



>> **ChatGPT – O desafio de controlar a imprevisibilidade**

Não restam dúvidas, mesmo para os mais cétricos.

O *ChatGPT* conquistou, definitivamente, o mundo, e veio para ficar.

São muitos, e cada vez mais, os que experienciaram e continuarão a experienciar as potencialidades deste novo e desafiante modelo de linguagem treinado pela *OpenAI* e que utiliza, numa abordagem muito simplista, um algoritmo de aprendizagem profunda (*deep learning algorithm*).

Treinado num ecossistema inundado por uma infindável quantidade de dados de texto que pululam no espaço virtual e noutras relevantes fontes, o *ChatGPT* prossegue, em primeira linha, o objetivo de compreender a linguagem natural (linguagem humana entenda-se) e, por essa via, ser capaz de gerar respostas semelhantes àquelas que poderíamos obter de um qualquer ser humano numa banal conversa de fim de tarde.

A arquitetura transformacional que suporta a utilização do *ChatGPT* assenta, sobretudo, na atenção que o modelo é capaz de dedicar, de forma concentrada e focada, a diferentes partes do texto de entrada e em momentos igualmente diferenciados, sendo que é precisamente este mecanismo de atenção e de maior capacidade de focagem que é depois utilizado para construir uma representação contextual dos diferentes textos (ou parte de textos), suscetíveis de permitir captar, a final, o significado e as múltiplas relações que se podem construir e divisar entre as diferentes palavras e frases identificadas.

Nesse contínuo processo de aprendizagem em que o algoritmo navega e mergulha intensivamente, o modelo é confrontado com um texto de entrada, sendo-lhe pedido que antecipe (ou preveja) a palavra ou sequência de palavras seguinte.

Este processo é, pasme-se, repetido milhões de vezes, permitindo ao modelo, de forma gradativa e sustentada, ir apreendendo novos conteúdos (dados) e, sobretudo, ir aprendendo a gerar respostas coerentes e gramaticalmente corretas a um vasto e diversificado leque de solicitações, pedidos e perguntas várias.

Em síntese, quando o utilizador faz uma pergunta ou uma afirmação ao *ChatGPT*, o modelo utiliza o conhecimento previamente apreendido e aprendido para gerar uma resposta que seja coerente e suficientemente importante para endereçar a pergunta formulada.

O *ChatGPT* constitui, nesta exata medida, um extraordinário exemplo que o poder da aprendizagem profunda e, sobretudo do processamento de linguagem natural é capaz de revestir na forma como nós, utilizadores, nos relacionaremos, no futuro, com os computadores e com os sistemas de inteligência artificial.

Sendo uma forma de inteligência artificial generativa, dada a sua vocação em gerar novos resultados com base nos dados (IA generativa) em que se foi desenvolvendo o seu treino, traz, e disso temos de estar absolutamente cientes, dificuldades acrescidas, dada a imprevisibilidade dos conteúdos e do novo conhecimento que o sua capacidade de aprendizagem vai gradativamente desenvolvendo.

Contrariamente às anteriores formas de IA, o *ChatGPT*, enquanto IA de natureza generativa, traz preocupações ao nível dos limites (ou da sua ausência) ao próprio processo de aprendizagem do algoritmo, em particular sobre a natureza e, refira-se, enorme extensão e volume de dados sobre os quais se desenvolve esse mesmo processo.

Esta sua marca de imprevisibilidade nos conteúdos e novos dados que vai gerando, converte o *ChatGPT*, enquanto IA de natureza generativa, numa IA de alto risco que escapa, de forma preocupante, à própria capacidade preditiva associada à atividade regulatória.

O *ChatGPT* é, numa palavra, aquilo que o seu utilizador quiser que ele seja. Se o utilizador quiser que ele responda a uma pergunta de baixo risco, ele responderá e, naquele momento, não será uma IA generativa de alto risco. No entanto, se um utilizador pretender que ele elabore sobre dados de maior risco, ele será considerado naquele momento uma IA de alto risco.

Esta sua imprevisibilidade baralha os que tem por missão regular e/ou mitigar riscos de uma utilização abusiva ou mesmo criminosa dessa IA generativa.

Veremos, em breve, como a União Europeia responde a este desafio, garantindo que o *ChatGPT* continue a sua caminhada de sucesso e, simultaneamente, que todos dele possam, em segurança e sem comprometer o bem comum, tirar o maior benefício.

Esse será, creio, o maior desafio que teremos pela frente.