



Universidad
del País Vasco

Euskal Herriko
Unibertsitatea

HiTZ

Hizkuntza Teknologiako Zentroa
Basque Center for Language Technology

www.hitz.eus

 @hitz_zentroa



Centre Basque des Technologies du Language

Rapport annuel
2022

INTRODUCTION



La compréhension du langage est l'un des principaux objectifs de l'intelligence artificielle. Les progrès impressionnants réalisés récemment dans l'utilisation des techniques d'apprentissage profond, combinées à de grandes quantités de texte et à la puissance de calcul, nous ont amenés à une nouvelle frontière. En effet, nous assistons à ce que certains chercheurs considèrent comme une nouvelle renaissance, dans laquelle de grands modèles de langage (LLM) facilitent l'exécution de tâches répétitives, permettant aux gens de se concentrer sur des tâches plus créatives et plus stimulantes. Comme toutes les technologies de rupture, il y a aussi des risques, tels que la désinformation, les préjugés nuisibles, les besoins élevés en énergie et l'élargissement de la fracture numérique avec les langues pauvres en ressources.

Pour relever les nouveaux défis et saisir les opportunités offertes par l'intelligence artificielle et la transformation numérique, il est impératif d'encourager la recherche ouverte et publique dans le domaine des technologies linguistiques. En outre, il est essentiel que la recherche soit effectuée au niveau local, afin d'éviter de devenir de simples consommateurs de technologies créées dans d'autres pays. Les avantages de cette approche sont évidents : la recherche et les résultats sont mis à la disposition de l'économie et de la société locales, et les risques peuvent être facilement vérifiés.

C'est un point de vue que nous prenons en main à HiTZ. L'année 2022 a été une année fructueuse au cours de laquelle nous avons vu notre centre se développer considérablement. Nous avons organisé

une conférence autour du PERTE pour la nouvelle économie de la langue qui a donné lieu à deux projets importants financés par le PERTE lui-même et par le Gouvernement Basque: NEL-GAITU et IKER-GAITU. Ces projets de trois ans, qui seront développés en étroite collaboration avec des acteurs clés de Galice, de Catalogne et de la Communauté Valencienne, donneront naissance à la prochaine génération de technologies linguistiques pour le basque, le galicien et le catalan. De plus, HiTZ a mené l'intégration des organismes de recherche basques et espagnols dans les infrastructures européennes de recherche CLARIN et DARIAH à travers le réseau de recherche INTELE.

Eneko Agirre (Directeur de HiTZ) et German Rigau (Directeur adjoint de HiTZ)

HITZ EN CHIFFRES

Recherche et transfert

35

Projets

5

Contrats de
transfert

47

Publications

4

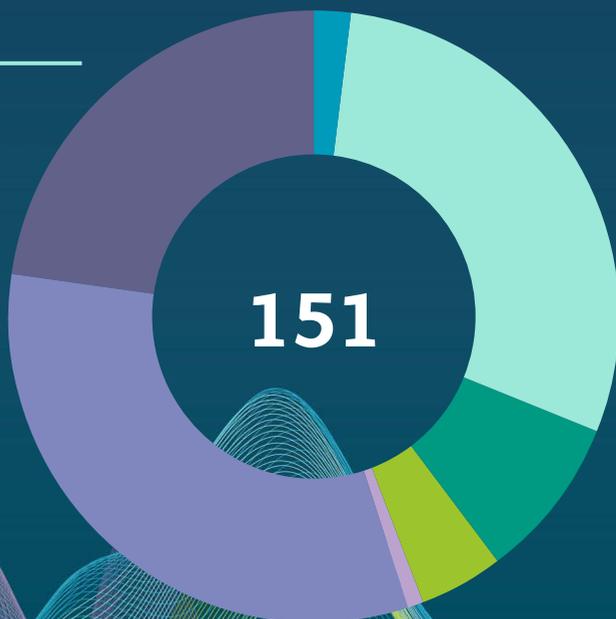
CENTRE BASQUE DES TECHNOLOGIES DU LANGUAGE



Personnes

Membres

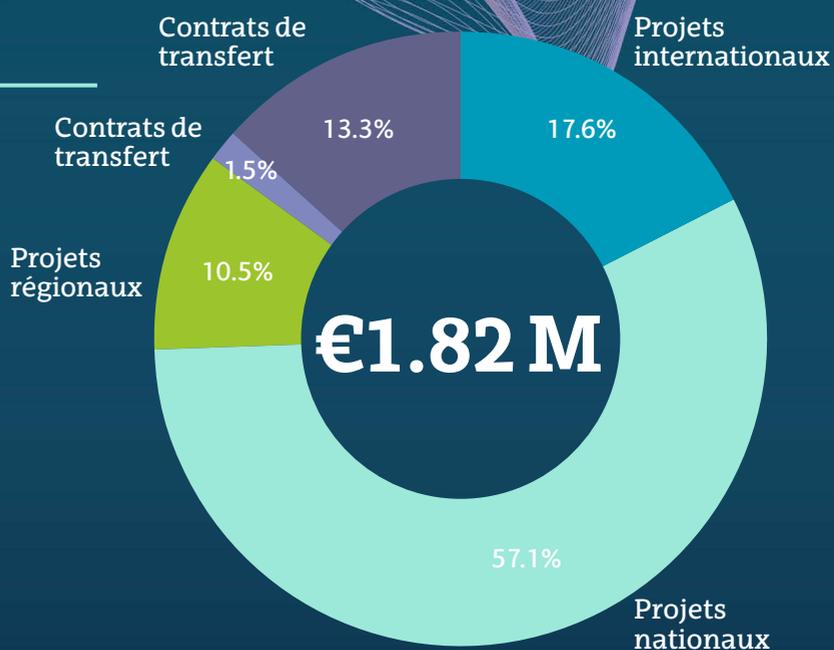
	Gestion et administration	3
	Enseignants	44
	Bourses de doctorat	13
	Chercheurs postdoctoraux	7
	Ramón y Cajal	1



Membres Associés

	Étudiants HAP/LAP	49
	Étudiants en doctorat	34

Budget



ORGANISATION

HiTZ est un centre de recherche multidisciplinaire sur **l'intelligence artificielle centrée sur la langue**, dont les membres proviennent de sept départements de l'Université du Pays Basque. L'objectif du centre est **d'étudier** les technologies du langage et de la parole, avec un effort important pour le **transfert** de connaissances et de technologies aux entreprises. Il comprend deux groupes de recherche, Aholab et Ixa, tous les deux dotés d'une grande expérience depuis 1993, qui effectuent de la recherche fondamentale, créent des ressources et des outils langagiers et lancent différents produits commerciaux sur le marché. HiTZ est également un **membre fondateur** du **Centre CLARIN-K**, membre de **CLAIRE** et membre à part entière de **BDVA** et **DAIRO**. Par l'intermédiaire de **CLAIRE** et de **BDVA**, nous participons également au European Partnership on Artificial Intelligence, Data and Robotics (partenariat européen sur l'intelligence artificielle, les données et la robotique).



Eneko Agirre

Directeur



German Rigau

Sous-directeur



Ander Salaberria

Membre du conseil



Maite Oronoz

Membre du conseil

Les membres du centre sont des référents internationaux dans leurs domaines scientifiques. Il est actuellement composé de **plus de 60 membres**, dont des informaticiens, des linguistes et 3 techniciens de recherche. Au cours des cinq dernières années, les chercheurs du centre ont publié plus de 200 publications scientifiques. Le groupe est un leader dans l'application des techniques d'apprentissage approfondi au traitement des langues, et ses travaux récents dans ce domaine ont été **cités plus de 4 000 fois** au cours des deux dernières années. Les membres du centre ont été **conseillers** dans la création du plan national pour les technologies de la langue espagnole et conseillent actuellement l'homologue du gouvernement basque.

Tant Ixa comme Aholab ont été évalués comme des

CLARIN K CENTRE

CLAIRE



BDV BIG DATA VALUE ASSOCIATION



ahō LAB



**Itziar
Aldabe**

Membre du
conseil



**Inma
Hernaez**

Membre du
conseil



**Esther
Miranda**

Membre du
conseil



**Aitor
Soroa**

Membre du
conseil

groupes de recherche performants lors du dernier exercice d'évaluation de la recherche par l'agence scientifique du gouvernement basque. Au cours de leur histoire, les groupes ont participé à plus de 200 **projets de recherche** allant de projets régionaux à des projets européens. Ils ont également participé à plus de 100 **contrats industriels** dans le but de transférer des technologies dans l'industrie.

HiTZ est également membre du programme **Erasmus Mundus+ European Masters Program** in Language and Communication Technologies (LCT). Il est conçu pour répondre aux demandes de l'industrie et de la recherche dans le domaine en pleine expansion des technologies linguistiques. HiTZ propose également un **programme doctoral** en analyse et traitement des langues,

et est membre à part entière de l'Académie Internationale des Études Doctorales en Intelligence Artificielle (AIDA).

L'Université du Pays basque (UPV/EHU) est le principal établissement d'enseignement et de recherche du Pays Basque, une région prospère qui s'étend le long de la côte atlantique du nord de l'Espagne. L'UPV/EHU fait partie des 400 meilleures universités du monde selon le classement de Shanghai, et a été reconnue comme un campus d'excellence international par le gouvernement espagnol. L'Université du Pays basque, une institution dynamique de 30 ans avec 45 000 étudiants, 5 000 membres du personnel académique de renommée mondiale et des installations de pointe réparties dans 20 centres sur ses trois campus.

DOMAINES DE RECHERCHE



**Extraction
et recherche
d'information**

Chercheur principal :



**Aitor
Soroa**



**Traduction
automatique**

Chercheur principal :



**Gorka
Labaka**



**Interaction
homme-
machine**

Chercheur principal :



**Gorka
Azkune**



**Ressources
vocales et
langagières**

Chercheuse principale :



**Ainara
Estarrona**

INFRASTRUCTURE

15

serveurs
multiprocesseurs
GNU/Linux

4

serveurs
SPARC
Solaris

1

cluster HPC
avec 128
noyaux

1

salle acoustiquement isolée
avec équipement audio
pour des enregistrements
professionnels



Analyse de textes

Chercheur principal :



Rodrigo Agerri



Technologies de la parole

Chercheuse principale :



Inma Hernaez



Domaines médical et juridique

Chercheuse principale :



Arantza Casillas



Sciences humaines numériques et éducation

Chercheur principal :



Mikel Iruskietia

10 serveurs avec

45

GPUs

Plus de

260_{TB}

de capacité de stockage brute en réseau

1

Behringer 4x4 audio/MIDI interface

1

Quiet PC Sentinel Fanless i10

RECHERCHE ET TRANSFERT

35

Projets de
recherche

5

Contrats de
transfert de
connaissances

5

Thèses de
doctorat
soutenues
(3 internationales)

16

Articles de
revue
(7 Q1)

25

Communications
de congrès
(8 A ou A+)

3

Chapitres
de livre

10

CENTRE BASQUE DES TECHNOLOGIES DU LANGAGE





49

Étudiants en master

34

Étudiants en doctorat

FORMATION

18

Thèses du master HAP/LAP soutenues

62

Étudiants dans 2 cours complémentaires d'apprentissage profond

8

Ikasiker

16

Stages internes et externes

ACTIVITÉS

21

Séminaires

8

Webinaires

3

Workshops

4

Prix

Prix Gladys
(décerné à Olatz Pérez de Viñaspre)

Aula Janssen UPV-EHU
(décerné à Nuria Lebeña)

Finaliste du prix Institución del Año
(décerné à HiTZ)

Persona TIC Asociación de Ingeniería de
Telecomunicación del País Vasco
(décerné à Inma Hernández)

