



HiTZ

Adimen Artifizial eta Hizkuntza
Teknologiako Katedra
Cátedra de Inteligencia Artificial
y Tecnología de la Lengua



HiTZ Chair in AI & LT

LREC INDUSTRY DAY, MAY 2026

Aitor Soroa

Director

hitz.eus



Financiado por la Unión Europea
NextGenerationEU



GOBIERNO
DE ESPAÑA

MINISTERIO
PARA LA TRANSFORMACIÓN DIGITAL
Y DE LA FUNCIÓN PÚBLICA

SECRETARÍA DE ESTADO
DE DIGITALIZACIÓN
E INTELIGENCIA ARTIFICIAL



Plan de
Recuperación,
Transformación
y Resiliencia



España | digital

20
26

How it came about

- 2022 call by the Ministry for Digital Transformation
 - Within ENIA (Spanish AI Strategy)
 - 14 thematic areas
 - University–industry chairs
 - Duration: 2023-2026
- Founding members
 - EHU research groups: HiTZ Center, ALDAPA, ISG, RSAIT
 - Avature (company)

Motivation

- NLP community immersed in a paradigm shift.
- Extremely short development cycles: from basic research to final product
 - Tech companies risk becoming mere resellers of third-party technology
 - Heavy dependence on large corporations
- Shortage of experts
 - Lack of computational resources to train experts
- Shortcomings of LLMs
 - Anglo-centric reference models
 - High compute needs, when smaller models can suffice
 - Transparency and bias issues
- Need to educate society about the risks and opportunities

Lines of work

- Research and transfer
 - New collaboration mechanisms to identify synergies inside and outside the chair.
 - Virtuous circle:
 - Excellent research is guided by industry needs
 - Latest advances and frontier knowledge are passed on and inspire the next generation of products and services.

Lines of work

- Training: Increase the number of experts:
 - Training of bachelor and master students through final projects on AI< topics.
 - Continuing education for technical staff, managers and social agents.
- Dissemination and outreach of the chair's activities, and of the impact of AI< on society.

Two main actors: HiTZ and Avature

- University of the Basque Country, Computer Science Faculty (UPV/EHU Spain)
 - 88 members, multidisciplinary with computer science core
- Owns strong infraestructure: 50+50 A100 GPUs
 - Largest LLM trained in Spain (70B Latxa)
- Yearly budget: ca. 3M € (plus overheads/professors covered by Univ.)
- Reference center in NLP research in the Basque country and Spain
 - Research & Applications on English, Spanish and other languages
 - Spanish coordinators of two European infrastructures CLARIAH-ES
 - European Language Equality (core partner)
 - Consultants and relevant actors in Basque and Spanish LT plans

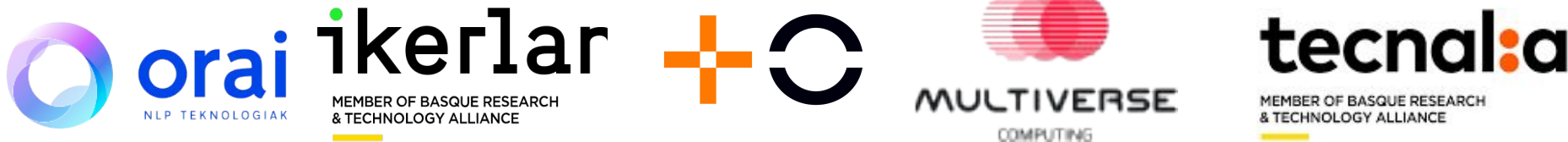
Two main actors: HiTZ and Avature

- Avature, a pioneer in recruitment technology (Human Capital Management)
- Highly configurable enterprise SaaS platform for talent acquisition and management
- HR strategies: more than 650 clients in 164 countries and 32 languages.
 - 110 of the Fortune 500 and 28 of the Forbes Global 100
- Strengthen the AI< Dept. (Dr. Rabih Zbib.)
 - Two PhD theses on efficiency and bias

<https://www.avature.net/>

Associated entities

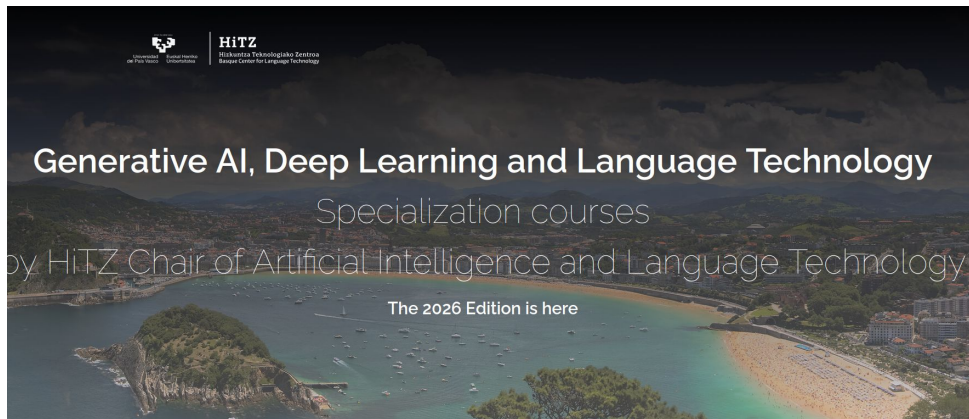
- The chair is open to other entities (maximum 8)
- Several entities joined the chair:
 - OraiNLP: <https://www.orai.eus/>
 - Ikerlan: <https://www.ikerlan.es/>
 - MasOrange: <https://masorange.es/>
 - Multiverse Computing: <https://multiversecomputing.com/>
 - Tecnalía: <https://www.tecnalia.com/>



Training activities: Continuing education

- Design and delivery of expert continuing-education courses
- First two editions: more than 250 students
- In-depth technical courses
 - Deep Learning for NLP
 - Language Technology applications
 - Large Language Models
- Technical course for non-programmers
 - Generative AI playground: using LLMs

<http://www.hitz.eus/training/>



Training activities: Bachelor and Master

- Definition of 16 Bachelor Theses (TFG) and 25 Master Theses (TFM)
- Internships at companies and institutions associated with the chair, as well as at other companies and institutions active in LT.
- Acquisition of a server with high compute capacity.

Training activities: Chair Classroom

- Chair classroom at the Faculty of Computer Science in San Sebastián
- Organisation of events to attract students.



Research and transfer activities

Incorporation of new entities (maximum 8):

- Four-year commitment
- We expect to add between 2 and 4 per year

A framework for the chair's entities to share their needs and capabilities in order to design disruptive processes and applications.

- Specialised high-level advisory
- The latest advances from the chair's entities will be presented
- Synergies between chair entities will be sought
- Several Lines of research will be chosen among the chair's entities, a chair-funded PhD thesis will be launched for each of them

Research and transfer activities

Chair Day:

- Exchange of ideas and debate on the Lines of research most relevant to industry
- These can later be developed through new doctoral theses
- Currently three theses:
 - Avature: Bias and explainability: user empowerment.
 - Avature: Green LT — reducing the computational cost of recommendation systems.
 - RAG systems

Dissemination and outreach activities:

- Conference sponsorship
 - SEPLN 2024, 2025 Doctoral symposium
 - CLEF 2025
 - LREC 2026
- Events for **outreach to industry and government**
 - LREC 2026, 2026

TecnoLing

En el marco del congreso se va a celebrar la muestra tecnológica **TecnoLing 2024 el día 25 de septiembre** coincidiendo con la inauguración del congreso. Se trata de la tercera edición de esta importante iniciativa que congrega a las empresas nacionales más importantes del sector de las tecnologías del lenguaje y creemos que es una buena oportunidad para que puedan mostrar su potencial. Startups y grupos de investigación de PLN fundados en nuestro país para poner a España como epicentro de estas tecnologías de Inteligencia Artificial.

Con el apoyo de:



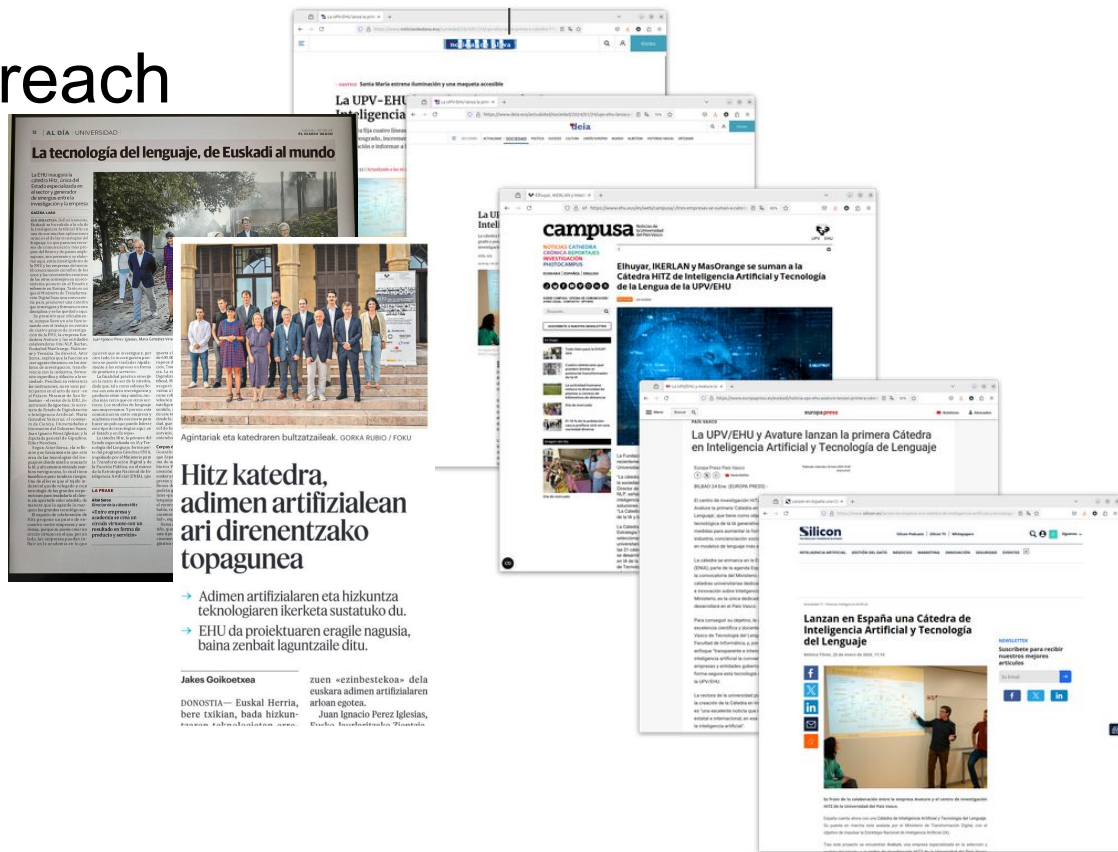
Adimen Artifiziala
Ita Hizkuntzen Teknologia
Katedra
Centro de Inteligencia Artificial
y Tecnología del Lenguaje

HiTZ
Hizkuntza Teknologia Zentroa
Basque Center for Language Technology



Dissemination and outreach

- Media presence



Dissemination and outreach

Events for **scientific outreach**

- Events for public outreach,
- Science Week
- European Researchers' Night
- Pint of Science
- Monthly webinars

<https://www.hitz.eus/en/webinars>

If you want to attend the next webinar, please complete [this](#) form and you will receive the link.

If you want to receive information about the next webinars, please complete [this](#) form.

future webinars



[Barbara Plank](#) (Ludwig-Maximilians-Universität München)

Title: The Emergence of Multilingual Representations: Tracing Linguistic Capabilities During Language Model Pretraining **(Thursday, May 21, 2026 - 15:00 CET)**▼

[Registration](#)



[Beyza Ermiş](#) (Cohere Labs)

Title: TBA **(Thursday, June 4, 2026 - 15:00 CET)**▼

[Registration](#)



[Marco Valentino](#) (University of Sheffield)

Title: TBA **(Thursday, July 2, 2026 - 15:00 CET)**▼

[Registration](#)



HiTZ

Adimen Artifizial eta Hizkuntza
Teknologiako Katedra
Cátedra de Inteligencia Artificial
y Tecnología de la Lengua



Thank you!

Aitor Soroa

ixa2.si.ehu.eus/katedra



Financiado por la Unión Europea
NextGenerationEU



GOBIERNO
DE ESPAÑA

MINISTERIO
PARA LA TRANSFORMACIÓN DIGITAL
Y DE LA FUNCIÓN PÚBLICA

SECRETARÍA DE ESTADO
DE DIGITALIZACIÓN
E INTELIGENCIA ARTIFICIAL



Plan de
Recuperación,
Transformación
y Resiliencia



España | digital ²⁰²⁶