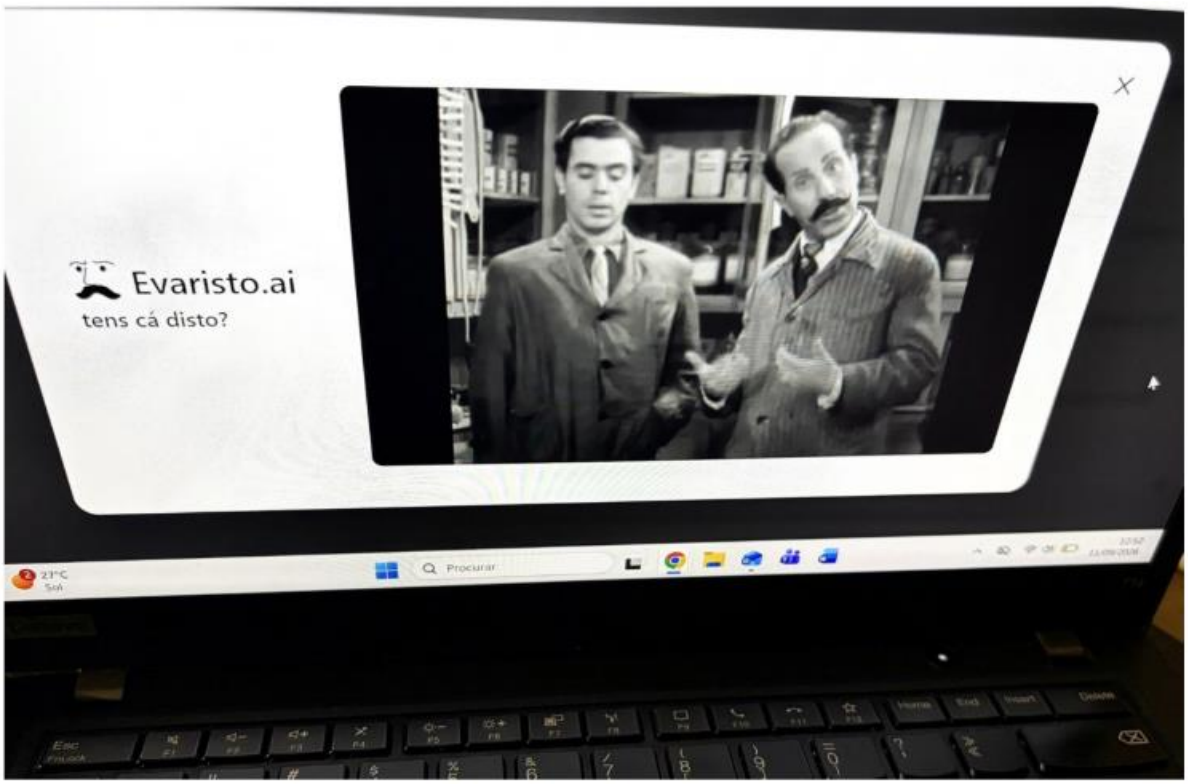
https://www.publico.pt/2026/09/11/enter/noticia/amalia-ja-evaristo-2184232

EXCLUSIVO INTELIGÊNCIA ARTIFICIAL

Antes do Amália já havia o Evaristo

Na verdade, quando o Governo apresentou o "ChatGPT português", uma equipa da FCUL já tinha desenvolvido vários modelos de inteligência artificial e chatbots para a língua portuguesa.

Laura Ferreira
11 de Setembro de 2026, 16:13



O nome do *chatbot* foi inspirado no clássico do cinema português, *O Pátio das Cantigas* DR

Antes de o Governo apresentar o Amália como o primeiro modelo de linguagem aberto desenvolvido em português europeu, já existia em Portugal um ecossistema de inteligência artificial (IA) pensado para a língua portuguesa.

O NLX – Grupo de Fala e Linguagem Natural do departamento de informática da Faculdade de Ciências da Universidade de Lisboa (FCUL) trabalha há 25 anos em tecnologia e IA adaptada ao Português e há três começou oficialmente a lançar sistemas baseados na língua.

António Branco, professor da Universidade de Lisboa e coordenador do NLX, reconhece a importância do Amália para o panorama da IA aberta em Portugal, mas afirma que não é nenhuma novidade. “Acho que o Amália podia ter sido uma melhor contribuição, trazendo alguma coisa complementar ao que já havia no panorama”, diz ao PÚBLICO.



António Branco, professor da Universidade de Lisboa e coordenador do NLX DR

A história do "ChatGPT português", desenvolvido pelo Instituto de Telecomunicações, o Instituto Superior Técnico da Universidade de Lisboa e a Universidade Nova de Lisboa, foi marcada por controvérsias desde o anúncio, no final de 2024, até ao lançamento, em Julho de 2026.

Houve desagrado em torno do processo de selecção dos centros de investigação que iam desenvolver o projecto, questões sobre os dados utilizados para treinar o modelo e dúvidas sobre o orçamento, que passou dos 5,5 milhões de euros inicialmente previstos para cerca de sete milhões.

Nas semanas que se seguiram ao lançamento, multiplicaram-se nas redes sociais os comentários e as críticas às respostas absurdas dadas pelo modelo a perguntas aparentemente simples. Numa delas, o Amália identificava Roberto Martínez como seleccionador da Bélgica, noutra, afirmava que a palavra morango tinha três vezes a letra R.

“A onda negativa nas redes sociais ao Amália, e, de arrasto, à IA generativa, era tudo o que nós não precisávamos”, sustenta António Branco.

Questionado sobre se o investimento de cerca de sete milhões de euros no Amália se justificou, António Branco responde: “O que posso dizer é que aquilo que nós

fizemos — oito modelos abertos, três chatbots e vários conjuntos de dados — ficou por cerca de um terço desse valor."

Albertinas, Gervásios, Serafins e Evaristos

Foi em 2023 que o NLX lançou o primeiro modelo de IA feito para a língua portuguesa: a Albertina. O LLM (large language model, ou grande modelo de linguagem) pertence à família dos codificadores, ferramentas especializadas na leitura, compreensão e classificação de texto. Numa aplicação de críticas de cinema, por exemplo, perante a frase "Gostei muito deste filme", a Albertina consegue interpretar o significado e identificar que se trata de uma avaliação positiva.

Depois veio o Gervásio, da família dos descodificadores. Ao contrário dos primeiros, estes modelos não se limitam a interpretar a informação que recebem: conseguem gerar texto a partir dela. São estes modelos que estão na base de muitos chatbots.

Em 2024, surgiu o Serafim, um vectorizador. Esse tipo de modelo é capaz de medir a proximidade de significado entre palavras, frases ou documentos. É uma tecnologia útil, por exemplo, num motor de pesquisa, que pode assim encontrar os conteúdos semanticamente mais próximos e dar a resposta a uma determinada pergunta.

"O passo seguinte era criar uma interface que permitisse interagir com estes modelos", lembra António Branco. E foi então, em Junho de 2025, que surgiu o Evaristo.ai.

Qual é a diferença entre um modelo e um *chatbot*?

Um modelo de linguagem é a tecnologia que aprende padrões a partir de grandes quantidades de dados e consegue, por exemplo, gerar ou compreender texto.

Um *chatbot* é uma aplicação que usa um ou mais modelos de IA para interagir com as pessoas através de uma conversa. É a "interface" que permite fazer perguntas e receber respostas.

O Amália é um modelo de linguagem; O Evaristo é um *chatbot*.

O nome foi inspirado no clássico do cinema português, O Pátio das Cantigas. No filme, Evaristo é o dono da mercearia do bairro, personagem que ficou associada à frase "Ó Evaristo, tens cá disto?". O projecto recupera a referência para a era da inteligência artificial sob a variante: "Ó Evaristo.ai, tens cá resposta para isto?"

“De certeza que o Evaristo também não sabe quem é o Roberto Martinez, mas tem a capacidade de pesquisar e dar a resposta certa”, afirma o professor, referindo-se às respostas do Amália.

Tanto o Amália como os modelos usados pelo Evaristo foram treinados com informação recolhida até uma determinada data — nenhum tem, por isso, conhecimento dos acontecimentos posteriores a essa data. No entanto, o Evaristo contorna esse problema através daquilo que se designa por agentividade: a capacidade que o chatbot tem para decidir quando precisa de recorrer a ferramentas externas, como a Internet, para completar uma resposta.

“Quando fizemos o Evaristo percebemos que tínhamos de o pôr a funcionar de acordo com as expectativas dos utilizadores”, explica António Branco. “E as expectativas das pessoas já são moldadas pelas experiências que tiveram com outros chatbots.”

A possibilidade de o Amália passar a pesquisar informação na web é, para João Magalhães, investigador da Universidade Nova de Lisboa e coordenador do projecto, uma evolução tecnicamente exigente. “A ‘evolução’ para utilizar pesquisa Web é na verdade uma aplicação do LLM muito desafiante porque as perguntas são completamente abertas”, afirmou. Ainda assim, admite que poderão surgir aplicações com essa capacidade: “Apesar de ser muito complexo e dispendioso do ponto de vista de engenharia e manutenção, alguém acabará por o fazer.”

O Amália é um modelo de linguagem aberto e de uso geral, pensado para ser integrado e adaptado a diferentes aplicações. Não foi lançado como um chatbot, como Evaristo. As interfaces que surgiram posteriormente e que permitiram aos utilizadores conversar com o modelo foram desenvolvidas por terceiros.



O caminho para a IA soberana

A diferença entre os dois projectos não está apenas na forma como podem ser utilizados. Está também na dimensão e na origem dos modelos em que assentam.

O Gervásio mais potente tem 70 mil milhões de parâmetros e o Amália conta com nove mil milhões. São números expressivos, mas modestos quando comparados

com os maiores modelos desenvolvidos pelas grandes tecnológicas.

Para serem treinados, os LLM precisam de quantidades colossais de texto, de capacidade de computação e investimentos que apenas algumas das maiores empresas tecnológicas do mundo conseguem suportar. “Com os dados livres que existem para a língua portuguesa nunca conseguiríamos obter os resultados que temos hoje”, sublinha António Branco.

Uma forma de contornar essa barreira é partir de um modelo já treinado e adaptá-lo a uma língua ou tarefa específica. Os investigadores da FCUL partiram do modelo Llama, da Meta, e continuaram o treino com dados em português europeu, de forma a ensinar ao sistema as particularidades gramaticais, o vocabulário e as expressões próprias da variante nacional.

Na base do Amália está o EuroLLM, uma família de LLM desenvolvida na Europa para várias línguas europeias. “Consideramos que a utilização de modelos de fora da Europa, com desenvolvimento fechado, não contribuem verdadeiramente para o ecossistema ou soberania nacional em IA”, sustenta João Magalhães. “O EuroLLM na base do Amália é completamente aberto, e houve uma grande preocupação em libertar um modelo cuja licença fosse verdadeiramente aberta e protegesse as aplicações que utilizassem o Amália.”

Para António Branco, a soberania não depende disso. “Para uma organização ter um serviço de IA soberana, o software respectivo não tem de ser desenvolvido por essa organização; nem no país dessa organização; nem no continente dessa organização”, defende. Para o professor da Universidade de Lisboa, a possibilidade de pegar num grande modelo já treinado e especializá-lo pode ser precisamente uma forma de a alcançar. “Ter um LLM especializado em português de Portugal a partir de um modelo de base pré-treinado não é uma desvantagem. É uma vantagem e até uma necessidade se se pretende ter um LLM com desempenho competitivo.”

A abertura é o que aproxima, apesar das diferenças, as duas abordagens. Tanto o Amália como os modelos desenvolvidos pelo NLX estão disponíveis na Hugging Face, um dos maiores repositórios mundiais de modelos de inteligência artificial, onde podem ser descarregados, estudados e reutilizados.

“Depois de termos libertado a Albertina, recebemos um email a dizer que tinha sido criada a MediAlbertina, uma aplicação médica”, recorda o professor António Branco. Segundo o investigador, os modelos do NLX foram descarregados “pelo menos largas centenas de milhares de vezes” desde o lançamento do primeiro LLM.



Membros do NLX trabalharam no desenvolvimento dos modelos de linguagem DR

Para o investigador, é precisamente essa possibilidade de reutilização que justifica disponibilizar os modelos de forma aberta.“A construção dos modelos é a parte mais custosa da IA generativa. Tivemos o privilégio de receber financiamento para os desenvolver. Agora disponibilizamo-los para quem os quiser reutilizar.”

As ferramentas de IA do NLX foram desenvolvidas no âmbito de um projecto financiado pelo PRR, que terminou no final de Junho. A manutenção e o funcionamento dos sistemas estão agora a cargo da Portulan Clarin, uma infra-estrutura de investigação que integra uma rede europeia mais ampla, criada para apoiar investigadores em ciências da linguagem.

Modelos que podem correr numa cave

“Não há dúvida nenhuma que o ChatGPT ou o Gemini são exponencialmente superiores ao Evaristo”, admite António Branco. Mas competir com a big tech nunca foi a missão do projecto.

Em vez de tentar construir um modelo cada vez maior e mais potente, a equipa do NLX procurou responder a outra pergunta: será possível desenvolver serviços de inteligência artificial úteis recorrendo a modelos mais pequenos, especializados e suficientemente baratos para serem utilizados por universidades, empresas ou organismos públicos?

> **“O Evaristo tem o propósito de ser uma demonstração empírica de que a IA soberana é viável tecnologicamente e comportável economicamente”, defende o investigador.**

“Os centros de dados das grandes tecnológicas custam milhares de milhões porque têm de responder a milhões de utilizadores ao mesmo tempo. Um chatbot como o Evaristo recebe, quando muito, algumas dezenas ou centenas de pedidos por dia”, detalha o professor. “Estes modelos podem correr numa cave”, diz António Branco — instalados nos servidores de uma universidade, de uma empresa ou de um organismo do Estado, sem depender de gigantescas infra-estruturas.

Por isso, defende, pequenos modelos especializados podem ser mais úteis do que um único sistema que tenta responder a tudo. Foi seguindo esta lógica que a equipa da FCUL começou a desenvolver sistemas para diferentes finalidades. Hoje, já existem três versões do chatbot Evaristo, cada uma pensada para um contexto específico.

O primeiro a ser lançado foi o Evaristo.ai – Generalista, pensado para utilização geral e com acesso a extensões como a pesquisa na Internet. Em Novembro de 2025, surgiu o Evaristo.ai – Serviços Públicos, um chatbot especializado nos serviços da Administração Pública portuguesa, baseado na informação disponibilizada no site gov.pt. Já em Maio de 2026, foi apresentado o Evaristo.ai – Linguagem Clara, criado para ajudar a reformular textos numa linguagem mais simples e acessível.

A equipa quis levar o trabalho que desenvolveu ao longo dos anos mais longe. Para isso, criou a *spin-off* Ousar.ai, para transformar a investigação em soluções de IA para organizações.

A empresa desenvolve agentes que podem funcionar directamente no hardware das próprias entidades, incluindo em áreas como as finanças, o turismo e a cibersegurança. O Evaristo e o Gervásio são alguns dos modelos que servem de demonstração. A ideia é que as instituições possam ter as suas próprias ferramentas de IA, em português, adaptadas às suas necessidades e sem abdicar do controlo sobre a tecnologia e os seus dados.