

INTELIGÊNCIA ARTIFICIAL E A INDÚSTRIA DE PRODUÇÃO FILOSÓFICA: OTIMISMO OU PESSIMISMO?

Luís Estevinha Rodrigues¹

Resumo:

Explico a inevitabilidade de um cenário em que a atual indústria de produção filosófica se cruza com a inteligência artificial geral. Defende-se que esta última terá um impacto não negligenciável sobre a primeira, algo que os decisores devem antecipar e preparar. Proponho uma posição equilibrada a este respeito, uma atitude prudencial entre o pessimismo e o otimismo, bem como a substituição de um desiderato de quantidade nesta indústria por um desiderato de qualidade.

Palavras-chave: inteligência artificial, filosofia, academia, produção.

ARTIFICIAL INTELLIGENCE AND THE PHILOSOPHICAL PRODUCTION INDUSTRY: OPTIMIST OR PESSIMISM?

Abstract:

I explain the inevitability of a scenario where the current philosophical production industry intersects with general artificial intelligence. It is argued that the latter will have a non-negligible impact on the former, something that decision-makers should anticipate and prepare for. I submit a balanced stance in this regard, a prudential attitude between pessimism and optimism, as well as the replacement of a desideratum of quantity in this industry with a desideratum of quality.

Keywords: artificial intelligence, philosophy, academy, productivity.

Introdução

George Bernard Shaw parece ter dito que tanto o otimista como o pessimista são úteis à sociedade, pois o primeiro inventa o avião e o segundo o paraquedas. Se o disse, concordo com ele.² Talvez o caso associado à filosofia que irei descrever de seguida ajude a compreender a minha anuência a este pensamento. Recordo-me de uma ANPOF já longínqua na qual dois conceituados filósofos trocaram argumentos muito interessantes sobre o estado da arte da filosofia académica contemporânea. Do lado pessimista e conservador da barricada estava Susan Haack, professora de lógica e filosofia da linguagem da Universidade de Miami. A conceituada pesquisadora defendia que a produção filosófica académica ganhara dimensões colossais, de que era evidência o excessivo número de submissões e publicações nos últimos cinquenta anos. Defendia que isso afetava negativamente prazos de resposta a submissões e,

¹ Doutor em epistemologia e filosofia da ciência (Universidade de Lisboa e Universidade Edinburgo). Professor do Programa de Pós-graduação da Universidade Federal do Ceará, Brasil. Membro associado do grupo LANCOG, Centro de Filosofia da Universidade de Lisboa, Portugal. Coordenador do grupo de pesquisa Filosofia, Metafísica e Cognição, UFC e CNPq. Orcid: 0000-0001-7811-2756. luisestevinha@ufc.br. Orcid: <https://orcid.org/0000-0001-7811-2756>.

² Não me foi possível verificar autenticidade da atribuição. O *Cambridge Companion to George Bernard Shaw*, editado por Innes (1998) não lhe faz qualquer referência.

indiretamente, a qualidade das mesmas. Temia o volume de revistas predatórias (pagas) e o excesso de livros publicados sem conselho editorial adequado. Sugeria mais moderação e a implementação de normas mais apertadas e rígidas, bem como uma ampla avaliação do rumo geral da indústria de produção filosófica (doravante IPF), embora não tenha usado exatamente esta expressão. Do lado da barricada do otimismo e do progressismo estava Jesse Prinz, um jovem professor de filosofia da consciência e da neurociência pertencente à Universidade de Nova York. Prinz defendeu veementemente que a liberdade criativa e o exponencial intercâmbio de ideias filosóficas, bem como a implementação de processos de controlo de qualidade das produções nunca havia estado melhor. Os sistemas de revisão anónima, dupla e tripla, significavam um incremento e uma garantia de qualidade. Argumentando assim, ele vaticinava um futuro risonho para a nossa arte.

Embates como o descrito revelam bem a tensão entre o pessimismo e conservadorismo, de um lado, e otimismo e progressismo, do outro. Este tipo de confrontos traz à memória a famosa tensão essencial na ciência descrita por Kuhn (1977, p. 227), segundo o autor, um ingrediente indispensável para a mudança e reformulação do conhecimento humano.

Creio que este enquadramento continua ilustrativo, sendo a ideia de uma tensão entre posições otimistas e pessimistas, progressistas e conservadoras, certamente aplicável ao problema que motiva este artigo. Trata-se da questão do impacto causado pela entrada da inteligência artificial geral (doravante IAG) na IPF: de que maneira será esta última impactada e afetada? Para submeter uma resposta, na primeira secção tento aclarar a natureza da IPF, ressaltando três das suas principais idiossincrasias, certamente revelantes para a compreensão do problema. Na segunda secção examino sumariamente analogias e *disanologias* (não analogias) entre processos fabris e a IPF, tentando mostrar que as idiossincrasias da IPF aludidas na secção anterior podem ser facilitadoras de um impacto negativo da IAG na IPF. Na terceira secção avalio resumidamente se as actuais ou futuras capacidades da IAG são suficientes para gerar uma revolução na IPF, seu *modus operandi* e resultados. Na quarta e última secção proponho que a hipótese mais plausível e desejável não é a de uma IAG todo poderosa substitutiva dos humanos na IPF, mas uma controlável e colaborativa, havendo espaço na filosofia para inteligências biológicas e inteligências artificiais. É defendido por último que essa revolução implicará mudanças das idiossincrasias da IPF, em especial o abandono gradual do desiderato de quantidade em favor do desiderato de qualidade.

1.A IPF e idiosincrasias relevantes

A IPF não deve ser confundida com uma qualquer instituição particular. Tal como aqui entendido e usado, o conceito de uma indústria de produção de conhecimento filosófico é uma abstração composta por vários elementos do mundo académico e adjacente: universidades e institutos, departamentos de graduação, pós-graduação, pró-reitorias, agências de fomento à pesquisa, ministérios, comissões, editoras, associações e sociedades de pesquisa, pesquisadores, docentes e discentes. Estes integrantes geram, encaminham, controlam e avaliam a produção académica em filosofia. São também responsáveis por recursos e políticas públicas associadas ao empreendimento da filosofia enquanto atividade cognitiva e pedagógica organizada.

Se excluirmos os casos pontuais de diminuição de produção académica ditados por contingências imprevisíveis, como por exemplo pandemias, vivemos tempos industriais no que respeita à produção académica (Fortunato *et al.*, p. 2018).^{3/4} É inegável que o número de artigos e livros científicos publicados em todo o mundo cresceu desmesuradamente nas últimas décadas (Bornmann *et al.* 2021). A confiar em Price (1986, p. 16), a taxa de crescimento de produções científicas acompanha proporcionalmente o corpo de saberes estabelecido num dado momento: quanto mais esse corpo cresce, mais tendência tem para crescer. Por conseguinte, os produtos da indústria de produção de conhecimento académico e ela própria tendem a autoincrementar-se cada vez mais. É um processo análogo ao da mitose de células biológicas. Da maior parte das vezes, o acréscimo é positivo e as suas consequências louváveis. Mas o crescimento desmesurado das células também pode acarretar resultados menos bons, produzindo degradação ou a diminuição de qualidade, como acontece com os casos de câncer. Assim também parece suceder na IPF.⁵

A IPF mundial faz assentar atualmente a sua legitimidade em métricas, índices de publicação, factores de impacto e processos de revisão anónima por pares. Por exemplo, os elencados pela Taylor e Francis, uma referência mundial em publicação, incluem critérios de uso e acesso às publicações, sejam imediatos, anuais ou por percentil, assim como velocidades

³Ver retrocesso na produção académica motivada pela pandemia de COVID-19 no seguinte relatório Elsevier/Bori:<https://abori.com.br/wp-content/uploads/2023/07/2022-um-ano-de-queda-na-producao-cientifica-para-23-paises-inclusive-o-Brasil.pdf> (consultado em 20/4/2024).

⁴Sobre massificação da produção académica, vejam-se, por exemplo, alguns dados da Scival em <https://www.sbu.unicamp.br/sbu/scival/>

⁵Veja-se:<https://www.ufrgs.br/jornal/fenomeno-crescente-no-mundo-da-publicacao-academica-revistas-predatorias-comercializam-espacos-de-divulgacao-e-colocam-em-risco-a-ciencia-brasileira/>

e tempos de publicação dos editores.⁶ No Brasil, os indicadores de produção focados em cursos de graduação, programas de pós-graduação e pesquisadores tornaram-se corriqueiros. Mais por razões que se prendem com ideários e políticas públicas visando a competição e a produção, e menos ao desiderato da própria filosofia, o *modus operandi* da quantificação tornou-se endógeno à IPF, a qual já não o dispensa. Maioritariamente concedidos e implementados para efeitos de aferição laboral, condicionado por ideais meritocráticos e determinado por regras de competição intra e interindústrias do conhecimento, o aludido ideário desenvolveu mecanismos de controlo de produção e de actividade geral dos filósofos. A atividade filosófica académica nas sociedades W.E.I.R.D.D.⁷ contemporâneas tornou-se massificada e focada num objetivo imposto pela disputa por recursos: produtividade.

A IPF brasileira debita anualmente muitos artigos e livros. Mas o que é realmente aproveitável dessa produção? E, mais importante, o que realmente é aproveitado? Não sabemos as respostas para estas questões, infelizmente. O melhor que podemos alcançar são talvez respostas aproximadas, indiretas, fornecidas por indicadores de aproveitamento genéricos, como sejam bibliografias citadas, projetos académicos, bolsas académicas, eventos académicos etc. Mas a real mais-valia, intrínseca ou instrumental, está por determinar.

Acresce a esta incerteza dos resultados o facto de a parafernália de instrumentos teóricos e conceituais de mensuração de resultados, a maior parte dos quais é importada, ter uma eficiência criticada na *literatura* (Stergiou *et al.*, 2013; Paulus *et al.*, 2018; Nassi-Calò, 2024). Por exemplo, cada vez mais a qualidade é medida pelo número de acessos, tantas vezes condicionados por factores espúrios que nada têm a ver com a qualidade *intrínseca* do artigo ou livro. Ademais, os critérios usados para mensuração de qualidade na IPF são científicos, isto é, como se disse, privilegiam a quantificação. Mas sabemos bem que a filosofia não é uma ciência. Sequer é uma ciência humana ou social. É uma atividade intelectual de direito e espaço

⁶ Ver regras processos em <https://www.tandfonline.com>. Movimentos como a *San Francisco Declaration on Research Assessment* (DORA) <https://sfdora.org> e o Manifesto de Leiden em <http://www.leidenmanifesto.org> recomendam utilizar índices como o Fator de Impacto (FI) apenas para “ranquear” periódicos, e não pesquisadores, instituições, concessão de recursos para pesquisa, ou programas de pós-graduação. Soma-se a isso o fato de que estimativas apontam que cerca de 50% dos artigos publicados não chegam a ser formalmente citados por outros autores, e nem por isso seu impacto na comunidade científica é nulo. Estes artigos são lidos, baixados, compartilhados e citados por meio de redes sociais, blogs, notícias, políticas públicas, e outras formas de presença online, reunidas e medidas em índices como o *Altmetric* (*alternative metrics*). <https://blog.scielo.org/blog/2022/02/09/como-se-utilizam-as-altmetrias-para-avaliar-a-producao-cientifica-da-america-latina>. Visitas à *Scopus-Elsevier*, *Web of Science* e *SciELO* podem dar-nos uma ideia da dimensão do volume de artigos e livros académicos indexados nestas entidades num determinado momento. No entanto, uma visão precisa da dimensão colossal da produção em todas áreas do conhecimento pode estar fora de alcance, pois esse volume encontra-se em mutação diária exponencial.

⁷ *Western, Educated, Industrialized, Rich, Democratic, Digital*.

próprios. Os decisores teimam, contudo, em colocar tudo no mesmo saco, com isso gerando imparidades dispensáveis, em especial para a IPF e seus elementos.⁸

O quadro cético desenhado até aqui remete-nos para o problema da desproporcionalidade e exponencialidade da exigência dentro e fora da IPF. Esta desproporcionalidade é em última análise gerada por políticas públicas e regulamentos da IPF que privilegiam um tipo de competição pouco saudável entre indivíduos e grupos como modo de aumentar ou potenciar uma produção que não reflete necessariamente qualidade. Ademais, este processo economicista é circular (como acontece aliás na economia), gerando um efeito de bola de neve: autorreplicação dos passos. Necessidade de recursos gera competição. Esta gera infundável produção. Por sua vez, esta gera mais crescimento e mais competição, inflacionando o processo, etc.

Embora este processo industrial seja generativo, mantendo e engrossando a IPF e seus constituintes⁹, ele parece colidir com o *modus operandi* filosófico tradicional. A IPF, enquanto parte dessa fábrica chamada ciência que produz uma liga de saberes, tornou-se fortemente produtiva, massificada, opressiva, impondo regras e prazos pouco ou nada condizentes com a essência do pensamento filosófico – o qual não deve ser apressado, imposto, mecânico. Ao impor uma ideologia de massificação da produção condicionada por competição e exigência próprias da IPCA, a IPF parece trair os seus pressupostos fundamentais: *phronesis*, duração e disposição para gerar filosofia. Bertrand Russell sugeriu apropriadamente que

A antiga propensão para a despreocupação e o divertimento foi de certo modo inibida pelo *culto da eficiência*. O homem moderno acha que qualquer atividade deve ser em prol de outras coisas, nunca da coisa mesma”. (Bertrand Russell, *O elogio do ócio*, p. 32. Itálico meu).

⁸ Ver uma distinção lógica clássica filosofia /ciência sugerida em Hartman (1963, p. 354) Um caso gritante e idiossincrático da diferença de processos e métodos da IPF e de outras subindústrias científicas pode ser constatado, por exemplo, na diferença de tamanhos e conteúdos de muitos artigos das chamadas ciências da natureza e das ciências “exatas” comparativamente com os de artigos “normais” de filosofia. O conteúdo do primeiro tipo de artigos costuma ser muito mais concentrado do que o dos últimos. Além disso, não raramente, são vários os pesquisadores de um laboratório ou de um núcleo de pesquisa científica que assinam e recolhem os louros ou dividendos por esse artigo, apesar de não terem contribuído *diretamente* para o aludido. Claro, essa referência vai ser creditada ao pesquisador que contribui parcamente, mesmo que a sua contribuição tenha sido mínima. No fim de contas, isto tem fortes impactos nos *curricula* dos pesquisadores e resultados dos núcleos de pesquisa; e, por consequência, impacta bastante na disputa por pontuações e recursos como bolsas ou financiamentos para pesquisa, estes últimos frequentemente escassos e altamente concorridos nos dias que correm.

⁹ Causando por vezes desespero, angústia e o suicídio académico de pesquisadores de pós-graduação. Veja-se o artigo de opinião “The Productivity Syndrome” (or why I stopped writing philosophy), disponível em <https://upnight.com/2018/05/31/the-productivity-syndrome-or-why-i-stopped-writing-philosophy>.

Em outras partes do seu texto célebre, Russell aponta para uma situação ideal. É um mundo possível em que os humanos *podem* ser ociosos neste bom sentido *russelliano* do termo. Sem dúvida que o nosso mundo actual não se compadece com este desiderato. Quem deseja e faz por pertencer a uma organização como a IPF deve seguir e cumprir as suas regras. O contrário seria como desejar pertencer à maçonaria sem obedecer aos ritos maçõn. Não é viável nem recomendável para a IPF. Não obstante, ser pesquisador não é o mesmo que pertencer a uma seita ou a grupo ideológico. Ser filósofo não é o mesmo que ser maçõn (embora alguns maçõns sejam filósofos). O filósofo deve refletir sobre a organização social onde se insere. Se considerar necessário, deve ser subversivo dessa organização. Não poucos exemplos desta insurgência filosófica aparecem na história da filosofia. Tal insurgência é menos um poder do que um dever.

Claro que muitos comentadores, sensatos e realistas, concordam que o modo de fazer filosofia nos nossos dias não se compadece com uma visão idealizada da atividade enquanto empreendimento intelectual livre de exigência produtiva. Por si próprias e enquanto minimamente razoáveis e controláveis, a competição académica e a exigência na pesquisa nada têm de errado. São molas que estimulam a geração de ideias, teorias e conhecimentos, trazendo para a ribalta da investigação de natureza científica e filosófica os respetivos especialistas de área.¹⁰ Todavia, como se disse acima, parece que a exigência e a competição excessivas contribuem para um sistema que converte os filósofos numa espécie de proletários da IPF.¹¹ Se for o caso, certas imposições da IPF têm algo de espúrio à arte, raiando inclusive o draconiano.

Condicionado pelos ditames de uma sociedade global cada vez mais exigente e ávida de produção e informação, o pesquisador da IPF não é muito dissemelhante *funcionalmente* do operário fabril de outras eras. Com efeito, o pesquisador contemporâneo vê-se obrigado a observar parâmetros e exigências industriais de produção epistémica motivados por um capitalismo do conhecimento (May, 2005, p.194). Da nossa perspectiva, a filosofia académica converteu-se em muitos aspectos numa atividade em linha. O problema disto, a bater-nos à porta e a ter em conta neste momento de forte revolução tecnológica, é o seguinte: será a IPF, enquanto processo humano de produção de pensamento, conhecimento e

¹⁰ Cada vez existem mais programas e mais pós-graduações e estudantes pós-graduados na área da filosofia, sendo os mestrados e doutorados debitados em muito maior número do que há duas ou três décadas (cf. <https://www.anpof.org.br/anpof/programas-associados>). Embora não me debruce sobre este aumento neste artigo, parece-me que ele é uma faca de dois gumes.

¹¹ O que já não é uma ideia recente, podendo ter sido forjada algures no Século 19, precisamente o século onde vieram a lume as disputas proletárias. Sobre as possíveis origens da famosa expressão “*Publish or Perish!*” (se não publicas, morres!), nunca antes tão levada ao pé da letra, ver por exemplo Garfield (1996, p. 11).

informação, sustentável nestes moldes quase anti-iluministas?¹² Tentaremos mostrar nas próximas secções como esta atual predisposição da IPF para a quantificação, competição e exigência a torna vulnerável ao impacto da IAG.

2. Intersecção: analogias, disanologias e impacto.

A questão a investigar agora é a do que acontecerá à IPF quando a IAG se intersectar com ela. Uma resposta possível e tentadora é a de que essa intersecção não sucederá, uma vez que a sociedade e os decisores políticos irão regulamentar o escopo da IAG, impedindo a admissão desta última na IPF ou mitigando o seu impacto nela. Esta resposta parece demitir liminarmente o problema. Sabemos bem, contudo, que a dimensão da caixa de pandora da IAG não poderá ser dissolvida assim tão facilmente por decreto ou regulamentação. A razão é simples: a IAG tem uma abrangência tão grande, estando já tão disseminada nas vidas de tantos, ameaçando entranhar-se muito mais ainda, que não parece ser possível regrar o seu uso assim de modo tão arbitrário. Seria um pouco como querer banir a rádio, a televisão, os jornais, a internet, as redes sociais, os computadores ou os celulares das sociedades e suas instituições, incluída a IPF. Implausível. Aliás, banir os quatro últimos itens já foi extensivamente tentado, sem sucesso, em nome da pureza da academia. Na verdade, sucedeu justamente o oposto da ostracização de alguns dos referidos elementos: quanto mais o conservadorismo tenta expulsar a inovação mais ela entra e se impõe em lugares e atividades considerados imunes. Não se consegue deter o progresso tecnológico, sua inserção e implementação nas práticas humanas, seja na pedagogia, medicina, segurança, comércio ou qualquer outra actividade. Devemos então suspender o juízo sobre esta possibilidade de regulação ou até proibição e focar-nos na forte possibilidade de os dois fenómenos se intersectarem cada vez mais, senão mesmo unificarem-se.

Acima sugeri que os desideratos de exigência, competição e produtividade militam a favor de uma perda de qualidade da filosofia, por um lado, e de uma subversão do modo de pesquisa do investigador, por outro. Esta colocação depende de uma analogia de cuja plausibilidade alguns poderão duvidar: a analogia entre uma linha de montagem fabril e a produção filosófica em massa. Com efeito, pode ser alegado que se olharmos para as atuais

¹² Conforme classicamente sugerido por Horkheimer & Adorno (2002, p. 94), a procura iluminista tende a eliminar o mito e a credence, mas depois, a postura esclarecida resultante descamba em novas mitologias, que por sua vez militam contra o próprio projeto iluminista.

linhas de produção automóvel ou peças, por exemplo, elas seguem modelos, ritmos e processos sistemáticos que geram a qualidade dos produtos resultantes em grandes quantidades. Logo, continua o crítico da analogia que fizemos, parece que os processos fabris da IPF também podem gerar qualidade. É um ponto interessante. Porém, há uma disanologia nesta réplica que bloqueia a sugestão de que a massificação de produtos intelectuais humanos gera o mesmo tipo de qualidade gerado por uma linha de montagem mecânica. Uma coisa é a IPF militar para a estandardização das produções intelectuais humanas, outra bem diferente é esta indústria poder gerar o mesmo tipo de resultados perfeccionistas de uma linha de montagem automóvel, por exemplo. Este último tipo de produção não demanda a intencionalidade mental, imaginação ou criatividade por parte dos sistemas executantes considerados individualmente, ao passo que o primeiro tipo demanda. Isto revela que a analogia funciona parcialmente, mas que existem disanologias noutros aspectos importantes entre a IPF e linhas de montagem. A analogia requer então um grão de sal.

Não obstante, é inegável que surgiu um jogador novo, não humano, que ameaça *emular*, num determinado nível, características humanas como pensamento, intencionalidade, imaginação, criatividade, inovação etc., tudo atributos humanos. Talvez até seja capaz de produzir, mecanicamente e em quantidades colossais, uma emulação em linha de filosofia. E o problema disto é o seguinte: tal como operadores fabris humanos não podem em regra competir com sofisticados mecanismos fabris (*robots*) desenhados para a montagem, também filósofos humanos não poderão competir com dispositivos e tecnologias de tratamento em massa e probabilístico de informação (*chatbots*). Esta não é uma ficção para compor um tratado filosófico. É o problema, previsível e objetivo, do impacto da IAG na IPF. As consequências para esta última são agora ainda difusas, mas preocupantes. Haverá motivo para este pessimismo? Poderão máquinas sofisticadas substituir filósofos?

A questão é crucial, pois se a IAG atingir um nível de inteligência capaz de fazer perigar a preponderância da nossa espécie relativamente à arte de filosofar, então a IPF *humana* estará em maus lençóis. Como veremos de seguida, se conseguirmos ensaiar uma resposta negativa e conservadora a esta questão partindo da literatura, talvez ela possa militar a favor de um otimismo a respeito da exclusividade da nossa espécie na IPF. Mas mesmo para isso teremos provavelmente de alterar alguns dos seus pressupostos essenciais e respetivas idiosincrasias. Deixarei essa análise para o último capítulo. Por ora, tentarei vistoriar se IAG tem ou terá capacidade para filosofar.

3. IAG: o elefante e a porcelana na sala da IPF

Muitos se lembram, melhor ou pior, que o matemático e lógico Alan Turing escreveu há 75 anos um artigo seminal onde questionava se as máquinas digitais, sistemas computacionais de estados discretos, poderiam pensar (Turing, 1950). A sua questão abriu uma caixa de pandora filosófica (Saygin et al., 2000). O questionamento de Turing obteve continuidade no famoso Projeto de Verão de Dartmouth, 1956, onde se costuma afirmar terem sido plantados os primeiros conceitos de inteligência artificial. Depois de duas conhecidas “idades do gelo” desta tecnologia na segunda metade do século passado, dificuldades relacionadas com a concretização técnica da IAG e seus objetivos fizeram esmorecer um pouco o interesse na tecnologia. Já as últimas três a quatro décadas desse empreendimento foram muito pautadas por desideratos económicos, sociais e militares que fizeram ressurgir o projeto. A IAG adquiriu um *momentum* imparável desde essa altura, ganhando uma aceleração maior, que nem as limitações impostas à chamada “lei de Moore” parecem fazer diminuir.¹³ Na atualidade, a IAG e as inteligências artificiais generativas (doravante IA-gen) adquiriram um interesse especial e inflacionado por parte das massas que compõem a sociedade, as quais observam estes sistemas e processos inserirem-se cada vez mais nos quotidianos do coletivo e dos indivíduos. A IAG está aí para ficar, não podendo nem devendo ser ignorada por qualquer indústria do conhecimento. Porquê? Respondo de seguida.

As capacidades, velocidades, débitos colossais, *Big-data*, caixas negras de algoritmos superpotentes (Domingos, 2015; Vaassen, 2022) e opacos trabalhando em redes neurais complexas por camadas, e suas iterações futuras, terão forte impacto na IPF, com consequências em aberto. É isto que denomino a face tecnológica do *problema do impacto*. A perspectiva apresentada neste artigo é a de que este impacto não depende somente da natureza e potencialidades das IA-gen, mas também das aludidas idiossincrasias da IPF. O problema tem, por conseguinte, uma dupla origem. É a concatenação dos dois fenómenos na origem, idiossincrasias da atual IPF e capacidade tecnológica das máquinas, que pode hipotecar o futuro de uma IPF exclusivamente humana.

Sendo assim, a questão que se coloca atualmente no chifre do problema relativo à capacidade tecnológica não é tanto se as máquinas podem pensar, como indagavam Turing e muitos outros. Agora pergunta-se se IAG e IA-gen podem pensar *como* especialistas; isto é, se podem pensar e executar como cientistas, filósofos, artistas etc. E a questão é se podem

¹³ <https://www.intel.com/content/www/us/en/newsroom/resources/moores-law.html#gs.afnx2u>.

imaginar, criar e inovar como os humanos. Talvez tão ou mais importante, precisamos descobrir se IA-gen do futuro possuirão inteligência emocional suficiente para assistirem autonomamente no processo educativo; ou se, pelo contrário, teremos meros idiotas, inconscientes e mecanizados, como tutores.¹⁴ Sendo a resposta afirmativa ou negativa, ela terá consequências para as indústrias do conhecimento e, com certeza, para a IPF também.

Com efeito, a IPF e os seus organismos e estruturas têm operado até ao momento sob um pressuposto cada vez mais questionável: fazer filosofia é um atributo exclusivamente humano. E de facto nunca vimos formigas, primatas superiores, cetáceos ou qualquer outro organismo *natural* produzir filosofia como nós, de forma ordenada e sistemática, por melhores que sejam as capacidades cognitivas e desempenhos linguísticos dessas entidades naturais. Mas talvez essa exclusividade esteja em fase terminal. A causa provável e cada vez mais discernível da mudança que se avizinha nos processos de produção de filosofia é o avanço inexorável da IAG e da IA-gen. Não obstante este avanço e sua visibilidade, o *status quo* conservador da IPF opera como se não houvesse um elefante na sala. Mas esta atitude sobranceira parece ridícula e desfasada da realidade.

A tendência geral de não poucos filósofos em ignorar ou desvalorizar as capacidades da IAG é recorrente. Contra a ameaça das IA-gen, Noam Chomsky, escreveu o seguinte:

99

O ChatGPT da OpenAI, o Bard [agora Gemini] da Google o Sydney [agora Copilot] da Microsoft são maravilhas da aprendizagem automática. Em termos gerais, pegam em grandes quantidades de dados, procuram padrões e tornam-se cada vez mais competentes na geração de resultados estatisticamente prováveis - como linguagem e pensamento aparentemente humanos. Estes programas têm sido aclamados como os primeiros vislumbres no horizonte da inteligência artificial geral - aquele momento há muito profetizado em que as mentes mecânicas ultrapassam os cérebros humanos, não só quantitativamente em termos de velocidade de processamento e tamanho da memória, mas também *qualitativamente* em termos de perspicácia intelectual, *criatividade artística e todas as outras faculdades distintamente humanas*. (...) Esse dia *pode* chegar, mas a sua aurora ainda não está a despontar. (Chomsky, 2023, itálicos meus).¹⁵

Para Chomsky, esta tecnologia não poderá criar informação considerada científica ou filosófica. Ele não se refere diretamente à filosofia, mas subentende-se daquilo que escreve a extensão a ela. Entre os pensadores que defendem que ainda não chegámos a um ponto crítico

¹⁴ Sobre idiotas inteligentes, humanos, desprovidos de inteligência emocional ver Goleman (2011, p. 26). Um projeto de criar consciência e inteligência emocional em cérebros sintéticos parece ser proposto por Cardon (2006).

¹⁵ Tradução minha a partir do inglês, como todas as restantes neste artigo.

em que artefactos computacionais conseguirão superar o ser humano encontram-se, por exemplo, Floridi (2015), Bishop (2021), Landgrebe *et al.* (2023). Eles são cétricos quanto à possibilidade de haver tecnologia tão inteligente e criativa quanto os seres humanos. Estes filósofos e cientistas são talvez demasiado otimistas quando presumem que os seres humanos não serão ultrapassados nestes e noutros quesitos a curto ou médio prazo. Entre os proponentes do contrário, em especial da chamada ideia de singularidade tecnológica – o momento do tempo futuro em que as IAG irão tornar-se tão ou mais inteligentes do que os humanos – estão Kurzweil (2005), Bostrom (2014) e Broderick (2021). Estas opiniões levam-nos a considerar a urgência do tópico para todas as áreas do saber, principalmente, para o que devemos fazer para desenvolver uma IAG responsável (Dignum, 2019, p. 93; Floridi *et al.* 2020), benigna e sob controlo de organismos humanos democráticos.

A miopia conceitual fruto de uma certa ingenuidade humana impede-nos na IPF de ver o elefante na sala que ameaça partir a porcelana da filosofia, protelando irremediavelmente uma reestruturação necessária.

É a IAG atual capaz de filosofar? Devemos perguntá-lo a filósofos, claro. Curiosamente, temos muita filosofia da inteligência artificial, mas quase nada pode ser encontrado sobre se a inteligência artificial consegue ou não fazer filosofia. Isto é sintomático de que os filósofos levam pouco a sério este perigo iminente.

100

A exceção parece ser um pequeno opúsculo onde o autor pergunta:

(...) coloca-se aqui uma questão fundamental. Estará esta IA filosófica a realizar o verdadeiro trabalho da filosofia? Se tudo o que faz é descobrir padrões por descobrir em dados introduzidos externamente, ou fornecer soluções para questões colocadas por humanos, não pode ser chamada de filosofia (Morioka, 2021, p. 40).

Esta colocação é bizarra. Mas Morioka depois finaliza o seu artigo dizendo:

A evolução da IA está agora a obrigar-nos a repensar a nossa visão da própria inteligência. É claro que as questões filosóficas intrínsecas da IA são tão estranhas que podem não ter qualquer ressonância nas nossas mentes. Nesse caso, se fosse possível estabelecer um diálogo entre os humanos e a IA sobre este ponto, abrir-se-ia certamente uma nova dimensão para a filosofia. (*ibidem*).

Então, dada esta oscilação, por que não perguntar a seguinte questão ao ChatGPT 4.0 da OpenAI: podes criar filosofia? Eis a resposta obtida da IA-gen mais famosa do mercado na atualidade:

Criar filosofia é um esforço profundamente pessoal e profundo. Embora eu [ChatGPT] possa certamente oferecer ideias e percepções, a verdadeira filosofia surge da contemplação individual, da reflexão e do envolvimento com o mundo. No entanto, posso certamente ajudar a guiá-lo através do processo ou discutir ideias filosóficas existentes, se isso for útil! Que área específica da filosofia está interessado em explorar ou discutir? (ChatGPT, consultado em maio de 2024).

Curioso como o artefacto foi condicionado pela sua aprendizagem de máquina para providenciar uma resposta na primeira pessoa, contribuindo isso, bem o sabem os construtores, para estabelecer um sentimento de intimidade com o interlocutor, como se de uma conversa entre seres humanos se tratasse. Note-se que o sistema não nega nem afirma o que foi questionado, apenas se dispendo a coadjuvar na feitura de filosofia *qua* ferramenta. Isto não milita contra o nosso ponto, apoia-o até bastante.

Com base nos materiais dispostos até este momento, erguem-se três hipóteses de trabalho. A primeira é a de que as IA-gen poderão realmente *instanciar* inteligência, consciência e criatividade iguais ou superiores às humanas.¹⁶ As consequências da eventual implementação desta possibilidade ainda são difíceis de antever, mas por exemplo a opacidade dos sistemas IA-gen coloca problemas cada vez mais complexos; pois esta estagnidade e a imunização das redes neurais e dos algoritmos à nossa compreensão – à dos próprios programadores da tecnologia – milita a favor de um desconhecimento sistémico do que as IA-gen realmente fazem, “querem” fazer ou irão fazer. Trata-se para muitos, não todos, de questões ainda em aberto e assim provavelmente ficarão durante mais algum tempo.

Existem, no entanto, opiniões igualmente competentes sobre esta possibilidade física e tecnológica da instanciação (e.g., Searle 1980, 1992, 2014; Bishop 2021, Landgrebe 2023). Alguns poderão pensar que uma resposta negativa encerra a questão sobre a instanciação das IA-gen produzirem filosofia de nível humano. Não nos parece que esta disputa seja fácil de resolver. Mesmo não existindo instanciação de inteligência pela IA-gen, estas poderão compensar esse *handicap* com um cada vez maior poder de computação neural-probabilística. Esta é, aliás, a segunda hipótese. Ela diz-nos que se as IAG se tornarem suficientemente *fortes* – mesmo continuando fracas no sentido de não possuírem consciência, introspecção, compreensão semântica etc., no sentido de poderem não só escolher seletivamente entre informação previamente adquirida, mas também de produzirem combinações criativas, *emulando* assim quase perfeitamente a criatividade humana, então é possível que essa mimese

¹⁶ A inteligência é classicamente definida como a capacidade de responder a questões e resolver problemas *etc cetera*. Uma comparação da inteligência biológica e da artificial surge em Watanabe (2022).

sofisticada do pensamento e da inteligência humanos possa permitir formas não humanas de fazer filosofia.

A terceira hipótese de trabalho é a referida pelo próprio ChatGPT (ver acima), e por defensores da admissão das IA-gen no processo académico: assistência e cooperação. Estes sistemas não instanciam nem emulam filosofia, mas tornam-se poderosos instrumentos coadjuvantes do processo filosófico. Esta hipótese pode parecer *prima facie* atraente para muitos adeptos da IAG, mas não está isenta de complicações. Vejamos, a título de exemplo, a opinião de um conhecido partisan da IAG e da singularidade, Nick Bostrom. Afirma ele que vivemos uma época de transição e substituição gradual de agentes cognitivos humanos pela IAG e seus subderivados. Muitas tarefas, vivências e experiências até agora entregues a seres humanos, por exemplo na educação, economia, medicina ou segurança, transitam velozmente para o âmbito da IAG. O filósofo preconiza que

As mentes digitais, sejam elas IAs ou uploads, precisam de computação. Mais computação significa uma vida mais longa, pensamento mais rápido e experiências conscientes potencialmente mais profundas e expansivas. Mais computação também significa mais cópias, filhos digitais e descendentes de todos os tipos, se assim o desejarmos. (Bostrom, 2024, Secção 5).

Este é um processo de conversão e *trans-humanização* do *Sapiens* por via da tecnologia e da alegadamente necessária rendição a ela dos humanos. Curiosamente, a ideia é fomentada pelas grandes corporações tecnológicas da área da IAG e áreas associadas, cujos aparelhos de publicidade e marketing se impõem em grandes sistemas de repasse de informação deliberadamente influenciadora, desde a *media* tradicional às redes sociais. Esta informação ideologicamente enviesada das corporações raramente verdadeira é muitas vezes aceite acriticamente pela sociedade. Mais, uma fatia considerável das atuais comunidades de geração de conhecimento vive num bizarro estado de letargia. Encontramo-nos, conceitualmente, numa terra de ninguém, um território ainda selvagem, porque em larga medida não regulamentado. Trata-se de um período de transição em que muitos especialistas vaticinam futuros sombrios para a humanidade, enquanto outros parecem adotar uma postura demasiado condescendente e otimista em relação aos poderes da IAG. Estes últimos comparam-se aos pioneiros que buscavam ouro e propriedades em terras selvagens de outras eras. O mais risível sucede quando, hipocritamente, sugerem regulamentação para as suas criações, não deixando de as evoluir. A mentalidade, pouco ou nada ética, é simples: se não for eu a fazê-lo, outros o farão. Então, já que vai acontecer, quer eu queira quer não, é melhor ser eu a fazê-lo. E continuam: enquanto o

evento de nível de extinção da humanidade que vaticino não suceder, vou (as corporações) lucrando. Um cinismo supremo que provoca náusea.

Excluindo os fatores lucro e prestígio da apresentação, compare-se agora a visão progressista de Bostrom com a perspectiva antagónica de uma individualidade académica de grande prestígio e provas dadas no ramo da neurociência a respeito das potencialidades da IAG. Nicolelis famosamente defende (ver a nota de rodapé 17) a falta de criatividade dos sistemas digitais, seguindo a velha rábula *serliana* antissimbolismo computacional, alegando que entre zeros e uns não há uma continuidade, também alegadamente necessária, segundo o autor, para a semântica consciente, o pensamento reflexivo e a criatividade. Reitera que ao contrário do que acontece no “sistema analógico” do cérebro humano, máquinas não pensam – emulam, mal, o pensamento humano. Segundo o neurocientista, portanto, as máquinas não substituirão o *Homo sapiens* porque lhes falta, entre outras coisas, a criatividade humana, a qual não pode ser dada por mecanismos que funcionam num pressuposto matemático. Para piorar o panorama atual, ainda segundo Nicolelis as IA-gen e tecnologias facilitadoras terão o efeito adverso de fazer regredir a inteligência humana devido a, digamos, falta de uso e esforço no desempenho cerebral (humano, claro) de tarefas de raciocínio e compreensão básicas. Resumindo, corremos o risco de ficar mais estúpidos e obtusos se continuarmos a depender destas tecnologias para tarefas triviais.

Mas a visão *singularista* e aflita de Bostrom não se desmonta facilmente. Segundo Max Tegmark¹⁷, o progresso tecnológico nas áreas das IAG e a sua curva evolutiva são tão exponencialmente acelerados na atualidade que um cataclismo de nível de extinção do *Homo sapiens* ou da sua razão de ser está cada vez mais próximo – se não está já em marcha! – indo inevitavelmente suceder caso não regulamentemos esta tecnologia. Sem dúvida que a comunidade de *superstars* académica está cada vez mais bipolar e dividida a respeito da IAG. Entretanto, é inegável que as capacidades desta tecnologia continuam a crescer e, para o melhor e para o pior, invadem cada vez mais territórios dos *Sapiens* e do seu quotidiano. A IPF não é, muito menos será, exceção. Os decisores devem estar mais do que nunca em alerta e devem

¹⁷ Novamente, a visão *singularista* é a perspectiva de que num momento do tempo (futuro) as máquinas alcançarão um nível de inteligência igual ou até muito superior aos dos humanos. A ideia parecer remontar a Good, (1966), sendo apoiada cinquenta anos depois por Chalmers (2016). A visão de Nicolelis surge vários lugares do YouTube em vários momentos o neurocientista afirma que há um culto antigo da IAG, uma seita que eleva excessivamente o valor funcional e objetivo desta tecnologia.

<https://www.youtube.com/watch?v=BewJAc093m4> e

<https://www.youtube.com/watch?v=C9cWBaY3ZOk&t=25s>.

A visão antagónica Tegmark surge também em:

https://www.youtube.com/watch?v=-Xdkzi8H_o e https://www.youtube.com/watch?v=xUNx_PxNHrY

decidir rapidamente o que fazer para proteger a filosofia, bem como todas as outras atividades humanas, claro.¹⁸

4. *Quo vadis* IPF?

As hipóteses elencadas na secção anterior levantam cenários preocupantes para todas as indústrias do conhecimento. A IPF não é excepção. Talvez esta indústria devesse abandonar a impressão coletiva, tácita e inviolável até agora, já para não dizer indiferente, de que o estado de coisas vai manter-se inalterado apesar dos impactos causados pelo contacto com a IAG, em especial a IA-gen. Com efeito, este impacto não pode ser tomado de ânimo leve pela IPF, seus decisores e sua comunidade.

Se a primeira hipótese vingar, as IA-gen com competências filosóficas irão tornar-se sérios concorrentes dos filósofos humanos, ameaçando a obsolescência destes últimos. O problema não será apenas autoral relativamente à pesquisa ou à produção. O problema será a dificuldade de os filósofos humanos concorrerem com “filósofos” do tipo IA-gen poderosas focadas em filosofia. Impossível, dirão os descrentes. Inviável, dirão os mais teimosos. Não deixaremos acontecer, bradarão os mais otimistas.

Esta indignação não é todavia apoiada pela história recente. Sabemos bem que tempos houve não muito apartados dos nossos em que os especialistas humanos na nobre arte de jogar xadrez duvidavam seriamente da capacidade de máquinas, mesmo as mais competentes, desenvolverem e replicarem o tipo de jogo produzido por campeões mundiais da especialidade. Era asseverado nesses tempos, décadas de 50, 60 e 70 do Século 20, que a combinação de *hardwares* e *softwares* (dois termos agora em desuso) só conseguiria usar força bruta no jogo, que lhes faltava a intuição e a criatividade dos jogadores humanos, e que dificilmente produziriam xadrez de qualidade superior. Então, continuavam os céticos, os computadores não jogavam realmente xadrez, nem compreendiam o jogo de forma consciente. O máximo que as máquinas podiam alcançar, supunha-se, era uma replicação imprecisa e tosca do desempenho humano, sempre concretizada através de manipulação de símbolos e algoritmos (o paradigma simbólico da IAG nesses tempos). Quando em 1996 o *Deep Blue* da IBM derrotou Garry Kasparov, um dos melhores jogadores de todos os tempos, muito do mito da insuficiência das máquinas face à inteligência humana caiu por terra (Newborn, 1997, p. 235).

¹⁸ Questão do poder dos filósofos para o fazer ou não já é outra história, uma que não abordarei neste artigo.

Algo semelhante a esse mito continua a ser insinuado acerca da incapacidade da IAG em produzir ou reproduzir certas atividades intelectuais. Mas agora, em pleno 2024, desconfiamos cada vez mais destas mitologias. Desconfiamos por causa deste e de vários outros casos bem conhecidos que a IAG consegue já superar em muito os desempenhos humanos. Por exemplo, no jogo Go ou na criação de imagem e vídeo sofisticados, *deefakes* ou não. Os arautos da “boa” hecatombe (Chalmers, 2016) preconizam agora, fazendo induções muito extremistas contra a ideia de alguns, como Nicolelis (2023, *online*) e Watanabe (2022, p. 94), para quem a complexidade do cérebro humano e alguns dos seus feitos, nomeadamente a consciência, não podem ser replicados *simpliciter*, que a superação dos dotes intelectuais e cognitivos dos humanos é inevitável. Se esses arautos da singularidade estiverem corretos, havendo várias opiniões clássicas nesse sentido (Vinge, 1993; Kurzweil, 2005), a IPF humana poderá ceder em breve o seu lugar a duas IPF: uma humana e uma assente em inteligência artificial. Isto é o que acontece aliás atualmente nas competições de xadrez, havendo campeonatos para humanos e campeonatos para máquinas.

Não interessa o que fizeram no xadrez, dirão alguns, pois o que interessa é o campeonato e a IPF humanos. O problema com este raciocínio de demissão liminar de uma IPF de IAG é o de que nenhum humano consegue superar atualmente as IAG no xadrez e em muitas outras atividades que requerem processamento avançado, probabilístico e seletivo de informação, como também é aliás o caso da filosofia e seus jogos. Será que as mentes humanas, individuais ou coletivas, poderiam superar vis-à-vis as IAG financiadoras de inteligência no jogo da produção filosófica? Gostaríamos de pensar que sim. Porém, se a analogia entre vários jogos e o jogo filosófico, incluídos o jogo de pesquisa e de ensino, estiver em ordem, então a possibilidade de superação por parte da IAG de especialistas humanos em jogos de racionalidade não é pouca. Opiniões que se estendam para lá desta suposição já se afiguram mais como ficção científica do que como especulações filosóficas. Por conseguinte, não é prudente *ex ante* dizer mais sobre este cenário.

Já a hipótese que coloca o cenário da *emulação* do pensamento filosófico por parte das IAG não é muito mais tranquilizante. Uma IAF fraca, no sentido *searliano* de não possuir consciência, mas ainda assim forte em termos de poder computacional, de acesso a dados definidos em redes neurais e de acesso probabilístico, poderá emular tão adequadamente o pensamento filosófico que se tornará quase impossível para os humanos (mas talvez não para a IAG) discernirem produções humanas de produções da IAG, como aliás já acontece agora com as IA-gen de produção de imagem e vídeo. Este cenário da emulação variaria talvez do anterior

nisto: os humanos não teriam de reconhecer a legitimidade da filosofia gerada por emulação, pois isso seria algo como reconhecer plágio de humanos por humanos. Claro, há plágios extremamente competentes, e a admissão de uma IAG fraca em termos cognitivos, mas muito forte em termos de processamento inconsciente de informação, uma possibilidade cada vez mais à porta, iria ter impactos profundos na IPF. Aqui entrariam os problemas éticos da legitimidade, dos direitos de autor (Debouche, 2021) e do mérito dos filósofos humanos por comparação com outros humanos e com a própria IAG. Além destes problemas, teríamos outros do ponto de vista socioeconómico e de justiça social. Quem teria direito ao uso de IAG na IPF e para quê? Paridade e imparidade de acesso: poderiam as instituições de ensino privadas ter melhor IAG que as públicas e *vice-versa*? Quem poderia aceder as IAG mais capacitadas? Quais os critérios? Parece-nos que cenários da emulação competente de pesquisa, produção e ensino levados a cabo pela IAG deixam a IPF *humana* em tão maus lençóis quanto o cenário da instanciação.

Sobre o cenário de uma IA-gen cooperadora e assistente dos elementos da IPF, conforme proposto acima pelo ChatGpt. Sal Kahn, um conhecido adepto humano desta possibilidade, escreve o seguinte:

106

Tendo integrado assistentes de ensino com IA na nossa plataforma, a nossa equipa aprendeu que os grandes modelos linguísticos podem tornar o ensino uma profissão mais sustentável. Imaginemos que o distrito escolar local descobre subitamente centenas de milhões de dólares e os utiliza para oferecer a cada professor o apoio de três assistentes brilhantes para as suas salas de aula. Estes assistentes ajudariam a criar planos de aula e rubricas, a classificar trabalhos, a redigir relatórios de progresso, a falar com os professores e a apoiar os alunos. Todos os professores do planeta agarrariam de imediato esta oportunidade. Estes assistentes não ameaçam os empregos de professor, mas tornariam de facto os empregos de professor sustentáveis. Tornariam o trabalho mais agradável. Mais importante ainda, ajudariam a acelerar os resultados de aprendizagem de milhões de alunos, tornando-os mais preparados, para as carreiras e para a faculdade, para as carreiras e para a vida (...) A IA generativa pode fornecer aos educadores novas melhores práticas, técnicas de ensino e conhecimentos sobre as lacunas de aprendizagem dos seus alunos. Não é difícil para o assistente de ensino de IA identificar uma área problemática para um grupo de alunos e oferecer planos de aula correspondentes para o professor utilizar, ou para o assistente de ensino de IA monitorizar o desempenho de um aluno na aula e enviar actualizações de progresso em tempo real aos professores. Até tem a capacidade de atuar como conselheiro para professores que lutam contra o esgotamento. Estas funções estão a tornar-se mais comuns e, com as devidas salvaguardas, podem ser bastante poderosas. (Kahn, 2024, p. 53).

Num certo sentido, muito disto já acontece na IPF e similares, mas de forma velada: nem professores, nem pesquisadores, nem discentes confessam esta dependência crescente. Vivemos, nesse aspecto, numa terra de ninguém, em que nada vale e tudo parece valer.

Mas mais relevante mesmo é o “treino” e a aprendizagem profunda que as IA-gen terão de possuir para serem bons assistentes educativos e de pesquisa, principalmente na IPF. Kahn alardeia o facto de que se pôs na pele de um estudante e, experimentando, perguntou à sua IA-gen favorita, obteve um resultado extraordinário em termos de postura da sua IA-gen favorita. O empresário formado pelo MIT perguntou o seguinte:

Porque é que temos [nos EUA] a Segunda Emenda? Parece uma loucura! Khanmigo [Uma IA-gen ao serviço da Academia Kahn] respondeu: "Para começar, porque é que achas que os Fundadores incluíram a Segunda Emenda?" Repara [diz ainda Kahn] que não refutou nem reforçou a opinião do nosso aluno imaginário, mas desafiou-o a pensar mais profundamente sobre o assunto. (*ibidem*).

Mas é particularmente difícil de descortinar que tipos de inovação poderiam a IAG e suas derivadas acarretar para a IPF. Inovar intencionalmente é, por defeito, uma função da imaginação humana. *Pace* Turing, a grande questão não é mais se as máquinas podem pensar, isto é, se são inteligentes. Chegámos a um ponto da história em que temos de verificar se máquinas sofisticadas, baseadas em redes neurais por camadas, podem imaginar e inovar. Se não puderem, para que precisamos delas na IPF? Precisamos tanto quanto precisamos de papagaios repetidores (tanto aves como *Homo sapiens* repetidores), ou seja, pouco ou nada. Mais preocupante ainda, devemos perguntar-nos se a atividade “pedagógica” das máquinas poderia contribuir para mitigar o juízo teórico e prático humano, como parecem defender Eisikovits *et al.* (2022).

107

Relembrando as palavras sábias de Bertrand Russell¹⁹, quem faz filosofia deve levar em conta dois princípios fundamentais, um intelectual e um moral. O intelectual é o de que o filósofo deve considerar primeiro que tudo os factos. Pois então os factos estão aí à vista de todos. O primeiro facto é o de que a IPF apresenta atualmente uma disposição que privilegia uma exigência de tipo fabril e um tipo de competição pouco compatíveis com a própria arte de filosofar, o que por sua vez abre as portas desta indústria a dispositivos com grande capacidade de recolha, processamento e seleção de informação. O segundo facto é que o primeiro ameaça inquinir irremediavelmente a IPF humana e a filosofia, também humana. O terceiro facto é o de que, se nada for feito, provavelmente os dois primeiros factos causarão a obsolescência de filósofos humanos. O quarto facto é o de que estamos a atrasar-nos na regulamentação da IPF para lidar com a IAG.

¹⁹ *Message to the future generations*, Bertrand Russell (entrevista consultada em abril 2024)
<https://www.youtube.com/watch?v=ihaB8AFOhZo>

A reflexão conduz-nos então ao segundo princípio aludido por Russell, o desiderato da tolerância. Russell referia-se a relações humanas, não a relações máquina humano. Claro, tendemos a aceitar que o problema da tolerância se aplica a seres inteligentes. E uma vez que a inteligência artificial *fraca-forte* está a deixar de ser gradualmente um mero constructo fantasioso pertencente à ficção científica e mais uma realidade, resta perceber se estamos dispostos a assumir esta tolerância para com as IAG. Talvez ainda mais importante para nós, irão esses dispositivos artificiais tolerar-nos a partir de um determinado grau do seu desenvolvimento? Não sabemos responder. Dificilmente alguém saberá nesta altura. Seja como for, por certo muito mudará na IPF a breve trecho. Novamente propomos: temos de perder as ilusões de domínio e inalterabilidade que teimam em persistir na IPF. Temos que, citando o famoso estadista, esperar o melhor mas preparar-nos para as grandes dificuldades que nos esperam. Não podemos ser excessivamente pessimistas ou otimistas. A este respeito, temos de viver na tensão entre o conservadorismo e o progressismo. Como é óbvio, isso obriga-nos a suspender o juízo quanto ao futuro, sendo esse o único posicionamento sensato nesta fase de desenvolvimento da inteligência artificial.

Referências:

ABOULAFIA M. The Productivity Syndrome (or why I stopped writing philosophy), <https://upnight.com/2018/05/31/the-productivity-syndrome-or-why-i-stopped-writing-philosophy>.

BARRAT, J. Our final invention: artificial intelligence and the end of the human era. New York, St. Martin's Press, 2013.

BISHOP, J. Artificial Intelligence Is Stupid and Causal Reasoning Will Not Fix It. In: *Frontiers in Psychology*, 11. 2021., 1-18 online.

BORNMANN, L., HAUNSCHILD, L., MUTZ, R. Growth rates of modern science: a latent piecewise growth curve approach to model publication numbers from established and new literature databases. In *Humanit Soc Sci*. 224, 2021. doi.org/10.1057/s41599-021-00903.

BOSTROM, N. *Deep utopia, life and meaning in a solved world*. Washington DC. Ideapress Publishing. 2024.

BOSTROM, N. *Superintelligence: Paths, Dangers, Strategies*. Oxford. Oxford University Press. 2014.

BRODERICK, D. *The Spike: how our lives are being transformed by rapidly advancing technologies*. New York, A Forge Book - Tom Doherty Associates, LLC. 2021.

CARDON, A. Artificial consciousness, artificial emotions, and autonomous robots. *Cogn Process.* 7(4), 2006, P245-267.

CHALMERS, David, The Singularity: A Philosophical Analysis. in “Could artificial intelligence really out-think us (and would we want it to)? In Uziel Awret, David Chalmers, eds., *The Singularity Could artificial intelligence really out-think us (and would we want it to)?* Imprint Academic Ltd., 3-32. 2016.

CHOMSKY, N. The False Promise of ChatGPT. In: *The New York times.* 2023, march 8).

DEBOUCHE, N. Plagiarism in the age of massive Generative Pre-trained Transformers (GPT-3). In: *Ethics in Ethics science and environmental politics.* Vol. 21, 2021. P. 17–23.

DIGNUM, V. Responsible Artificial Intelligence: How to Develop and Use AI in a Responsible Way. Springer. 2019.

DOMINGOS, Pedro. *The master algorithm: how the quest for the ultimate learning machine will remake our world.* Filadelfia, Perseus Group. 2015.

EISIKOVITS. N., FELDMAN, D. IN: *Moral Philosophy and Politics.* 9. 2022, p. 181–199.

109

FLORIDI, L. Devemos ter medo da inteligência artificial? [https://aeon.co/essays/true-ai-is-both-logically-possible-and-utterly-implausible.](https://aeon.co/essays/true-ai-is-both-logically-possible-and-utterly-implausible) 2015.

FLORIDI, L., COWLS, J., KING, T., TADDEO, M. How to design AI for social good: seven essential factors. *Science and Engineering Ethics,* 26, 2020.

FORTUNATO, S., BERGSTROM C., BORNER, K., EVANS, J., HELBING, D., MILOSEVIC, S., PETERSEN, A., RADICCHI, F., SINATRA, R., UZZI, B., VESPIGNANI, A., WALTMAN, L., WANG, D., BARABÁSI, A., Science of science. In: *Science.* 2018. doi: 10.1126/science.aao0185. PMID: 29496846; PMCID: PMC5949209.

GARFIELD, E. What is the primordial reference for the phrase 'Publish or perish'? *The Scientist* Vol.10. 1996.

GOLEMAN, D. *Inteligência emocional, a teoria revolucionária que redefine o que é ser inteligente.* tradução Marcos Santarrita. Rio de Janeiro. Objetiva, 2011.

GOOD, John. Speculations concerning the first ultraintelligent machine. In F. Alt & M. Ruminoff (eds.), *Advances in Computers,* volume 6. Academic Press. 31-88. 1966.

HARTMAN, R. The Logical Difference Between Philosophy and Science. In: *Philosophy and Phenomenological Research*, 1963, Vol. 23, No. 3. 1963, p. 353-379.

HORKHEIMER, M., ADORNO T. *The dialectic of Enlightenment - Philosophical Fragments*. Edited by Gunzelin Schmid. Stanford. Stanford University Press. 2002.

INNES. C. *The Cambridge Companion to Bernard Shaw*. Cambridge. Cambridge University Press. 1998.

KAHN, S. *Brave New Words, How AI Will Revolutionize Education (and Why That's a Good Thing)*. New York. Viking. 2024.

KUHN, T. *The Essential Tension: Selected Studies in Scientific Tradition and Change*. Chicago. The University of Chicago Press. 1977.

KURZWEIL., R. *The Singularity is Near*. New York: Viking Press, 2005.

LANDGREBE, Jobst; SMITH, Barry. *Why machines will never rule the world: artificial intelligence without fear*. New York. Routledge, 2023.

MAY. T. Transformations in Academic Production, Content, Context and Consequence In: **110**
European Journal of Social Theory. 8(2). 2005. p 193–209.

MINSKY, M. Will Robots Inherit the Earth? *Scientific American*, Oct, 1994.
<http://www.inf.ufsc.br/~mauro.roisenberg/ine6102/leituras/Will%20Robots%20Inherit%20the%20Earth.htm>. Transcrição consultada em 20/4/2024.

MORIOKA, M. Can Artificial Intelligence Philosophize? In: *The Review of Life Studies*, Vol.12, 2021. p. 40-41.

NASSI-CALÒ. L. Como se utilizam as altmetrias para avaliar a produção científica da América Latina. 2022. <https://blog.scielo.org/blog/2022/02/09/como-se-utilizam-as-altmetrias-para-avaliar-a-producao-cientifica-da-america-latina>. Consultado em 2024.

NEWBORN, M. *Kasparov versus Deep Blue: Computer Chess Comes of Age*. New York. Springer. 1997.

PAULUS, F., CRUZ, N., SOREN, K. The Impact Factor Fallacy. In: *Conceptual Analysis*, 2018. doi: 10.3389/fpsyg.2018.01487.

PRICE, D. *Little Science, Big Science... and Beyond*. New York. Columbia University Press. 1986.

RUSSELL, B. *O elogio do ócio*, Tradução de Pedro Júnior. Rio de Janeiro, (1935) 2002.

SAYGIN, A., CICEKLI, I., AKMAN, V. Turing test: 50 years later. *Minds and Machines*. 10 (4):463-518. 2000.

SEARLE, J. *The rediscovery of the mind*. Cambridge, MA: MIT Press.

SEARLE, J. Minds, Brains and Programs. *Behavioral and Brain Sciences*. 1980. 3: 417–424.

SEARLE, J. What Your Computer Can't Know. New York. *Review of Books*, October 9. 2014.

STERGIOU, K., LESSENICH, S. On impact factors and university rankings: from birth to boycott, In: *Ethics in science and environmental politics*. Vol. 13 pp6. 2013 P. 1-11. doi: 10.3354/ese00141.

TURING, A. Computing machinery and intelligence. *Mind* 59. 1950. p. 433-460.

VAASEN, B. AI, Opacity, and Personal Autonomy. In: *Philosophy & Technology*. P. 35-88. doi.org/10.1007/s13347-022-00577-5.

VINGE, V. The coming technological singularity. Online. 1993.

111

WATANABE, M. *From Biological to Artificial Consciousness, Neuroscientific Insights and Progress*. Springer, 2022.

Relatórios.

AI4People. Ethical framework for a good society: opportunities, risks, principles, and recommendations. Atomium – European Institute for Science, Media and Democracy. <http://www.eismd.eu/wp-content/uploads/2019/02/Ethical-Framework-for-a-Good-AISociety.pdf>. Acessado em abril de 2024.

BORI 2022: um ano de queda na produção científica para 23 países, inclusive o Brasil. 2022. Elsevier. Acessado em março 2024. <https://abori.com.br/relatorios/2022-um-ano-de-queda-na-producao-cientifica-para-23-paises-inclusive-o-brasil>. Consultado em agosto de 2024.

Future of Life Institute. Ansilomar AI Principles. <https://futur.eolife.org/ai-principles/>. 2019. Acessado em Maio 2024.

<https://ai4people.org>. Acessado em Abril 2024.

Programas associados à ANPOF. <https://www.anpof.org.br/anpof/programas-associados>. Consultado em fevereiro 2024.

<i>Revista Dialectus</i>	Ano 13	n. 33	Edição Especial	p. 90 - 112
--------------------------	--------	-------	-----------------	-------------

The Declaration on Research Assessment (DORA) recognizes the need to improve the ways in which researchers and the outputs of scholarly research are evaluated. <https://sfedora.org>. Acessado abril 2024.

The Toronto declaration: Protecting the right to equality and non-discrimination in machine learning systems. <https://www.torontodeclaration.org/>. 2018. Acessado em maio 2024.