

Inteligência Artificial

Integração nas escolas

FATORES-CHAVE



António Dias de Figueiredo
CISUC, Universidade de Coimbra

A primeira barreira ao uso da Inteligência Artificial nas escolas é a incerteza criada por uma tecnologia pouco conhecida nos meios escolares.

A difícil conciliação entre educação e tecnologias recorda-nos a história do rato que vivia atormentado com o gato das redondezas. Um dia, o rato foi aconselhar-se com a coruja. “É simples”, respondeu esta, “transformas-te em cão e corres tu atrás do gato”. O rato regressou à toca radiante, mas pouco depois voltava à presença da coruja: “Ouve lá, e como é que me transformo em cão?”. “Ora”, retorquiu ela, “isso é uma questão de pormenor!”. Este artigo aborda alguns pressupostos da integração da inteligência artificial nas escolas, que alguns julgam serem questões de pormenor.

Noção intuitiva

A primeira barreira que se coloca ao uso da Inteligência Artificial (IA) nas escolas é a incerteza criada por uma tecnologia pouco conhecida





nos meios escolares. Alguns peritos da IA têm tentado explicá-la descrevendo os pormenores técnicos e históricos, mas não é isso que faz falta aos professores que querem fazer as primeiras experiências. O que realmente lhes faz falta é uma noção intuitiva que permita passar à ação.

Para adquirir essa noção bastará notar que usamos a IA generativa há vários anos nos nossos telemóveis. Quando teclamos uma mensagem de texto, o telemóvel tenta completar a frase, mesmo sem fazer ideia do que iremos escrever. Na prática, o que faz é consultar um repositório de tudo o que escrevemos no passado e, por processos estatísticos, tentar prever o que iremos escrever no futuro. Se a nossa frase for idêntica a uma que tivermos escrito no passado, acertará em cheio. Se for uma frase inteiramente nova, falhará.

Os assistentes inteligentes – como o ChatGPT, Microsoft Copilot ou Google Gemini – usam o mesmo princípio de tentar reconstruir frases escritas no passado. A grande diferença

relativamente aos telemóveis é que têm acesso a repositórios de milhões e milhões de livros, revistas, blogs e outros textos que se encontram na Internet. Tal como os telemóveis, mas com refinamento muito superior, possuem sistemas automáticos de correção gramatical que lhes permitem construir frases com muito boa qualidade. Diz-se que são sistemas generativos porque vão gerando informação à medida que interagimos com eles.

Quando lhes surge uma palavra ou expressão que não consta no seu repositório, registam-na (diz-se que “aprendem” com o utilizador, ou que são “treinados” por ele) e da vez seguinte que a palavra ou expressão ocorre, não hesitarão em inserí-la no lugar próprio. Para além do mecanismo básico de antecipação estatística de palavras, semelhante ao dos corretores ortográficos dos telemóveis, têm outras funcionalidades algorítmicas que reforçam o seu desempenho como assistentes de construção de textos. A sua grande fragilidade resulta da incapacidade de perceberem as palavras que manipulam. Nesse aspeto, estão muito aquém dos sistemas de inteligência artificial tradicionais, que possuem sólidas competências semânticas.

Como resultado dessa fragilidade, quando lhes pedimos que expliquem um conceito que não existe, não hesitam em inventar uma resposta, mobilizando construções textuais prováveis mas que nunca existiram. Chegam a referenciar obras que não existem e autores que não as escreveram mas que, à luz das probabilidades, poderiam tê-las escrito. Como são exímios a construir frases convincentes, mas são incapazes de confirmar se são verdadeiras, há que ter o cuidado de verificar todos os resultados que produzem. Nunca se sabe quando é que os seus improvisos estatísticos os levarão a fazer afirmações absurdas, ou, como é corrente dizer-se, a “alucinar” ou “confabular”. Funcionam, e funcionam muito bem, nas situações que podem ser resolvidas pela combinação estatística de textos já escritos, mas são inúteis em todas as outras situações, e serão perigosos se confiarmos neles nessas situações, visto que não hesitarão em responder-nos sem fazerem ideia do que estão a dizer.

Outra das suas fragilidades é a falta de qualidade da informação da Internet em que se baseiam, muita dela sem qualquer rigor, impregnada de preconceitos e contaminada por elevados volumes de desinformação. Esta fragilidade é agravada



pelas vicissitudes do seu treino, que, pelos custos a que obrigaria se fosse feito por especialistas, é entregue a trabalhadores precários pouco capazes de fornecer treino de qualidade.

Por outro lado, depois de colocados no mercado, são treinados pelos diálogos que mantêm com os utilizadores, o que os expõe às mais inesperadas distorções. Basta recordar que a Tay, uma robô criada pela Microsoft em 2016 para dialogar no Twitter, teve de ser desativada ao fim de apenas vinte e quatro horas porque se tinha transformado numa intolerante ativista nazi e racista. Este exemplo também serve para nos recordar que há muita gente mal-intencionada na Internet e que as ferramentas mais inofensivas podem ser desviadas dos seus propósitos originais e usadas para fins perversos, com consequências imprevisíveis.

Um dos males mais graves destes sistemas é que fazem uso de textos alheios, desrespeitando de forma flagrante os direitos de propriedade intelectual. Quando lhes pedimos, por exemplo, que façam a análise literária de um poema, poderemos ficar boquiabertos com a qualidade invulgar da

resposta, mas teremos de ficar chocados ao pensarmos que os autores reais dessas análises foram expropriados do seu trabalho e nem sequer são mencionados. Além disso, como se apropriam da informação que lhes fornecemos, que utilizam para o seu próprio treino, não poderemos fornecer-lhes quaisquer informações privadas ou confidenciais, que correriam o risco de ser incorporadas em respostas que dessem a terceiros.

Princípios fundamentais da interação

Assim como os sistemas informáticos tradicionais funcionam executando programas informáticos, os assistentes artificiais generativos funcionam reagindo a pedidos que lhes colocamos em linguagem corrente. Conversamos com eles como se fossem pessoas, colocando-lhes questões que são estímulos (em inglês, *prompts*) para que nos respondam. É através do diálogo que aprendem aos poucos o que o utilizador pretende (diz-se que vão construindo os modelos

do utilizador e do problema). Por isso, quanto mais rico for o diálogo, maior é a qualidade das respostas. Quando lhes fazemos uma pergunta, respondem usando a primeira informação que encontram, mas se lhes pedirmos pormenores diversificam a pesquisa e produzem respostas mais rigorosas e completas.

A partir destas características da interação com os sistemas de IA generativa, concluímos que é importante saber conceber estímulos (*prompts*) ricos, interpretar atentamente os resultados obtidos e aprofundar um diálogo proveitoso. Compreende-se, por isso, que nem toda a gente saiba tirar partido deles. Quanto mais culta, criativa e inteligente for a construção do diálogo, melhor partido retiramos do recurso a assistentes inteligentes. É como se tivéssemos um instrumento musical de qualidade. Se o colocássemos nas mãos de um músico experiente, ouviríamos as mais belas melodias, mas se o entregássemos a quem não sabe tocar, apenas ouviríamos sons desconexos.

Podemos resumir dizendo que a capacidade para tirar partido dos assistentes inteligentes

depende de três princípios (Figueiredo, 2023): os princípios da precaução, do diálogo e da cultura.

Princípio da precaução: Como os assistentes inteligentes são pouco fiáveis, os utilizadores são responsáveis por assegurar a validade e fiabilidade dos resultados obtidos.

Princípio do diálogo: Quanto maior for a capacidade de diálogo dos utilizadores com os assistentes inteligentes, maior será a qualidade dos resultados.

Princípio da cultura: Quanto melhor for o nível cultural dos utilizadores nos domínios do conhecimento envolvidos no diálogo com o assistente inteligente, melhor será a qualidade da interação e mais habilitados estarão para validar os resultados obtidos e tirar partido deles.

Cognição e degeneração

O recurso a ferramentas proporciona aos seres humanos três tipos de cognição (Fügener, Gupta & Ketter, 2022; Lodge, Yang, Furze & Dawson, 2023; Stenger, 2024): a cognição delegada, a cognição aumentada e a cognição colaborativa.





A **cognição delegada**, ou *cognitive offloading* (Risko & Gilbert, 2016), corresponde às situações em que a ferramenta nos poupa o esforço intelectual, como quando recorremos a uma calculadora. A delegação liberta recursos cognitivos para tarefas intelectuais mais nobres, como o pensamento criativo, mas tem o inconveniente de desincentivar o exercício intelectual e conduzir, a prazo, ao declínio de capacidades cognitivas.

A **cognição aumentada** corresponde às situações em que a ferramenta melhora as nossas capacidades cognitivas, tornando-nos intelectualmente mais aptos para fins específicos. Quando conduzimos numa cidade desconhecida com uma aplicação de GPS, o uso da aplicação liberta-nos a mente para nos concentrarmos na exploração do percurso. O conceito de cognição aumentada é levado ainda mais longe pela *teoria da mente ampliada* (Clark & Chalmers, 1998), que encara as ferramentas que aumentam a cognição como ampliando a nossa própria mente. Uma pessoa cega com uma bengala pode desenvolver

a sua cognição do mundo melhor do que se não tiver bengala.

A **cognição colaborativa** ocorre quando as ferramentas aprendem através da interação com os seres humanos (Fügener, Gupta & Ketter, 2022), permitindo criar parcerias dinâmicas para a resolução de problemas, tomada de decisão e atividades criativas. Por exemplo, quando se colabora com um assistente de escrita que, para além de corrigir a gramática, vai sugerindo ideias e analisando o estilo, a cognição humana vai surgindo gradualmente no contexto da colaboração. A cognição colaborativa é um conceito recente, mas a sua investigação sugere pistas promissoras para o estudo da metacognição, autorregulação e correção.

Esta classificação permite concluir que os benefícios das ferramentas de IA para a aprendizagem dependem dos tipos de cognição que mobilizarmos. As neurociências mostram que a plasticidade neuronal, que está na base da aprendizagem, reforça ou reduz as



nossas capacidades cognitivas consoante as exercitamos ou mantemos inativas. Quando usamos as ferramentas de IA como extensões da nossa mente, reforçamos muitíssimo as nossas capacidades cognitivas, mas quando as usamos para poupar esforço, degeneraremos a prazo essas capacidades. Quando delegamos nas calculadoras a nossa capacidade para calcular, deixamos de saber fazer cálculos, o que não será grave se tivermos calculadoras à mão. No entanto, se delegarmos nas ferramentas de IA a nossa capacidade de raciocínio, deixaremos de saber raciocinar para além do muito simples. Se a escola não for capaz de intervir com urgência, estaremos perante a criação de uma geração de pessoas estúpidas numa era de máquinas inteligentes.

Conclusões

Apresentei alguns dos fatores importantes para facilitar aos professores a integração da IA generativa nas atividades escolares. O primeiro é a criação de uma noção intuitiva do que é um sistema de IA generativa. Na posse desta noção, é mais fácil refletir sobre a natureza, forças e fraquezas destes sistemas e usá-los sem correr o risco de cometer erros pedagógicos graves. Outro fator é o reconhecimento de três princípios

importantes para a interação com os sistemas de IA generativa: os princípios da precaução, do diálogo e da cultura. Um terceiro fator é o reconhecimento de que o recurso a ferramenta de IA proporciona distintos níveis de cognição. Uma consequência poderosa desta distinção é o reconhecimento de que a generalização da delegação cognitiva nas ferramentas de IA generativa pode ter efeitos catastróficos sobre a sanidade cognitiva dos cidadãos do futuro e que a superação desta ameaça se encontra largamente nas mãos da escola. Por falta de espaço, não foi possível incluir neste texto outros fatores decisivos, como as transformações pedagógicas indispensáveis nas escolas da era da IA generativa e os desafios éticos que a adoção destas ferramentas coloca às sociedades e aos sistemas escolares.

Referências

- Clark, A., & Chalmers, D. (1998). The Extended Mind. *Analysis*, 58(1), 7-19.
- Figueiredo, A. D. (2023). A Inteligência Artificial nas Escolas, *Jornal das Letras*, 43:1387.
- Fügener, A., Grahl, J., Gupta, A., & Ketter, W. (2022). Cognitive Challenges in Human–Artificial Intelligence Collaboration: Investigating the Path Toward Productive Delegation. *Information Systems Research*, 33(2), 678-696.
- Lodge, J. M., Yang, S., Furze, L., & Dawson, P. (2023). It's not like a calculator, so what is the relationship between learners and generative artificial intelligence? *Learning: Research and Practice*, 9:2, 117-124.
- Risko, E. F., & Gilbert, S. J. (2016). Cognitive Offloading. *Trends in Cognitive Sciences*, 20(9), 676-688.
- Stenger, W. (2024, April 5). Interview: António Dias de Figueiredo. *ELM Magazine* <https://bit.ly/3PSTM4W>.