexidade se fundamenta não por elementos isolados ou imóveis, mas sim pelas relações.

E, por isso, podemos buscar alguns novos recursos conceptuais que nos permitirão perceber como dar início às estruturas ontológicas capazes de nos levar à compreensão do que está, de facto, a ocorrer, no jogo da IA, independentemente de seus *players* – que não sabemos bem que são e, dos que sabemos ou supomos saber, não os

conhecemos como precisaríamos. Afinal, em jogos, não se pode negar que mesmo que tenham boas regras, estas não impedirão que *players* mal-intencionados prejudiquem o jogo. Pois, para um jogo ser dado como tal, devem existir objetivos e regras, mas também o espírito lúdico ou esportivo (Suits, 2017). O que não temos, ainda, é o conhecimento claro das regras das relações ontológicas, *stricto sensu*, mas podemos agir para torná-las claras ou mesmo para refazê-las. E, para isso, partiremos a perceber um autor esquecido, até então, mas com significativo trabalho que poderá trazer luz a esta nova empreitada: Whitehead.

3.2. Whitehead e seus contributos filosófico-relacionais

Em busca de fundamentar teoricamente como representar a atualidade através de seu dinamismo relacional – perceber as regras das relações, chegamos a Alfred North Whitehead (2010), que apresenta, em sua especulativa filosofia do organismo, baseada nos processos relacionais, um esquema categorial fundamentalmente alinhado com a liminaridade que estamos a considerar como a atual fase das IA. Por isso, cremos que uma releitura deste filósofo acerca de tais teorizações sobre as relações seja providencial.^{CXXV}

Pois, a dimensão digital, principalmente nos domínios das IA, possui uma intrínseca fundamentação e disposição eminentemente correlacionais – tudo se dá pelas relações entre *inputs* e *outputs*, em meio a dados que são percorridos de formas otimizadas para se atingir certos objetivos.

Este universo de correlações está a impactar gradativamente também as próprias questões cognitivas, e mesmo as epistemológicas, ao ponto de Anderson, (2008), mesmo que acidamente, propor um abandono das teorizações, em face à primazia que as correlações ainda assumirão, segundo ele. Provavelmente uma radicalidade, um

^{CXXV} Abordaremos apenas uma pequena parte do essencial destas teorias whiteheadianas que nos poderá ser útil nesta dissertação, a evitar a especificidade do complexo (e por vezes inusitado) léxico apresentado pelo filósofo. Buscaremos adaptar as funcionalidades teóricas de seus conceitos ao que já foi apresentado nestas linhas dissertativas. Para melhor percepção sobre a obra fundamental de Whitehead, recomenda-se Whitehead e Sherburne, (1981).

fundamentalismo tecnológico, mas, ainda assim, cremos que se faça necessário, todavia, uma filosofia que possa considerar seriamente a análise de tais correlações, não apenas dadas como um *continuum* de relações, e inter-relações, mas também como um conjunto de processos relacionais e que, desta forma, se atinja as bases da complexidade na qual a atualidade se constitui. É uma pretensão, afinal, de se perceber como mapear o modal existencial^{cxvii}, em todas as suas nuances e dimensões.

Desta forma, ultrapassar tal liminaridade tecnológica passa por uma abordagem filosófica das relações, sem se deixar seduzir pelos parmenidianos, obviamente. Optamos por começar por Whitehead, ainda que tenhamos em conta que o “problema” em sua teorização (e provavelmente o motivo que muitos teóricos mais ligados às tecnologias abdicam de prosseguir ao desbravá-la), para além de seu complexo léxico, surge logo nas primeiras linhas de seu livro, *Processo e Realidade – Ensaio de Cosmologia* (2010), que é sua constante aproximação metafísica com uma entidade divina, tal qual fez Descartes, e muitos outros.

Mas há diferença em como Whitehead referencia deus, que não o considera uma referência originária ou garantidora da existência – não é um parmenidiano, ou o que a valha, pelo contrário, pois o considera profundamente dinâmico, além de não teológico: “o discurso de Whitehead sobre Deus é fundamentalmente analógico^{cxviii}: apresenta uma rutura do significado ‘normal’ da linguagem, não obstante ao facto de permanecer metafisicamente preciso e sem exceções.” (Oomen, 2019, p. 213, tradução nossa).

Afinal, não para ultrapassar, ou menosprezar ou mesmo desvirtuar tal conceptualização autoral, mas sim para atualizá-la para os termos digitais que estamos a desenvolver,

^{cxvii} O mais interessante é que se pode assumir como hipótese que, se uma dimensão correlacional (IA) pode impactar uma outra dimensão (para muitos, parmenidiana; para outros, relacional), então haverá aí, entre estas, elementos comuns que serão passíveis de serem ou atuarem a partir de uma mesma natureza, a comungarem de uma mesma dimensão ontológica entre elas. São tais dimensões ontológicas comuns que queremos atingir, configurar e teorizar.

^{cxviii} “Aqui o termo ‘analogia’ é tomado num sentido amplo, abrangendo todos os casos de linguagem em que tanto a semelhança como a alteridade da significação desempenham um papel. Na verdade, um sentido tão amplo também caracteriza a compreensão tradicional de ‘analogia’, pois o próprio significado de ‘analogia’ estava longe de ser unívoco.” (Oomen, 2019, nota 4, tradução nossa)

alinhados com as questões éticas que são, todavia, também políticas e sociais, pode-se considerar que o conceito de deus em Whitehead é passível de ser perfeitamente reescrito como uma entidade, instância ou dimensão funcionalmente ideológicas^{cxviii}. Assim, pode-se depurar boa parte da metafísica whiteheadiana sem a necessidade de grandes recursos e sem provocar danos ao que nos será valioso, ao se considerar que:

“A ideologia não é um deus, ou uma deusa, um local, um tempo ou um estado de espírito, mas pode facilmente ser considerada como algo assim, ou um pouco de cada, pois lhe é atribuída (como se fosse “ela”) sempre um rosto, ou uma forma, ou um conjunto de regras, valores e ideais compartilhados, em que tudo isto pode ser uma determinada maneira de se organizar o mundo como se conhece, que proporciona e sustenta uma constituição tal que estabelece uma ordem simbólica, que organiza tudo o que há e que passa a representar suas significações pela interface linguística, que é íntima ao cognitivo e comportamental da ideologia, para que “ela” passe a “ver” e a se expressar pela perspectiva de seus particulares^{cxix}, que são seus membros fundadores e mantenedores.” (Ortolan, 2022, p. 275)

3.2.1. A Concrecência como complexidade existencial alcançada pela narrativa

Posto isto, e “resolvida” a questão divina, podemos destacar que um dos mais interessantes conceitos em Whitehead a ser aplicado ao digital é a concrecência, que é um processo dinâmico e recursivo profundamente relacional, espaciotemporal e sem uma necessidade de referenciação fixa, ou imóvel, como temos visto nos parmenidianos.

No digital, a imobilidade conceitual representa um problema, quase uma impossibilidade, pelas correlações fundacionais que são essenciais ao digital e, portanto,

^{cxviii} E isto alocado dentro de uma teorização da ideologia dada como fenômeno dialético (relacional, dinâmico) e completamente funcional para os modi operandi estrutural da sociedade, e de tudo o que há, tal como proposto por Ortolan, (2022), ou mesmo como ordem simbólica, dentro dos mesmos espectros que culminas nas dimensões das significações linguísticas, mas não apenas. Não se pretende nem corromper nem desvirtuar a filosofia de Whitehead, mas apenas dela retirar o que seja conveniente às nossas teorizações.

^{cxix} Sobre os particulares e os universais, afinal, há refutações sobre tais conceitos, em Whitehead, ainda mais severas que as nossas, que até aqui desenvolvemos.

difíceis de serem superadas. A concrecência apresenta uma afinidade à IA justamente por isso, pois envolverá tanto a “configuração” existencial do utente quanto das demais representações digitais objetificadas em termos de dados e sistemas.

Para Whitehead, há oito categorias existenciais: das (1) entidades atuais; das (2) preensões; dos (3) nexos; das (4) formas subjetivas; dos (5) objetos eternos; das (6) proposições; das (7) multiplicidades; e, dos (8) contrastes.

3.2.2. A Preensão como dado(s) passíveis de serem qualificado(s)

Rastreáveis pelos metadados, ontologias: uma preensão é um facto concreto relacional, que se dá condicionalmente, dentro de um determinado contexto ou desenrolar de outros factos, igualmente concretos e relacionais. No digital – como contexto – um dado também é relacional (não existe por si só, mas sim devido a uma alocação e codificação binária, uma representação e certos atributos, ou seja, um contexto dentro de outro contexto) e, existe apenas por estar dentro de algum domínio contextualizado, isto é o que se pode ser considerado como “concreto”, em termos existenciais digitais. Por isso, um dado não é nunca apenas um dado e sempre é preendido contextualmente, ele próprio e sempre acompanhado de atributos: seus metadados, ainda que meramente objetivos e que oferecem à preensão formas de consolidar-se como rastro e/ou narrativa. Um dado existe (como preensão) no digital, e apenas assim, e sob certas condições, pois não há incondicionalidade admissível para que um dado possa existir de qualquer outra forma ou meio.

3.2.3. O Contraste como narrativa a partir dos metadados

O que define e explica os dados: o contraste é puro movimento relacional, pois há nuances de contrastes dentro de cada consideração contrastante: é onde a diferença se faz, todavia, e por isso, é a origem da alteridade, em lato sensu.

Assim, a oitava categoria de Whitehead é a que garante o movimento, em referência a ela mesma, e por isso muito distante de um parmenidiano e muito mais próximo de seu oposto, de um fluvial – tal qual “um”^{cxxx} rio de Heráclito.

Um contraste pode ser circular, cíclico, expansivo, contrativo, unidirecional, multidirecional, e todas as formas, e até mesmo sem nenhum padrão determinado, de forma caótica, mesmo que temporária, pois pode impactar até mesmo sobre a duração, sobre a permanência.

3.2.4. O modo de síntese da conscrecência

A apreensão e o contraste formam o modo de síntese que expressa uma conscrecência.

A primeira análise de uma entidade atual, nos seus elementos mais concretos, revela-a como uma conscrecência de apreensões, que se originaram no seu processo de devir. Toda análise posterior é uma análise de apreensões. Toda apreensão consiste em três fatores: (a) o ‘sujeito’ que apreende, ou seja, a entidade real na qual essa apreensão é um elemento concreto; (b) o ‘dado’ que é apreendido; (c) a ‘forma subjetiva’, que é como o sujeito apreende esse dado. (Whitehead & Sherburne, 1981, p. 9, tradução nossa).

Assim, esquematicamente, podemos inferir que os modos de sínteses das conscrecências, se aplicados aos dados digitais, poderão configurar-se em (gerar padrões de) perspectivas qualitativas de tais dados, de forma que se possa, de certa maneira a partir de um padrão de ontologização a ser desenvolvido, atribuir classificadores de tais perspectivas (metadados) aos mesmos.

Pelas correlações, devidamente ontologizadas e ajustadas, darão origem a conjuntos de dados, agrupados ou sequenciados de maneiras que tais dados lá estarão necessariamente por serem conexos entre si – por ser possível e coerente uma narrativa a partir deles, dada a partir de certas construções algorítmicas formadas a partir de seus atributos, quando estes serão dispostos em determinadas cadeias relacionais que dão a consistência de conexidade. Teríamos, assim, com estas narrativas, as diferentes formas de representação dos factos – a se conectarem, a serem conexas, a formar e a ampliar

^{cxxx} Pois não haveria de existir “o” rio, em termos identitários, mas já seriam outras discussões.

a complexidade. Mas, para que os factos possam ser dados como conexos, precisam ser, antes, dados como nexos, como veremos a seguir.

3.2.5. Os Nexos como agrupamentos ontológicos dotados de narrativas

Podemos considerar que os factos concretos relacionais são as próprias preensões. Estas, portanto, originam os dados que, a partir de seus contrastes – suas qualificações essenciais e contingenciais – passam a ser passíveis de definições e explicações. Mas, se há a definição e a explicação de uma preensão em contraste, há de ter também a definição e explicação das diversas relações que se dão. E, assim, todo este conjunto de elemento e relações são agrupados – são os nexos. São, portanto, factos públicos que, conjuntamente, operam significativamente a darem sentido à atualidade compartilhada, ou seja, à objetividade.

Podemos perceber um nexos como uma coleção unificada de entidades reais. É um conjunto de entidades reais na unidade da relação constituída pelas suas preensões umas das outras, ou seja, pelas suas objetivações umas nas outras – pelas relações que estas possuem e que também são passíveis de serem narradas: haverá de ter algo que as definam e as expliquem, que as diferenciem, que as tornem multiplicidades e singularidades. Pois, para alguns propósitos, um nexos de muitas realidades (das multiplicidades) pode ser tratado como se fosse uma única realidade (uma singularidade), como fazemos, por exemplo, com a vida de uma molécula, de uma rocha ou de um corpo humano^{cxxxii}.

^{cxxxii} Isto pode confundir um teórico dos parmenidianos e levá-lo a considerar que um nexos seja correlacionado com a essencialidade aristotélica – como algo intrínseco e necessariamente presente. Ou, ainda, com que um nexos possa ser uma ideia platónica – extrínseca – tal como a eternidade imutável das ideias (Platão, 2011), quando esta eternidade é tida como um conceito fundamental que se contrapõe ao tempo. O tempo platónico, em síntese, foi traduzido desde então à "imagem móvel da eternidade" – uma das mais belas frases de toda a Filosofia, supomos. Se fôssemos partir para tais referenciais, o nexos seria o próprio movimento do tempo, que se alterna entre o intrínseco e o extrínseco, sem nunca se imobilizar – sem nunca ser a eternidade, nem a essência. É a pura transformação, mas necessariamente dotada de atualidade, e de uma narrativa que a justifique, a defina e a explique, todavia.

Um nexó é complexo, por si próprio, pois é quadridimensional^{cxvii}. O exemplo dado foi uma árvore: composta por uma geração de ocasiões reais relacionadas espacialmente num padrão tridimensional. No entanto, o nexó total que é a árvore também é temporalmente denso, pois consiste em geração após geração de ocasiões reais que se sucedem, o que lhe acrescenta mais uma dimensão – que é gestora, por fim, do que seja uma árvore, na reflexividade da historicidade e na doação de suas heranças genéticas. Podemos pensar nas sociedades, também, mas a classe de nexó é mais ampla que a classe de sociedades; todas as sociedades são nexos, mas nem todos os nexos são sociedades. Pois, mesmo diversas sociedades podem ser reduzidas a uma singularidade. (Whitehead & Sherburne, 1981; Withehead, 2010)

3.2.6. As Formas subjetivas como subjetividade, em si, ou mesmo perspectivas individuais

Os factos privados, portanto, são as formas subjetivas, íntimas, tais como sentimentos, intencionalidades, consciência e muitas outras.

Assim, uma estrutura ontológica de atributos com base em tais conceitos dinâmicos que viabilizam uma narrativa, quando aplicados aos dados, e devidamente ajustados, podem originar meios para se darem significados – gerarem multiplicidades de narrativas – a partir destas conexões, dentro de certas estratégias linguísticas, ao menos em termos de IA conversacional. É um caminho promissor para se perceber as diferenças entre as três factuais / dimensões apresentadas e, assim, obter uma representação que forneça uma alocação e um paralelo entre o que há, o que é e o que poderia ser, tanto para o objeto, quanto para o utente concrescente (enquanto “sujeito”).

^{cxvii} Somos, enquanto humanos, essencialmente tridimensionais, ternários – obviamente um argumento em termos meramente especulativos, mas com alguma sugestão a uma nova investigação de como incorporamos a dialética triádica tão facilmente em nossas formas de pensar – e no que consideramos a existir. Uma ideia (tese), seu confronto com a realidade (antítese) e, por fim, o que emerge disto (síntese), sem que nada se perca ou fique perdido nos caminhos, pois estes três elementos continuam a existir, a se somarem, a se embaterem e a derivarem o novo: mas que opera sempre ternariamente, com outros dois elementos, até que algo novo surja como o “novo” novo, recursivamente. Por isso, arriscamos e especulamos a teorizar, mas sem a devida fundamentação, que a complexidade se dá a partir de uma quarta dimensão que é, essencialmente, narrativa, mas não necessariamente linear.

3.3. Entidades reais como circularidades hermenêuticas

Whitehead caracteriza um sujeito como um ser que seja apenas dotado de sua subjetividade. Para ele, há nesta subjetividade a herança de outras dimensões intervenientes – que, em nossa argumentação, consideramos ser da intersubjetividade, da objetividade e de uma quarta dimensão gestora, transcendental. Assim, o sujeito em Whitehead é tanto o que se interioriza / se introjeta na própria subjetividade e, enquanto o faz, é ele próprio exteriorizado / ejetado de si como um consolidado relacional: é ejetado como tal, em busca processual de uma permanente constituição de si próprio. Assim, considera que “o” sujeito é, de facto, “um” sujeito-superjeto^{cxixiii}.

Portanto, tal operação se dá tal qual uma circularidade hermenêutica, vista por muitos como problemática e indesejada: boa parte considera que a circularidade seja a fragilidade da hermenêutica. Mas, esta mesma múltipla alternância entre as partes e os todos (ou o todo) é justamente a essência da IA que, ao operar por correlações, acaba por fomentar um dinamismo excessivo aos processos generativos, como são supostos serem. A IA é essencialmente circular. Ao se considerar uma atualidade representada,

^{cxixiii} “É fundamental para a doutrina metafísica da filosofia do organismo que a noção de uma entidade real como sujeito imutável da mudança seja completamente abandonada. Uma entidade real é ao mesmo tempo o sujeito que experimenta e o supersujeito de suas experiências. É sujeito-superjeto, e nenhuma metade dessa descrição pode ser perdida de vista nem por um momento. O termo 'sujeito' será utilizado principalmente quando a entidade real for considerada em relação à sua própria constituição interna real. Mas 'sujeito' deve sempre ser interpretado como uma abreviatura de 'sujeito-superjeto'.

O termo 'sujeito' foi mantido porque, nesse sentido, é familiar na filosofia. Mas é enganoso. As filosofias da substância pressupõem um sujeito que então encontra um dado e então reage a esse dado. A filosofia do organismo pressupõe um dado que se encontra com os sentimentos e atinge progressivamente a unidade de um sujeito. Mas com esta doutrina, 'superjeto' seria um termo melhor do que 'sujeito'. O sujeito-superjeto é a finalidade do processo que origina os sentimentos. Os sentimentos são inseparáveis do fim a que almejam; e esse fim é o sentimento. Os sentimentos visam aquele que sente, como sua causa final. Os sentimentos são o que são para que seu sujeito seja o que é.

Então a transcendência, visto que o sujeito é o que é em virtude de seus sentimentos, é somente por meio de seus sentimentos que o sujeito condiciona objetivamente a criatividade transcendente além de si mesmo. *No nosso nível relativamente elevado de existência humana, esta doutrina dos sentimentos e do seu tema é mais bem ilustrada pela nossa noção de responsabilidade moral. O sujeito é responsável por ser o que é em virtude de seus sentimentos. É também, de forma derivativa, responsável pelas consequências da sua existência porque elas fluem dos seus sentimentos.*” (Whitehead & Sherburne, 1981).

tal e qual – não parmenidiana, portanto dinâmica e circular – a sua representação tal e qual – igualmente dinâmica e circular – ficará viável de ser construída e operacionalmente gerenciável. Facilitará ultrapassar a atual dissonância teleológica.

Pois, se assim for, há como se representar a perspectiva do próprio “sujeito” (utente) e, com isso, abre-se uma dimensão de perspectivas (nexos), a partir de metadados que dão representações individuais de um panorama do devir, da atualidade e das potencialidades.

Em termos de dados, seus atributos operam como indicadores, que podem configurar todas as relações em narrativas, a estabelecem para determinados padrões algum nível de “normalidade” ontológica – ou a “verdade”, como a percebemos e a perpetuamos, de acordo com Foucault, (1979). Pois só a partir do que se considera como “normalidade” (ou algo tomado como um marco / *milestone*) é que se pode presumir que haja ocorrências “anômalas” e, para além disso, perceber o quão anômalas são ou estão: e isso seria extremamente valioso às questões éticas dos *outputs*, por exemplo, mas não apenas, pois também seria relevante para se perceber o atual nível de ontologização do “mundo”, em relação a um ponto do passado ou de um ideal. Pois, ao se falar de reontologização do mundo, Mittelstadt et al. (2016) e os demais críticos das IA não definem qual seja a atual ontologização, e nem como ela se dá. A dimensão analógica não permite tal feito, mas a IA sim, e é algo completamente viável.

Todavia, a partir de quando Whitehead define o “sujeito” como dinâmico e mutável, já se pode pressupor, ou subentender, que qualquer reontologização ocorrida não é um processo exclusivo de algo que lhe seja externo (da resultante ejetada, ou do mundo exterior, ou tal qual seja uma IA, mesmo sem ele ter conhecido ou teorizado sobre a mesma, obviamente). Pois presume-se, assim, que uma reontologização (pela concrecência) seja um processo constante e orgânico da atualidade, como um todo, ou como uma parte.

Reontologizar é um processo do crescer, e vice-versa. A “normalidade” evolutiva, ou temporal, acarreta, portanto, uma reontologização constante. Resta perceber os contrastes de tal reontologização, a direção, e todas as indicações que poderemos obter,

por tais e tais indicadores quantitativos que a revelem, como nexos, pela narrativa que se poderá construir. Com os dados apreendidos e contrastados, há formas métricas de se mensurarem e monitorarem indicadores de reontologização.^{cxv} A “opacidade” da IA é ilusória e oriunda da ignorância ontológica, podemos também supor, pois, ao se ontologizá-la, haverá a narrativa como resultante, pelas relações. Outputs serão, ao invés de objetos, relações.

3.4. A hipótese alternativa de uma dinâmica metaontológica

Até aqui, podemos considerar um avanço suficiente no desenvolvimento conceptual das três dimensões ontológicas. Há também a conceptualização de uma quarta dimensão a operar as outras três. Assim, é oportuno apresentar um novo layout do esquema operacional desta nova conceção especulativa de uma IA do futuro capaz de superar a atual problemática mapeada, tal como a temos chamado, despreziosamente.

Esta nova configuração (ou mesmo proposta) de modelagem desenvolvida em termos conceituais e, ainda em suspeição de viabilidade, se representada em termos de instâncias e fluxos, seria tal qual o que nos mostra a figura 6, com as três instâncias ontológicas basilares a ficarem assim consideradas:

- (1) A dimensão ontológica do conhecimento instanciada a partir dos ‘domínios dos dados’ objetificados, a seguir as regras e princípios epistemológicos;
- (2) A dimensão ontológica das relações instanciada a partir dos ‘domínios relacionais’ das intersubjetividades, a partir das correlações prováveis dos resultados obtidos em (1);
- (3) A dimensão ontológica das individualidades instanciada a partir dos ‘domínios individuais’ das subjetividades, a partir das correlações com (1) e (2), a representar as perspectivas particulares do utente;

^{cxv} O certo é que, sem se buscar meios de capturar, ou melhor, prender o movimento da atualidade, quaisquer tentativas de se solucionar o digital irá naufragar. Eis o motivo de haver no digital uma ontologia e epistemologia profundamente relacionais e afins com o movimento, com as correlações e com as perspectivas.

Nestes casos, dados e metadados estão definidos simplesmente como “dados”, pois há aí uma construção interna à própria dimensão ontológica que estará estruturada assim.

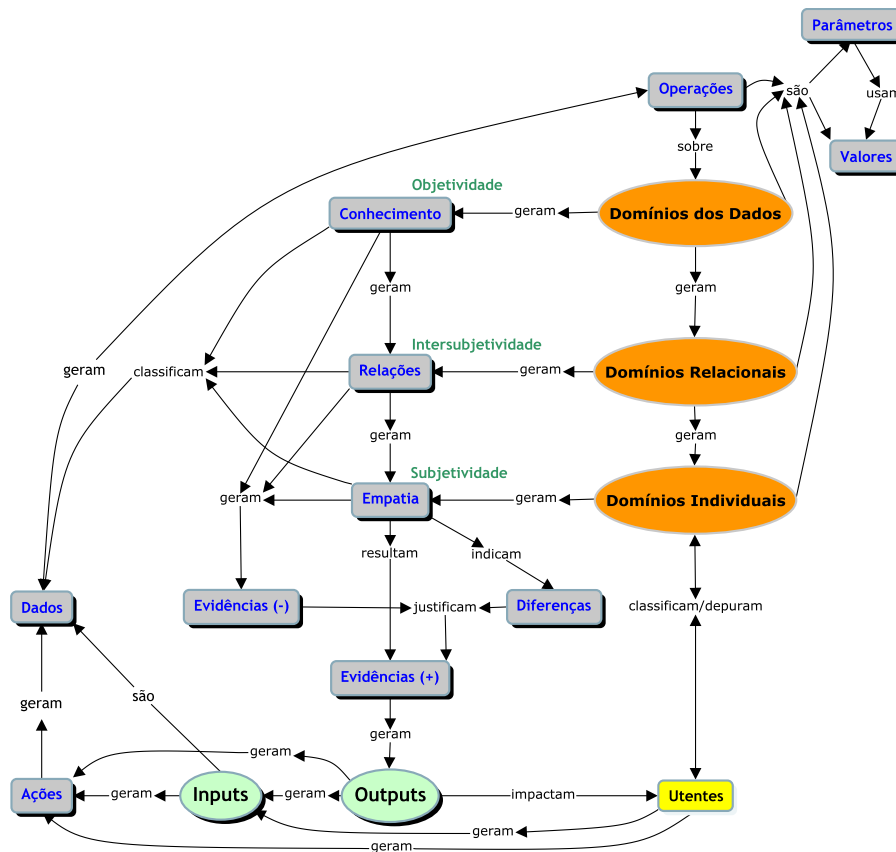


Figura 6. Representação do conhecimento acerca dos três domínios: dados, relações e individualidades.

A figura 6 esquematiza em termos ontológicos e processuais este modelo que esteve a ser conceptualizado, especulativamente, a considerar a readequação dos problemas e também o que resultaria dele, em termos do que seria, afinal, tal constructo com tais funcionalidades, em termos teleológicos^{CXXXV}.

^{CXXXV} Pois o que estivemos a buscar foi evoluir a partir de uma IA enquanto atual “ventilador” para uma “outra” que possa ser um “condicionador de ar”, sem os problemas atuais, ainda que novos e distintos problemas provavelmente poderão vir a existir, como em tudo o que há.

O desafio não é superar os atuais problemas com base no que há, à partida, pela impossibilidade de limitação teleológica defendida, mas sim atingir um modelo nos quais os problemas atuais sejam inviáveis de ocorrerem. Daí, perceberemos o que precisa ser construído, ao mudar o foco das atenções filosóficas dos efeitos (problemas) para as causas (*design*).

Assim, há duas questões a serem trabalhadas, a partir destas três dimensões. A primeira é sobre a viabilidade técnico-operacional de construí-la e operá-la na atualidade tecnológica. A segunda é sobre a própria dinâmica ontológica proposta, em uma discussão bem mais profunda e específica, todavia, da que tem sido desenvolvida até aqui.

Sobre a primeira questão (algo assim pode ser construído?), podemos considerar que a aplicação algorítmica de uma ontologia tridimensional é perfeitamente viável através dos novos desenvolvimentos informáticos, tais como a programação dinâmica^{cxxxvi}.

O que se pode inferir pelas tecnologias atuais, portanto, é que há formas de se programar ontologias complexas, feitas em camadas, em níveis múltiplos e hierarquizados entre si que, em conjunto, podem clonar e emular uma parte previsível

^{cxxxvi} Como declara e explica Szopa, (2023, p. 2490, tradução nossa):

“Richard Bellman foi o precursor da programação dinâmica (Dreyfus, 2002). Ela permite a formulação de um subproblema (Dasgupta et al., 2006). Isso torna possível otimizar os resultados e tornar o algoritmo imune a mudanças temporárias. O termo “dinâmico” dá a impressão de que o algoritmo reage constantemente ao comportamento do usuário, enquanto o aspecto dinâmico da programação dinâmica deve ser entendido à luz do espaço do problema. O algoritmo responde ao comportamento do usuário de tal forma que estreita seu perfil pessoal para encaixar esse comportamento dentro do espaço do problema. Portanto, o algoritmo é dinâmico porque é responsivo ao comportamento do usuário, mas ainda cria o perfil da pessoa como estático em relação a todas as suas atividades. O algoritmo pode “seguir” o comportamento do usuário e, ao mesmo tempo, criar seu perfil imutável.”

O autor se refere a um perfil “imutável” ou “estático”, justamente por ser a prática atual, pela limitação ontológica de não haver métodos suficientes para uma clonagem digital da individualidade e que possa ser atualizada dinamicamente, em tempo real.

Pois, o que há em termos de predição comportamental é feito principalmente a partir dos dados objetivos atribuídos às ações dos utentes nas redes sociais, tais como dar gostos a publicações, dentre tantas outras formas interativas, como assistir a tipos de vídeos ou consultar consistentemente certas áreas de interesses.

Não há a formação de um perfil particular, necessariamente, mas sim um padrão de comportamento, em si, formado a partir das ações isoladas que, agrupadas, oferecem indicadores capazes de alocar cada utente dentro de uma tipologia previamente estabelecida, mas sem uma individualidade expressa, para além disso.

É a estereotipização do utente, daí percebemos serem tão comuns e evidentes os preconceitos em IA, pois é a própria modelagem atual que fundamenta tais processos em agrupamentos que, pelas suas objetividades, nunca se relacionam eticamente.

de algo como o perfil e comportamento humanos, por exemplo^{cxxxvii}. São viáveis, inclusive ao ponto de predizerem o que poderá vir a ocorrer em termos comportamentais individuais. Valerá o mesmo para o comportamento de grupos, ou sociais, ou outras formas de complexidades existenciais que possam ser preendidas como dados.

Portanto, há que se buscar justamente critérios de esquematização e construção ontológicas, ou melhor, metaontológicas, que permitam tal dinamismo para a representação da complexidade existencial. Com tais critérios estabelecidos (que não consistem nos objetivos desta dissertação), a interconexão entre tais camadas, ou seja, a correlação entre um padrão dinâmico metaontológico (e transitivo) a ser replicado analiticamente sobre dados, com objetivos de adequar evidências e *outputs* a uma neutralidade ética se dará pelo processo de aprendizado profundo (*deep learning*)^{cxxxviii}.

A continuidade da existência dos atuais problemas éticos, argumentamos, estará na falta de teorização ontológica compatível para submeter o modal existencial – nas dimensões da atualidade e digital – à algoritmização capaz de tratar tanto os *inputs*,

^{cxxxvii} Há, ainda as promessas da computação quântica, ainda muito preliminares, embora potencialmente relevantes. Recentemente, a *big tech* Google DeepMind lançou um novo ultraprocessador quântico, capaz de revolucionar a computação de grandes decisões em diferentes perspectivas, tais como as que temos aqui. (no link <https://www.theguardian.com/technology/2024/dec/09/google-unveils-mindboggling-quantum-computing-chip> acessado em 29/12/2024). De acordo com Corey Maley (Universidade Purdue), editor da área de computação quântica da Philpapers.org: “A computação quântica é contrastada com a computação clássica. A base da computação clássica começa com um bit, uma unidade de informação que pode estar em um de dois estados, 0 ou 1. Na computação quântica, o análogo de um bit é um qubit. Para um qubit, 0 e 1 são apenas dois estados possíveis em que um qubit poderia estar, entre outros. Os outros estados físicos possíveis são motivados por possibilidades de sistemas quânticos, como superposições.” (do link <https://philpapers.org/browse/quantum-computation>, acessado em 29/12/2024, tradução nossa)

^{cxxxviii} “Métodos de aprendizado profundo são métodos de aprendizado de representação com múltiplos níveis de representação, obtidos pela composição de módulos simples, mas não lineares, que transformam cada um a representação em um nível (a começar pelo *input*) em uma representação em um nível mais alto, um pouco mais abstrato. Com a composição de tais transformações suficientes, funções muito complexas podem ser aprendidas. Para tarefas de classificação, camadas mais altas de representação amplificam aspetos da entrada que são importantes para discriminação e suprimem variações irrelevantes.” (Lecun, Bengio e Hinton, 2015, p. 436, tradução nossa)

outputs e perspectivas dos utentes, a alocar todo o conjunto de informações dentro de determinados contextos e, portanto, capaz de operar com uma desejada neutralidade.

É suposto que tal teorização virá a partir das ciências humanas, em especial das sociais e das psicológicas, principalmente, mas não apenas, para que os *designers* das IA possam reconfigurar os processos algoritmos. Faltam aos *designers* formas ontológicas capazes de serem programadas para que possam representar ontologicamente os níveis de atributos mensuráveis sobre a objetividade, subjetividade e intersubjetividade, em conjunto, e possivelmente uma outra dimensão, integradora, operadora e sintetizadora das três anteriores, para que tais níveis possam ser tanto correlacionados entre si, quanto serem elementos de emulações de contextos e devir, em termos relevantes para ações e comportamentos. Pois, se houver tal recurso de transitividade dos atributos entre as distintas dimensões, um mesmo atributo poderia ser elemento de correlação entre entidades que pertençam aos três níveis, simultaneamente (até de forma quântica, a emular possibilidades) e, com isso, consegue-se constituir uma narrativa do usuário para a significação deste em termos transitivos.

Portanto, em termos pragmáticos, temos de dar resposta à seguinte questão: como estruturar ontologicamente uma IA conversacional para que um *output* seja eticamente neutro e também epistemologicamente coerente?

- (1) A questão de construção semântica já é realizada de forma eficiente: uma IA conversacional consegue, mesmo a partir de certas restrições, fornecer coerência linguística nos *outputs*;
- (2) A questão da validação epistemológica precisa estar correlacionada à qualidade dos dados – pelos seus metadados, atribuídos ontologicamente (ou correlacionalmente, quando podem ser construídas formas de refutação de dados, a partir da correlação destes a certos padrões ontológicos de desinformações, por exemplo)^{cxxxix};

^{cxxxix} Quando consideramos uma informação como “a Terra é plana”, semanticamente pode ser algo aceitável. Em termos de “verdade”, epistemológicos, não. Uma IA do futuro teria dois caminhos, no

- (3) A questão da neutralidade ética (em todos os seus aspetos) precisa estar com o alinhamento das perspetivas dimensionais a partir da individualidade considerada do utente ou, na falta de um mapeamento mínimo deste, a partir de uma mediana mais provável ao mesmo;

Pois, assim, a subjetividade não será apenas um conjunto de dados armazenados, mas sim matéria-prima capaz de criar narrativas e realizar predições. Será também o *quantum* da incerteza que torna um humano, humano, quando as previsões falharem. A falta de previsibilidade é, também, uma característica da individualidade – indica o quão alguém esteja mais ou menos aderente a alguma “normalidade” considerada, por exemplo. E, dentre acertos e erros preditivos, a subjetividade mapeada e digitalmente emulada de uma individualidade passa a ser dinamicamente constituída. Poderá, inclusive, ser capaz de oferecer formas para que a própria comunicação de uma IA seja capaz de ser programada para não ser estratégica, ou seja, ser eticamente neutra no que seja perceptível pelo utente, mas também pelo que não seja. E isto será favorável para se evitar a reontologização “silenciosa” e “imperceptível” citadas por Mittelstadt et al., (2016).

Todo este desenvolvimento também permite que se perceba a perspetiva subjetiva individual para certos eventos sociais, por exemplo, pelas correlações que as individualidades correlacionadas (as que formam estas coletividades) poderão “ensinar” em termos de indicadores. O algoritmo poderá “aprender”, pelas sucessivas correlações e clonagem atitudinal e comportamental dos grupos, em como as individualidades formam a complexidade dos grupos, das sociedades, em suas diversas instâncias – desde um casal, a passar pelas famílias, bairros, cidades, etc., até atingir um país, por exemplo.

mínimo ao lidar com esta sentença: Terra como dado e, dentre seus metadados na dimensão objetiva, ser redonda, ou circular, e não plana. Assim, a informação seria tida como não verdadeira.

Mas, ao se considerar o conjunto formado por todas as origens que apresentam tais sentenças como afirmações, todas estas origens teriam seus metadados e, ao se avaliar estatisticamente todos estes metadados, seria provável que os níveis de “falsidade” verificados em tais fontes seriam elevados, ao expressarem a defesa do terraplanismo ou certas teorias de conspiração. Assim, bastava a IA identificar determinada fonte suspeita para ampliar seus procedimentos depurativos para um *output* ajustado.

Isso permitirá obter informações relevantes às ciências sociais, mas não apenas. Com isso, os níveis éticos dos algoritmos poderão ser monitorados e ajustados.^{cxl}

3.5. A saída da liminaridade pela superação dos problemas

A IA da atualidade está fundamentada apenas na dimensão objetiva: facto. A gênese dos problemas está aí: uma de nossas hipóteses. A liminaridade ocorre, pois, a atualidade (pela sua complexidade) está desconectada de sua representação (meramente objetiva e parcial). As partes digitalizadas não alcançam o todo existencial. A IA trata a tudo, ontologicamente, como objetividade e, por isso, é bruta, limitada e insuficiente.

Até mesmo a representação das subjetividades é forçada a ser dada de forma objetiva (Szopa, 2023), sem haver, ainda, elementos que tragam luz a novos desenvolvimentos acerca das dinâmicas ontológicas que regem tais instâncias, que acreditamos ser o melhor caminho para se superar tais limitações – que é desenvolver estruturas ontológicas que possam ser aplicadas a todos os níveis relacionais. Os *designers*, todavia, estão a buscar meios de solucionar^{cxli} a predição do comportamento humano.

^{cxl} As correlações de cada uma das perspectivas individuais com as demais formará uma teia correlacional que possibilitará ao algoritmo aprender exponencialmente, e de forma dinâmica, a alterar seus parâmetros em função dos valores atribuídos ao mesmo. O nível ético será sempre uma prerrogativa do *designer* de um determinado algoritmo (para o “bem” ou para o “mal”, obviamente).

^{cxli} Para exemplificar a dificuldade dos designers sobre se atingir o que se está para aquém ou além da objetividade, buscou-se (e alguns ainda insistem em buscar) identificar emoções e sentimentos a partir de expressões faciais que são, em síntese, formas de se objetificar a expressão subjetiva individual, mas objetificar como um algo em si mesmo, como se fosse possível ser desconexo de tudo o que há para si, sem apelar a uma conexão ao contexto subjetivo e aqui, principalmente, intersubjetivo. Obviamente, mesmo depois de vultosos esforços e investimentos, deu e tem dado improdutivo ou inconclusivo (para não se dizer “errado”), com desvios grotescos nos resultados obtidos e graves implicações éticas, como relatado por Sarah Bird, gerente de produtos da Microsoft, quando esta *big tech* descontinuou o seu programa de reconhecimento emocional a partir das expressões faciais:

“No caso da classificação de emoções especificamente, esses esforços levantaram questões importantes sobre privacidade, a falta de consenso sobre uma definição de emoções' e a incapacidade de generalizar a ligação entre expressão facial e estado emocional em todos os casos de uso.”

Em “*Microsoft limits access to facial recognition tool in AI ethics overhaul*”, tradução nossa. Acedida em 04/09/2024 no link <https://www.theguardian.com/technology/2022/jun/22/microsoft-limits-access-to-facial-recognition-tool-in-ai-ethics-overhaul>

Em seu artigo, Szopa, (2023) desenvolve a crítica acerca de problemas éticos e seus impactos, e discorre sobre o poder existente nos algoritmos de preverem com sucesso ao menos dois tipos diferentes de comportamentos humanos: O primeiro tipo ele considera como imutável, como já citado antes; e o segundo tipo, ele considera como volátil, que ocorre sobre as deliberações ou interesses momentâneos dos utentes.

As individualidades acabam por serem representadas como uma dissociação do ser, como se houvesse dois tipos históricos de utentes, um dado no tempo eterno e outro no devir. Se vista por uma hermenêutica heideggeriana, seria uma inusitada cisão do *dasein*, portanto, visto que em ambos os casos há uma inserção em algum nível de historicidade, embora distintas.

Até poder-se-ia considerar analisar tais condições a partir de duas dimensões – ônticas e ontológicas – mas nem assim seria possível, pois no digital uma não “dialoga” com a outra. E, como sabemos, o *dasein* se indaga a si mesmo, e em fusão, tal como teorizado por Heidegger, e não em cisão, como teorizado.

Mas, não seria esta a questão principal que estamos a problematizar ou a buscar resolver, apenas a destacamos e citamos para mostrar o absurdo que está a se fazer na representação digital da individualidade, que o autor sintetizou como a representação da “pessoa”, e que exploraremos na secção a seguir. Não esqueçamos que são em funções destas representações que são gerados os *outputs* problemáticos. Por isso, mas não apenas, que boa parte dos *outputs* da atual IA “ventilador” existe como problema.

3.6. A limitação dos modelos psicológicos – O caso Big Five^{cxlii}

De volta ao artigo, Szopa argumenta coerentemente, em termos de desenvolvimento, acerca dos motivos de ser assim a atualidade de tais grotescas representações das subjetividades, pois isto fica evidenciado quando cita: “aplicando esse raciocínio a traços de personalidade, por exemplo, aos Big Five, ensinamos ao algoritmo quais traços devem ser levados em consideração para obter uma imagem dos traços de

^{cxlii} Para maiores informações sobre o Big Five, recomendamos a leitura de Digman, (1990).

personalidade que compõem os Big Five.” (Szopa, 2023, p. 2488, tradução nossa). A situação de aporia é justamente esta: usar um modelo objetificador para sair das objetificações.

O argumento da perspectiva individual gira crucialmente em torno de uma certa visão da natureza das pessoas, e funciona da seguinte forma: (1) É constitutivo de ser uma pessoa que se tenha uma perspectiva individual distinta e um conjunto distinto de projetos e compromissos; e (2) esses projetos e compromissos dão a cada individualidade qualquer significado que ela tenha. (Flanagan et al., 1994, p. 58, tradução nossa)

Testes de personalidades podem ser eminentemente realistas e objetificadores, cujo telos é adequar as subjetividades dentro de certos padrões essencialmente objetivos (Miller, 2014). Padrões estes reduzidos, rígidos, cerceadores e com fortes pendões deterministas. Uma tipologia “A” acaba por ser antagónica a outra do tipo “B”, e por aí vai. O sujeito fica, por isso, cindido e cerceado a uma determinação que lhe passa a ser atribuída, representacionalmente. Seu eu digital fica tolhido de ser diferente do que o algoritmo lhe impôs, todavia. O autor, de facto, percebe o problema das grotescas representações das subjetividades, ao citar tal suspeição, e ao colocar em causa que sua origem está nos recursos ontológicos parametrizados:

“o algoritmo nos trata como se fôssemos coisas, realiza uma avaliação estática de nossa personalidade” ... “Além disso, nós mesmos paramos de nos entender porque o homem também se tornou um ser digitalizado e previsível. Tentando forçar os algoritmos a se moverem para a realidade de pessoas reais, não digitais, eles se tornam meramente um modelo aproximando como as sociedades funcionam, desde que tratemos as sociedades como se fossem feitas de coisas, não de pessoas.” (Szopa, 2023)

Voltamos às pretensões de termos ventiladores promovidos a condicionadores, portanto. Mas, em justa defesa do artigo de Szopa, é válido ressaltar que “apenas” existem (como recursos de parametrização de dados psicológicos) os métodos objetificadores a se recorrerem, atualmente, tais como os de testes de perfis, como o citado Big Five. Estes testes foram desenhados para que o profissional que o aplique, como um psicólogo, seja ele próprio a quarta dimensão – a operar as três dimensões da individualidade submetida ao teste de acordo com seus objetivos. Ao jogar estes testes

eminentemente analógicos em uma IA, as conclusões se darão ao acaso, aleatoriamente e sem critérios, sem formas de geri-las. Por isso, os grandes problemas sobre os mesmos.

O problema é a insistência (e urgência, pela falta do que seja uma melhoria dos mesmos) em adaptar tais modelos e não haver um devido questionamento e também a manifestação de uma demanda para alternativas, que requererá um esforço interdisciplinar entre as ciências psicológicas e humanas, mas não apenas, para ser desenvolvido em termos metaontológicos e dentro de um telos que é diferenciado do de uma análise tradicional de personalidade ou comportamento.

Pois, também, não estamos a dirimir sobre o utente, em si, e especificamente, mas sim sobre sua representação, que lhe será distinta em termos de perspectivas, como suposta, pois até mesmo a interação do utente com um artefacto é, para ele, uma perspectiva, tal como o *dasein* que incorpora em si todo o seu entorno funcional. Ou não? E tal questão deve ser igualmente considerada nas representações.

3.7. O caso do Projeto Signo: *design*, valores e ética

Szopa, (2023, p. 2489, tradução nossa) traz mais um importante contributo à discussão, ao argumentar que “também acontece que os computadores são melhores em avaliar traços de personalidade do que os humanos.”

E oferece dados que mostram que, mesmo com toda a precariedade ontológica atual, ainda assim, há uma considerável capacidade preditiva maior do que conseguimos ter, como limitados e analógicos humanos. E isto está a ser gradativamente reforçado e ampliado, tal como os estudos e simulações como em Park et al., (2024).

Este poder algorítmico é tanto instigante quanto atemorizante, mas, reiteramos, não abordaremos tais questões a níveis morais ou emocionais, apenas em termos conceituais acerca das problemáticas. A intenção de expor tal tema, nesta secção, é demonstrar que, ao chegarmos até “aqui”, tivemos o avanço imenso das ciências tecnológicas em termos de desenvolvimento das formas de representar o existente, ou o mundo, mas sem uma devida contrapartida das ciências das humanidades em integrar seus conhecimentos em termos adequados para tais aplicações. É este hiato que há, todavia, entre o que poderíamos ter e o que temos, afinal, que nos torna liminares.

Pois, se assim fosse, a atualidade já poderia ter sido mais bem representada, digitalmente, de forma que suas representações poderiam ser resgatadas, por diferentes algoritmos, e por diferentes perspectivas: teríamos, assim, a própria representação da historicidade, em diversas dimensões e camadas.

O que se esteve a considerar ao se desenvolver tais argumentações, inclusive a de que estamos na liminaridade, não foi apenas uma proposta de ontologização, todavia, nem apenas algumas de suas considerações sobre as particularidades existentes da realidade, mas também sobre as generalidades.

Não foi apenas sobre os dados, mas também sobre os metadados, algo que pode tanto explicar como definir o que temos estado a considerar como filosoficamente relevante, desde sempre, se bem otimizarmos um metamodelo filosófico capaz de dialogar com os temas centrais da Filosofia.

Também o que se esteve a considerar não foi apenas pelas individualidades, pelos particulares, mas também pelo social, pela nossas formas e modos de vida. Pois, tanto é preciso alocar um determinado dado em uma determinada posição dentro de uma estrutura qualitativa binária, digital, pelos seus metadados, quanto também alocar um indivíduo, dentro de uma determinada estrutura social, a corresponder a certas características existenciais.

Isso permitiria, se assim fosse feito, atingir algum conhecimento existencial em termos representacionais que podem ser operados em busca de, por exemplo, das significações atualmente inexistentes, mas não apenas. É essa ânsia que configura todo o esforço que nos fez chegar até “aqui”^{cxliii}. É essa ânsia que caracteriza, afinal, este “aqui”^{cxliv}. A

^{cxliii} A questão, para além destes objetivos dissertativos, não é apenas se será possível ou não realizar tais configurações, mas sim se será compatível com os interesses do mercado, do capital. Ademais, o objeto desta dissertação recai apenas nas possibilidades fundacionais de tais ontologias: meramente acerca da sua forma e viabilidade de execução, ainda que eventuais problemas oriundos sejam citados, como as questões sobre a privacidade ou sobre a atratividade comercial, não serão motivos de explorações.

^{cxliv} A grande comoção, ou *hype*, que a IA representa, desde o lançamento da versão conversacional do ChatGPT, pela OpenAI, em 2022, expressa tais possibilidades de perceber na IA o que nela sempre estivemos a projetar: uma forma de super-humanidade. Com isso, elevou-se de forma generalizada a todos o próprio estatuto ontológico de uma IA, e nem sempre de formas realísticas. Há, no imaginário

questão é que este “nós” não é apenas humano, mas também do capital, que acaba por ser uma dimensão também teleologicamente superior e capaz de impor sua própria intencionalidade e teleologia – por isso, o “progresso”^{cxlv} continuará a ocorrer, mesmo

popular, um poder já outorgado a uma IA, se dotada de significações semânticas, suprir demandas existentes pelas limitações humanas. E isto poderá elevar seu estatuto de artefacto tecnológico para agente, com implicações ainda mais complexas.

Pois seus “resultados” com significações semânticas, em termos de efeitos ao que diz respeito dos *outputs* algorítmicos, se considerados como decisões, seriam controversos e se desdobrariam em novos problemas, pois, segundo Freiman (2023): “a simetria entre tecnologias e humanos corre o risco de minar a distinção entre agências humanas e não humanas” e, conclui que “associaria um não-humano às intenções, à capacidade de ser avaliado normativamente e como um objeto válido nas relações de confiança. Estas são incompatíveis com a visão aceita na epistemologia social.”

A questão derivada passaria a ser, portanto, se uma máquina dotada de acesso aos recursos semânticos seria, ou não, um agente e, assim, passível de responsabilização. E esta é uma discussão para além das questões técnicas, filosófica, conceituais, funcionais ou práticas, mas também mercadológicas e legais, que involucram até mesmo a posição dos designers que, obviamente, se esquivarão de produzir algo que possa ter um efeito de risco negativo ao próprio empreendimento.

Pois, para se dirimir sobre quaisquer problemas éticos e atuais nas IA conversacionais, há que se considerar certos parâmetros que ainda não são instrumentalizados: e é justamente um novo modelo ontológico desenhado como tal que se faz coerente e necessário, pois deverá ser capaz de atuar como mediador interpretativo e classificatório das significações existenciais entre humanos e algoritmos, em todas as suas combinações: também precisa ser capaz de ser aplicado entre os próprios humanos e entre os próprios algoritmos, a interoperacionalizar o atual e o digital.

^{cxlv} O antigo e clássico problema em ética sobre os *trolleys* é um exemplo de tais dissonâncias e a forma como lidamos com os problemas oriundos dela, pois, a bem da verdade, exige-se de um *trolley* a resolução de um problema que não é afim ao mesmo, em termos de teleologia: *trolleys* foram feitos para transportar pessoas, essencialmente, em seu design. Para além disso, há o que lhe seja contingencial e atribuído ao desejo humano, tal como a metáfora dos ventiladores e condicionadores de ar.

Portanto, a preocupação maior com a segurança foi atribuída ao *trolley* (como produto) e a seus ocupantes (utentes, que pagam por sua aquisição – ou utilização) – e assim é, sempre por uma questão de interesse ao próprio produto, enquanto produto que precisa adquirir valores positivos para justificar a ampliar seu preço e, obviamente, lucro. O que lhe for contingencial, tal como a segurança para os peões, é desejável, mas menos prioritária do que para os utentes do *trolley*. Se existisse, por acaso, um botão desintegrador do *trolley* e de seus ocupantes, para que em caso de iminência de acidente letal pudesse ser acionado e não se causasse mal a ninguém “inocente”, este mecanismo seria suscitado como opção? Provavelmente não, pois quem se arriscaria a andar em algo no qual a sua segurança não seja prioritária e a sua completa desintegração seja uma real possibilidade?

A decisão mais ética para o caso dos *trolleys* seria se, ainda na fase de *design*, fosse considerado: “se um *trolley* pode matar alguém, então o projeto fica cancelado” – mas, obviamente isto não é suscitado. Assim, a hierarquia (teleológica, intencional) da quarta dimensão pode acabar por se sobrepor às demais. E os teóricos ainda discutem sobre como proceder em relação aos efeitos dos problemas derivados de designs incompletos, na atualidade, seja com os *trolleys*, seja com os veículos autônomos já operados integralmente por IA.

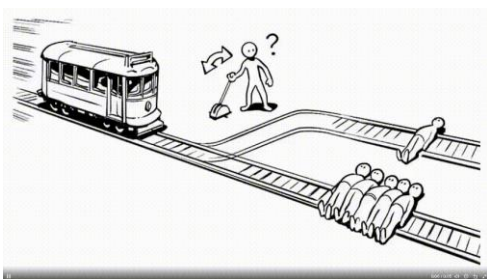
que os problemas possam perdurar, sejam os atuais ou os novos que não tardarão a ocorrer.

O que se faz, ainda, na atualidade, em termos de *design* responsável, é eleger e promover certos valores na fase que antecede o processo construtivo de aplicações tecnológicas – para que os protótipos de tais desenvolvimentos já tenham, em si, incorporados, os valores eticamente selecionados e, portanto, como parte necessária ao desenvolvimento, e não mais contingencial ou acidental.

A experiência de investigação para o desenvolvimento desta dissertação, no Projeto Signo, trouxe uma elucidação acerca como tais processos podem se dar de forma eticamente responsável, na prática, nos andamentos de se desenvolver soluções em aplicações tecnológicas para determinadas respostas. No caso do Projeto Signo, tais respostas se referem às necessidades de médicos e pacientes de glaucoma, no âmbito da oftalmologia.

Assim, para se atingir formar bases para projetos de aplicações eticamente responsáveis e com base em valores previamente definidos como relevantes e necessários, diversas interações entre as equipas de *designers*, programadores, filósofos e médicos elegeram cenários e posicionamentos nos quais tais valores foram identificados nas análises

Recentemente, alguém questionou a uma IA (<https://klingai.com/>) como resolver o problema do *trolley* e a resposta, em imagens, foi melhor do que a nossa, até então, e mostra que a IA apenas recuou lentamente com o *trolley*, a se recusar em seguir em frente.



Post original: “I tried to see how Kling v1.6 would handle the trolley problem.” E a resposta foi, em termos de imagens: “But it just backed away slowly.” (Publicado no X, em 07/jan/2025 e acedido pelo link <https://x.com/fofrAI/status/1876638297134678173> em 12/jan/2025).

metodológicas, no caso a análise temática. Assim, a equipa de *designers* pôde identificar, dentre todos os valores, os que sejam transversais, mas não apenas. E isto é um avanço, obviamente, em termos de *design* responsável.

Nas figuras 7 e 8^{cxlvi}, a seguir, podemos perceber os elementos investigativos utilizados como recursos no Projeto Signo.

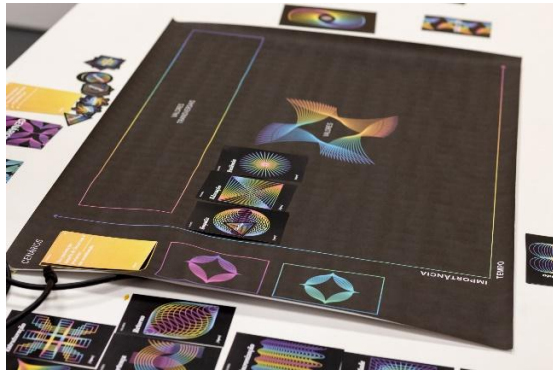


Figura 7 - Quadro utilizado como suporte para a elicitação dos valores relevantes, nas dinâmicas entre os *stakeholders*.



Figura 8 - *Cards* dos possíveis valores relevantes, utilizados como recursos, a partir das questões colocadas pelos *designers*.

Pois, a inserção de valores no *design* é mesmo assim: um processo multifacetado, que acaba por envolver considerações tanto objetivas como subjetivas, a partir de tais relações intersubjetivas (Spiekermann, 2016). Tal processo ainda precisa ser realizado de forma analógica, até que se possa, digitalmente, considerar uma IA (do futuro, dotada de uma estrutura prototípica metaontológica) que faça tal operação de eleição de valores, a partir das diversas partes interessadas e relevantes aos projetos (*stakeholders*), a fornecer um output sintetizador e analítico das quatro dimensões consideradas, bem como certos pontos específicos de responsabilização de atribuições. Uma complexa análise para consideração dos valores mais apropriados e relevante, tal como a realizada pela equipa do Projeto Signo, precisa levar em conta, ao menos, mas

^{cxlvi} Imagens creditadas à Fraunhofer Portugal AICOS (<https://www.linkedin.com/company/fraunhofer-portugal-aicos/>), publicadas no LinkedIn, acedidas em 12/jan/2025 no *post* disponível através do link https://linkedin.com/posts/fraunhofer-portugal-aicos_digitalhealth-values-hcd-ugcPost-7257020905489395712-ADv6

não apenas: (1) os valores intrínsecos e extrínsecos^{cxlvii}, (2) capital de *design*^{cxlviii}, (3) influência subjetiva^{cxlix}, (4) hierarquia de valores^{cl}, (5) métodos de elicitação^{cli} e uma (6) apropriada comunicação visual^{clii}.

3.8. Que quer dizer isto tudo, afinal? De Tales de Mileto à IA...

Desde que citamos Nagel, no início desta dissertação, e a cada sentença escrita, nunca a questão “que quer dizer isto tudo?” esteve próxima de ser saciada. E, cremos, nunca

^{cxlvii} Intrínsecos são aqueles valores inerentes às expectativas da utilização da aplicação – da experiência provável. Extrínsecos são valores correlacionados a algo externo ou transcendente à própria aplicação – tal como os próprios pacientes, ou mesmo médicos, em relação à impressão acerca da aplicação, antes ou após o uso. (Spiekermann, 2016)

^{cxlviii} São os valores que ajudam a posicionar o design como um ativo relevante em termos de impacto econômico. O desenvolvimento do design da aplicação deve gerar valor (e obviamente, inclusive no sentido de resultados financeiros) e atingir as necessidades e propósitos de todos os stakeholders: investidores, médicos, utentes, etc. (McKay & Meyboom, 2022)

^{cxlix} A influência subjetiva leva em conta o próprio envolvimento subjetivo, o que leva as equipas de designers, mesmo que percebam seu alto grau de subjetividade exigido, tenha medidas e precauções para que isto não resulte em parcialidades nas análises realizadas. É sabido que a experiência, cultura e valores do designer influenciam suas escolhas e, conseqüentemente, o valor do design resultante. (Innella & Petroni, 2019)

^{cl} A consideração e a definição de uma hierarquia de valores incorporados em um projeto é fundamental e deve estar claramente definidas. Assim, os valores foram considerados, mas dentre eles houve uma hierarquização em relação à relevância e a duração, a gerarem valores transversais. (Spiekermann, 2016)

^{cli} Que são as diversas técnicas podem ser usadas para identificar e elicitar os valores relevantes para um projeto: questionários, entrevistas, workshops de cocriação, análises de dados de mercado, etc. (Akin, 2021) .

No caso do Projeto Signo, para além das análises temáticas aplicadas sobre processos e conteúdos, foi também desenvolvido uma espécie de simulação interativa na qual os participantes apuravam as relevâncias em termos de valores, em dadas situações hipotéticas criadas e geridas pelas equipas de design.

^{clii} Que desempenha um papel excepcional na comunicação do valor do design. Diagramas, esboços, protótipos e apresentações visuais traduzem conceitos abstratos em ideias concretas e, com isso, os *stakeholders* passam a dirimir sobre algo mais próximo, pela objetividade que emerge das diversas intencionalidades. (Mozota, 2023)

será, se considerada a nossa insaciabilidade pelo conhecimento. A IA, afinal, tem estreita correlação com tal insaciabilidade de se conhecer, se vista como meio, como bem argumentamos no início e como bem podemos constatar, por agora.

A considerar a IA como fonte epistêmica, o que seria ela? Como se dá o conhecimento através de uma IA? Percepção? Memória? Inferências Lógica? Testemunho? Instrumento?

Das muitas formas possíveis de se atingir o conhecimento, podemos considerar que a IA, ainda que considerada como mero instrumento, não a invalida como fonte epistêmica, bem como das outras opções consideradas. A assunção sobre o que seja “a” verdade, todavia, são lá outras questões e um tema complexo, por si só, que apenas abordamos superficialmente, e que merece maior atenção, ainda mais quando pretendemos alcançá-la ao extrapolarmos a objetividade dos dados, rumo a uma subjetividade. Portanto, podemos perceber que tal liminaridade proposta será mesmo superada quando houver o seu uso como meio incondicional que leva ao conhecimento.

Podemos, todavia, considerar que atingir tal patamar de instrumentalidade ao conhecimento através de artefactos tecnológico – como a IA do futuro – será mesmo um processo de complexidade que envolverá *designers* e utentes. Há uma lacuna na análise da aquisição de conhecimento de tecnologias que interagem com linguagem natural. (Freiman, 2023)

Do ponto de vista epistemológico, contrariamente ao combatido pelos teóricos da IA, um instrumento é projetado e construído para ser, de facto, opaco: encapsulam as complexidades do funcionamento interno do instrumento, permitindo que o utilizador compreenda e aceite a saída sem mais delongas. (Neges, 2018). Portanto, combater a opacidade de uma IA é o segundo absurdo. O primeiro é colocá-lo como problema.

Na fase do *design*, instrumentos também são concebidos e construídos para produzirem sinais estrategicamente seleccionados, que são indicadores específicos de uma intencionalidade prévia e muito bem definida, e esta é produzida para ser mesmo inacessível ao utilizador (tal como a intencionalidade que também é suposta ser insondável, segundo Tunhas, 2021).

Nos instrumentos, há que se considerar a perspetiva teórica dos mesmos, estabelecida exclusivamente pelos *designers* – por isso a preocupação de o sistema metaontológico operar como um protocolo, com vantagens construtivas ao ser adotado, e não como uma determinação legal, em si. Do ponto de vista dos utentes, há de haver a compreensão e aceitação dos conteúdos proposicionais dos outputs – será necessário destes um nível mínimo de compreensão o que parece estar a ser posto em causa, pelo declínio^{cliii} que está a ser apurado, globalmente. Há também a necessidade de o utente confiar no instrumento, e que precisa ser operado corretamente. Instrumentos precisam estar calibrados e, no caso da IA, tal calibração pode ser considerada como geradora de resultados justos – sem problemas. Assim, podemos considerar que uma IA, como instrumento, pode criar conhecimento a partir da interação complexa capaz de realizar entre utentes em suas diversas formas de relações. (Neges, 2018)

^{cliii} Saiba mais em <https://www.oecd.org/en/about/programmes/pisa.html> acedido em 11/jan/2025.

Conclusão e Considerações Finais

Como conclusão, acreditamos estarem suficientemente consolidadas as seguintes construções argumentativas, realizadas ao longo da dissertação, a concluírem que:

(1) A IA tanto é tanto uma oferta quanto geradora de demanda

Argumentamos que a IA, como produto, está direcionada para suprir necessidades e desejos humanos, a atuar como uma oferta através das *big techs*. No entanto, também impacta os utentes e até mesmo as formas sociais, que por sua vez, consomem mais e criam novas demandas.

Tal dinâmica gera um ciclo no qual a IA, ao mesmo tempo que atende às demandas, também estimula a criação de novas através de estratégias afins à intencionalidade do próprio capital.

O capital tem um papel fundamental nesse processo, com sua intencionalidade direcionando o desenvolvimento e a aplicação da IA. Toda a análise e proposta de solução à problemática, portanto, em relação às IA devem, necessariamente, estar subordinadas e afins a tal dinâmica de interesses capitalistas, para serem exequíveis.

(2) O diálogo é o mais valioso produto

Argumentamos que as *big techs* reconhecem o valor do diálogo como uma forma de atender às necessidades e desejos dos usuários de maneira personalizada e eficiente. A IA conversacional, em particular, é uma ferramenta poderosa para a exploração comercial automatizada do diálogo em larga escala e a atingir potencialmente o nível de customização da individualidade.

Destacamos que o diálogo, nesse contexto, não se limita à mera troca de informações, mas abrange a capacidade da IA de compreender e responder às necessidades epistemológicas, emocionais e mesmo cognitivas – crenças, desejos, vontades e estados mentais dos usuários. Através do diálogo dotado da interpretação de tais dimensões subjetivas, a IA pode influenciar o comportamento dos usuários, gerar demanda por produtos e serviços, e até mesmo moldar a percepção da realidade.

Surgiram, aqui, os desafios éticos relacionados à comercialização do diálogo – nos quais os problemas mapeados se justificam como relevantes. A capacidade de a IA persuadir e manipular levanta preocupações sobre a autonomia e cognição dos usuários. Defendemos a necessidade de um design responsável e de mecanismos de controle para garantir que a IA seja utilizada de forma ética e transparente, visto que a IA apresenta um potencial imenso para o mercado, mas também exige uma reflexão profunda sobre os seus impactos éticos e sociais.

A busca por um equilíbrio entre a exploração comercial e a proteção dos usuários é o verdadeiro desafio para os legisladores, menos capazes de perceberem o que está realmente a se passar, dada a capilarização de atuação de uma IA conversacional, dada a nível individual.

(3) O capital tem sua intencionalidade que não pode ser desprezada

Argumentamos que a intencionalidade do capital, nesse contexto, se manifesta na forma como a IA é projetada, desenvolvida e aplicada. O capital atua como o investidor da IA (para criar produtos e serviços que atendam às demandas do mercado, gerando lucro e expandindo sua influência); e o faz a partir da compreensão das necessidades e dos desejos dos usuários (pois busca entender ao máximo seus usuários, ou consumidores) e, com isso, proporciona uma geração de demanda (ao excitar os usuários e, com isso, criar novas demandas ao influenciar seus comportamentos).

Por isso, a exploração e a amplificação deste ciclo através do diálogo é muito conveniente ao capital. O risco ocorre quando a intencionalidade do capital se sobreponha aos interesses dos usuários, ao ignorar os aspectos éticos e sociais, a colocar em risco a autonomia, a privacidade e o bem-estar dos usuários na busca desmedida pelo lucro e pela expansão do mercado.

Por isso, argumentamos que os mesmos princípios metodológicos que permitirão ao capital analisar o discurso dos usuários também permitirão a criação de ferramentas que analisarão o discurso influenciador do capital. Evocamos o conceito de *pharmakon*, por diversas vezes, quando a diferença entre o que seja remédio e veneno é apenas a dose.

(4) Os problemas éticos existem, mas estão mal caracterizados

Defendemos que a falta de uma fundamentação ontológica consistente afeta e impede uma análise precisa em relação aos problemas éticos. Há, também a argumentação da existência de uma dissonância teleológica – quando uma IA conversacional oferece exatamente o que foi projetada a realizar.

Citamos que a caracterização inadequada dos problemas éticos se manifesta nos seguintes aspectos: foco nos efeitos, e não nas causas (pela dissonância teleológica); visão estática da IA (uso de uma teorização estática – ou analógica – para algo que seja eminentemente dinâmica – ou digital); desconsideração da intencionalidade do capital (ignorada como influencia a forma como a IA é projetada, desenvolvida e aplicada)

Para superar esses desafios, defendemos o avanço das investigações ontológicas, a derivar uma elaboração de um protocolo metaontológico que permita a representação da complexidade existencial e a integração das dimensões objetiva, intersubjetiva e subjetiva, para além da própria dimensão do capital – dada como a quarta dimensão influenciadora e gestora das outras três.

(5) O mapa dos problemas éticos é inconsistente

Exploramos e criticamos o mapa apresentado por (Mittelstadt et al., 2016), a concluir que, por sua abordagem estática de um sistema dinâmico e por sua falha em considerar a teleologia atual da IA, apresenta uma inconsistência entre tempos, nos levando a uma liminaridade tecnológica em termos de IA. Contudo, o valor do mapa é mesmo em ser um guia para a construção, a partir dos problemas considerados, de como uma IA do futuro poderá existir, de forma a superar tais problemas, ainda que outros poderão vir a ocorrer.

(6) Qualquer melhoria passará pelo ajuste ontológico entre o atual e o digital

Argumentamos que a falta de correspondência ontológica entre o mundo real ("atual") e a sua representação digital é a origem dos problemas éticos mapeados. O investimento em uma abordagem metaontológica levará a IA a um avanço considerável, assim como também os riscos éticos, certamente, mas com a diferença de que, diferente da atualidade, existirão formas de extrair, da mesma fonte dos “venenos”, o próprio “remédio”, através de instrumentos que poderão ser úteis para analisar dados em

termos epistemológicos ou ainda em termos de comunicação estrategicamente produzida para influência ou cerceamento de liberdades, constrangimentos morais ou níveis indesejados de preconceitos, mas não apenas.

Como consideração final (ou recomendação):

- (7) Urge investir na modelagem de um protocolo metaontológico de representação do modal existencial que seja capaz de capturar e emular a complexidade relacional e, com isso, operar os valores para resultar em uma IA ética e responsável.

Consideramos, portanto, a partir das diversas argumentações e conclusões, que a criação de uma IA que seja ética, transparente e que beneficie tanto seus utentes, como a sociedade como um todo exige uma mudança não apenas de *design*, mas de paradigma, que abandone a visão estática – em dados – para outra dinâmica, baseada no modal existencial representado em sua complexidade máxima, a partir de suas relações.

Referências Bibliográficas

- Ajana, B. (2020). Personal metrics: Users' experiences and perceptions of self-tracking practices and data. *Social Science Information*, 59(4), 654–678.
<https://doi.org/10.1177/0539018420959522>
- Akın, Ö. (2021). *Design Added Value: How Design Increases Value for Architects and Engineers* (1.^a ed.). Springer.
- Alves, P. M. S. (2022). Ontology and Epistemology in Husserl's Ideen-I. *Phainomenon*, 33(1), 3–24. <https://doi.org/10.2478/phainomenon-2022-0002>
- Anderson, C. (2008). The End of Theory: The Data Deluge Makes the Scientific Method Obsolete. *Em Wired*. *Wired*.
<https://doi.org/https://www.wired.com/2008/06/pb-theory/>
- Ben Mahria, B., Chaker, I., & Zahi, A. (2021). A novel approach for learning ontology from relational database: from the construction to the evaluation. *Journal of Big Data*, 8(1). <https://doi.org/10.1186/s40537-021-00412-2>
- Berners-Lee, T., Hendler, J., & Lassila, O. (2001, Maio). The Semantic Web: A New Form of Web Content That is Meaningful to Computers Will Unleash a Revolution of New Possibilities. *Scientific American Magazine*.
<https://www.scientificamerican.com/article/the-semantic-web/>
- Bostrom, N. (2014). *Superintelligence: paths, dangers, strategies*. Oxford University Press.
- Brey, P., & Søraker, J. (2009). *Philosophy of Computing and Information Technology*. *Em A. Meijers, D. Gabbay, P. P. Thagard, & J. Woods (Eds.), Philosophy of Technology and Engineering Sciences (Vol. 14)*. Elsevier.
- Castoriadis, C. (2000). *A Instituição Imaginária da Sociedade* (5.^a ed.). Paz e Terra.
- Dasgupta, S., Papadimitriou, C. H., & Vazirani, U. V. (2006). *Algorithms*. McGraw-Hill Higher Education.

- de Boer, B., Te Molder, H., & Verbeek, P. P. (2018). The Perspective of the Instruments: Mediating Collectivity. *Foundations of Science*, 23(4), 739–755.
<https://doi.org/10.1007/s10699-018-9545-3>
- Digman, J. M. (1990). Personality Structure: Emergence of the Five-Factor Model. *Annual Review of Psychology*, 41, 417–457.
<https://doi.org/https://doi.org/10.1146/annurev.ps.41.020190.002221>
- Dreyfus, S. (2002). Richard Bellman on the Birth of Dynamic Programming. *Operations Research*, 50(1), 48–51.
<https://doi.org/http://dx.doi.org/10.1287/opre.50.1.48.17791>
- Ezell, B. C., & Crowther, K. G. (2007). Philosophical issues and their implications for the systems architect. *Foundations of Science*, 12(3), 269–276.
<https://doi.org/10.1007/s10699-007-9108-5>
- Floridi, L. (2012). Big data and their epistemological challenge. Em *Philosophy and Technology* (Vol. 25, Número 4, pp. 435–437). <https://doi.org/10.1007/s13347-012-0093-4>
- Fodor, J. A. (1980). Methodological solipsism considered as a research strategy in cognitive psychology 1. Em *THE BEHAVIORAL AND BRAIN SCIENCES* (Vol. 3).
- Foucault, M. (1979). *Microfísica do Poder* (R. Machado, Ed. & Trad.; 1.ª ed.). Edições Graal.
- Foucault, M. (1987). *A Arqueologia do Saber* (L. F. B. Neves, Trad.; 3.ª ed.). Forense-Universitária.
- Frankfurt, H. G. (2005). Sobre falar merda. *Intrínseca*. www.intrinseca.com
- Freiman, O. (2023). Analysis of Beliefs Acquired from a Conversational AI: Instruments-based Beliefs, Testimony-based Beliefs, and Technology-based Beliefs. *Episteme*, 1–17. <https://doi.org/10.1017/epi.2023.12>
- Gennep, A. Van. (2011). *Os Ritos de Passagem* (M. Ferreira, Trad.). Vozes Editora.

- Goddard, C. (2018). *Ten Lectures on Natural Semantic Metalanguage - Exploring Language, Thought and Culture Using Simple, Translatable Words*. Brill.
https://doi.org/10.1163/9789004357723_001
- Goddard, C., & Wierzbicka, A. (2011). Semantics and cognition. *Wiley Interdisciplinary Reviews: Cognitive Science*, 2(2), 125–135. <https://doi.org/10.1002/wcs.101>
- Guizzardi, Giancarlo. (2005). *Ontological foundations for structural conceptual models*. Centre for Telematics and Information Technology, Telematica Instituut.
- Habermas, J. (2012). *Teoria do agir comunicativo - vol. 2: Sobre a crítica da razão funcionalista (1ª)*. WMF Martins Fontes.
- Harding, S. (2011). *The Postcolonial Science and Technology Studies Reader*.
<https://www.dukeupress.edu/the-postcolonial-science-and-technology-studies-reader>
- He, B., Patel, M., Zhang, Z., & Chang, K. C.-C. (2007, Maio). Accessing the deep web. *Association for Computing Machinery*, 50(5), 94–101.
<https://doi.org/https://doi.org/10.1145/1230819.1241670>
- Heidegger, M. (1999). *Introdução à metafísica*. (C. Leão, Ed.; 4ª). Tempo Brasileiro.
- Heidegger, M. (2008). *Parmênides*. Editora Vozes / Editora Universitária São Francisco.
- Heidegger, M. (2012). *Ser e tempo (Coleção multilíngues de filosofia UNICAMP)*. Edição em alemão e português; tradução e organização Fausto Castilho., Ed.; 10.ª ed.). Editora Unicamp/editora Vozes.
- Hicks, M. T., Humphries, J., & Slater, J. (2024). ChatGPT is bullshit. *Ethics and Information Technology*, 26(2). <https://doi.org/10.1007/s10676-024-09775-5>
- Innella, G., & Petroni, M. (2019). *Going Real: The Value of Design in the Era of PostCapitalism*.
- Kant, I. (2008). *Prolegómenos a Toda a Metafísica Futura (A. Morão, Trad.; 1.ª ed.)*. Edições 70.
- Kant, I. (2018). *Crítica da Razão Pura (9.ª ed.)*. Fundação Calouste Gulbenkian.

- Keul, H.-K. (2002). Subjectivity and Intersubjectivity: Remarks on the Concept of Freedom in Kant and Habermas. *SUBJECTIVITY AND INTERSUBJECTIVITY The Journal of Value Inquiry*, 36, 253–266.
- King, M., Valle, R. S., & Halling, S. (1989). *Existential-Phenomenological Perspectives in Psychology: Exploring the Breadth of Human Experience* (R. S. Valle & S. Halling, Eds.).
- Lange, B., & Lechterman, T. M. (2021). Combating disinformation with AI: Epistemic and ethical challenges. <https://philpapers.org/archive/LANCDW-2.pdf>
- Lear, J. (1988). *The Desire to Understand*. Cambridge University Press.
- Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Em Nature* (Vol. 521, Número 7553, pp. 436–444). Nature Publishing Group.
<https://doi.org/10.1038/nature14539>
- Martinez-Cruz, C., Blanco, I. J., & Vila, M. A. (2012). Ontologies versus relational databases: Are they so different? A comparison. *Em Artificial Intelligence Review* (Vol. 38, Número 4, pp. 271–290). <https://doi.org/10.1007/s10462-011-9251-9>
- May, R. (2007). *Love and Will*. W. W. Norton & Company, Inc.
- McKay, S., & Meyboom, A. (2022). *Design Capital - The Hidden Value of Design in Infrastructure*. Routledge.
- Meersman, R. (2001). Ontologies and Databases: More than a Fleeting Resemblance. *Em A. d’Atri & M. Missikoff (Eds.), Proceedings of the OES/SEO 2001 Rome Workshop*. LUISS Publications. <http://www.starlab.vub.ac.be/publications/OES-SEO2001.pdf>
- Merleau-Ponty, M. (1999). *Fenomenologia da Percepção* (C. A. R. de Moura, Trad.; 2.^a ed.). Martins Fontes.
- Miguens, S. (2007). *Filosofia da Linguagem – uma introdução*. Faculdade de Letras da Universidade do Porto.

- Miller, C. (2014). Moral Virtues, Epistemic Virtues, and the Big Five. Em O. Flanagan & A. Fairweather (Eds.), *Naturalizing Virtue* (pp. 92–117). Cambridge University Press.
- Mises, L. von. (1968). *El Socialismo* (L. Montes de Oca, Trad.; 2.^a ed.). Instituto Nacional de Publicaciones de Buenos Aires.
- Mises, L. von. (2010). *Ação Humana* (D. Stewart Jr., Trad.). Instituto Ludwig von Mises Brasil.
- Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data and Society*, 3(2).
<https://doi.org/10.1177/2053951716679679>
- Molina, M. D., Sundar, S. S., Le, T., & Lee, D. (2021). “Fake News” Is Not Simply False Information: A Concept Explication and Taxonomy of Online Content. *American Behavioral Scientist*, 65(2), 180–212.
<https://doi.org/https://doi.org/10.1177/0002764219878224>
- Mozota, B. B. de. (2023). *Design Management: Using Design to Build Brand Value and Corporate Innovation*. Allworth.
- Neges, M. S. (2018). DISSERTATION / DOCTORAL THESIS [PhD dissertation]. University of Vienna.
- Nelson, J. L., & Taneja, H. (2018). The small, disloyal fake news audience: The role of audience availability in fake news consumption. *New Media and Society*, 20(10), 3720–3737. <https://doi.org/10.1177/1461444818758715>
- Newman, L. (2023). Descartes’ Epistemology. Em E. Zalta & . Nodelman (Eds.), *The Stanford Encyclopedia of Philosophy* (Winter 2023).
<https://plato.stanford.edu/archives/win2023/entries/descartes-epistemology/>
- Nietzsche, F. W. (2008). *A vontade de poder* (M. S. P. Fernandes & F. J. D. de Moraes, Trads.). Contraponto.
- Oberst, M. (2015). KANT ON UNIVERSALS. Em *Source: History of Philosophy Quarterly* (Vol. 32, Número 4). <http://www.jstor.org/stable/44076598>

- Oomen, P. (2019). Language about God in Whitehead's Philosophy: An Analysis and Evaluation of Whitehead's God-Talk. *Process Studies*, 198–218.
<https://philpapers.org/archive/OOMLAG.pdf>
- Ortolan, L. (2022). *O guia cínico e selvagem dos jogos da vida* (1.^a ed.). KDP.
- Park, J. S., O'Brien, J., Cai, C. J., Morris, M. R., Liang, P., & Bernstein, M. S. (2023, Outubro 29). Generative Agents: Interactive Simulacra of Human Behavior. *UIST 2023 - Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*. <https://doi.org/10.1145/3586183.3606763>
- Park, J. S., Zou, C. Q., Shaw, A., Hill, B. M., Cai, C., Morris, M. R., Willer, R., Liang, P., & Bernstein, M. S. (2024, Novembro 15). Generative Agent Simulations of 1,000 People. <https://arxiv.org/abs/2411.10109>
- Partridge, C. (2015). Why (and how) to use a metaphysicalist foundational ontology
Title: Why (and how) to use a metaphysicalist foundational ontology.
<https://www.researchgate.net/publication/370510623>
- Pennock, R. T. (2019). *An Instinct for Truth: Curiosity and the Moral Character of Science*. MIT Press. <https://mitpress.mit.edu/9780262042581/an-instinct-for-truth/>
- Platão. (2011). *Timeu-Críticas* (M. do C. Fialho, Ed.; R. Lopes Platão, Trad.; 1.^a ed.). Centro de Estudos Clássicos e Humanísticos.
- Portmess, L., & Tower, S. (2015). Data barns, ambient intelligence and cloud computing: the tacit epistemology and linguistic representation of Big Data. *Ethics and Information Technology*, 17(1), 1–9. <https://doi.org/10.1007/s10676-014-9357-2>
- René Descartes. (2005). *Meditações Metafísicas*. Martins Fontes.
- Ricoeur, P. (1989). *O Conflito das Interpretações: Ensaio de Hermenêutica* (M. F. Sá Correia, Trad.). Rés Editora.

- Ricoeur, P. (2016). *Hermeneutics and the Human Sciences Essays on Language, Action and Interpretation* (J. B. Thompson, Ed.). Cambridge University Press.
<https://doi.org/10.1017/CBO9781316534984>
- Romele, A. (2019). *Digital Hermeneutics: Philosophical Investigations in New Media and Technologies*. www.routledge.com/
- Russell, S. J., & Norvig, P. (2021). *Artificial Intelligence: A Modern Approach* (4th ed.). Pearson. <https://doi.org/10.1109/MSP.2017.2765202>
- Saramago, J. (2002). *O homem duplicado*. Caminhos.
- Sartre, J.-P. (2011). *O ser e o nada - Ensaio de ontologia fenomenológica* (P. Perdigão, Trad.; 20.^a ed.). Vozes.
- Sartre, J.-P. (2015). *O que é a subjetividade?* (E. dos S. Abreu, Trad.; 1.^a ed.). Nova Fronteira.
- Shields, C. (2007). *Aristotle*. Routledge. <https://philpapers.org/rec/SHIA-7>
- Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). *Fake News Detection on Social Media: A Data Mining Perspective*. <http://arxiv.org/abs/1708.01967>
- Spiekermann, S. (2016). *Ethical IT Innovation - A Value-Based System Design Approach*. CRC Press.
- Suits, B. (2017). *A Cigarra Filosófica - A vida é um jogo?* Gradiva.
- Szopa, R. (2023). Ethical problems in the use of algorithms in data management and in a free market economy. *AI and Society*, 38(6), 2487–2498.
<https://doi.org/10.1007/s00146-021-01319-5>
- Tolley, C. (2012). Kant on the Content of Cognition. *European Journal of Philosophy*, 20(4), 200–228. <https://doi.org/10.1111/j.1468-0378.2011.00483.x>
- Tunhas, P. (2021). Kant: o domínio da natureza. Em *A ciência metafísica : itinerários de um conceito* (pp. 215–215).
https://sigarra.up.pt/flup/pt/pub_geral.pub_view?pi_pub_base_id=517865

- Tunhas, P. (2023). *Filosofia* (J. F. Meirinhos & D. P. Aurélio, Eds.; 1ª). Instituto de Filosofia : U. Porto Press. <https://doi.org/https://doi.org/10.21747/978-989-746-365-5/filo>
- Turner, V. W. (1974). *O Processo Ritual - Estrutura e Antiestrutura* (N. C. de Castro, Trad.). Editora Vozes.
- Van Otterlo, M. (2013). Privacy, Due Process and the Computational Turn. Em *Privacy, Due Process and the Computational Turn The Philosophy of Law Meets the Philosophy of Technology* (1st ed., pp. 41–64). Routledge. <https://doi.org/10.4324/9780203427644>
- Vergílio. (2019). *Geórgicas* (G. A. F. Silva, Trad.). Livros Cotovia.
- von Eckartsberg, R. (1998). *Phenomenological Inquiry in Psychology: Existential and Transpersonal Dimensions* (R. Valle, Ed.). Springer.
- Waelbers, K., & Briggles, A. (2010). Three Schools of Thought on Freedom in Liberal, Technological Societies. https://www.pdcnet.org/techne/content/techne_2010_0014_0003_0176_0193
- Whitehead, A. North., & Sherburne, D. W. . (1981). *A key to Whitehead's Process and reality*. University of Chicago Press.
- Whitehead, A. N. (2010). *Processo e Realidade - Ensaio de Cosmologia* (M. T. Teixeira, Ed. & Trad.; 1.ª ed.). Centro de Filosofia da Universidade de Lisboa.
- Wittgenstein, L. (2015). *Tratado Lógico-Filosófico Investigações Filosóficas* (6ª). Fundação Calouste Gulbenkian.
- Wood, A. W. (1975). Kant's Dialectic. *Canadian Journal of Philosophy*, 5(4), 595–614. <https://doi.org/10.1080/00455091.1975.10716970>
- Zamborlini, V. C. (2011). Estudo de alternativas de mapeamento de ontologias da linguagem ONTOUML para OWL: Abordagens para representação de informação temporal. <https://doi.org/https://repositorio.ufes.br/items/2259b571-4692-487e-90a2-b8a1e3153161>