



Evaluating and Programming Intelligent Chatbots for Any Language

Fact Sheet

Project Information

EPICAL

Grant agreement ID: 101141712

DOI

[10.3030/101141712](https://doi.org/10.3030/101141712) 

EC signature date

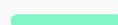
26 November 2024

Start date

1 March 2025

End date

28 February 2030



Funded under

European Research Council (ERC)

Total cost

€ 2 498 200,00

EU contribution

€ 2 498 200,00

Investment in EU policy priorities

Digital agenda



Clean air



Artificial
Intelligence



Climate action



Biodiversity



Coordinated by

TECHNISCHE UNIVERSITAET

MUENCHEN

 Germany

Objective

Intelligent chatbots (ICs) such as ChatGPT have revolutionized the generation of content for a few languages such as English, but there are 7099 currently spoken languages in the world. EPICAL will, for the first time, determine how to add new low

resource languages (LRLs) to ICs. We will make six advances to revolutionize the capabilities of ICs, unifying different areas of research that are incorrectly studied separately. We will: 1) determine how to generate hallucination-free text using ICs, and how to enter a virtuous cycle where LRL text is created using cross-lingual knowledge from ICs and then quickly post-edited and trained upon, resulting in a better LRL representation in the IC. 2) develop more powerful encoding and language adaptation approaches which combine the benefits of fine-tuning and adapters, taking full advantage of linguistically related languages to model LRLs. 3) enable ICs to reason about their own LRL capabilities and determine what they know and do not know. 4) unify research on machine translation and ICs to obtain ICs which can translate to LRLs with state-of-the-art accuracy. 5) enable high quality text-to-speech and automatic speech recognition of LRLs with ICs, thereby unifying the research on low resource speech processing with research on LRL text processing. 6) develop a novel evaluation methodology including a robust method for automatically measuring fact hallucination. My research group is well-known for LRL research, which differs from large commercial labs focusing only on the top 200 languages. Our work is critical for a multilingual Europe which values the role of minority languages, culture and heritage. Our innovations will benefit natural language processing beyond text generation and machine translation and strongly impact other areas of machine learning research suffering from data bottlenecks.

Keywords (i)

[intelligent chatbots](#)

[large language models](#)

[machine translation](#)

[natural language processing](#)

Programme(s) (i)

HORIZON.1.1 - European Research Council (ERC) MAIN PROGRAMME

[See all projects funded under this programme](#)

Topic(s) (i)

[See all projects funded under this topic](#)

Funding Scheme [\(i\)](#)

HORIZON-ERC - HORIZON ERC Grants

[See all projects funded under this funding scheme](#)

Call for proposal [\(i\)](#)

[ERC-2023-ADG](#) ↗

[See all projects funded under this call](#)

Host institution



TECHNISCHE UNIVERSITAET MUENCHEN

Net EU contribution [\(i\)](#)

€ 2 498 200,00

Total cost [\(i\)](#)

€ 2 498 200,00

Address

Arcisstrasse 21

80333 Muenchen

 **Germany** 

Region

Bayern > Oberbayern > München, Kreisfreie Stadt

Activity type

Higher or Secondary Education Establishments

Links

[Contact the organisation](#)  [Website](#) 

[Participation in EU R&I programmes](#) 

[HORIZON collaboration network](#) 

Beneficiaries (1)



TECHNISCHE UNIVERSITAET MUENCHEN

 Germany

Net EU contribution 

€ 2 498 200,00

Address

Arcisstrasse 21
80333 Muenchen 

Region

Bayern > Oberbayern > München, Kreisfreie Stadt

Activity type

Higher or Secondary Education Establishments

Links

[Contact the organisation](#)  [Website](#) 

[Participation in EU R&I programmes](#) 

[HORIZON collaboration network](#) 

Total cost 

€ 2 498 200,00

Last update: 12 December 2024

Permalink: <https://cordis.europa.eu/project/id/101141712>

European Union, 2026