

Part A. INFORMACIÓN PERSONAL		Fecha CV		01/09/2023
Nombre y apellidos	RUBEN URIZAR ENBEITA			
DNI		Edad		
Números de investigador	Researcher ID	R-3820-2018		
	Código Orcid	0000-0001-9332-5151		

A.1. Puesto actual

Nombre de Universidad/Institución	Universidad del País Vasco / Euskal Herriko Unibertsitatea		
Departamento	Didáctica de la Lengua y la Literatura Facultad de Educación, Filosofía y Antropología		
Dirección	Oñati Plaza, 3 20018 Donostia - San Sebastián		
Nº teléfono	943 01 7009	Correo electrónico	ruben.urizar@ehu.eus
Puesto actual	Titular de Universidad	Desde	29/11/2019
Espec. cód. UNESCO	570107, 570111, 570104, 570504, 570513, 570505		
Palabras clave	Lengua, Didáctica de la Lengua, Procesamiento del Lenguaje Natural, Fraseología		

A.2. Educación

Grado/Doctorado	Universidad	Año
Licenciado en Filología Vasca	Universidad de Deusto	1989
Doctor en Filología Vasca	Universidad del País Vasco	2012

A.3. Artículos JCR, h Index, tesis dirigidas...

Google Scholar

	All	Since 2018
Citations	914	278
h-index	16	7
i10-index	18	6

Parte B. CV RESUMIDO (máx. 3500 caracteres, incluidos espacios)

He desarrollado la mayor parte de mi labor investigadora dentro del grupo de investigación IXA (ixa.si.ehu.es), que ha sido designado en repetidas ocasiones grupo de alto rendimiento por el Gobierno Vasco y la Universidad del País Vasco y actualmente es grupo de investigación consolidado de tipo A'(2010-2015) por el Gobierno Vasco. IXA tiene como objetivo el desarrollo de herramientas para el tratamiento automático de las lenguas, especialmente del euskera.

Soy autor de más de 40 publicaciones y 6 patentes, he participado en más de 30 proyectos de investigación en el campo de la lingüística computacional.

Participé en el proyecto europeo NewsReader (www.newsreader-project.eu) cuyo objetivo era la creación de índices de eventos estructurados de gran volumen (de datos financieros y económicos) creando storylines que ayuden en la toma de decisiones. Mi principal tarea fue la creación de un corpus goldstandard (en inglés y español) para la evaluación de sistemas automáticos.

Dentro del área de la lingüística computacional, mi interés se centra en el tratamiento computacional de las locuciones vascas (sobre las cuales realicé mi tesis doctoral), la morfología computacional, la lexicología...

Parte C. MÉRITOS RELEVANTES

C.1. Publicaciones (incluidos libros)

Autores : Conde A., Arruarte A., Larrañaga M., Jon A. Elorriaga J.A., Urizar R.

Título: Testu-liburuetatik domeinu-modulu eleaniztunak eraikitzen

Ref. ISSN: 0214-9001. e-ISSN: 2444-3255. Revista: Ekaia. Book:

Clave: A Volumen: 31 Páginas, primera: 117 última: 132 Fecha: 2017

Editorial (si se trata de un libro): UPV/EHU Press

Lugar de publicación:

Autores (en orden alfabético): Aduriz I., Aranzabe M.J., Arriola J.M., Díaz de Ilarraza A., Gonzalez-Dios I., Urizar R.

Título: Building the Gold Standard for the Surface Syntax of Basque

Ref. ISSN (edición impresa): 1135-5948, ISSN (edición electrónica): 1989-7553. Revista:

Procesamiento del Lenguaje Natural. Book:

Clave: A Volumen: 56 Páginas, primera: 125 última: 132 Fecha: 2017

Editorial (si se trata de un libro): Sociedad Española para el Procesamiento del Lenguaje Natural (SEPLN)

Lugar de publicación:

Autores (en orden alfabético): Aduriz, I., Urizar R. (ed.)

Título: Euskal hizkuntzalaritzaren egungo zenbait ikerlerro. Hizkuntzalari euskaldunen I. topaketa (ISBN: 978-84-8438-524-0).

Clave: Libro Volumen: Páginas, primera: última: Fecha: 2014

Editorial (si se trata de un libro): Udako Euskal Unibersitatea (UEU)

Lugar de publicación: Bilbao

Autores :Urizar R.,

Título: TESIS: Euskal Lokuzioen Tratamendu Konputazionala / Tratamiento Computacional de las Locuciones en Euskara

Clave: Libro Volumen: Páginas, primera: última: Fecha: 2012

Editorial (si se trata de un libro):

Lugar de publicación: Departamento de Filología Vasca - Universidad del País Vasco

Autores (in orden de firma): R. Urizar, I. Alegria, J.C. Odriozola, N. Ezeiza

Título: Euskarazko hitz anitzeko unitate lexikalen tratamendu konputazionala

Ref. Revista: Beñat Oihartzabali Gorazarre. Anuario del Seminario de Filología Vasca Julio Urquijo

Book:

Clave: A Volumen: XLIII:1-2 Páginas, primera: 891 última: 908 Fecha: 2009

Editorial (si se trata de un libro):

Lugar de publicación: Vitoria-Gasteiz

C.2. Proyectos de investigación y becas

Título del proyecto: DEEPTXT

Entidad financiadora: Gobierno Vasco. Departamento de Industria.

Entidades participantes: EHU/UPV

Duración, desde: 28-02-2020 hasta: 31-12-2021 Cuantía de la subvención: 153.496 €

Investigación principal: Aitor Soroa

Número de investigadores participantes: 25 (grupo IXA)

Título del proyecto: TANDO: Métodos y Sistemas de Traducción Automática Neuronal Coherente

Entidad financiadora: Gobierno Vasco

Entidades participantes: VicomTech; UPV/EHU; AMETZAGAIÑA A.I.E.; ISEA S. COOP; Elhuyar fundazioa

Duración, desde: 28-02-2020 hasta: 31-12-2021 Cuantía de la subvención: 79.208 €

Investigación principal: Kepa Sarasola

Número de investigadores participantes: 14 (grupo IXA)

Título del proyecto: DOMINO: Traducción Automática Neuronal, en DOMInio, NO supervisada

Entidad financiadora: Ministerio de Ciencia, Innovación y Universidades

Entidades participantes: EHU/UPV

Duración, desde: 01-01-2019 hasta: 31-12-2021 Cuantía de la subvención: 81.070 €

Investigación principal: Eneko Agirre, Kepa Sarasola

Número de investigadores participantes: 13 (grupo IXA)

Título del proyecto: LINGUATEC Desarrollo de la cooperación transfronteriza y la transferencia de conocimiento en tecnologías de la lengua

Entidad financiadora: Comisión Europea

Entidades participantes: Elhuyar fundazioa, Lo congrés permanent de la lenga occitana, EHU/UPV, CNRS, Euskaltzaindia, Sociedad de Promoción del Turismo Aragonés, SLU

Duración, desde: 01-01-2018 hasta: 31-12-2020 Cuantía de la subvención: 797.875 €

Investigación principal: Kepa Sarasola

Número de investigadores participantes: 8 (grupo IXA)

Título del proyecto: QUALES: Aprendizaje Automático mediante Supervisión Modulable para Estimación Automática de Calidad de Traducción

Entidad financiadora: Gobierno Vasco

Entidades participantes: VicomTech; UPV/EHU; AMETZAGAIÑA A.I.E.; ISEA S. COOP; Elhuyar fundazioa

Duración, desde: 01-07-2017 hasta: 31-12-2018 Cuantía de la subvención: 336.000 €

Investigación principal: Thierry Etchegoyhen, Kepa Sarasola

Número de investigadores participantes: 10 (grupo IXA)

Título del proyecto: Grupo de investigación UPV/EHU

Entidad financiadora: Universidad del País Vasco (UPV/EHU)

Entidades participantes: EHU/UPV

Duración, desde: 22-12-2016 hasta: 21-12-2019 Cuantía de la subvención: 135.000 €

Investigación principal: Arantza Díaz de Ilarraza

Número de investigadores participantes: 32

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

Entidad Financiadora: Comisión Europea (ICT Call 8 FP7-ICT-2011-8-316404)

Entidades participantes: VU University Amsterdam (VUA), Universidad del País Vasco, Fondazione Bruno Kessler (FBK), LexisNexis (LN), ScraperWiki (SW), SynerScope (SYN)

Duración, desde: 01/01/2013 hasta: 31/12/2015 Cuantía de la subvención: 2.800.000 €

Investigación principal: Piek Vosen

Número de investigadores participantes: 18 (Subproyecto UPV)

C.3. Contratos

C.4. Patentes

Inventores (p.o. alfabético): Aldezabal I., Aranzabe M.J., Arriola J.M., Artola X, Atutcha, A., Díaz de Ilarraza A., Estarrona, A., Ezeiza N., Fernandez-Terrones, E., Urizar R.

Título: EPEC-DEP (BDT): Corpus anotado a nivel sintáctico basándose en la Gramática de Dependencias

Aplicación número: SS-131-12 País de prioridad: España Fecha de prioridad: 25-04-2012

Entidad: Universidad del País Vasco

Países a los que se ha extendido:

Empresas que la explotan:

Inventores (p.o. alfabético): Alegría I., Aldezabal I., Aranzabe M.J., Arregi O., Arriola J.M., Artola X, Bengoetxea F.J., Ceberio K., Díaz de Ilarraza A., Ezeiza N., Gojenola K., Lersundi M., Oronoz M., Urizar R.

Título: IXAti: Detector de "Chunkers". Analizador sintáctico superficial para el euskara

Aplicación número: 8 País de prioridad: España Fecha de prioridad: 07-03-2008

Entidad: Universidad del País Vasco

Países a los que se ha extendido:

Empresas que la explotan: Eleka

C.5, C.6, C.7... (por ej., responsabilidades institucionales, pertenencia a sociedades científicas...)

Participación en comités de evaluación

Título del Comité: Joint Workshop on Linguistic Annotation, Multiword Expressions and Constructions (LAW-MWE-CxG-2018) COLING 2018

Entidad de la que depende: SIGLEX (a Special Interest Group on the Lexicon of the Association for Computational Linguistics ACL)

Área: Multiword Expressions (Expresiones Multipalabra)

Título del Comité: The 13th Workshop on Multiword Expressions (MWE 2017)

Entidad de la que depende: European Chapter of the Association for Computational Linguistics (EACL)

Área: Multiword Expressions (Expresiones Multipalabra)
