Egoitz Laparra

School of Information

University of Arizona

1103 E. 2nd St., 85716

Tucson, AZ, USA

☎: +1 (520) 247 7459

✉: laparra@arizona.edu

⌂: https://egolaparra.github.io

𝔤: sCY2JkMAAAAJ

Research Interests

Natural Language Processing, Information Extraction, Domain and Cross-lingual adaptation, Machine Learning, Deep Learning

Education

2015	Ph.D. in Computer Science.
	University of the Basque Country. Computer Science Faculty.
2009	M.S. in Language Analysis and Processing
	University of the Basque Country. Computer Science Faculty.
2007	B.S. in Computer Science
	University of the Basque Country. Computer Science Faculty.
2004	B.S. in Computer Management
	University of Mondragon. Business Science Faculty.

Work Experience

2023–	Adjunct Instructor.
	University of Arizona. School of Information.
2021–	Researcher/Scientist.
	University of Arizona. School of Information.
2017–2021	Post-doctoral research associate.
	University of Arizona. School of Information.
2015–2017	Post-doctoral research associate.
	University of the Basque Country. Computer Languages and Systems.
2012–2015	Part-time research associate.
	University of the Basque Country. Computer Languages and Systems.

Teaching Experience

Mentoring

2023	Maddalen Lopez de Lacalle	PhD. Thesis	Co-advisor

Courses

2023	INFO 555: Applied Natural Language Processing University of Arizona (in-person)	Lecturer
2023	INFO 555: Applied Natural Language Processing University of Arizona (online)	Lecturer
2021	INFO 507: Information Research Methods	Guest Lecturer

2016	University of Arizona	
2016	HAP5: Programming for NLP (Spanish)	Instructor
	University of the Basque Country	
2016	LAP5: Programming for NLP (English)	Instructor
	University of the Basque Country	
2013	16282: Software Engineering	Teaching Assistant
	University of the Basque Country	

Grants and Awards

2011	Complementary Grant for International Stay. Spanish Ministry of Science and Innovation.
2008–2012	Research Staff Training (FPI) Grant. Spanish Ministry of Science and Innovation.
2008	Accommodation Grant for "Information extraction and visualization techniques" course. Dukes of Soria Foundation.
2008	Research Staff Training Grant. University of The Basque Country.
2006	General University Scholarship. University of the Basque Country.

Grant Writing Experience

[1] 2021   Call: PA20-185   US National Institutes of Health
Automated Domain Adaptation for Clinical Natural Language Processing
Principal investigator: Timothy A. Miller

[2] 2017   Call: H2020-SC6-CULT-COOP-2016-2017   European Commission
Deep and Efficient Library Processing for Historical Inquiry
Principal investigator: German Rigau

[3] 2009   Call: Complementary Actions to Research Projects (2009)   Spanish Government
Multilingual Central Repository 2.0: IXA
Principal investigator: German Rigau

Research Activity

Research Stays

2011/04 – 2011/07   Faculteit der Letteren. Vrije Universiteit. Amsterdam
Supervisor: Piek Vossen

Participation in Research Funded Projects

[1] 2018/09 - 2022/12   1831551   US National Science Foundation, SBE, SMA
A Data Science Platform and Mechanisms for Its Sustainability
Principal investigator: Laura Lopez-Hoffman

[2] 2017/11 - 2022/04   World Modelers program   DARPA
Eidos: An Open-Domain Machine Reading System
Principal investigator: Mihai Surdeanu

[3] 2017/02 - 2018/09   R01LM010090   US National Institutes of Health, NLM
Temporal Relation Discovery for Clinical Text
Principal investigator: Guergana Savova

[4] 2013/01 - 2015/12   FP7-ICT-2011-8-316404   European Comission
NewsReader: Building structured event indexes of large volumes of financial and economic data for decision making
Principal investigator: Piek Vossen

[5] 2013/01 - 2015/12   TIN2012-38584-C06-02   Spanish Government
SKaTeR: Scenario Knowledge Acquisition by Textual Reading
Principal investigator: Horacio Rodríguez

[6] 2013/01 - 2015/12   PCIN-2013-003-C02-02   CHIST-ERA, MINECO
READERS: evaluation and development of reading systems
Principal investigator: Anselmo Peñas

[7] 2012/07 - 2014/06   FP7-ICT-2011-SME-DCL-296451   European Commission
OpeNER: Open Polarity Enhanced Named Entity Recognition
Principal investigator: Seán Gaines

[8] 2010/01 - 2012/12   TIN2009-14715-C04-01   Spanish Government
KNOW2: Tecnologías de compresión del lenguaje para el acceso multilingüe a la información orientada a dominios
Principal investigator: Eneko Agirre

[9] 2010/01 - 2011/01   TIN2009-06135-E   Spanish Government
TIMM: Red temática de tratamiento de información multilingüe y multimodal
Principal investigator: Alfonso Ureña

[10] 2010/01 - 2010/6   INF10/65   University of the Basque Country
EHU-Azpiegitura2010: Prozesadore linguistikoetarako zerbitzari dedikatuaren berritzea
Principal investigator: Eneko Agirre

[11] 2009/11 - 2011/11   TEC2009-06876-E/TEC   Spanish Government
RTTH: Red temática en tecnologías del habla
Principal investigator: Ruben San Segundo Hernándes

[12] 2009/11 - 2011/12   IE09-262   Basque Government
Berbatek: Hizkuntz industria sustatzeko tresnak eta teknologiak
Principal investigator: Xabier Arregi Iparagirre

[13] 2008/03 - 2011/03   ICT-2007-211423   STREP European Community
KYOTO: Knowledge yielding ontologies for transition-based organizations
Principal investigator: Piek Vossen

Publications

Journals

[1] Laparra, E., Binford-Walsh, A., Emerson, K., Miller, M., López-Hoffman, L., Currim, F. and Bethard, S. 2023. **Addressing structural hurdles for metadata extraction from environmental impact statements.** *Journal of the American Society for Information Science and Technology.*

[2] Laparra E., Mascio A., Velupillai S. and Miller T. 2021. **A Review of Recent Work in Transfer Learning and Domain Adaptation for Natural Language Processing of Electronic Health Records.** *IMIA Yearbook of Medical Informatics, Volume 30.*

[3] Laparra E., Bethard S. and Miller T. 2020. **Rethinking domain adaptation for machine learning over clinical language.** *JAMIA Open, Volume 3.*

[4] Laparra E., Xu D. and Bethard S. 2018. **From Characters to Time Intervals: New Paradigms for Evaluation and Neural Parsing of Time Normalizations.** *Transactions of the Association for Computational Linguistics, Volume 6.*

[5] Laparra E., Aldabe I., Agerri R. and Rigau G. 2017. **Multi-lingual and Cross-lingual TimeLine Extraction.** *Knowledge-Based Systems, Volume 133.*

[6] Lopez de Lacalle M., Laparra E., Aldabe I. and Rigau G. 2016. **Predicate Matrix. Automatically extending the semantic interoperability between predicate resources.** *Language Resources and Evaluation, Volume 50(2).*

[7] Vossen P., Agerri R., Aldabe I., Cybulska A., van Erp M., Fokkens A., Laparra E., Minard A., Palmero A., Rigau G., Rospocher M. and Segers R. 2016. **NewsReader: using knowledge resources in a cross-lingual reading machine to generate more knowledge from massive streams of news.** *Knowledge-Based Systems, Volume 110.*

[8] Agerri R., Agirre E., Aldabe I., Altuna B., Beloki Z., Laparra E., Lopez de Lacalle M., Rigau G., Soroa A. and Urizar R. 2015. **NewsReader project.** *Procesamiento del Lenguaje Natural, Volume 53.*

[9] Alonso L., Castellón I, Laparra E. and Rigau G. 2009. **Evaluación de métodos semi-automáticos para la conexión entre FrameNet y SenSem.** *Procesamiento del Lenguaje Natural, Volume 43.*

[10] Agirre E., Casillas A., Díaz de Ilarraza A., Estarrona A., Fernandez K., Gojenola K., Laparra E., Rigau G. and Soroa A. 2009. **The KYOTO Project.** *Procesamiento del Lenguaje Natural, Volume 43.*

Conferences

[1] Laparra E. and Bethard S. 2020. **A Dataset and Evaluation Framework for Complex Geographical Description Parsing.** *In Proceedings of the 28th International Conference on Computational Linguistics (COLING'20).*

[2] Xu D., Laparra E. and Bethard S. 2019. **Contextualized Character Embeddings Lead to Major Improvements in Time Normalization: a Detailed Analysis.** *Proceedings of the 8th Joint Conference on Lexical and Computational Semantics (*SEM'19).*

[3] Sharp R., Pyarelal A., Gyori B., Alcock K., Laparra E., Valenzuela-Escarcega M., Nagesh A., Yadav V., Bachman J., Tang Z., Lent H., Luo F., Paul M., Bethard S., Barnard K., Morrison C. and Surdeanu M. 2019. **Eidos, INDRA, & Delphi: From Free Text to Executable Causal Models.** *Proceedings of the 17th Annual Conference of the North American Chapter of the Association for Computational Linguistics (Demo Track) (NAACL'19).*

[4] Lopez de Lacalle M., Laparra E., Aldabe I and Rigau G. 2016. **A Multilingual Predicate Matrix.** *Proceedings of the 10th Language Resources and Evaluation Conference (LREC'16).*

[5] Segers R., Rospocher M., Vossen P., Laparra E., Rigau G. and Minard A. 2016. **The Event and Implied Situation Ontology (ESO): Application and Evaluation.** *Proceedings of the 10th Language Resources and Evaluation Conference (LREC'16).*

[6] Segers R., Rospocher M., Vossen P., Laparra E., Rigau G. and Minard A. 2016. **The Predicate Matrix and the Event and Implied Situation Ontology: Making More of Events.** *Proceedings of the 8th Global WordNet Conference (GWC 2016).*

[7] Laparra E., Aldabe I. and Rigau G. 2015. Document Level Time-anchoring for TimeLine Extraction. *Proceedings of 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing, (ACL-IJCNLP'15).*

[8] Lopez de Lacalle M., Laparra E. and Rigau G. 2014. Predicate Matrix: extending SemLink through WordNet mappings. *Proceedings of the 9th Language Resources and Evaluation Conference (LREC'14).*

[9] Lopez de Lacalle M., Laparra E. and Rigau G. 2014. First steps towards a Predicate Matrix. *Proceedings of the 7th Global WordNet Conference (GWC'14).*

[10] Laparra E. and Rigau G. 2013. ImpAr: A Deterministic Algorithm for Implicit Semantic Role Labelling. *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics, (ACL'13).*

[11] Laparra E. and Rigau G. 2013. Sources of Evidence for Implicit Argument Resolution. *Proceedings of the 10th International Conference on Computational Semantics, (IWCS'13).*

[12] Laparra E. and Rigau G. 2012. Exploiting Explicit Annotations and Semantic Types for Implicit Argument Resolution. *Proceedings of the 6th IEEE International Conference on Semantic Computing, (ICSC'12).*

[13] Laparra E. and Rigau G. and Vossen P. 2012. Mapping WordNet to the Kyoto ontology. *Proceedings of 8th international conference on Language Resources and Evaluation, (LREC'12).*

[14] Gonzalez-Agirre A., Laparra E. and Rigau G. 2012. Multilingual Central Repository version 3.0. *Proceedings of 8th international conference on Language Resources and Evaluation, (LREC'12).*

[15] Gonzalez-Agirre A., Laparra E. and Rigau G. 2012. Multilingual Central Repository version 3.0: upgrading a very large lexical knowledge base. *Proceedings of the 6th Global WordNet Conference, (GWC'12).*

[16] Laparra E., Rigau G. and Cuadros M. 2010. Exploring the integration of WordNet and FrameNet. *Proceedings of the 5th Global WordNet Conference, (GWC'10).*

[17] Laparra E. and Rigau G. 2010. eXtended WordFrameNet. *Proceesdings of 7th international conference on Language Resources and Evaluation, (LREC'10).*

[18] Vossen P., Bosma W., Cuadros M., Laparra E. and Rigau G. 2010. Integrating a large domain ontology of species into WordNet. *Proceedings of 7th international conference on Language Resources and Evaluation, (LREC'10).*

[19] Alecha M., Álvez J., Hermo M. and Laparra E. 2009. A New Proposal for Using First-Order Theorem Provers to Reason with OWL DL Ontologies. *Proceedings of Spanish Conference on Programming and Computer Languages, (PROLE'09).*

[20] Laparra E. and Rigau G. 2009. Integrating WordNet and FrameNet using a knowledge-based Word Sense Disambiguation algorithm. *Proceedings of Recent Advances in Natural Language Processing, (RANLP09).*

[21] Álvez J., Atserias J., Carrera J., Climent S., Laparra E., Oliver A. and Rigau G. 2008. Complete and Consistent Annotation of WordNet using the Top Concept Ontology. *Proceedings of 6th international conference on Language Resources and Evaluation, (LREC'08).*

Workshops

[1] Surdeanu M., Hungerford J., Chan Y.S., MacBride J., Gyori B., Zupon A., Tang Z., Qiu H., Min B., Zverev Y., Hilverman C., Thomas M., Andrews W., Alcock K., Zhang Z., Reynolds M., Bethard S., Sharp R., Laparra E. 2022 Taxonomy Builder: a Data-driven and User-centric Tool for Streamlining Taxonomy Construction. *Proceedings of the Second Workshop on Bridging Human–Computer Interaction and Natural Language Processing (HCI+NLP'22).*

[2] Laparra E., Su X., Zhao Y., Uzumer Ö., Miller T. and Bethard S. 2021. SemEval-2021 Task 10: Source-Free Domain Adaptation for Semantic Processing. *Proceedings of The 15th International Workshop on Semantic Evaluation (SemEval'21).*

[3] Miller T., Laparra E. and Bethard S. 2021. Domain adaptation in practice: Lessons from a real-world information extraction pipeline. *Proceedings of the Second Workshop on Domain Adaptation for NLP (AdaptNLP'21).*

[4] Yadav V., Laparra E., Wang T., Surdeanu M. and Bethard S. 2019. University of Arizona at SemEval-2019 Task 12: Deep-Affix Named Entity Recognition of Geolocation Entities. *Proceedings of The 13th International Workshop on Semantic Evaluation (SemEval'19).*

[5] Bethard S., Laparra E., Wang S., Zhao Y. and Al-Ghezi R. 2019. Inferring missing metadata from environmental policy texts. *Proceedings of The 3rd Joint SIGHUM Workshop on Computational Linguistics for Cultural Heritage, Social Sciences, Humanities and Literature (LaTeCH-CLfL'19).*

[6] Laparra E., Xu D., Elsayed A., Bethard S. and Palmer M. 2018. SemEval 2018 Task 6: Parsing Time Normalizations. *Proceedings of The 12th International Workshop on Semantic Evaluation (SemEval'18).*

[7] Bell D., Laparra E., Kousik A., Ishihara T., Surdeanu M. and Kobourov S. 2018. Detecting Diabetes Risk from Social Media Activity. *Proceedings of the 9th International Workshop on Health Text Mining and Information Analysis (LOUHI'18).*

[8] Laparra E., Aldabe I. and Rigau G. 2015. From TimeLines to StoryLines: A preliminary proposal for evaluating narratives. *Proceedings of 1st Workshop on Computing News StoryLines at ACL-IJCNLP 2015, (CNewS'15).*

[9] Vossen P., Laparra E., Aldabe I. and Rigau G. 2015. Semantic Interoperability for Cross-lingual and cross-document Event Detection. *Proceedings of 3rd Workshop on EVENTS at NAACL-HLT 2015, (EVENTS'15).*

[10] Segers R., Vossen P., Rospocher M., Serafini L., Laparra E. and Rigau G. 2015. ESO: A Frame based Ontology for Events and Implied Situations. *Proceedings of the MAPLEX workshop, (MAPLEX'15).*

Presentations

[1] 2021. 15th International Workshop on Semantic Evaluation, (SemEval-2021). Online. *SemEval-2021 Task 10: Source-Free Domain Adaptation for Semantic Processing.*

[2] 2020. 28th International Conference on Computational Linguistics, (COLING-2020). Barcelona, Spain. *A Dataset and Evaluation Framework for Complex Geographical Description Parsing.*

[3] 2018. 56th Annual Meeting of the Association for Computational Linguistics, (ACL-2018). Melbourne, Australia. *From Characters to Time Intervals: New Paradigms for Evaluation and Neural Parsing of Time Normalizations.*

[4] 2018. 12th International Workshop on Semantic Evaluation, (SemEval-2018). New Orleans, USA. *SemEval 2018 Task 6: Parsing Time Normalizations.*

[5] 2015. 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing, (ACL-IJCNLP'15). Beijing, China. *Document Level Time-anchoring for TimeLine Extraction.*

[6] 2013. 10th International Conference on Computational Semantics, (IWCS'13). Potsdam, Alemania. *Sources of Evidence for Implicit Argument Resolution.*

[7] 2013. 51st Annual Meeting of the Association for Computational Linguistics, (ACL'13). Sofia, Bulgaria. *ImpAr: A Deterministic Algorithm for Implicit Semantic Role Labelling.*

[8] 2012. 6th IEEE International Conference on Semantic Computing, (ICSC'12). Palermo, Italia. *Exploiting Explicit Annotations and Semantic Types for Implicit Argument Resolution.*

[9] 2010. 5th Global WordNet Conference, (GWC'10). Mumbai, India. *Exploring the integration of WordNet and FrameNet.*

[10] 2010. 7th international conference on Language Resources and Evaluation, (LREC'10). Valleta, Malta. *Integrating a large domain ontology of species into WordNet.*

[11] 2010. 7th international conference on Language Resources and Evaluation, (LREC'10). Valleta, Malta. *eXtended WordFrameNet.*

[12] 2009. Recent Advances in Natural Language Processing, (RANLP09). Bovorets, Bulgaria. *Integrating WordNet and FrameNet using a knowledge-based Word Sense Disambiguation algorithm.*

Services

Academic Service

2022	Area Chair, International Conference on Computational Linguistics (COLING)
2021, 2022	Reviewer, ACL Rolling Review
2021	Organizer, Source-Free Domain Adaptation Shared Task (SemEval)
2020	Senior Program Committee, European Conference on Artificial Intelligence (ECAI)
2020	Conference Reviewer, European Conference on Artificial Intelligence (ECAI)
2020	Conference Reviewer, AAAI Conference on Artificial Intelligence (AAAI)
2020	Workshop Reviewer, Clinical Natural Language Processing Workshop (ClinicalNLP)
2019	Journal Reviewer, Expert Systems With Applications (ESWA)
2019	Journal Reviewer, Language Resources and Evaluation (LRE)
2018, 2019, 2021	Task Proposal Reviewer, SemEval
2018	Organizer, Parsing Time Normalizations Shared Task (SemEval)
2017, 2018, 2021	Journal Reviewer, Knowledge-Based Systems
2018	Journal Reviewer, JASIST
2018	Journal Reviewer, Fontes Lingvae Vaconvm
2018, 2020	Conference Reviewer, International Conference on Language Resources and Evaluation (LREC)
2016, 2018	Conference Reviewer, International Conference on Computational Linguistics (COLING)
2013, 2016, 2017, 2018	Conference Reviewer, Annual Meeting of the Association for Computational Linguistics (ACL)
2017	Grant Reviewer, Netherlands Organization for Scientific Research

2011   Conference Reviewer, International Joint Conference on Natural Language Processing
                (IJCNLP)

Departmental Service

2009-2010   Permanent Commission. Graduate Student Representative.
                University of the Basque Country. Computer Languages and Systems.