A Era da Inteligência Artificial
e o nosso futuro humano

Armando Malheiro da Silva

Faculdade de Letras da Universidade do Porto/CITCEM
armando.malheiro@gmail.com



Estamos a viver tempos de transição. Isto é certo e não é inédito. Os contemporâneos das revoluções industrial, americana e francesa, ao longo do séc. XVIII, sentiram a mesma sensação de forte "aragem de mudança" e de inquietação face ao inusitado. Atravessado o turbilhão dos acontecimentos, caberia aos historiadores arrumar cronologicamente (tarefa para eles natural e aparentemente simples) essas mudanças, consideradas profundas, em épocas, designadas de "idades" – a Idade Média, a Moderna e a Contemporânea. Esta última abrangia uma espécie de presente extenso, pelo séc. XX dentro.

Outros olhares, porém, pousaram no devir social e o conceito de Modernidade ganhou maior amplitude, abarcando desde o Renascimento e as Descobertas até meados do séc. XX. Fim de longo ciclo e início de...? Filósofos, sociólogos, ensaístas e "artífices" das Humanidades, atentos ao impacto da política, da economia e dos fatores sociais e mentais nas artes e nas letras, expressaram nova mudança e nova periodização de forma tão sugestiva, como vaga - a Pós-Modernidade. Na década de setenta e seguintes de novecentos anuncia-se mais do que um novo ciclo ou Era; anuncia-se um estado "revolucionário" de atos e de intenções. Considerando-a em retrospetiva percebe-se que correspondeu à necessidade de vazio que urgia periodizar, tanto mais que uma poderosa revolução tecnológica (muito para lá de técnica) irrompera no pós-guerra e que carecia de caracterização – sociedade pós-industrial, revolução informacional, sociedade da informação... As expressões foram-se sucedendo, mas de recorte mais sociológico que histórico. E de facto, parece, até hoje, que os historiadores não se atrevem a nomear declarativa e assertivamente o decurso de um novo período, de uma nova Era, para a qual, no entanto, circulam substantivos e adjetivos – digital, informação e, muito recentemente, inteligência artificial!...

Três personalidades de proa no universo político-diplomático e tecnológico subscreveram implicitamente a proposta nominativa para o novo ciclo estrutural já em curso – a Era da Inteligência Artificial. Henry A. Kissinger (56º Secretário de Estados dos EUA entre 1973 e 1977 e prémio Nobel da Paz em 1973), Eric Schmidt (tecnólogo, empresário e filantropo, foi CEO da Google responsável pela transformação desta *startup* de Silicon Valley em empresa tecnológica global) e Daniel Huttenlocher (reitor inaugural do MIT Schwarzaman College of Computing e presidente do Conselho de Administração da John D. and Catherine T. MacArthur Foundation) associaram-se à escrita colaborativa para nos revelarem e apresentarem "a uma era que começa agora a despontar: a da coabitação entre humanos e a inteligência artificial" (na contracapa).

No Prefácio o projeto editorial é explicado de forma clara e sucinta:

> *Há quatro anos que nós, juntamente com Meredith Potter, que ajuda Kissinger nas suas pesquisas intelectuais, nos reunimos para a análise destas e outras questões, procurando abranger as oportunidades e os riscos colocados pela ascensão da IA. Em 2018 e 2019, Meredith ajudou-nos a alinhavar as nossas ideias em artigos, que nos convenceram de que devíamos – e de que com a sua ajuda permanente podíamos - expandi-los para escrever este livro.*
> *O derradeiro ano dos nossos encontros coincidiu com a pandemia da COVID 19, que nos obrigou a reunir-nos por videoconferência – outra tecnologia que era considerada extravagante ainda há pouco tempo, e agora é omnipresente. À medida que o mundo se ia fechando, sofrendo perdas e mudanças como só sofrera um século antes durante os períodos das guerras mundiais, as nossas reuniões semanais foram-se transformando num fórum dos atributos humanos que a IA não exerce: amizade, empatia, curiosidade, dúvida e preocupação (p.11).*

Aí é assumido que os Autores divergiam quanto ao otimismo com que encaravam a IA, mas convergiram "no sentido em que a tecnologia está a mudar o nosso pensamento, o conhecimento, a perceção e a realidade – e ao fazê-lo, está a mudar o curso da história humana" (p. 11). Concordam neste ponto essencial e isto leva-os a seguir uma via compreensiva, equidistante da celebração ou da rejeição da IA. Os Autores escreveram este livro para analisarem as consequências e implicações de uma tecnologia que já é omnipresente e em desenvolvimento acelerado.. E ao longo de sete capítulos (1 – Onde estamos; 2 –Como chegamos aqui – tecnologia e pensamento humano; 3 – De Turing aos dias de hoje – e mais além; 4 – Plataforma de rede globais; 5 – Segurança e ordem mundial; 6 . IA e identidade humana; 7 – IA e o futuro) perecem-se dois movimentos; um inicial de contextualização; e um segundo de problematização prospetiva.

Os três primeiros capítulos oferecem-nos uma síntese da realidade em que estamos imersos e em que a IA está operante somando sucessos e impressionando o seu criador – o Homem. A partir do Alpha Zero, programa de IA desenvolvido pela Google DeepMind que derrotou em maio de 2017 o Stockfish em jogo de xadrez, iniciando uma "revolução silenciosa", os programas e as proezas sucedem-se nas páginas do capitulo primeiro até ao ponto da constatação incontornável: "Até agora, a razão tem sido uma prerrogativa da humanidade e, desde o Iluminismo, o seu atributo definidor. O advento de máquinas de aliar uma aproximação da razão humana ao processamento refletir-se-á tanto nos humanos como nas máquinas (p. 31). Abre-se uma interação máquina-humano com virtualidades e riscos: "Ajudada pelos programas e difusão da IA da IA, a mente humana está a aceder a novas perspetivas, permitindo vislumbrar objetivos antes inatingíveis.. Estes incluem novos modelos com que prever e mitigar desastres, conhecimentos matemáticos mais profundos, e completo conhecimento do Universo e da realidade da sua envolvência" (p. 33). Mas todos estes novos horizontes tangíveis trazem consigo uma alteração da relação humana com a razão e a realidade e para isto não estamos nem filosófica, nem politicamente preparados.

No entanto, o que estamos a viver resulta de uma trajetória que tem algum lastro e isso é explicado no capitulo segundo, em que os desenvolvimentos filosóficos e científicos conduziram a um patamar tecnológico superador do plano estritamente técnico. Para melhor se compreender esse lastro é tecida uma exposição, neste capítulo, de teor filosófico e histórico-científico, que culmina na assunção de que nos encontramos num ponto de viragem, em que algumas das inovações obtidas não podem ser entendidas como extensões do conhecimento até agora consolidado: "Ao comprimir a medida de tempo em que a tecnologia altera a experiência da vida, a revolução do digital e o avanço da IA geraram fenómenos que são genuinamente novos, e não meras versões mais potentes ou eficazes de coisas passadas" (p. 55). E neste quadro novo o volume crescente e o acesso vertiginoso à informação surge como inovação pura e transformadora, sendo entendida deste modo, que nos parece ambíguo e simplista, por consequência frustrante: "Quando é contextualizada, a informação torna-se conhecimento. Quando suscita convicções, o conhecimento torna-se sabedoria. No entanto, a Internet inunda os utilizadores com as opiniões de milhares de pessoas, privando-os da solidão necessária à reflexão atenta, que foi, historicamente, aquilo que conduziu ao desenvolvimento de convicções" (p. 57).

O tópico da informação tem, dentro desta temática, uma particular importância e prende-se com o essencial do capitulo 3 do livro – "De Turing aos dias de hoje – e mais além". Inevitavelmente o matemático e descodificador inglês Alan Turing é evocado como referencia basilar, através da solução por ele proposta em 1950:

> *Num ensaio despretensiosamente intitulado "Máquinas de Computação e Inteligência", Turing propunha que se pusesse definitivamente de lado a questão da "inteligência" das máquinas. O importante, postulava Turing, não era o mecanismo, mas sim a manifestação de inteligência. Tendo em conta que as vidas interiores de outros seres continuam indecifráveis, explicava ele, a única forma de medirmos a inteligência terá de ser o comportamento exterior. Com este postulado, Turing marginalizava séculos de debates sobre a natureza da inteligência. O "Jogo da Imitação que introduziu defendia que que se alguma máquina operasse de forma tão proficiente que os observadores não conseguissem distinguir o seu comportamento do dos humanos, então a máquina devia ser considerada "inteligente"!. Nasceu o "teste de Turing" (pp. 59-60).*

Não nasceu, apenas, um "teste", nasceu um rumo que tornou possível a IA dos dias de hoje, como é mostrado nas páginas desse terceiro capítulo, mas nsceu, também, um reducionismo lógico-prático que produziu avanços tecnológicos surpreendentes e que continuam a evoluir, inutilizando o debate conceitual sobre inteligência, cognição e consciência no seio dos laboratórios onde essas "máquinas inteligentes" são criadas e aperfeiçoadas. Ora, uma óbvia vantagem com benefícios práticos tem sido obtida de forma unidimensional, desprezando a complexidade integral. Esta conceção unidimensional está patente ao longo do referido terceiro capítulo e é evidente no último ítem intitulado "O sonho da inteligência geral artificial", onde se afirma que a IA transformará as nossas vidas e o nosso futuro, alimentando-se de quantidades colossais de dados e de informação (entendidos ambos os termos num sentido mecanicista e fisicamente mensurável), mas onde, também, se reconhece que "Devemos todos garantir que a IA não seja criada isoladamente – e, consequentemente, ter em atenção tanto os seus benefícios, como os seus potenciais riscos" (p. 95). Ou seja, a IA precisa ser controlada pelos humanos e para que este controlo seja considerado possível e eficaz está pressuposto, inevitavelmente, que a inteligência das máquinas gerível por uma "inteligência" forçosamente diferente e capaz dessa função gestionária.

Os Autores do livro deixaram-se cair numa contradição inevitável; por um lado aceitaram o reducionismo mecanicista e filosófico imposto, com visível sucesso tecnológico, por Turing, mas, por outro, admitem, implicitamente, que a IA tem de estar ontologicamente num patamar distinto da humana. Trata-se de uma contradição que não prejudica o propósito da obra, porquanto todos quantos desejem manter a indispensável sanidade crítica não podem ignorar que, no final, a IA será dirigida humanamente para a tragédia total, como poderá, humanamente, ser travada em favor de exigências éticas multiseculares.

Os Autores confiam, acima de tudo, na Humanidade e, por isso, orientam os restantes capítulos do livro no sentido de mostrar exemplos e colocar questionamentos que, partindo da evidência irreversível, de que estamos imersos numa realidade em que a IA é um elemento integrante e ativo, é possível preservar o livre arbítrio para a escolha de um futuro onde a espécie humana prossiga com a sua exclusiva e indeclinável responsabilidade na sobrevivência própria e de todo o planeta e tudo o que nele está e vive.

Este livro não foi escrito para ser compreendido pelas "máquinas de computação" previstas por Turing, mas ajudar os humanos a ajustarem-se a uma realidade que, como sempre, eles estão a criar e, como convém, saibam manter sob seu controlo pela positiva. Por exemplo e no que toca às chamadas "redes globais", os Autores pretendem salientar que "as plataformas de rede assentam cada vez mais em IA, dando origem a uma interação entre humanos e IA, cuja importância alcança o grau de marco civilizacional" (p. 99). No entanto, os operadores dessas redes são pessoas e a maioria delas originárias dos EUA, o que diz muito sobre o modo como essa interação é feita e desenvolvida. Daí a

importância em compreender essas plataformas de rede, como elas se mesclam com as comunidades e a vida quotidiana, como se articulam com as empresas e as nações, como elas se implicam no atualíssimo problema da "desinformação" (pp.118-121), como se encontram já envolvidas nos dilemas a enfrentar por governos e regiões, como se posicionam na definição atual e futura da geopolítica, enfim como compreender e encarar o futuro da Humanidade interligada às plataformas de rede. Vale a pena registar este extrato: "Os operadores de plataformas de rede defrontarão escolhas que ficam para além do serviço ao cliente e do êxito comercial. Até agora não foram obrigados a definir um objetivo "racional" ou de serviço que fosse para além da melhoria do produto, do mais largo alcance e de servir os interesses de acionistas e utilizadores. (..) Necessitarão não só de participar na definição da capacidade e dos objetivos derradeiros dos reinos virtuais que criaram, como terão ainda de contribuir para delinear os princípios segundo os quais interagem entre si e com outros sectores da sociedade" (pp.135-136)..

Diante da segurança e da ordem mundial, matéria de um outro extenso capitulo gizado, desta feita, por Henry Kissinger, os problemas, dilemas e desafios são postos e analisados tendo o foco na guerra cibernética, nas armas nucleares "inteligentes", no controlo de todo o armamento, nos conflitos digitais, na segurança convulsa em que vivemos, etc. A gestão da IA neste cenário desafiador tem de ser feita a partir de uma negociação humana vital, como se exemplifica: " A IA aumenta o risco de ataque preventivo e do escalar prematuro para o conflito. Um país que receie que outro esteja a desenvolver determinada capacidade poderá tentar prevenir-se contra ela: se esse ataque for "bem-sucedido", não haverá depois maneira de descobrir se este se justificava. As negociações devem concentrar-se em limitar esta corrida, mas também em garantir que cada lado tem conhecimento, em termos gerais, do que o outro está a fazer" (p. 166).

E chegamos, enfim, aos dois capítulos finais do livro que melhor explicitam a posição dos Autores e que, tendo em vista o escopo desta recensão, trazem para pauta as questões centrais do presente e futuro imediato: como se manterá a identidade humana sob o impacto cada vez mais profundo da IA e como esta continuará a evoluir pelo futuro afora (questão futurológica que os Autores tratam com muita contenção).

No capítulo 6 vale a pena destacar que, segundo os Autores, a colaboração entre a IA e as pessoas exigirá que estas se ajustem a "um mundo em que a sua razão deixa de ser a única via de conhecimento ou exploração da realidade. Isto prognostica uma viragem na experiência humana mais significativa do que qualquer outra ocorrida há quase seis séculos – desde o advento da imprensa" (p. 182). Em tal quadro as vias que se abrem são duas: as sociedades reagem e adaptam-se progressivamente à mudança ou as sociedades abrem um debate em que seja determinado o papel da IA e, consequentemente, o papel dos humanos. Um debate interdisciplinar e tão urgente, quanto difícil e complexo. E nos tópicos que se seguem os Autores encadeiam tópicos para esse debate: como pode a IA transformar positivamente a experiência humana em diversos domínios de atividade, com a advertência dos riscos associados; como a IA estará presente na tomada de decisões e isso terá tanto de promissor como de gerador de tensões com o entendimento humano; como ela se imiscuirá mais e mais nos processos de descoberta científica; como se tornará peça inseparável dos processos de educação e aprendizagem vitalícias; como pode contribuir para a produção de perceções da realidade diferentes das que os humanos poderiam ter imaginado; e como a IA tem condicionado cada vez mais o espaço informativo, surgindo novos intermediários da informação – e dentro destes tópico os Autores deixam impresso o seguinte:

> *Os efeitos da IA no conhecimento humano são paradoxais.*
> *Por um lado, as IA intermediárias conseguem navegar e analisar conjuntos de dados mais vastos do que uma mente humana entregue a si própria alguma vez poderia ambicionar.. Por outro lado, esse mesmo poder, , essa mesma capacidade de lidar com grandes quantidades de dados pode favorecer formas de manipulação ou erro. Quando dirigida por um humano, a IA tem capacidade para explorar as paixões humanas de forma mais eficaz do que a propaganda tradicional. (p. 194)*

E o rol de efeitos perversos continua nessa e nas páginas seguintes, no entanto os Autores formulam um conjunto de princípios fundamentais necessários para que a autonomia humana e a imunidade a interferências nocivas da atividade desregulada da iA, especialmente no domínio governativo e político (pp. 198 a 201).

Este plano de "proteção" prossegue no capítulo final – "IA e o futuro" - e transforma-se numa espécie de manifesto com tópicos fortes que merecem ser realçados até para que no fecho desta recensão possamos introduzir que papel tem a Ciência da Informação num futuro com IA. E os tópicos são sugestivos por si mesmos. Os Autores abrem com a comparação entre as duas "galáxias" que marcam a evolução da Humanidade no Ocidente; "a de Gutenberg" no séc XV (cunhada por MacLuhan) e a "da Internet" a partir de meados do séc. XX (cunhada por Castells). Ambas representam mudanças não apenas culturais, mas sociais, económicas, políticas, sendo que a "galáxia" ou revolução que estamos a viver com desenvolvimentos inusitados e imprevisíveis comporta alterações no modo de viver, de raciocinar e de sentir nunca antes experienciados, assim como, destacando especificamente a IA, o desafio posto é que ela tanto pode melhorar a trajetória humana em devir, como pode piorá-la. E esta perspetiva ambivalente é mais escalpelizada no tópico "Inteligência Geral Artificial", prosseguindo no desafiante "IA, Informação Livre e Pensamento Independente" e , para rematar, na "IA e Ordem Internacional" (ver para todos eles pp, 213 a 228).

Diante do quadro aberto que os Autores nos oferecem nesta obra várias reflexões e posicionamentos podem ser elaborados. Aqui e a fechar esta nossa "longa leitura anotada" interessa-nos colocar a Ciência da Informação, cuja natureza epistemológica e perfil de aplicação não cabe, agora, sequer aflorar, como eventual parceira disciplinar no complexo processo de retificar a perspetiva negativa que os Autores não excluem e dar contributo significativo, com outras ciências sociais e naturais (tecnologia inclusa), Trata-se de um assunto crucial que tem de ser trabalhado doravante dentro do campo ainda pouco articulado e amadurecido da Ciência da Informação, mas para o qual nesta recensão vale a pena deixar, em jeito de premissas a analisar e testar, o seguinte:

a) Assumir a CI como uma ciência social aplicada que integra a tecnologia digital como componente intrínseca, mas subordinada a um aparato teórico-metodológico procedente das ciências sociais;

b) Aprofundar a relação com a interdisciplina denominada "Ciência dos Dados" por forma que a atividade desta não se converta num "ritual" estatístico e mecanicista gerador de "certezas" que não se adequam ao âmago complexo da vida humana e social, mas que se torna, apenas, um "dictat" rígido e "frio" condicionador e dominador do agir humano.

c) Decorrente da premissa anterior sobressai a exploração que tarda a ser feita de como a CI deve participar no desenvolvimento da IA e como pode a partir daí, dessa participação, ser um "agente científico" positivo para sua humanização.