



**Josu Goikoetxea Salutregi**

Generated from: Editor CVN de FECYT

Date of document: 11/06/2019

**v 1.4.0**

ae9a262a76416d750dcbecf662171581

This electronic file (PDF) has embedded CVN technology (CVN-XML). The CVN technology of this file allows you to export and import curricular data from and to any compatible data base. List of adapted databases available at: <http://cvn.fecyt.es/>



## Summary of CV

This section describes briefly a summary of your career in science, academic and research; the main scientific and technological achievements and goals in your line of research in the medium -and long- term. It also includes other important aspects or peculiarities.

The summary of my CV is comprised by the following sections: academic field, research, teaching and working experience.

In the academic field I yielded the Superior Engineering degree in Industrial Electronics and Automation in the University of Deusto in 2003. After a considerable experience in technological centers and private companies I coursed the Language Analysis and Processing Master in the UPV/EHU, finishing it in September 2014. Next I started my PhD in the same field as in the Master, finishing it in July 2018 with Cum Laude qualification.

Regarding the research field, I developed my Degree Final Project in ROBOTIKER, in the field of industrial robotics. After finishing the degree, I achieved a research grant from "Fundación Centros Tecnológicos", and I took part in several research projects in the area of regulation and control of mechatronic systems in IKERLAN, from October 2003 and October 2005. After spending 7 years in several private companies, I coursed the abovementioned master, publishing my first paper in the COGALEX Workshop and opening the research field for the upcoming PhD. As a result of my research in the PhD I had 5 publications; one of them a JCR Q1, and two important conferences of level CORE A and A+. Currently I'm part of a collaboration research with UNED and SFAX Team, and our paper is under a minor revision in a journal JCR Q1. During the PhD I participated in the QTleap project, and after the PhD, as a postdoc, in LIHLITH, BERBAOLA and HISPAVISTA projects.

As a teacher, I have given theoretical and practical classes in the degree of Computer Sciences in the Faculty of Informatics of Donostia in 2016/2017 and 2018/2018, two credits each course. I also have given two seminars in 2018: one in the ADITU course, and the other one in the Master of Language Analysis and Processing. Recently I achieved the accreditation from UNIBASQUE that certifies my suitability for giving classes in the University.

As a complement to my academic and research career, I have a wide working experience in private companies, working as a R+D engineer in various companies. I have been flexible enough to adapt myself to several knowledge areas in the industry; in 2006/2007 I worked in the consultancy company GFI as a programmer of knowledge-bases; in 2007 I entered SIEMSA as a R+D engineer, programming windmills; from 2007 to 2013 I worked in INGETEAM as a programmer



of frequency converters and attending to several commissionings in naval and industrial applications.

All my academic and professional career has been complemented with courses and seminars in order to update and extend my knowledge and capabilities. I also have C1 level in english (as well as accreditation for teaching in English) and C2 (equivalent to Proficiency) in Basque.



## General quality indicators of scientific research

This section describes briefly the main quality indicators of scientific production (periods of research activity, experience in supervising doctoral theses, total citations, articles in journals of the first quartile, H index...). It also includes other important aspects or peculiarities.

It is important to emphasize that in the NLP field the high level **SCIE** (<http://gii-grin-scie-rating.scie.es/ratingSearch.jsf>) conferences are as equivalent to the **JCR** Q1 journal papers. The paper "A Reproducible Survey ..." is under a "minor review" currently.

Bilingual embeddings ... -> JCR 4.396 ( **Q1** )

A Reproducible Survey... -> JCR 2.819 ( **Q1** ) UNDER MINOR REVIEW

Similitud entre palabras... -> SJR 0.201 (Q2 Linguistics, Q3 Computational Sciences)

Single or Multiple? ... -> GGS 1 (Core **A++** , excellent quality)

Random Walks... -> GGS 1 (Core **A+** , excellent quality)



## Josu Goikoetxea Salutregi

Surname(s): **Goikoetxea Salutregi**  
 Name: **Josu**  
 DNI: **78892323T**  
 ORCID: **0000-0001-5568-4014**  
 Land line phone: **(+34) 687607333**  
 Email: **josu.goikoetxea@ehu.eus**

### Current professional situation

**Employing entity:** UPV/EHU **Type of entity:** University  
**Department:** Faculty of Informatics (Donostia)  
**Professional category:** Doctoral researcher  
**Start date:** 01/08/2018  
**Type of contract:** Temporary employment contract **Dedication regime:** Full time  
**Primary (UNESCO code):** 330406 - Computer architecture; 330499 - Other  
**Performed tasks:** Non doctoral researcher in LIHLITH (LSI department), dialogue systems training. Doctoral researcher in BERBAOLA (KZAA department), design and development of training data, experimentation. Doctoral researcher for S3FORKIDS (ATC department), design of multilingual classifier and integration to the S3FORKIDS browser.

**Employing entity:** UPV/EHU **Type of entity:** University  
**Department:** LSI, Faculty of Informatics (Donostia)  
**Professional category:** Non doctoral  
**Start date:** 06/07/2015  
**Type of contract:** Grant-assisted student (pre or post-doctoral, others) **Dedication regime:** Full time  
**Primary (UNESCO code):** 330406 - Computer architecture; 330499 - Other  
**Performed tasks:** Grant of Basque Vicerealty of UPV/EHU. PhD research.

**Employing entity:** UPV/EHU **Type of entity:** University  
**Department:** Faculty of Informatics (Donostia)  
**Professional category:** Non doctoral researcher  
**Start date:** 01/10/2014  
**Type of contract:** Temporary employment contract **Dedication regime:** Part time  
**Primary (UNESCO code):** 330406 - Computer architecture; 330499 - Other  
**Performed tasks:** Non doctoral researcher in QTLEAP; design of NMT experiments. Development of PhD.

**Employing entity:** UPV/EHU **Type of entity:** University  
**Department:** LSI, Faculty of Informatics (Donostia)  
**Professional category:** Non doctoral researcher  
**Start date:** 01/03/2014  
**Dedication regime:** Part time



**Type of contract:** Grant-assisted student (pre or post-doctoral, others)

**Primary (UNESCO code):** 330406 - Computer architecture; 330499 - Other

**Performed tasks:** Cooperation grant with the IXA Team. Development of Master Final Project.

**Employing entity:** Robotiker - Tecnalía

**Type of entity:** Technological Centre

**Professional category:** R+D grant holder

**Start date:** 10/2001

**Type of contract:** Grant-assisted student (pre or post-doctoral, others) **Dedication regime:** Full time

**Primary (UNESCO code):** 330603 - Electric motora

**Performed tasks:** R+D grant holder in teh Fabrication Department (1 year). Development of Final Degree Project.

### Previous positions and activities

	Employing entity	Professional category	Start date
1	INGETEAM TECHNOLOGY SA	R+D engineer	08/2007
2	SIEMSA	R+D engineer	02/2007
3	GFI Norte	Senior Programmer	10/2006
4	Ikerlan	R+D grant holder	10/2003
5	FUNDACION ROBOTIKER	R+D grant holder	10/2001

**1** **Employing entity:** INGETEAM TECHNOLOGY SA

**Professional category:** R+D engineer

**Start-End date:** 08/2007 - 09/2013

**Duration:** 6 years

**2** **Employing entity:** SIEMSA

**Professional category:** R+D engineer

**Start-End date:** 02/2007 - 08/2007

**Type of entity:** Private Company

**Duration:** 6 months

**3** **Employing entity:** GFI Norte

**Professional category:** Senior Programmer

**Start-End date:** 10/2006 - 01/2007

**Duration:** 4 months

**4** **Employing entity:** Ikerlan

**Type of entity:** Innovation and Technology Centres

**Professional category:** R+D grant holder

**Start-End date:** 10/2003 - 10/2005

**Duration:** 2 years

**5** **Employing entity:** FUNDACION ROBOTIKER

**Professional category:** R+D grant holder

**Start-End date:** 10/2001 - 10/2002

**Duration:** 1 year



## Education

### University education

#### 1st and 2nd cycle studies and pre-Bologna degrees

- 1 **University degree:** Higher degree  
**Name of qualification:** Superior Engineering in Automation and Industrial electronics  
**Degree awarding entity:** University of Deusto  
**Date of qualification:** 01/10/2003
- 2 **University degree:** Higher degree  
**Name of qualification:** Tecnical Engineering of Automation and Industrial Electronics  
**Degree awarding entity:** Universidad de Deusto      **Type of entity:** University  
**Date of qualification:** 01/10/2000

#### Doctorates

**Doctorate programme:** Natual Language processing  
**Degree awarding entity:** UPV/EHU      **Type of entity:** University  
**Date of degree:** 14/07/2018

#### Other postgraduate university studies

**Postgraduate qualification:** Master in Natural language Processing  
**Degree awarding entity:** UPV/EHU      **Type of entity:** University  
**Faculty, institute or centre:** Faculty of Informatics (Donostia)  
**Date of qualification:** 09/2014

### Specialised, lifelong, technical, professional and refresher training (other than formal academic and healthcare studies)

- 1 **Training title:** Deep Learning Summer School  
**Awarding entity:** Deusto University, Universitat Rovira  
**End date:** 07/2017      **Duration in hours:** 50 hours
- 2 **Training title:** Philosophical Anthropology  
**Awarding entity:** Instituto Virtual de Ciencias Humanas  
**End date:** 09/2009      **Duration in hours:** 480 hours
- 3 **Training title:** C# Language  
**Awarding entity:** CONSULTEC      **Type of entity:** Empresa privada  
**End date:** 06/2007      **Duration in hours:** 25 hours



**4 Training title:** Fiber optics  
**Awarding entity:** CONSULTEC  
**End date:** 04/2007

**Type of entity:** Empresa privada  
**Duration in hours:** 25 hours

**5 Training title:** Objetc Oriented Programming  
**Awarding entity:** INEM  
**End date:** 07/2006

**Type of entity:** Servicio publico de empleo estatal  
**Duration in hours:** 284 hours

## Language skills

Language	Listening skills	Reading skills	Spoken interaction	Speaking skills	Writing skills
English	C1	C1	C1	C1	C1
Basque	C2	C2	C2	C2	C2

## Teaching experience

### General teaching experience

**1 Name of the course:** Hizkuntzaren Prozesamendua  
**University degree:** Computer Engineering  
**End date:** 02/2017  
**Entity:** UPV/EHU **Type of entity:** University  
**Faculty, institute or centre:** Faculty of Informatics (Donostia)

**2 Name of the course:** Hizkuntzaren Prozesamendua  
**University degree:** Computer Engineering  
**End date:** 02/2016  
**Entity:** UPV/EHU **Type of entity:** University  
**Faculty, institute or centre:** Faculty Of Informatics (Donostia)

### Teaching experience in courses and seminars for university teacher training

**Type of event:** Seminar  
**Name of the event:** Unixtools for Natural Language Processing  
**Organising entity:** UPV/EHU **Type of entity:** University  
**Hours of teaching:** 3  
**Teaching date:** 27/11/2018





## Other activities/achievements not included above

- 1 Description of the activity:** Accreditation for teaching in university  
**Organising entity:** UNIBASQ **Type of entity:** Agency  
**End date:** 05/2019
- 2 Description of the activity:** Seminar in course "Aditu, hizkuntzak biziberritzeko estrategiak"  
**Organising entity:** GARABIDE **Type of entity:** NGO  
**End date:** 10/2018

## Scientific and technological experience

### Scientific or technological activities

#### R&D projects funded through competitive calls of public or private entities

- 1 Name of the project:** BERBAOLA: Tecnologías de la Lengua y el Habla aplicadas a los ámbitos del RIS3  
**Entity where project took place:** Eusko Jaurlaritza  
**Name principal investigator (PI, Co-PI....):** Aitor Soroa Echave  
**Start-End date:** 01/12/2018 - 31/03/2019  
**Total amount:** 104.154 €
- 2 Name of the project:** IXA: grupo de investigación consolidado  
**Name principal investigator (PI, Co-PI....):** Arantza Diaz de Ilarraza  
**Start-End date:** 22/12/2016 - 21/12/2018  
**Total amount:** 135.000 €
- 3 Name of the project:** LILITH: Aprendiendo a interactuar con personas mediante interacción continuada con personas  
**Entity where project took place:** Ministerio de Economía, Industria y Competitividad (Chistera)  
**Name principal investigator (PI, Co-PI....):** Eneko Agirre Bengoa  
**Start-End date:** 01/08/2018 - 30/01/2018  
**Total amount:** 105.000 €
- 4 Name of the project:** QTLEAP: quality translation by deep language engineering approaches  
**Name principal investigator (PI, Co-PI....):** Eneko Agirre Bengoa  
**Start-End date:** 01/10/2014 - 06/07/2015  
**Total amount:** 385.270 €



## R&D non-competitive contracts, agreements or projects with public or private entities

**Name of the project:** SF3FORKIDS: Hurrei zuzendutako soluzio adimendun eta unibertsal berria web informazio segurua bilatzeko

**Degree of contribution:** Researcher

**Participating entity/entities:** HISPAVISTA SL; IXA Group

**Start date:** 04/2019

**Duration:** 4 months

## Scientific and technological activities

### Scientific production

#### Publications, scientific and technical documents

- 1** Josu Goikoetxea Salutregi. Similitud entre palabras: aportaciones de las técnicas basadas en bases de datos. Sociedad Española para el Procesamiento del Lenguaje Natural. 62, pp. 103 - 106. 03/2019.

**Type of production:** Scientific paper **Format:** Journal

**Corresponding author:** Yes
- 2** Juan Lastra-Díaz; Josu Goikoetxea Salutregi; Mohamed Ali Hadj Taieb; Ana García-Serrano; Mohamed Ben Aouicha; Eneko Agirre bengoa. A reproducible survey on word embeddings and ontology-based methods for word similarity: linear combinations outperform the state of the art. Engineering Applications for Artificial intelligence. Elsevier, 2019.

**Type of production:** Scientific paper **Format:** Journal

**Corresponding author:** No
- 3** Josu Goikoetxea Salutregi; Eneko Agirre bengoa; Aitor Soroa Echave. Bilingual embeddings with random walks over multilingual wordnets. Knowledge-based Systems. 150, pp. 218 - 230. Elsevier, 15/02/2018.

**Type of production:** Scientific paper **Format:** Journal

**Corresponding author:** Yes
- 4** Josu Goikoetxea Salutregi; Eneko Agirre bengoa; Aitor Soroa Echave. Single or Multiple? Combining Word Representations Independently Learned from Text and WordNet. Proceedings of the 30th Association for the Advancement of Artificial Intelligence. 30, pp. 2608 - 2614. 02/2016.

**Type of production:** Scientific paper **Format:** Scientific and technical document or report

**Corresponding author:** Yes
- 5** Josu Goikoetxea Salutregi; Eneko Agirre bengoa; Aitor Soroa Echave. Random Walks and Neural Network Language Models on Knowledge Bases. Proceedings of the North American chapter of the Association of Computational Linguistics. 15, pp. 1434 - 1439. 05/2015.

**Type of production:** Scientific paper **Format:** Scientific and technical document or report

**Corresponding author:** Yes



## Works submitted to national or international conferences

- 1** **Title of the work:** Creating vocabulary exercises through NLP  
**Name of the conference:** Digital Humanities in Nordic Countries  
**Corresponding author:** No  
**City of event:** Copenhagen, Denmark  
**Date of event:** 06/03/2019  
**End date:** 08/03/2019  
**Organising entity:** Digital Humanities in Nordic Countries  
Manex Agirrezabal; Begoña Altuna; Lara Gil-Vallejo; Josu Goikoetxea Salutregi; Itziar Gonzalez-Dios. "Random Walks and Neural Network Language Models on Knowledge Bases".
- 2** **Title of the work:** Konbinatu eta Irabazi! Hitzen Semantikaren komunikazioa Errepresentazio Osoagoaren Bila  
**Name of the conference:** IKERGAZTE, nazioarteko ikerketa euskaraz  
**Corresponding author:** Yes  
**City of event:** Durango, Spain  
**Date of event:** 13/05/2015  
**End date:** 15/05/2015  
**Organising entity:** IKERGAZTE, nazioarteko ikerketa euskaraz  
Josu Goikoetxea Salutregi; Eneko Agirre bengoa; Aitor Soroa Echave. "Random Walks and Neural Network Language Models on Knowledge Bases".
- 3** **Title of the work:** Testu-loturen labirinto semantikoan barna, esanahi-bektoreak lagun!  
**Name of the conference:** IKERGAZTE, nazioarteko ikerketa euskaraz  
**Corresponding author:** Yes  
**City of event:** Iruñea, Spain  
**Date of event:** 11/05/2015  
**End date:** 13/05/2015  
**Organising entity:** IKERGAZTE, nazioarteko ikerketa euskaraz  
Josu Goikoetxea Salutregi; Iñigo Lopez-Gazpio; Eneko Agirre bengoa; Montse Maritxalar; Aitor Soroa Echave. "Random Walks and Neural Network Language Models on Knowledge Bases".
- 4** **Title of the work:** Exploring the Use of Word Embeddings and Random Walks on Wikipedia for the Cogalex Shared task  
**Name of the conference:** COGALEX workshop  
**Corresponding author:** Yes  
**City of event:** Dublin, Ireland  
**Date of event:** 23/08/2014  
**End date:** 23/08/2014  
**Organising entity:** COLING conference  
Josu Goikoetxea Salutregi; Eneko Agirre bengoa; Aitor Soroa Echave. "Random Walks and Neural Network Language Models on Knowledge Bases".



## Other achievements

### Obtained grants and scholarships

- 1** **Name of the grant:** Basque Vicerectory of UPV/EHU  
**Aims:** Pre-doctoral  
**Awarding entity:** Basque Vicerectory of UPV/EHU **Type of entity:** University  
**Conferral date:** 07/2015 **Duration:** 3 years  
**End date:** 07/2018  
**Entity where activity was carried out:** UPV/EHU  
**Faculty, institute or centre:** Faculty of Informatics (Donostia)
- 2** **Name of the grant:** Fundacion centros tecnológicos  
**Aims:** Pre-doctoral  
**Awarding entity:** Fundacion centros tecnológicos  
**Conferral date:** 01/10/2003 **Duration:** 2 years  
**End date:** 30/09/2005  
**Entity where activity was carried out:** Ikerlan