

Challenges
proposed by

Accent obert |

ANTEDECENTES



Accent obert

Accent Obert x IA is a Catalan initiative working to make Catalan a fully operational language in artificial intelligence – across text, speech and video – and a European reference case for medium-sized languages.

Accent Obert x IA connects linguistic data and quality, open infrastructure, sectoral adoption and governance. Accent Obert acts as the institutional and ecosystem–orchestration umbrella, while BSC/AINA is the reference scientific and technological partner, contributing infrastructure, models, data and technical validation to specific projects. The goal is to turn technological capabilities into real products and services, particularly in education, public administration, law, healthcare, culture and social inclusion.

AREAS OF INTEREST

LANGUAGE TECH / AINA

DATA & ANALYTICS

ARTIFICIAL INTELLIGENCE

SUPERCOMPUTING / HPC



CHALLENGE 1

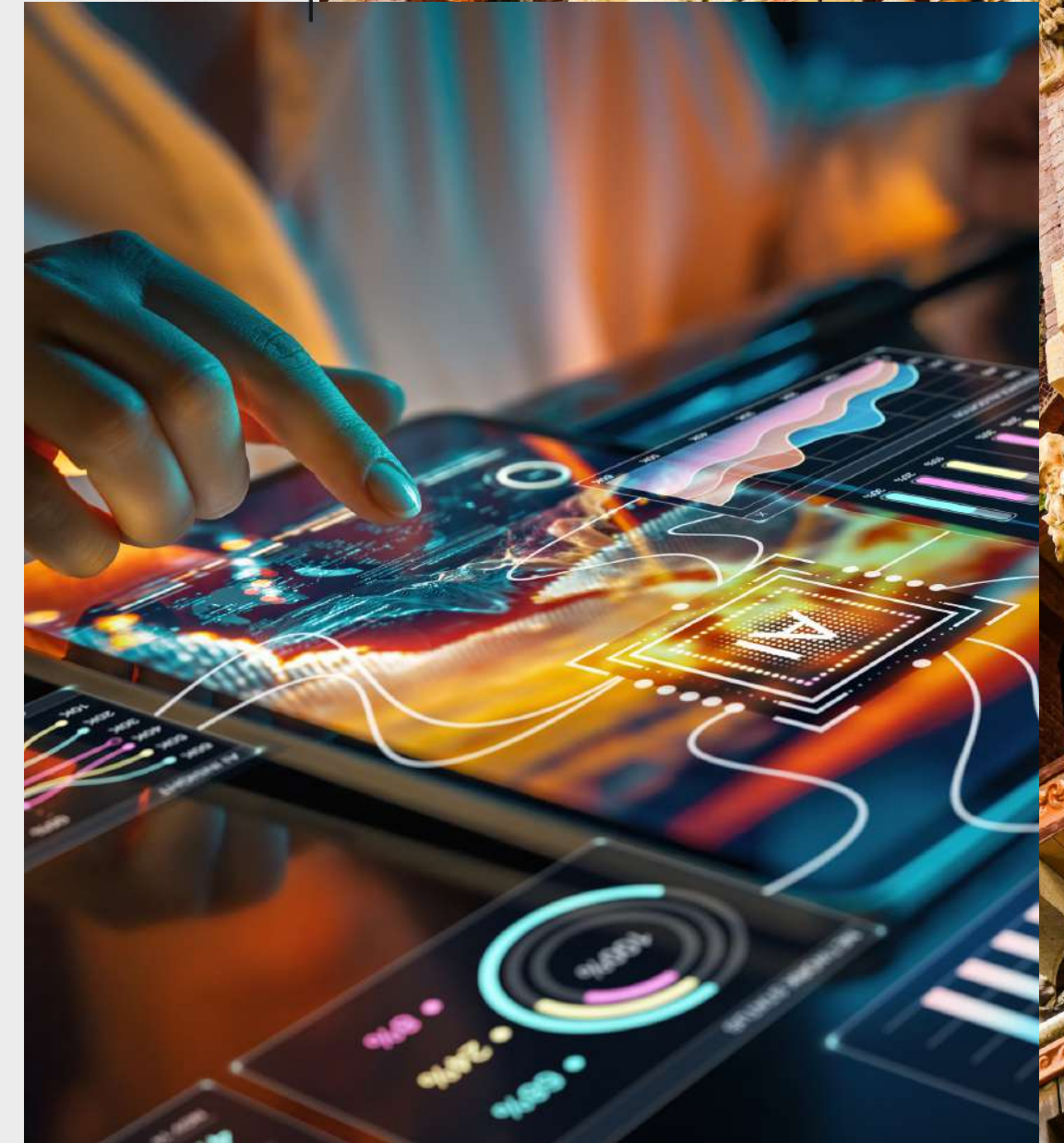
TÍTULO

Catalan Debate Agents (CDA) – Domain–specialised multi–agent AI for high–quality Catalan data

DESCRIPCIÓN

General–purpose models still have significant limitations in professional contexts where terminological precision, conceptual consistency and normative appropriateness are critical. We propose a platform of specialised agents that debate a question from opposing positions — affirmative and negative — in domains such as law, education, healthcare and finance. Each debate is generated in Catalan, submitted for expert validation and converted into transcripts and annotated corpora, in text and audio, that can be reused to train and evaluate models. The system should combine specialised agents, retrieval from authoritative sources, source traceability, professional validation, and linguistic and factual quality metrics.

Accent
obert



This challenge focuses on language technologies; the solution you propose must incorporate the resources from the Aina project. Please find all of the Aina's resources linked [here](#) and on the website.

CHALLENGE 1

Accent
obert

OBJECTIVE

Build a functional prototype with four agents, produce an initial corpus of curated debates and integrate it with AINA/BSC infrastructure to improve coverage of critical domains in Catalan.

EXPECTED OUTCOME

A safe and auditable multi-agent environment that produces high-quality professional data and can scale to other sectors, formats and language communities.

POTENTIAL ROADMAP

Initial prototype; generation of a pilot corpus of approximately 1,000 curated debates; integration and evaluation on AINA/BSC.

KEY WORDS

multi-agent AI · specialised LLMs · Catalan · domain adaptation · RAG · expert validation · data generation · NLP · ASR/TTS · language quality · legal/health/education AI



This challenge focuses on language technologies; the solution you propose must incorporate the resources from the Aina project. Please find all of the Aina's resources linked here and on the website.

CHALLENGE 2

Accent
obert

TÍTULO

Catalan for Newcomers A 24/7 GenAI mentor for Catalan language learning and multimodal data generation.

DESCRIPCIÓN

The limited availability of accessible Catalan courses for newcomers restricts linguistic inclusion. We propose a web or mobile self-learning platform based on GenAI that acts as a digital mentor available 24/7, initially for adults at A1 level.

The product will offer five-to-ten-minute micro-lessons, interactive dialogues, guided practice, activities generated from a curated curriculum, automated feedback and adaptive progression. It will include support languages to facilitate access – for example Arabic, Amazigh, Urdu or Romanian – and will focus on real-life situations: work, housing, public services, healthcare, shopping and social interaction.

With appropriate consent and data governance, user interactions will generate authentic multimodal data – text, speech and potentially video – that can be curated and reused to strengthen Catalan corpora and speech and language models. Quality will be controlled through automatic filters, normative references, user feedback, human review and benchmarks developed by the Accent Obert x IA Observatory.



This challenge focuses on language technologies; the solution you propose must incorporate the resources from the Aina project. Please find all of the Aina's resources linked here and on the website.

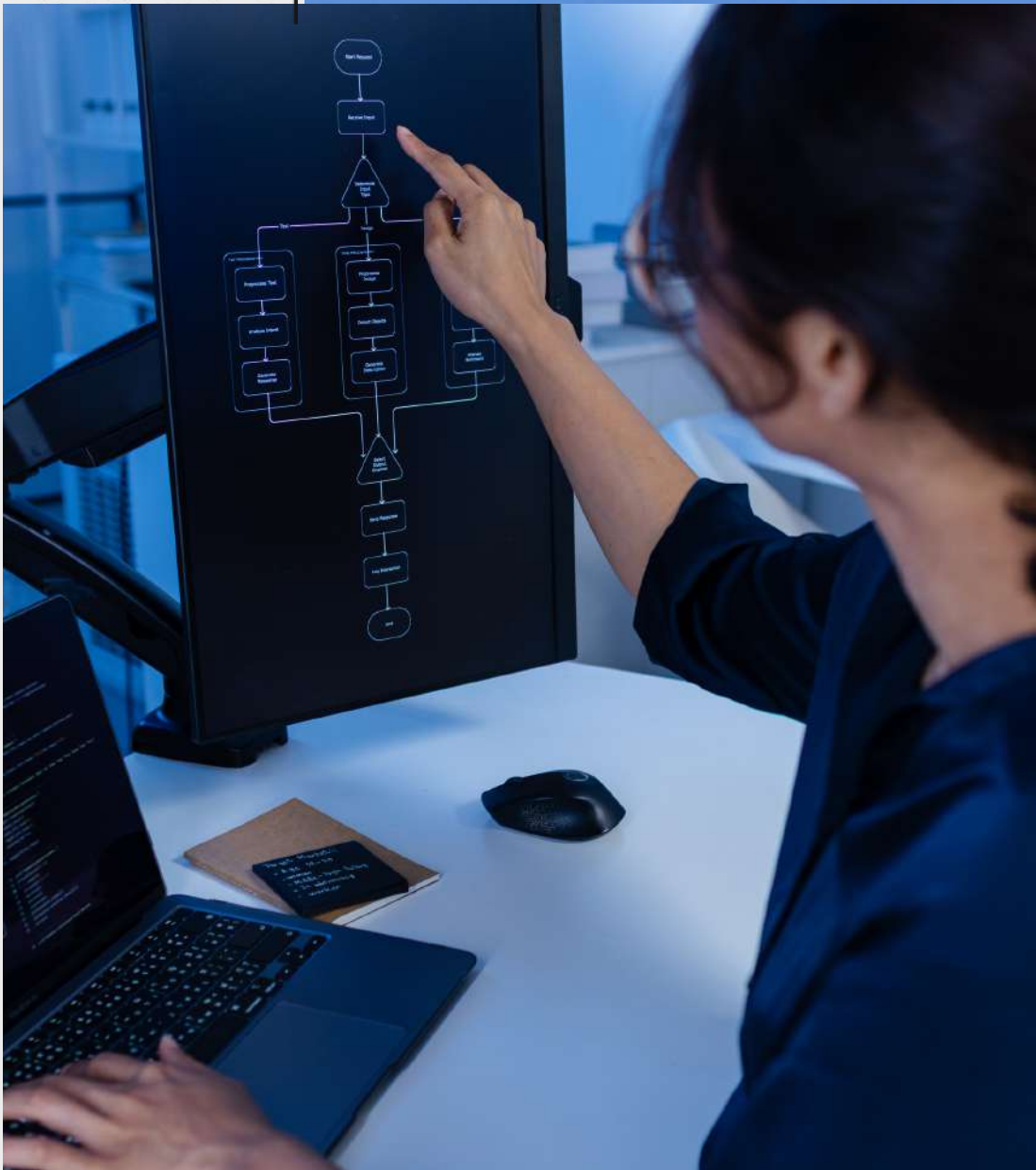
CHALLENGE 2

Accent
obert

OBJECTIVE Move from zero to a minimum viable product in three to six months, starting with six to eight A1 units and validating it through a pilot of approximately 50–100 users recruited through a newcomer–support organisation.

EXPECTED OUTCOME A tool with immediate social impact that makes Catalan learning more accessible while creating a sustainable flow of high–quality multimodal data.

KEY WORDS GenAI education · language learning · Catalan · newcomers · adaptive learning · conversational AI · multimodal data · speech technology · inclusion · digital mentor · AINA



This challenge focuses on language technologies; the solution you propose must incorporate the resources from the Aina project. Please find all of the Aina's resources linked here and on the website.

Accent obert |

[Submit your proposal here](#)

[Aina Kit](#)



Generalitat
de Catalunya



Barcelona
Supercomputing
Center
Centro Nacional de Supercomputación

 connects@bsc.es

 <https://x4hpc.cat/bsc-connects-awards>