CVA date	8/9/2023

Part A. PERSONAL INFORMATION

First name	Eneko		
Family name	Agirre Bengoa		
Gender (*)	Male	Birth date	16/03/1968
ID number	15377993D		
e-mail	e.agirre@ehu.eus	URL Web	http://ixa2.si.ehu.eus/eneko
Open Researcher and Contributor ID (ORCID) (*)		0000-0002-0195-4899	

A.1. Current position

Position	Full professor (Profesor Pleno aka Catedrático Contratado)		
Initial date	17/10/2017		
Institution	Universidad del País Vasco/Euskal Herriko Unibertsitatea		
Department/Center	Lenguajes y Sistemas Informáticos		
Country	Spain	Teleph.	943015019
Keywords	Natural Language Processing (NLP), Artificial Intelligence (AI)		

A.2. Previous positions (research activity interuptions, art. 14.2.b))

Period	Position/Institution/Country/Interruption cause
12/1995-10/2017	Profesor asociado, then agregado (contratado doctor) - UPV/EHU

A.3. Formación Académica

PhD, Licensed, Graduate	University/Country	Year
Licenciado en Informática	UPV/EHU (Spain)	30/07/1990
Master of Science KBS	Edinburgh University (UK)	30/09/1991
Doctorado en Informática	UPV/EHU (Spain)	15/01/1999

Parte B. CV SUMMARY

Professor since 1995, now full professor. Director of the LSI department (2017-2019). He directs the university's HiTZ research center (hitz.ehu.eus), which has 60 researchers and an annual funding of 1.4 million euros, and is the PI of the Ixa Group (ixa.ehu.eus), recognized as a High Performance Group by the Basque Government. He has been visiting researcher/professor at New Mexico State University, University of Southern California, Melbourne University, Stanford University (twice), and is currently visiting New York University (second time). He is a member of the European Lab for Learning and Intelligent Systems (ELLIS), a society that brings together the best European researchers in Artificial Intelligence. He has received the National award of Research in Informatics (Aritmel) in 2021, and has been elected Fellow by the ACL in 2021, the prime scientific association in his area, entering a selected set of 74 worldwide researchers, only 15 from Europe and he is first in Spain. He has won five awards for the best article in SCIE class 1 or 2 conferences, three Google research awards and two victories in relevant scientific competitions. Several PhD theses under his superision have been awarded: SEPLN 2018 and 2021, EurAI 2021, National Award 2021. He has four "sexenios".

Scientific contributions: He is the most prolific national author in SCIE class 1 and 2 conferences in Natural Language Processing (NLP). He is also the most cited national author on Google Scholar in his area. He has more than 200 articles, 18 in Q1 / Q2 journals and 60 in SCIE conferences (classes 1 and 2). Over 50% of his top-tier publications are co-authored with international researchers.

Note about conferences: Conferences are the most influential publication outlets in NLP. All the articles mentioned in this CV have been reviewed by 3 peers with acceptance rates lower than 30%, and recognized by the Spanish Computer Science Society (SCIE) in classes 1 and 2 (4th and 13th percentiles, respectively, of the 1811 indexed congresses). Therefore: 1) The articles in conferences of said classes are as relevant or more than those of quartiles 1 and 2 among journals; 2) WoS is not a suitable referent for counting citations, as it ignores all citations made in such conferences. Google indicators do include conferences, and Scopus also does it to a lesser extent

Citations according to Google http://scholar.google.es/citations?hl=en&user=kSuqts0AAAAJ: 18038 citations, h-index 55. According to the number of citations, I occupy position 49 in the area of "computational linguistics" among scientists from all over the world, 126 in NLP, and 604 in AI. I occupy the 15th position among the authors of "University of the Basque Country" of any discipline. According to guide2research.com I occupy the position 48 according to the h-index among all computer science researchers in Spain.

Citations in Scopus: Total citations 6311, h-index 36. **WoS**: Total citations 791, h-index 15

Recognized as **one of the most influential authors worldwide (top 2%)** based on Scopus citations, both in his career and in 2020, 21 and 22 (https://doi.org/10.1371/journal.pbio.3000918).

He has participated in **50 research projects**. In 23 of them he has been IP in competitive calls from the United States, Europe, Spain and the Basque Country (€2,536,612).

Contributions to society: He has had 9 transfer contracts with companies and administrations (€423,000). He has organized and taught in the 7 editions of the DL4NLP course at Carlos III University and UPV/EHU, with 50-60 students on average from companies and research centers, 50% of them international. Co-founder of the HiTZ center, PI of the Ixa group. He has been interviewed in the press, radio and television on numerous occasions.

Contributions to training: He has supervised 11 doctoral theses in the last 10 years (17 in total). Two of them defended at University of Zurich and University of Darmstadt. He has given 32 invited international talks (7 national), and six international courses and tutorials (2 nationally). His last keynote was at the prestigious SIGIR conference in 2022 (SCIE Class 1 conference). He is or has been a member of the editorial committee in the three reference journals in his field (Computational Linguistics, TACL, JAIR), and Secretary and President of the Association of Computational Linguistics Special Interest Group in the LEXicon. Co-founder of the Joint Conference on Lexical and Computational Semantics (*SEM), now in the 9th edition. He is a regular reviewer for the most relevant journals. He is a senior member, area chair or regular member on the program committees of the major conferences. He was a member of the expert committee of the national plan to promote language technologies launched by the Secretary of State for Telecommunications and for Society of Information in 2016. Regular reviewer, he participated in the panel that evaluates projects of the national plan in 2020.

Parte C. RELEVANT MERITS

Given that the recent adhesion to DORA opens some uncertainty, I selected the most cited publications (mostly conferences SCIE class 1-2), as well as those in journals with high IF at JCR. I also added complementary bibliographic statistics in part B.

C.1. *Publications*

Note: In my area it is common that the senior researcher heading the research is listed as last author.

1. Arantxa Otegi, Iñaki San Vicente, Xabier Saralegi, Anselmo Peñas, Borja Lozano, and Eneko Agirre. *Information retrieval and question answering: A case study on COVID-19*. Knowledge-Based Systems , vol. 240, 2022. DOI: 10.1016/j.knosys.2021.108072. ISSN: 0950-7051
   Citas Scholar: 8 JCR IF (2020): 8,038 (Q1 CS, AI)

2. Jan Deriu, Alvaro Rodrigo, Arantxa Otegi, Guillermo Echegoyen, Sophie Rosset, Eneko Agirre, Mark Cieliebak. *Survey on Evaluation Methods for Dialogue Systems,* Artificial Intelligence Review, Volume online, pages1-56, 2020. DOI 10.1007/s10462-020-09866-x ISSN: 1573-7462
   Citas Scholar: 185 JCR IF (2019): 5.095 (Q1 CS, AI)

3. lñigo Lopez-Gazpio, Montse Maritxalar, Mirella Lapata, Eneko Agirre. *Word n-gram attention models for sentence similarity and inference*. Expert Systems with Applications, Volumen 132, paginas1-11, 20219. DOI 10.1016/j.eswa.2019.04.054 ISSN: 0950-7051.
   Citas Scholar: 47 JCR IF (2018): 4.292 (Q1 CS, AI)

4. Juan J. Lastra-Díaz, Josu Goikoetxea, Mohamed Ali Hadj Taieb, Ana García-Serrano, Mohamed Ben Aouicha, Eneko Agirre. *A reproducible survey on word embeddings and ontology-based methods for word similarity: linear combinations outperform the state of the art.* Engineering Applications of Artificial Intelligence, vol. 85, pp 645-665, 2019. DOI 10.1016/ j.engappai.2019.07.010 ISSN: 0952-1976. Citas Scholar: 80 JCR IF (2018): 3.526 (Q1 CS, AI)

5. Artetxe, M., Labaka, G., & Agirre, E.. An Effective Approach to Unsupervised Machine Translation. In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics (pp. 194-203). Citas Scholar: 140. GGS-SCIE conference rating Class 1 Rating A++

6. Artetxe, Mikel, Gorka Labaka, Eneko Agirre, and Kyunghyun Cho. "Unsupervised Neural Machine Translation." In International Conference on Learning Representations. 2018.
Citas Scholar: 744. GGS-SCIE conference rating Class 1 Rating A++

7. Mikel Artetxe, Gorka Labaka and Eneko Agirre. *Generalizing and Improving Bilingual Word Embedding Mappings with a Multi-Step Framework of Linear Transformations*. Proceedings of the AAAI Conference on Artificial Intelligence, pages 1-8, 2018.
Citas Scholar: 217. GGS-SCIE conference rating Class 1 Rating A++

8. Artetxe, Mikel, Gorka Labaka, and Eneko Agirre. "Learning bilingual word embeddings with (almost) no bilingual data." In Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pp. 451-462. 2017.
Citas Scholar: 524 . GGS-SCIE conference rating Class 1 Rating A++

9. Mikel Artetxe, Gorka Labaka, and Eneko Agirre. "A robust self-learning method for fully unsupervised cross-lingual mappings of word embeddings." In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 789-798. 2018. Citas Scholar: 545. GGS-SCIE conference rating Class 1 Rating A++

10. Josu Goikoetxea, Aitor Soroa and Eneko Agirre. "Bilingual Embeddings with Random Walks over Multilingual Wordnets". Knowledge-Based Systems, Vol. 150, pp. 218-230, 2018. Citas Scholar: 23 JCR IF (2018): 5.1 Q1 (CS, AI) DOI: 10.1016/j.knosys.2018.03.017, ISSN: 0950-7051

C.2. Conferences

I have participated in many conferences, including the talks corresponding to the papers in SCIE Class 1 and 2 conferences. I have given 12 invited talks in international conferenes (TSD Brno 2022, AKBC London 2022, SIGIR Madrid 2022, NLDB Valencia 2022, Neurips EXPO workshop, Vancouver 2020, AMLD Laussane 2020, CCCI Hendaye 2019, SEPLN Bilbao 2019, CWMT Wuyishan 2018, PROPOR Lisbon 2016, WSAP Amsterdam 2014, JSSP Trento 2013). I have also given 43 invited talks, 6 courses and 2 tutorials, mostly in international research centers.

C.3. Research projects.

1. Reference: TED2021-131617B-I00 Title: AWARE: Commonsense for a new generation of natural language understanding applications . Entidad financiadora: MCINN, Proyectos TED
   PI and entity: Eneko Agirre and Gorka Azkune, UPV/EHU
   Dates: 13/10/2022-12/10/2024   Funding at UPV/EHU: 311.070 €
   Participation: Principal Co-Investigator of co-ordinated project

2. Reference: PID2021-123988OB-C31 Title: TRAIN: EXTREMELY LOW-RESOURCED MACHINE TRANSLATION . Entidad financiadora: MCINN, Proyectos TED
   PI and entity: Eneko Agirre and Gorka Labaka, UPV/EHU
   Dates: 1/9/2022-31/8/2025   Funding at UPV/EHU: 185.000 €
   Participation: Principal Co-Investigator of project

3. Reference: BETTER. Title: CLEVER consortium for the BETTER program.
   Financing entity and call: US government, IARPA 2019.
   PI and entity: Eneko Agirre, UPV/EHU
   Dates: 1/11/2019-31/5/2023 Funding at UPV/EHU: 399.000 $
   Participation: Principal Investigator at UPV/EHU

4. Reference: IT1343-19. Title: Grupo de Investigación Ixa.
   Financing entity and call: Gobierno Vasco, Grupos de Investigación, 2018.
   PI and entity: Eneko Agirre, UPV/EHU
   Dates: 1/1/2019-31/12/2021 Funding at UPV/EHU: 282.900 €
   Participation: Principal Investigator of project

5. Reference: PCIN-2017-118. Title: LIHLITH: Aprendiendo a interactuar con las personas mediante la interacción continuada con personas.
   Financing entity and call: ERANet CHIST-ERA (MINECO), 2017.
   PI and entity: Eneko Agirre, UPV/EHU
   Dates: 01/01/2018-31/12/2020   Funding at UPV/EHU: 105.000 €
   Participation: Principal Investigator of coordinated CHIST-ERA European

6. Reference: PGC2018-102041-B-I00 Title: DOMINO: Traducción Automática Neuronal, en DOMInio, NO supervisada. Entidad financiadora: MINECO, Proyectos nacionales

PI and entity: Eneko Agirre, UPV/EHU
Dates: 01/07/2019-31/6/2021   Funding at UPV/EHU: 81.070 €
Participation: Principal Investigator of project
7. Reference: FP7-ICT-2013.4.1-610516. Title: BigKnowledge for Text Mining
Financing entity and call: Fundación BBVA
PI and entity: Eneko Agirre, UPV/EHU
Dates: 01/03/2019-28/01/2021   Funding at UPV/EHU: 71.500 €
Participation: Principal Investigator of group at UPV/EHU
8. Reference: FP7-ICT-2013.4.1-610516. Title: QTLeap: Quality Translation by Deep Language
Engineering Approaches. Financing entity and call: Comisión Europea (FP7), 2012.
PI and entity: Antonio Branco, FCUL
Dates: 01/11/2013-31/10/2016   Funding at UPV/EHU: 385.270 €
Participation: Principal Investigator of group at UPV/EHU
9. Reference: FP7-ICT-2011-270082. Title: PATHS - Personalised Access To cultural Heritage
Spaces. Financing entity and call: Comisión Europea (FP7), 2011
PI and entity: Mark Stevenson, Universidad de Sheffield
Dates: 01/11/2011-31/12/2013   Funding at UPV/EHU: 374.144 €
Participation: Principal Investigator of group at UPV/EHU
10. Title: VIGICOVID: Vigilancia automatizada de cuestiones clave sobre la COVID-19.
Financing entity and call: Banco Santander, CRUE, 2019
PI and entity: Anselmo Peñas, UNED
Dates: 1/10/2020-31/10/2021   Funding at UPV/EHU: 47.000 €
Participation: Principal Investigator of group at UPV/EHU

C.4. Participación en actividades de transferencia y explotación de resultados
1. Title: Chatbot the pregunta-respuesta. Company or entity: IZFE
PI: Eneko Agirre Bengoa. UPV/EHU. Dates: 9/12/2021-14/03/2022. Funding: 18,089 €
2. Title: Sistema de pregunta-respuesta. Company or entity: EJIE
PI: Eneko Agirre Bengoa. UPV/EHU. Dates: 26/04/2021-25/07/2021. Funding: 12,390 €
3. Title: Asesoría y proyectos de PLN (varios contratos)
Company or entity: SHERPA EUROPE, S.L.
PI: Eneko Agirre Bengoa. UPV/EHU. Dates: 15/01/2015-31/02/2021. Funding: 223,500 €
4. Title: Conversational agents that learn after deployment (Google faculty research awards 2019)
Company or entity: Google
PI: Eneko Agirre Bengoa. UPV/EHU. Dates: 2019-2020. Funding: 70.000$
5. Title: Informe del estado del arte de sistemas de diálogo. Company or entity: IKERLAN
PI: Eneko Agirre Bengoa. UPV/EHU. Dates: 1/12/2019-31/12/2020. Funding: 18.150 €
6. Title: Accessing FAQ and CQA sites via dialogue (Google faculty research awards 2018)
Company or entity: Google
PI: Eneko Agirre Bengoa. UPV/EHU. Dates: 2018-2019. Funding: 70.700 euros
7. Title: Learning Interlingual Representations (Google faculty research awards 2015)
Company or entity: Google
PI: Eneko Agirre Bengoa. UPV/EHU. Dates: 2016-2017. Funding: 52754$
8. Title: ICTReader: Vigilancia tecnológica automática del sector TIC.
Company or entity: Secretaría de Estado de Telecomunicaciones y Sociedad de la Información
PI: Eneko Agirre Bengoa. UPV/EHU. Dates: 21/09/2014-21/12/2014. Funding: 21.096,70 €
9. Title: Ondarebideak. Company or entity: Elhuyar Fundazioa.
PI: Eneko Agirre Bengoa. UPV/EHU. Dates: 15/01/2019-14/01/2021. Funding: 10.689,14 €