

A língua portuguesa a tentar resistir ao furacão da IA

O futuro do português joga-se num contexto em que a inteligência artificial pode ditar a forma como falamos, pensamos e sonhamos. *Terceiro de uma série de trabalhos sobre a língua portuguesa.*

João Pedro Pincha (texto) e José Alves (ilustração)

28 de Fevereiro de 2025, 7:04

Estamos em 2045. O Pedro, de 20 anos, discutiu com a namorada, Inês, de 22, por uma coisa sem importância. Quer pedir-lhe desculpa e decide escrever uma mensagem que seja romântica, mas não lamechas; madura, mas não chata.

Claro que não é ele que vai escrever - pedirá ao assistente virtual que o faça, seguindo aquelas instruções. Aliás, o Pedro e a Inês nunca aprenderam a escrever (<https://www.publico.pt/2025/02/26/sociedade/reportagem/leitura-escrita-precisam-espaco-aulas-portugues-2123618>), porque o assistente devolve-lhes todos os textos de que precisam em poucos segundos e executa uma série de tarefas por eles. É assim com tudo: os trabalhos para a faculdade, as mensagens para os amigos, as listas de compras, as cartas de motivação, as opiniões sobre restaurantes, até as cartas de amor. Também não sabem ler, porque o assistente consegue relatar-lhes tudo o que acabou de escrever, mais tudo o que esteja na Internet e fora dela.

Utópico? Distópico? Pura ficção científica? Num prazo de 20 anos? Mais? Menos?

O exercício é propositadamente radical e baseia-se numa reflexão de António Branco, professor da Faculdade de Ciências da Universidade de Lisboa. As crianças que nascerem de ora em diante já vão crescer num mundo em que a [inteligência artificial](https://www.publico.pt/inteligencia-artificial) (IA) generativa entrou em todos os domínios da vida quotidiana, de formas que talvez ainda nem consigamos imaginar. “Eles vão aprender a escrever e a ler? Se calhar, nunca aprenderão a escrever e a ler”, comenta o professor. “Porque não precisam. Da mesma maneira que nós, hoje em dia, para um cálculo complicado, não precisamos de exercitar o nosso algoritmo mental da soma, da divisão, da multiplicação. Agarramos uma máquina de calcular e fazemos”, prossegue.

Se isso viesse a acontecer - o que está longe de ser consensual entre os especialistas com quem o PÚBLICO falou -, provavelmente haveria consequências directas na “capacidade de pensar” das futuras gerações, porque é o ler e o escrever que têm permitido “abrir fronteiras” no pensamento humano, diz António Branco.

Aqui estamos no campo do longo ou do muito longo prazo, mas agora imaginemos que as ferramentas de IA são fornecidas apenas por um punhado de empresas a nível mundial e que, de um dia para o outro, os serviços ficam indisponíveis por qualquer razão. Como é que uma sociedade de tal forma dependente destas tecnologias conseguiria ser funcional? E em que estado ficaria a [língua portuguesa](https://www.publico.pt/2025/02/26/infografia/lingua-tese-mundo-853) (língua-tese-mundo-853)?

Isto já não é do domínio da ficção científica, alerta Branco, que há muitos anos acompanha os avanços no campo das tecnologias da linguagem que estão na base da actual IA e que em 2012 - há uma eternidade, portanto - foi o primeiro co-autor de um livro branco sobre *A Língua Portuguesa na Era Digital* (<https://www.di.fc.ul.pt/-ahb/pubs/2012BrancoMendesPereiraEtAl.pdf>). Naquela época, a principal preocupação era a de que “os falantes da língua portuguesa fossem empurrados para usarem a língua inglesa”, porque a tecnologia então existente não se adaptava facilmente a todas as línguas.

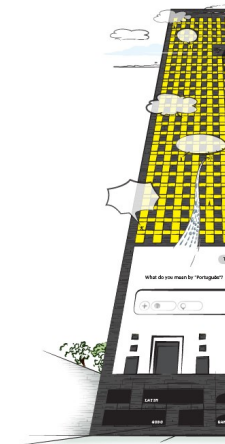
Agora, no [paradigma neuronal](https://www.publico.pt/2025/01/09/tecnologia/noticia/ia-lado-tempo-2118207) (ia-lado-tempo-2118207) da IA, isso parece ter deixado de ser um problema. Desde que haja um volume de dados com suficiente dimensão para treinar os modelos (lá voltaremos), é possível obter resultados em português semelhantes aos do inglês - quando conversamos com um *chatbot*, por exemplo.

A questão agora é que “a IA generativa está concentrada nas mãos de quatro ou cinco *players*, todos do mesmo país, e esta tecnologia é altamente dispendiosa, exige recursos imensos que deixaram de estar ao alcance de equipas de investigação nas universidades”, sublinha António Branco. “E isso cria uma dependência terrível, até do ponto de vista geoestratégico. Se nós, de repente, estivermos no lado errado do tabuleiro político, tudo pode acontecer.”

Uma boa pegada digital

Conversar com um ChatGPT, Copilot, Gemini, DeepSeek (<https://www.publico.pt/2025/01/28/tecnologia/noticia/startup-chinesa-abalou-conviccoes-ia-2120479>) ou qualquer outro grande modelo de linguagem (LLM) foi coisa que rapidamente passou de revolucionária a corriqueira e em breve será arcaica. “A evolução vai ser a ‘agentificação’ da inteligência artificial. É tornar a IA como um agente nosso, um [assistente pessoal](https://www.publico.pt/2024/12/29/ciencia/entrevista/jorge-gomes-miranda-esperemos-robos-futuro-humanos-certos-humanos-2116894) (assistente pessoal-2116894), uma inteligência artificial que não se limita a responder a perguntas, aos nossos pedidos, mas que começa a tomar iniciativas”, contextualiza Paulo Dimas, que lidera o Centro para a IA Responsável, uma entidade nascida no seio de uma *startup*, Unbabel, e que está a autonomizar-se como um consórcio sem fins lucrativos.

Eis um exemplo que nem sequer é dos mais distantes. “Qualquer telefonema vai ter uma intermediação tecnológica de processamento de linguagem natural e, quando terminarmos, há-de haver um resumo, de tal forma que, quando agarrarmos o telefone da próxima vez para ligar à mesma pessoa, [aparece uma mensagem a dizer] ‘da última vez, ficou combinado fazer isso e tal’”, explica António Branco. Na situação actual, “essa intermediação é feita de uma maneira completamente concentracionária, através de três ou quatro ou cinco *big tech* (<https://www.publico.pt/2021/11/15/tecnologia/entrevista/vera-jourova-quer-consertar-big-tech-dinheiro-tecnologia-vaio-chegar-legislacao-1984172>), e isso cria umas dependências imensas porque a linguagem humana é a maneira mais básica de comunicar e, estando dependente de agentes terceiros, eles têm na sua mão a possibilidade de nos fecharem a comunicação com qualquer outra pessoa.”



É por isso que o docente e investigador universitário tem vindo a defender “um contrabalço a essa concentração” que permita garantir “a soberania tecnológica”, a “soberania linguística” e “a soberania cultural” dos Estados onde se fala português. E isso, advoga, depende de investimentos que só os próprios Estados têm capacidade de fazer.

Se hoje pedirmos a um modelo que nos mostre um poema de Camões como o *Descalça vai para a fonte/Lianor pela verdura*, exemplifica Paulo Dimas, “ele cita os primeiros versos, aquilo até está fidedigno”, mas a certa altura já não segue nem formoso, nem seguro, “começa a alucinar”, a inventar poemas inteiros, até nomes de autores que não existem. “Nós temos de ter dados que sejam representativos da nossa cultura, da nossa língua, que permitam criar uma massa crítica de dados para nós conseguirmos ter controlo linguístico sobre a geração do texto”, defende.

Neste momento, o português não se sai mal no panorama digital. O Observatório da Diversidade Linguística e Cultural na Internet coloca a língua portuguesa como a oitava com maior percentagem de conteúdos *online* (3,09%, segundo [dados \(https://www.obdilci.org/projetos/principal/\)](https://www.obdilci.org/projetos/principal/) de Novembro passado), atrás do inglês, do mandarim, do espanhol, do hindi, do russo, do árabe e do francês. Isto deve-se sobretudo ao Brasil, onde uma larga maioria dos mais de 200 milhões de falantes de português está conectada à Internet e a produzir conteúdo - que, por sua vez, pode ser usado para alimentar os modelos de IA.

“A língua franca do mundo digital não será o inglês, mas a tradução assistida por ferramentas poderosas de IA”, considera Daniel Pimienta, que coordena o observatório, antevendo, num [texto \(https://imagingthefuture.org/reports-and-publications/the-impact-of-artificial-intelligence-by-2040/hopes-highlighting-expected-positives-in-years-to-come/\)](https://imagingthefuture.org/reports-and-publications/the-impact-of-artificial-intelligence-by-2040/hopes-highlighting-expected-positives-in-years-to-come/) para a Universidade Elon (EUA), que “em 2040 o mundo estará certamente muito mais propenso a lidar com o multilinguismo e a diversidade cultural”.

Paulo Dimas não enjeita a possibilidade de “pensar num LLM para a lusofonia e não apenas um LLM para o português europeu”, mas argumenta que é premente agir em Portugal, porque o volume de dados em português do Brasil supera largamente o das variantes europeia e africana. Pelo contrário, António Branco considera que, independentemente da sua origem, os falantes de português “estão todos no mesmo barco” face aos desafios imensos que por aí vêm. Desde logo o dinheiro. “Basta pensar neste [anúncio \(https://www.publico.pt/2025/01/22/mundo/noticia/trump-perdoa-martir-bitcoin-anuncia-investimento-inteligencia-artificial-criticado-musk-2119784\)](https://www.publico.pt/2025/01/22/mundo/noticia/trump-perdoa-martir-bitcoin-anuncia-investimento-inteligencia-artificial-criticado-musk-2119784) de que a OpenAI vai, em consórcio com um banco e com outra empresa, criar um *data center* de 500 mil milhões de euros. A [intervenção da] *troika* andou na casa dos 80 mil milhões de euros. Portanto, são não sei quantas vezes de *troikas* só para construir um *data center*”, sublinha.

Diferentes velocidades

Em Portugal, existem pelo menos duas famílias de modelos de linguagem, Albertina e Gervásio, precisamente desenvolvidos pela equipa liderada por António Branco, em código aberto e acessíveis a qualquer pessoa ou empresa. “A inteligência artificial tem de ser democratizada, e isso faz-se libertando modelos treinados especificamente para a língua portuguesa e totalmente abertos”, considera o cientista. Mas “há certos serviços que o Estado deve sentir que é sua responsabilidade fornecer e garantir”, acrescenta.

O Governo anunciou no ano passado o desenvolvimento de um LLM próprio (<https://www.publico.pt/2024/11/11/economia/noticia/web-summit-abre-clima-morno-promessa-ia-portugues-2111420>), a que chamou Amália, cuja versão *beta* está prometida para o final do primeiro trimestre, num investimento que ronda os 5,5 milhões de euros. Esse modelo, anunciou o Executivo em Novembro, será a primeira peça de uma Agenda Nacional de Inteligência Artificial, igualmente prometida até ao fim de Março. O Amália “permitirá contribuir para a preservação da soberania nacional, distinguir as diferentes variantes da língua portuguesa, reconhecer elementos da cultura e história de Portugal, permitir o controlo dos dados utilizados para a sua aprendizagem e assegurar condições de armazenamento de dados sensíveis”, explicou o Governo em [comunicado \(https://www.portugal.gov.pt/pt/gc24/comunicacao/noticia?i=modelo-de-linguagem-em-grande-escala-para-a-lingua-portuguesa\)](https://www.portugal.gov.pt/pt/gc24/comunicacao/noticia?i=modelo-de-linguagem-em-grande-escala-para-a-lingua-portuguesa).

Uns meses antes, em Julho de 2024, o Governo brasileiro [anunciou \(https://www.gov.br/lnc/pt-br/assuntos/noticias/ultimas-noticias-1/plano-brasileiro-de-inteligencia-artificial-pbia-2024-2028\)](https://www.gov.br/lnc/pt-br/assuntos/noticias/ultimas-noticias-1/plano-brasileiro-de-inteligencia-artificial-pbia-2024-2028) igualmente o seu plano para a IA a quatro anos, com investimentos na ordem dos 23 mil milhões de reais (cerca de quatro mil milhões de euros).

A par destes anúncios sonantes, a língua corre a diferentes velocidades consoante a geografia. “Eu não percebo como é que uma instituição como a CPLP, que se diz uma comunidade que visa o desenvolvimento dos países de língua portuguesa, não tem como uma das suas primeiras metas contribuir para a erradicação do analfabetismo”, aponta a linguísta Margarita Correia.

Em grande medida, por via do Brasil, que “é um dos países com maior número de produções científicas” no mundo, com “revistas próprias em português” e que “tem já há muito tempo uma política de divulgação, de ciência aberta”, a língua portuguesa “tem algum prestígio internacional”, o que se reflecte no “crescimento exponencial” do seu ensino como língua estrangeira, diz a também docente da Faculdade de Letras de Lisboa. E, portanto, entre as cerca de sete mil línguas que hoje existem, é uma das que estarão mais bem apetrechadas para sobreviver na era da IA.

Mas que português digital será esse? “Os países africanos não têm acesso à Internet em todo o seu território”, recorda Margarita Correia, exemplificando: “Eu tenho uns colegas que estão a fazer um projecto para um dicionário do português de Moçambique e que têm imensas dificuldades em manter contacto digital com os colaboradores das diversas áreas do país.”

Avessa a futurologias, a linguísta não vai tão longe como António Branco - “eu acho que a leitura vai continuar a ser feita; em relação à escrita, já ponho algumas reservas” - mas reconhece que a IA, sendo comparável às “grandes revoluções tecnológicas que foram a invenção da escrita e da imprensa”, “vai trazer profundas mudanças de comportamentos”. Ou seja, “é evidente” que as crianças nascidas a partir de agora, como o Pedro e a Inês, “vão ler de uma

maneira completamente diferente, vão escrever de uma maneira completamente diferente”.

Isso fará do português uma coisa diferente e de quem o fala também. A língua de uma pessoa, lembra Margarita Correia, é algo “mais íntimo do que o corpo”, porque é com ela que os humanos “falam, pensam e sonham”.

PATROCINADOR PRINCIPAL



<https://www.big.pt>

APOIO



<https://www.cplp.org/>



<https://example.com>

P

Abrir portas onde se erguem muros

Siga-nos

- ✉ Newsletters
- 🔔 Alertas
- f Facebook
- × X
- 📷 Instagram
- in LinkedIn
- 📺 Youtube
- 📡 RSS

Sobre

- Provedor do Leitor
- Ficha técnica
- Autores
- Contactos
- Estatuto editorial
- Livro de estilo
- Publicidade
- Ajuda

Serviços

- Aplicações
- Loja
- Meteorologia
- Imobiliário

Assinaturas

- Edição impressa
- Jogos
- Newsletters exclusivas
- Estante P
- Opinião
- Assinar

Informação legal

- Principais fluxos financeiros
- Estrutura accionista
- Regulamento de Comunicação de Infracções
- Política para a prevenção da corrupção e infracções conexas
- Plano de Prevenção de Riscos de Corrupção

- Gerir cookies
- Ajuda
- Termos e condições
- Política de privacidade