

### Potencialidades e ameaças da IA

Hoje, nenhuma Inteligência Artificial (AI) pensa. Toda a AI existente, que é notoriamente menos inteligente que a inteligência humana, “apenas”, apoiada em gigantescas bases de dados, observa o modo como as múltiplas palavras de uma frase se relacionam tentando induzir o seu significado. Baseando-se sempre em dados preexistentes.

Contudo, num futuro que se perspetiva não muito distante, uma Inteligência Artificial Geral (AGI) vai, eventualmente, ser capaz de pensar. Assim, nós, *homo sapiens*, vamos ter, pela primeira vez, de partilhar o universo conhecido com outra identidade capaz de aprender novas capacidades. Teremos de alargar os limites éticos para além da nossa espécie. Esta tarefa exigirá não só inovação, como extremo cuidado ético. Não precisará de amarras legais estranguladoras do desenvolvimento.

Os modelos de AI tornaram-se melhores que os humanos em áreas específicas do conhecimento há cerca de 30 anos, com o *match* entre Garry Kasparov e o computador Deep Blue. No entanto, mesmo jogando incrivelmente xadrez, Deep Blue não era sequer capaz de jogar damas. Tal comportamento pode ser explicado com uma analogia simples: um castor é especialista a construir represas, uma abelha é extraordinária a produzir favos de mel, mas um castor não consegue produzir favos de mel e uma abelha não constrói represas. No entanto, um humano aprende a executar ambas as tarefas. Uma AGI sabe pesquisar extremamente bem, como as melhores LLM's, mas também saberá conduzir e desviar-se de peões ou aconselhar um banco a aceitar uma candidatura de empréstimo.

Deste modo, uma AGI será parte fundamental da sociedade, sendo preciso inovar para se alcançar uma integração sublime. Porém, esta AGI terá de ser justa. Não poderá discriminar racialmente candidatos a um empréstimo bancário. É legítimo pensar que este seria um problema fácil de resolver: bastaria não fornecer as etnias dos candidatos à AGI. Não é assim tão simples: dada a enormíssima informação disponibilizada ao modelo, este pode filtrar os candidatos quanto à origem, eliminando ótimas candidaturas por estas terem sido efetuadas por pessoas originárias de bairros com menores rendimentos *per capita*.

É deste modo que surgem as mais variadas propostas de estrangulamento da inovação tecnológica: “é preciso assegurar a privacidade dos cidadãos, os dados são pessoais e devem ser protegidos. Nenhum modo de AI deve ser treinado com estes dados”.

Não renego a pertinência ética deste argumento: a privacidade é algo a valorizar.

---

Contudo, esta abordagem não satisfaz um dos aspetos fundamentais: não permite o treino destes modelos.

Considero, pelo contrário, que o melhor é, não abdicando desta ferramenta poderosíssima, ensinar-lhe os nossos valores sociais universalmente aceites nas sociedades ocidentais. Afinal de contas, uma AGI consegue aprender!

Embora sejam ferramentas notáveis, os atuais modelos de AI consistem numa mera amostra do potencial desta categoria de tecnologias. Não são o principal problema: medidas como a implementação de etiquetas identificadoras dos conteúdos que estas geram são capazes de contornar a maior parte dos dilemas éticos. No caso de uma AGI tal não acontece. É necessário encarnarmos o papel de educadores e ensinarmos, pacientemente, à pequena criança os nossos valores.

José Afonso Conceição 12º C  
Maria Coelho dos Santos 12º A  
Carina Moutinho 12º F  
Inês Matias 12º F  
Gonçalo Luzio 12º A