

# Portugal apresenta o AMALIA, o primeiro modelo de linguagem aberto desenvolvido em português europeu

- Novo LLM nacional reforça a soberania digital, acelera a transformação do Estado e cria uma plataforma de inovação para empresas, academia e sociedade
- Desenvolvimento inicial representou um investimento de 5,5 milhões de euros, financiado pelo PRR. Está prevista uma nova fase até 2027, com um investimento adicional de 1,5 milhões de euros para evolução do modelo e infraestrutura soberana
- Mais de 60 investigadores portugueses unem-se para desenvolver um modelo nacional de inteligência artificial, transformando conhecimento científico em inovação com impacto na sociedade.

Portugal apresenta hoje o AMALIA – Assistente Multimodal Automático de Linguagem com Inteligência Artificial, o primeiro modelo de linguagem (LLM) aberto desenvolvido em português europeu, um projeto estratégico que posiciona o país entre o grupo de nações com capacidade para desenvolver e supervisionar tecnologia própria de inteligência artificial.

Desenvolvido por um consórcio de universidades e centros de investigação portugueses, o AMALIA constitui um novo ativo estratégico nacional para a transformação digital do Estado, o fortalecimento do ecossistema científico e tecnológico e a competitividade da economia.

O modelo é disponibilizado em código aberto, permitindo que Administração Pública, empresas, universidades, centros de investigação e cidadãos possam desenvolver aplicações de inteligência artificial adaptadas à língua, ao contexto jurídico e à realidade portuguesa.

*"Este projeto representa uma escolha estratégica: investir na soberania digital, fortalecer a nossa capacidade científica e criar as condições para que o Estado, as empresas e a academia desenvolvam inteligência artificial em português, ao serviço das pessoas e da economia. É um investimento na autonomia tecnológica do país e na competitividade das próximas décadas."*

**Ministro Adjunto e da Reforma do Estado, Gonçalo Matias**

Num contexto em que a inteligência artificial se torna uma infraestrutura crítica para Estados, economias e sociedades, o AMALIA representa um passo decisivo para reduzir a dependência de plataformas proprietárias desenvolvidas fora da Europa.

Ao contrário dos modelos fechados, o AMALIA poderá ser auditado, adaptado e executado em infraestruturas soberanas, permitindo maior controlo sobre os dados, transparência tecnológica e alinhamento com os princípios europeus de proteção de dados e IA de confiança.

Esta capacidade permitirá ao Estado desenvolver serviços públicos baseados em inteligência artificial mantendo o controlo sobre tecnologias críticas, dados e evolução futura do modelo.

O lançamento do AMALIA integra-se na estratégia do Governo para construir um Estado mais simples, mais inteligente e mais digital.

O modelo será progressivamente integrado em serviços públicos, começando pelo desenvolvimento de novos assistentes digitais para atendimento aos cidadãos, apoiando igualmente a automatização de processos administrativos, pesquisa documental, gestão do conhecimento e novas soluções de apoio à decisão.

Esta visão enquadra-se na agenda de modernização da Administração Pública, baseada na utilização responsável da inteligência artificial para prestar melhores serviços, reduzir burocracia e aumentar a produtividade do Estado.

O AMALIA constitui também uma das bases tecnológicas para a evolução para um Estado cada vez mais preditivo, inteligente e orientado por dados.

O AMALIA resulta de um esforço conjunto da Universidade NOVA de Lisboa, Instituto Superior Técnico, Universidade de Coimbra, Universidade do Porto, Universidade do Minho e Fundação para a Ciência e a Tecnologia (Arquivo.pt/FCCN), envolvendo mais de 60 investigadores especializados em inteligência artificial, linguística computacional e engenharia informática.

Este projeto demonstra a capacidade científica nacional para desenvolver tecnologia de fronteira e reforça o papel das universidades portuguesas na criação de conhecimento com impacto económico e social.

Ao ser disponibilizado em código aberto, o modelo continuará a evoluir através da colaboração entre academia, centros de investigação e comunidade tecnológica.

Esta é uma oportunidade de transferir conhecimento da academia para a sociedade, gerando valor acrescentado e impacto social.

O Ministério da Educação, Ciência e Inovação (MECI) considera que não pode haver soberania tecnológica sem capacidade computacional, nem uma inteligência artificial nacional sem uma infraestrutura científica robusta. Por isso, nos últimos anos, tem vindo a investir em infraestruturas estratégicas, como o supercomputador MareNostrum 5, a iniciativa EuroHPC e o supercomputador Deucalion, instalado em Guimarães. Também na área da Educação existe um elevado potencial para o desenvolvimento de soluções inovadoras que estão em curso.

Para além da Administração Pública, o AMALIA constitui uma nova plataforma para inovação empresarial. Empresas, startups e centros tecnológicos poderão desenvolver produtos e serviços assentes num modelo otimizado para português europeu, reduzindo custos de desenvolvimento e acelerando a criação de soluções nas áreas da saúde, justiça, educação, finanças, indústria, turismo ou atendimento ao cliente.

Ao disponibilizar uma infraestrutura nacional de inteligência artificial aberta, Portugal cria condições para aumentar a competitividade do seu ecossistema digital, atrair investimento e desenvolver propriedade intelectual de elevado valor acrescentado.

### **Sobre o AMALIA**

O AMALIA foi desenvolvido ao longo de 18 meses utilizando infraestruturas de computação nacionais e europeias, incluindo o Deucalion, o MareNostrum 5 e a rede EuroHPC.

O modelo é multimodal, compreendendo texto, documentos, imagens e fala, tendo sido treinado especificamente para português europeu e desenvolvido segundo elevados padrões de segurança, transparência e conformidade com o AI Act europeu.

A disponibilização pública do modelo permitirá que investigadores, empresas, organismos públicos e programadores contribuam para a sua evolução contínua, reforçando um ecossistema nacional de inovação em inteligência artificial.

O projeto foi financiado através do Plano de Recuperação e Resiliência (PRR), representou um investimento de 5,5 milhões de euros, e integra a estratégia nacional para reforçar a soberania digital, acelerar a transformação digital da Administração Pública e posicionar

Portugal como produtor de tecnologia de inteligência artificial. Está prevista uma nova fase até 2027, com um investimento adicional de 1,5 milhões de euros para evolução do modelo e infraestrutura soberana.

## Perguntas e Respostas (Q&A)

### 1. O que é um LLM?

Um LLM (Large Language Model) é um modelo de inteligência artificial capaz de compreender e gerar texto em linguagem natural. É a tecnologia que suporta assistentes virtuais, motores de pesquisa, chatbots e sistemas automáticos de resposta.

### 2. O LLM português será equivalente ao ChatGPT?

Não. O ChatGPT é uma aplicação desenvolvida sobre um modelo de linguagem da OpenAI. O AMALIA é o próprio modelo de linguagem. Trata-se de um modelo aberto, desenvolvido em Portugal, que pode ser utilizado para criar diferentes aplicações de IA.

Ao contrário dos modelos proprietários, o AMALIA é auditável, documentado e pode ser adaptado pela comunidade científica, empresas e Administração Pública. O projeto reforça ainda a soberania digital, reduzindo a dependência de tecnologia estrangeira.

### 3. Porque é importante ter um LLM português?

Porque os modelos internacionais continuam a privilegiar o inglês e nem sempre compreendem corretamente o português europeu, a cultura e o contexto nacional. O AMALIA melhora a qualidade da IA em português, valoriza a língua e cria uma base tecnológica própria para cidadãos, empresas e Estado.

### 4. Quem desenvolveu o AMALIA?

O projeto reúne investigadores da Universidade NOVA de Lisboa, Instituto Superior Técnico, Universidade de Coimbra, Universidade do Porto, Universidade do Minho e Fundação para a Ciência e Tecnologia (Arquivo.pt).

A equipa envolve cerca de 30 especialistas no desenvolvimento do modelo base e mais de 50 investigadores em áreas específicas, demonstrando a capacidade científica nacional na área da inteligência artificial.

### 5. Qual a infraestrutura tecnológica utilizada?

O treino do modelo utilizou supercomputadores nacionais e europeus, incluindo o Deucalion, o MareNostrum 5 e infraestruturas EuroHPC.

## 6. Como foi treinado o modelo?

O desenvolvimento decorreu em duas fases. Na primeira, foi criado um modelo com 9 mil milhões de parâmetros, treinado com cerca de 4 biliões de palavras, recorrendo a dados em português, incluindo conteúdos do Arquivo.pt.

Na segunda fase, o modelo passou a ser multimodal, integrando capacidades para compreender texto, imagens e fala.

O desenvolvimento prevê novas evoluções até 2027, incluindo uma versão com 22 mil milhões de parâmetros e novas capacidades para suportar agentes inteligentes.

## 7. Como foi assegurada a propriedade intelectual dos conteúdos?

O treino utilizou dados públicos e legalmente acessíveis, enquadrados na legislação europeia aplicável à investigação científica.

O modelo será disponibilizado sob licença Apache 2.0 e existe um mecanismo para reporte de dúvidas ou preocupações através da plataforma IA.gov.pt.

## 8. Qual foi a cronologia do projeto?

O projeto foi desenvolvido em 18 meses.

Depois de uma versão beta em abril de 2025, o modelo foi sendo testado em ambiente real através de casos de uso na Administração Pública e da plataforma IA Edu. Em 2026 foi concluída a fase multimodal e de validação.

## 9. Como se enquadra no AI Act?

O AMALIA é classificado como um modelo de IA de uso geral (General Purpose AI), sem risco sistémico.

Ainda assim, foi desenvolvido seguindo elevados padrões de segurança, transparência e conformidade regulatória.

## 10. Quais os principais impactos para a sociedade?

O AMALIA reforça o ecossistema nacional de IA, disponibilizando um modelo aberto que pode ser utilizado por cidadãos, empresas, universidades e Administração Pública para desenvolver soluções em português europeu.



### **11. Porque é importante para Portugal e para a economia?**

O AMALIA reforça a soberania digital, reduz a dependência tecnológica externa e cria uma infraestrutura estratégica para a transformação digital do Estado.

Ao ser aberto, permite igualmente que empresas e centros de investigação desenvolvam novos produtos e serviços, aumentando a competitividade da economia portuguesa.

### **12. Como pode ser utilizado?**

O modelo será disponibilizado publicamente através do portal IA.gov.pt, permitindo a qualquer entidade utilizá-lo, adaptá-lo ou desenvolver novas aplicações.

### **13. Quais os primeiros casos de utilização?**

Os primeiros casos de uso centram-se em áreas como cultura, ciência, educação e Administração Pública, servindo simultaneamente para testar e melhorar o modelo em contexto real.

### **14. Onde poderá ser utilizado no futuro?**

O AMALIA poderá apoiar serviços públicos, automatizar processos administrativos, melhorar o atendimento ao cidadão e servir de base a aplicações em áreas como saúde, justiça, educação, finanças ou investigação.

### **15. Como será assegurada a evolução do modelo?**

O projeto continuará até ao final de 2027.

Sendo um modelo aberto, a sua evolução dependerá também dos contributos da comunidade científica, empresas e utilizadores.

### **16. Quais são hoje as capacidades do AMALIA?**

O AMALIA é um modelo multimodal otimizado para português europeu.

Compreende texto, documentos digitalizados, imagens e fala, suporta tarefas como resposta a perguntas, tradução, pesquisa documental e descrição de imagens, e inclui mecanismos de segurança para reduzir respostas inadequadas.

### **17. Haverá uma interface para o público?**

O primeiro grande caso de utilização será no portal gov.pt, através de um assistente virtual para atendimento aos cidadãos.



### **18. Como se compara com outros modelos?**

O AMALIA não pretende competir diretamente com os maiores modelos internacionais.

Os testes mostram um desempenho muito competitivo em português europeu, especialmente em compreensão da língua, cultura nacional e execução de instruções, superando outros modelos abertos em várias tarefas.

### **19. Que medidas de segurança foram adotadas?**

O modelo foi desenvolvido com mecanismos de mitigação de risco, filtragem de dados, testes de robustez e exercícios/testes de cibersegurança para reduzir a geração de conteúdos inadequados.

### **20. O AMALIA vai competir com o ChatGPT, Claude ou Gemini?**

Não.

Esses são produtos comerciais desenvolvidos por grandes empresas tecnológicas. O AMALIA é um modelo aberto, desenvolvido para responder às necessidades de Portugal, da Administração Pública e do ecossistema nacional de inovação.

### **21. O que significa ser um modelo soberano?**

Significa que Portugal pode desenvolver, adaptar e utilizar tecnologia crítica de IA mantendo controlo sobre os dados, a infraestrutura e a evolução do modelo, reduzindo a dependência de fornecedores externos.

### **22. Como é garantida a segurança dos dados?**

O AMALIA pode funcionar em infraestruturas soberanas, em Portugal ou na União Europeia, cumprindo o RGPD e permitindo às organizações controlar onde os dados são armazenados e processados.

### **23. Até que data vão os dados de treino?**

O modelo foi treinado com dados disponíveis até junho de 2024.

### **24. Foram pagos direitos de autor pelos dados utilizados?**

O projeto enquadra-se nas exceções previstas na legislação europeia para investigação científica.

A utilização futura do modelo em aplicações comerciais deverá cumprir a legislação aplicável a cada caso.

## 25. Os dados utilizados têm valor comercial jornalístico?

Não. Os dados de treino têm como data limite junho de 2024 e não incluem informação atualizada ou em tempo real.

## 26. O que é a licença Apache 2.0?

É uma licença de código aberto que permite utilizar, modificar e distribuir o modelo, incluindo para fins comerciais, desde que sejam respeitadas as condições da licença.

## 27. O AMALIA pode gerar respostas erradas ou inadequadas?

Como qualquer sistema de IA, o risco não é nulo.

O modelo foi sujeito a testes rigorosos de segurança e dispõe de mecanismos para reduzir respostas incorretas, discriminatórias ou perigosas. Os utilizadores poderão também reportar problemas através do portal IA.gov.pt.

## 28. Qual foi o investimento realizado?

O desenvolvimento inicial representou um investimento de 5,5 milhões de euros, financiado pelo PRR.

Está prevista uma nova fase até 2027, com um investimento adicional de 1,5 milhões de euros para evolução do modelo e infraestrutura soberana.

## 29. Como foi desenvolvido o AMALIA?

O AMALIA parte do modelo europeu EuroLLM-9B, posteriormente expandido e otimizado para português europeu.

Ao longo do projeto foram desenvolvidos novos dados de treino, mecanismos de segurança, capacidades multimodais e ferramentas de avaliação, que serão disponibilizados em código aberto.

## 30. Quais serão os primeiros serviços públicos a utilizar o AMALIA?

O AMALIA tem sido desenvolvido com um foco adicional na utilidade para a Administração Pública, com a prossecução de vários casos de uso em diferentes áreas e domínios, nomeadamente nas seguintes áreas:

- Cultura, através de assistentes para museus e monumentos;
- Educação, com ferramentas de apoio aos professores;
- Atendimento público digital, incluindo o portal gov.pt;
- Soberania, com aplicações de apoio à decisão na Marinha.

Está previsto que o AMALIA sirva de suporte a várias aplicações da Administração Pública. O **portal gov.pt** será um dos seus primeiros casos de uso, através da funcionalidade de assistente virtual ao atendimento público.

### 31. Que outros LLM soberanos foram e estão a ser desenvolvidos na Europa?

Para além do AMALIA, existem várias iniciativas de Estados europeus, a desenvolver modelos de IA em código aberto, apoiados por financiamento estatal. De entre estes, destaca-se:

País	Modelo(s)	Financiamento (aprox.)
Alemanha	OpenGPT-X / Teuken 7B	15 M€
Polónia	PLLuM	8 M€
Holanda	GPT-NL	14 M€
Suíça	Apertus	21 M€