

Parte A. DATOS PERSONALES

Fecha del CVA	18/06/2018
----------------------	------------

Nombre y apellidos	Aitor Soroa Etxabe		
DNI/NIE/pasaporte	35768408G	Edad	45
Núm. identificación del investigador	Researcher ID		
	Código Orcid	0000-0001-8573-2654	

A.1. Situación profesional actual

Organismo	Universidad del País Vasco/Euskal Herriko Unibertsitatea		
Dpto./Centro	Facultad de Informática		
Dirección	Manuel de Lardizabal 1, 20018 Donostia		
Teléfono	943015040	correo electrónico	a.soroa@ehu.eus
Categoría profesional	Profesor Agregado	Fecha inicio	01/10/2008
Espec. cód. UNESCO	570104 120304		
Palabras clave	Inteligencia Artificial, Lingüística Informatizada		

A.2. Formación académica (título, institución, fecha)

Licenciatura/Grado/Doctorado	Universidad	Año
Licenciado en informática	Facultad de Informática de San Sebastián	1993
Doctor en Informática	Facultad de Informática de San Sebastián	2004

A.3. Indicadores generales de calidad de la producción científica (véanse instrucciones)

Número de sexenios de investigación: 2

Fecha del último sexenio concedido: 01/01/2016

Citas totales: 1215 (Scopus) 2836 (Google scholar)

Promedio de citas/año durante los últimos 5 años: 883 (Scopus) 1949 (Google scholar)

Publicaciones totales en primer cuartil: 7

Índice H: 13 (Scopus) 22 (Google scholar)

Parte B. RESUMEN LIBRE DEL CURRÍCULUM (máximo 3500 caracteres, incluyendo espacios en blanco)

El profesor Dr. Aitor Soroa es licenciado en Informática (1993) y recibió su doctorado en Informática en 2004. Actualmente es profesor agregado en el departamento de Ciencias de la Computación e Inteligencia Artificial, cargo que ocupa desde 2008. Acredita 12 años de investigación hasta 2015 (2 sexenios), 15 años de experiencia docente hasta 2013 (3 quinquenios).

Ha publicado más de 100 artículos en revistas y conferencias internacionales en el área del Procesamiento de Lenguaje Natural (PLN), y, más específicamente, en el procesamiento semántico, desambiguación de palabras, desambiguación de entidades nombradas, similitud semántica, adquisición de conocimiento, integración de sistemas PLN y procesamiento NLP a gran escala. Las contribuciones incluyen 9 publicaciones en revistas según el the Journal Citation Report (JCR), 15 artículos en congresos de alto nivel (CORE A y A+), y 2 capítulos de libros.

Es revisor habitual en revistas internacionales de gran prestigio como AI, JNLE, LRE, EAAI. Además, ha participado en programas de comité de congresos internacionales de prestigio tales como WWW, EACL, NAACL, IJCAI, IJCNLP, *SEM, HLT, AAAI e JCAI.

Actualmente, es investigador responsable en el proyecto europeo "MUSTER: Procesamiento multi-modal de expresiones espaciales y temporales: Hacia el entendimiento del espacio y tiempo en el lenguaje ayudado por la visión" (código PCIN-2015-226, programa CHIST-ERA 2014) y, y ha participado como investigador en 6 proyectos europeos (LoCloud, KYOTO, PATHS, NEWSREADER, OPENER, MEANING). Además, ha participado en más de 10 proyectos de ámbito estatal (KNOW, KNOW2, BERBATEK, SKATER, READERS, ...).

Parte C. MÉRITOS MÁS RELEVANTES *(ordenados por tipología)*

C.1. Publicaciones

(Nota: las publicaciones estan ordenadas por fecha)

Autores (p.o. de firma): Josu Goikoetxea, Aitor Soroa eta Eneko Agirre

Título: Bilingual Embeddings with Random Walks over Multilingual Wordnets

Ref: Knowledge-Based Systems (KNOSYS). Volume 150, 15 June 2018, Pages 218-230. ISSN: 0950-7051.

Fecha: 2018

Assessment: Impact Factor 4.396 (Q1)

Autores (p.o. de firma) Zuhaitz Beloki and Xabier Artola and Aitor Soroa

Título: A scalable architecture for data-intensive natural language processing

Ref: Natural Language Engineering, 1-23, ISSN: 1351-3249

Fecha: 2017

Assessment: Impact Factor 0.8 (Q4)

Autores (p.o. de firma) Ander Barrena, Aitor Soroa, Eneko Agirre

Título: Alleviating Poor Context with Background Knowledge for Named Entity Disambiguation

Ref: Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics, pages 1903–1912, . ISBN:9781510827592

Fecha: 2016

Assessment: CORE A+

Autores (p.o. de firma): Marco Rospocher and Marieke van Erp and Piek Vossen and Antske Fokkens and Itziar Aldabe and German Rigau and Aitor Soroa and Thomas Ploeger and Tessel Bogaard

Título: Building Event-Centric Knowledge Graphs from News

Ref: Journal of Web Semantics. ISSN: 1570-8268

Fecha: 2016

Assessment: Impact Factor 1.075 (Q4)

Autores (p.o. de firma): Goikoetxea J., Agirre E., Soroa A.

Título: Single or Multiple. Combining Word Representations Independently Learned from Text and WordNet

Ref: Proc. of the Thirtieth AAAI Conference on Artificial Intelligence. pp. 2608-26014

Fecha: 2016

Assessment: CORE A+

Autores (p.o. de firma): Roller R., Agirre E., Soroa A., Stevenson M. (2015)

Título: Improving distant supervision using inference learning

Ref: Proceedings of the 53rd Annual Meeting of the Association of Computational Linguistics (ACL), pages 273-278. ISBN: 978-1-941643-73-0.

Fecha: 2015

Assessment: CORE A+

Autores (p.o. de firma): R. Agerri, X. Artola, Z. Beloki, G. Rigau, A. Soroa

Título: Big data for Natural Language Processing: A streaming approach

Ref: Knowledge-Based Systems. Vol.79, pages 36-42.

Fecha: 2014

Assessment: Impact Factor 2.947 (Q1)

Autores (p.o. de firma): D. Martinez, A. Otegi, A. Soroa, E. Agirre

Título: Improving search over Electronic Health Records using UMLS-based query expansion through random walks

Ref: Journal of Biomedical Informatics, vol. 51, pages 100-106, Elsevier. ISSN 1532-0464

Fecha: 2014

Assessment: Impact factor: 2.194 (Q1)

Autores (p.o. de firma): M. Hall, S. Fernando, P. Clough, A. Soroa, E. Agirre, M. Stevenson

Título: Evaluating hierarchical organisation structures for exploring digital libraries

Ref: Information Retrieval. Volume 17, Issue 4, Page 351-379. ISSN: 1386-4564

Fecha: 2014

Assessment: Impact Factor 2013: 1.220 (Q3)

Autores (p.o. de firma): Agirre E., Lopez de Lacalle O., Soroa A.

Título: Random Walks for Knowledge-Based Word Sense Disambiguation

Ref: Computational Linguistics. ISSN 0891-2017

Fecha: 2013

Assessment: Impact Factor 1.468 (Q2)

C.2. Proyectos

Referencia: KK-2017/00043

Título del proyecto: BERBAOLA: Tecnologías de la Lengua y el Habla aplicadas a los ámbitos del RIS3

Entidad financiadora: Gobierno Vasco, Dpto de Industria.

Investigador responsable: Aitor Soroa

Duración, desde: 2017 hasta: 2018

Cuantía de la subvención: 187.620

Tipo de participación: investigador Principal

Referencia: PCIN-2015-226

Título del proyecto: MUSTER

Entidad financiadora: Ministerio de Economía y Competitividad (programa CHIST-ERA 2015)

Duración, desde: 2016 hasta: 2019

Cuantía de la subvención: 75.000

Tipo de participación: investigador principal

Referencia: 325099

Título del proyecto: LoCloud: Local content in a Europeana cloud

Entidad financiadora: European Community (CIP Best Practice Network)

Duración, desde: 2013 hasta: 2016

Cuantía de la subvención: 3.400.000

Tipo de participación: investigador principal

Referencia: FP7-ICT-2011-8-316404

Título del proyecto: NewsReader: Building structured event indexes of large volumes of financial and economic data for decision making

Entidad financiadora: European Comission (STREP)

Investigador responsable: German Rigau (EHU)

Duración, desde: 2013 hasta: 2016

Cuantía de la subvención: 2.800.000

Tipo de participación: investigador

Referencia: FP7-ICT-2011-SME-DCL-296451

Título del proyecto: OpeNER: Open Polarity Enhanced Named Entity Recognition

Entidad financiadora: European Comission (STREP)

Investigador responsable: German Rigau (EHU)

Duración, desde: 2012 hasta: 2014

Cuantía de la subvención: 1.930.000

Tipo de participación: investigador

Referencia: ICT-2011-270082

Título del proyecto: PATHS – Personalized Access to Cultural Heritage Spaces

Entidad financiadora: European Commission (STREP)

Investigador responsable: Eneko Agirre (EHU)

Duración, desde: 2011 hasta: 2013

Cuantía de la subvención: 2.300.002

Tipo de participación: investigador

Referencia: TIN2012-38584-C06-02

Título del proyecto: SKaTeR: Scenario Knowledge Acquisition by Textual Reading

Entidad financiadora: Ministerio de Economía y Competitividad

Investigador responsable: German Rigau (EHU)

Duración, desde: 2013 hasta: 2015

Cuantía de la subvención: 372.915

Tipo de participación: investigador

C.3. Contratos

Título: Openminted: Sharing IXA pipes in the OpenMinTeD platform

Empresa o entidad: H2020 OPENMINTED.

Nombre del investigador principal y entidad de afiliación: Rodrigo Agerri. UPV/EHU.

Fecha de inicio y fin: 2018/02/15-2018/06/30

Cuantía: 17200,00 € (sin iva)

Título: : Servicio para la realización de un estudio sobre diferentes herramientas técnicas, métodos y arquitecturas existentes para desarrollar sistemas de procesamiento lingüístico de alto rendimiento.

Empresa o entidad: Secretaría de Estado de Telecomunicaciones y Sociedad de la Información.

Nombre del investigador principal y entidad de afiliación: Aitor Soroa. UPV/EHU.

Fecha de inicio y fin: 01/04/2016-31/07/2016

Cuantía: 17200,00 € (sin iva)

Título: ICTReader: Vigilancia tecnológica automática del sector TIC.

Empresa o entidad: Secretaría de Estado de Telecomunicaciones y Sociedad de la Información.

Nombre del investigador principal y entidad de afiliación: Eneko Agirre. UPV/EHU.

Fecha de inicio y fin: 21/09/2014-21/12/2014

Cuantía: 17425,31 € (sin iva)

C.4. Patentes

C.5 Miembro de Comité de programa

Congresos CORE A*: IJCAI

Congresos CORE A: COLING, WWW, EACL, EMNLP, NAACL

Revistas indexadas: AI, JNLE, LRE, EAAI