

# Curriculum Vitæ - Horacio Saggion

## Personal Information

Dr. Horacio Saggion  
Department of Information and Communication Technologies  
Universitat Pompeu Fabra  
C/Tanger 122  
Barcelona (08018)  
Spain  
Webpage: <http://www.dtic.upf.edu/~hsaggion>

## Education

- Ph.D. Computer Science, University of Montreal, Department of Computer Science and Operational Research (DIRO) Montréal, Québec, Canada. 1997-2000. Thesis: “Génération Automatique de Résumés par Analyse Sélective”.
- Master in Computer Science, University of Campinas (UNICAM), Computer Science Department (DCC), Mathematics, Statistics, and Computer Science Institute (IMECC), Campinas, São Paulo, Brazil. 1993 -1995. Master’s Thesis: “Análise automática de sumários em língua portuguesa: uma aproximação ao tratamento da estrutura de um texto”.
- Bachelor in Computer Science, University of Buenos Aires (UBA), Facultad de Ciencias Exactas y Naturales (FCEyN), Buenos Aires, Argentina. 1983-1988. Bachelor’s Thesis: “Un sistema de texto a habla para el español destinado a controlar un sintetizador fonémico”.

## Academic Positions

- 2015-present: Department of Information and Communication Technologies (DTIC), Universitat Pompeu Fabra, Barcelona, Spain  
Associate Professor (i.e. Senior Lecturer).  
Deputy director for Masters (since November 2016) and PhD (since January 2018)
- 2010-2015: Department of Information and Communication Technologies (DTIC), Universitat Pompeu Fabra, Barcelona, Spain  
Ramón y Cajal Research Professor (successfully passed evaluation for tenured position on March 2014).
- 2000-2010: University of Sheffield, UK  
Research Fellow (previously research associate) for the Natural Language Processing Group at the Department of Computer Science. Funded by EPSRC and EC.
- 2008: Universidad Nacional de Buenos Aires, Buenos Aires, Argentina  
Invited Professor for the Escuela de Ciencias Informáticas ECI -2008 (Computer Science School).
- 2004: Universidad de Río Cuarto, Córdoba, Argentina  
Invited Professor for the RIO 2004 summer school in Computer Science.

- 2001: Johns Hopkins University, USA  
Invited Researcher. 6-week Research Project on Summarization. National Science Foundation.
- 1998-1999: University of Montreal, Canada  
Teaching Assistant, part time, University of Montreal.
- 1986-1987 & 1991-1996: Computer Science Department, Facultad de Ciencias Exactas y Naturales, University of Buenos Aires, Argentina (Note that positions at University of Buenos Aires are awarded in public competition).  
1995-1996: Assistant Researcher, Project: "Construction of an empiric data base for research in Spanish language acquisition". Directors: Prof. Lic. Alejandro Raiter and Prof. Lic. Jose Alvarez.  
1995-1996: Assistant Professor, full time.  
1993-1994: Assistant Professor, full time.  
1992-1993: Teaching Assistant level 1, part time.  
1991-1992: Teaching Assistant level 2, part time.  
1986-1987: Teaching Assistant level 2, part time.
- 1992-1993: Science and Technical Secretary, University of Quilmes, Argentina  
Assistant Professor, part time.
- 1991-1992: Comisión Argentina de Informática (CAI, Post Secondary Institute), Buenos Aires, Argentina  
Teacher.

## Research Funding Obtained since 2010

- 2018 Minería de trabajos finales de grado: Un mapa de la relación entre la docencia y la investigación en el DTIC. Grant from Plan de Ayudas de Apoyo a la Calidad y a la Innovación en Aprendizaje y Conocimiento. Ref. 53 PlaCLIK (EUR 3.000,00)
- 2018 Scientific Summarization Services. Grant from OpenMinTed Call for Tenders. (EUR 17.000,00)
- 2018-2019 Metodologías y soluciones tecnológicas en el ámbito de las tecnologías del lenguaje. Grant from Spanish State Secretary / Red.es REF. 022/18-OT LOTE 1 (EUR 16.890,00)
- 2016-2018 TUNER-UPF - Técnicas Semisupervisadas de Extracción y Estructuración de Contenidos Semánticos para el Procesamiento del Lenguaje Natural. (TIN2015-65308-C5-5-R). Ministry of Economy and Competitiveness, Spain. UPF Principal Investigator (with Nuria Bel from IULA as co-PI). (EUR 72.600,00).
- 2014-2017 ABLE-TO-INCLUDE - An accessibility layer for the inclusion in the information society of people with intellectual disabilities. Funded under the Competitiveness and Innovation Programme of the European Commission (CIP-ICT-PSP-2013-7 /621055). UPF Principal Investigator. (EUR 144.365,00)
- 2014-2016 Dr. Inventor - Promoting Scientific Creativity by Utilising Web-based Research Objects. Funded under the 7th Framework Program of the European Commission (FP7-ICT-2013.8.1 / 611383). UPF Principal Investigator. (EUR 405.120,00)
- 2013-2015 Multisensor: Mining and Understanding of multilingual content for Intelligent Sentiment Enriched context and Social Oriented interpretation (FP7-ICT-610411). Involved in proposal preparation and writing. (EUR 408.640).

- 2013-2015 Scenario Knowledge Acquisition by Textual Reading: Knowledge-based Concise Summarization (TIN2012-38584-C06-03). Ministry of Economy and Competitiveness, Spain. UPF Principal Investigator. (EUR 12.402,00)
- 2013-2014 Development of tools and resources for Web Intelligent Applications (Ref. 65 PLAQUID 2013-2014), Pompeu Fabra University, Centre for Teaching Quality and Innovation. Principal Investigator. (EUR 1.500,00)
- 2010-2013 Simplext: A Text Simplification System for Spanish (TSI-020302-2010-84). Plan Avanza 2. Ministry of Industry, Tourism, and Trade. Spain. UPF Principal Investigator and Project Scientific Coordinator. (EUR 104.567,00)
- 2012-2013 Knowledge Acquisition and Clustering (SOL-00049025). The Centre for the Development of Industrial Technology (Centro para el Desarrollo Tecnológico Industrial, CDTI), Principal Investigator. (EUR 29.200,00)
- 2011-2012 Automatic Summarization Tools for Easy Access to Textual Information (Ref. 47 PLAQUID 2010-2011 1), Pompeu Fabra University, Centre for Teaching Quality and Innovation. Principal Investigator. (EUR 1.000,00)
- 2010-2015 Methods and Tools for Abstractive Text Summarization. Ramón y Cajal Advanced Research Fellowship (RYC-2009-04291). Ministry of Economy and Competitiveness, Spain. Individual Grant. (EUR 192.480 + EUR 15.000)
- 2009-2011 Natural Language Processing for Discourse Analysis in Psychology. Royal Society, United Kingdom. Principal Investigator. (aprox. EUR 10.000)
- 2010 Natural Language Processing Tools Applied to Discourse Analysis in Psychology (Herramientas de procesamiento del lenguaje natural aplicadas al análisis del discurso en psicología). FONSOFT/FONCYT, Argentina. Co-researcher. (aprox EUR 2.000)

## **Participation in Research Projects - in addition to above PI activities**

- 2012-2013 TOPAS - Tool platform for intelligent Patent Analysis and Summarization. Funded under the 7th Framework Program of the European Commission (FP7-SME-2011 286639). Researcher.
- 2010-2012 PESCaDO - Personalized Environmental Service Configuration and Delivery Orchestration. Funded under the 7th Framework Program of the European Commission (IST FP7-248594). Researcher.
- 2006-2008 MUSING (6 Framework Program Integrated Project of the European Commission). University of Sheffield, UK. Sheffield Team Manager and Researcher.
- 2003-2005 CUBREPORTER. Engineering and Physical Sciences Research Council ( (GR/R91465/01). University of Sheffield. Researcher.
- 2003: CLARITY: Cross-language Information Retrieval and Organisation of Text and Audio Documents (IST-2000-25310). University of Sheffield, UK. Researcher.
- 2000-2004 GATE: General Architecture for Text Engineering. Engineering and Physical Sciences Research Council (GR/M31699/01) . Researcher. University of Sheffield.
- 2001 SummBank. Johns Hopkins University, USA. Researcher.

- 2000-2002 MUMIS: Multimedia Indexing and Searching Environment. FP4 European Commission (IST-1999-10651). University of Sheffield, UK. Researcher.
- 2000-2002 SOCIS: Scene of the Crime Information System. Engineering and Physical Sciences Research Council (EPSRC/RA002260). University of Sheffield, UK. Researcher and Sheffield Coordinator.
- 1997-2000 SumUM: Summarization at Université de Montréal. University of Montreal. Canada. PhD project.

## Research Quality Indicators

- Catalan Research Quality Agency - Accreditation 26/02/2015
- I3 Evaluation for Researchers - State Secretary for Research and Development, Spain, 08/05/2014.
- Accreditation as Professor by the French Ministry of Education 04/02/2016 (Qualif. Number 16127157811).
- Researcher of the Ramón y Cajal program 2009.
- Web of Science. Docs: 25, h-index: 5, citations: 73, citations per document: 3,17.
- Google Scholar. h-index: 26, citations: 2,829
- Scopus. Number of documents: 51, h-index: 7, number of citations: 174, citations per document: 3,41.

## Research Profiles on the Web

- Google Scholar Profile: <http://scholar.google.es/citations?user=WMrCFCIAAAAJ&hl=en>
- DBLP: <http://www.informatik.uni-trier.de/~ley/pers/hd/s/Saggion:Horacio.html>
- Microsoft Academic: <http://academic.research.microsoft.com/Author/1927997/horacio-saggion>
- ORCID: <http://orcid.org/0000-0003-0016-7807>
- Research Gate: [https://www.researchgate.net/profile/Horacio\\_Saggion](https://www.researchgate.net/profile/Horacio_Saggion)
- ACM: [http://dl.acm.org/author\\_page.cfm?id=81100371947](http://dl.acm.org/author_page.cfm?id=81100371947)

## Journal Papers

NOTES: With asterisk (\*) I indicate the publications for which I am one of the lead authors. With *italics*, I indicate the most significant papers. Journals appear in JCR Science or JCR Social Sciences.

1. (\*) *Improving Lexical Coverage of Text Simplification Systems for Spanish*. Sanja Stajner, **Horacio Saggion** and Simone Paolo Ponzetti. Expert Systems With Applications (In Press).
2. *A Text Summarization Method based on Fuzzy Rules and applicable to Automated Assessment* Fábio Bif Goulartea, Silvia Modesto Nassara, Renato Fileto and **Horacio Saggion**. Expert Systems With Applications. Volume 115, January 2019, Pages 264-275
3. (\*) *Using genre-specific features for patent summaries*. Joan Codina-Filbla, Nadjat Bouayad-Agha, Alicia Burga, Gerard Casamayor, Simon Mille, Andreas Muller, **Horacio Saggion**, Leo Wanner. Inf. Process. Manage. 53(1): 151-174 (2017)

4. *Information extraction for knowledge base construction in the music domain.* Sergio Oramas, Luis Espinosa Anke, Mohamed Sordo, **Horacio Saggion**, Xavier Serra. *Data Knowl. Eng.* 106: 70-83 (2016)
5. (\*) *Simplifying words in context. Experiments with two lexical resources in Spanish.* **H. Saggion**, S. Bott, L. Rello. *Computer Speech & Language.* Volume 35, January 2016.
6. (\*) *Text Simplification Resources for Spanish.* S. Bott, **H. Saggion**. *Journal of Language Resources and Evaluation.* 2014.
7. (\*) *Making Numerical Information more Accessible: Implementation of a Numerical Expression Simplification System for Spanish.* S. Bautista, **H. Saggion**. *International Journal of Applied Linguistics.* 2014.
8. (\*) *Automatic Summarization: An Overview.* **Saggion, H.** *Revue Française de Linguistique Appliquée.* XIII(1). 2008.
9. Context-Based Generic Cross-Lingual Retrieval of Documents and Automated Summaries. Lam, W. & Chan, K. & Radev, D. & **Saggion, H.** & Teufel, S. *Journal of the American Society for Information Science and Technology.* 56(2), 2005, pp129-139.
10. (\*) *Multimedia indexing through multi-source and multi-language information extraction: the MUMIS project.* Saggion, H. & Cunningham, H. and & Bontcheva, K. & Maynard, D. & Hamza, O. & Wilks, Y. *Data & Knowledge Engineering.* Vol 48, Num 2, Feb. 2004, pp247-264.
11. (\*) *Extracting relational facts for indexing and retrieval of crime-scene photographs.* Pastra, K. & **Saggion, H.** & Wilks, Y. *Knowledge-Based Systems.* 16(2003). pp313-320.
12. (\*) *Intelligent Indexing of Crime Scene Photographs.* Pastra, K. & **Saggion, H.** & Wilks, Y. *IEEE Intelligent Systems.* Jan-Feb 2003, 18(1), pp55-61.
13. (\*) *Generating Indicative-Informative Summaries with SumUM.* **Saggion, H.** & Lapalme, H. *Computational Linguistics.* December 2002, 28(4), pp497-526.
14. *Architectural Elements of Language Engineering Robustness.* Maynard, D. & Tablan, V. & Cunningham, H. & Ursu, C. & **Saggion, H.** & Bontcheva, K. & Wilks, Y. *Journal of Natural Language Engineering – Special Issue on Robust Methods in Analysis of Natural Language Data.* 2000.

## Other Journal Publications (all peer reviewed)

NOTES: With asterisk (\*) I indicate the publications for which I am one of the lead authors. With *italics*, I indicate the most significant papers. *Computación y Sistemas* appears in the Thomson Reuters ISI - Master Journal List and SCOPUS. The *Journal of the Spanish Society for Natural Language Processing* is a CARHUS PLUS+ B journal and appears in SCOPUS.

1. *Savana: Re-using Electronic Health Records with Artificial Intelligence.* Ignacio Hernández Medrano, Jorge Tello Guijarro, Cristóbal Belda, Alberto Ureña, Ignacio Salcedo, Luis Espinosa Anke, **Horacio Saggion**. *IJIMAI* 4(7): 8-12 (2018)
2. (\*) *Is this Tweet Satirical? A Computational Approach for Satire Detection in Spanish.* F. Barbieri, F. Ronzano, **H. Saggion**. *Procesamiento del Lenguaje Natural* 55: 135-142 (2015) - **Best Paper Award**.
3. (\*) *A Web-based Text Simplification System for English.* D. Ferrés, M. Marimon, **H. Saggion**. *Procesamiento del Lenguaje Natural* 55, pp191-194, 2015.

4. (\*) *Summarization and Information Extraction in your Tablet*. F. Barbieri, F. Ronzano, **H. Saggion**. *Procesamiento del Lenguaje Natural* 55, pp203-206, 2015.
5. (\*) *Making It Simplext: Implementation and Evaluation of a Text Simplification System for Spanish*. **H. Saggion**, S. Stajner, S. Bott, S. Mille, L. Rello, B. Drndarevic. *ACM Transactions on Accessible Computing*. Volume 6 number 4. June 2015.
6. (\*) *Spanish Morphological Generation with Wide-Coverage Lexicons and Decision Trees*. Daniel Ferrés, Ahmed AbuRa'ed, **Horacio Saggion**. *Procesamiento del Lenguaje Natural* 58: 109-116 (2017)
7. (\*) *Towards content-oriented patent document processing: Intelligent patent analysis and summarization*. S. Brüggmann, N. Bouayad-Agha, A. Burga, S. Carrascosa, A. Ciaramella, M. Ciaramella, J. Codina-Filba, E. Escorsa, A. Judea, S. Mille, A. Müller, **H. Saggion**, P. Ziering, H. Schütze, L. Wanner. *World Patent Information*, Volume 40, March 2015, pp.30-42.
8. *Corpus-based Sentence Deletion and Split Decisions for Spanish Text Simplification*. Sanja, S. & Drndarevic, B. & **Saggion, H.** *Computación y Sistemas*. 17(2). 2013.
9. (\*) *Reducing Text Complexity through Automatic Lexical Simplification: an Empirical Study for Spanish*. Drndarevic, B. & **Saggion, H.** *Journal of the Spanish Society for Natural Language Processing*. 49. pp13-20. 2012.
10. (\*) *Spanish Text Simplification: An Exploratory Study*. Bott, S. & **Saggion, H.** *Journal of the Spanish Society for Natural Language Processing*. 47(2). pp87-95.
11. (\*) *Text Simplification in Simplext. Making Text More Accessible*. **Saggion, H.** & Gómez Martínez, E. & Etayo, E. & Anula, A. Bourg, L. *Journal of the Spanish Society for Natural Language Processing*. 47(2). pp341-342.
12. *Análisis de la Simplificación de Expresiones Numéricas en Español mediante un estudio Empírico*. Bautista, S. & Drndarevic, B. & Hervás, R. & **Saggion, H.** & Gervás, P. *Linguamática* 4(2). pp27-41. 2012.
13. *Summary Evaluation With and Without References*. Torres Moreno, JM. & **Saggion, H.** & da Cunha, I. & Velazquez-Morales, P. & SanJuan, E. & Molina, A. *Polibits* 42. Thematic Issue in Artificial Intelligence. pp13-19. 2010.
14. *Extracting Opinions and Facts for Business Intelligence* **Saggion, H.** & Funk, A. *RNTI Journal*, E(17), pp119-146. 2009.
15. (\*) *SUMMA: A Robust and Adaptable Summarization Tool*. **Saggion, H.** *Revue Traitement Automatique des Langues*. 49(2). 2008.
16. *English Information Extraction from Football Reports in the Context of the MUMIS Project*. Saggion, H & Cunningham, H. & Wilks, Y. & Declerck, T. & Wittenburg, P. *ELSNews* 11(2) Summer 2002.

## Conference Papers (all peer reviewed)

NOTES: With asterisk (\*) I indicate the publications for which I am one of the lead authors. With *italics*, I indicate the most significant papers. According to the Computer Science Conference Rankings (CORE) conferences where I have published have the following quality indicators: AAAI A+, ACL A+, IJCAI A+, COLING A, EACL A, NAACL A, INTERACT A, ISWC A, ECAI A, CICLING B, IJCNLP B, EPIA B, LREC C, NLDB C, ICCHP C, TALN C.

1. (\*) *PDFdigest: an Adaptable Layout-Aware PDF-to-XML Textual Content Extractor for Scientific Articles*. Daniel Ferrés, **Horacio Saggion** and Francesco Ronzano, Àlex Bravo. LREC 2018.

2. *Are Emojis Predictable?* Francesco Barbieri, Miguel Ballesteros, **Horacio Saggion**. EACL (2) 2017: 105-111
3. *What Sentence are you Referring to and Why? Identifying Cited Sentences in Scientific Literature.* Ahmed AbuRa'ed, Luis Chiruzzo, **Horacio Saggion**. RANLP 2017: 9-17
4. (\*) *MultiScien: a Bi-Lingual Natural Language Processing System for Mining and Enrichment of Scientific Collections.* Horacio Saggion, Francesco Ronzano, Pablo Accuosto, Daniel Ferrés. BIRNDL@SIGIR (1) 2017: 26-40
5. (\*) *Able to Read My Mail: An Accessible e-Mail Client with Assistive Technology.* Horacio Saggion, Daniel Ferrés, Leen Sevens, Ineke Schuurman, Marta Ripolles, Olga Rodriguez. W4A 2017: 5:1-5:4 (Best Communication Award)
6. *Extending WordNet with Fine-Grained Collocational Information via Supervised Distributional Learning.* Luis Espinosa Anke, Jose Camacho-Collados, Sara Rodriguez-Fernandez, **Horacio Saggion**, Leo Wanner. COLING 2016: 3422-3432
7. *Supervised Distributional Hypernym Discovery via Domain Adaptation.* Luis Espinosa Anke, Jose Camacho-Collados, Claudio Delli Bovi, **Horacio Saggion**. EMNLP 2016: 424-435
8. (\*) *ExTaSem! Extending, Taxonomizing and Semantifying Domain Terminologies.* L. Espinosa-Anke, F. Ronzano, **H. Saggion**, R. Navigli. Accepted for publication in AAAI-2016.
9. (\*) *A Deeper Exploration of the Standard PB-SMT Approach to Text Simplification and its Evaluation.* S. Stajner, H. Béchara, **H. Saggion**. In Proceedings of ACL, 2015, pp823-828.
10. Automatic Text Simplification for Spanish: Comparative Evaluation of Various Simplification Strategies. Stajner, S., Calixto, I. and **Saggion, H.** 2015. In Proceedings of the Recent Advances in Natural Language Processing (RANLP), Hissar, Bulgaria, 5-11 September 2015. pp. 618-626.
11. Translating from Original to Simplified Sentences using Moses: When does it Actually Work? Stajner, S. and **Saggion, H.** 2015. In Proceedings of the Recent Advances in Natural Language Processing (RANLP), Hissar, Bulgaria, 5-11 September 2015. pp. 611-617.
12. How Topic Biases Your Results? A Case Study of Sentiment Analysis and Irony Detection in Italian. F. Barbieri, F. Ronzano, **H. Saggion**. In Proceedings of the Recent Advances in Natural Language Processing (RANLP), Hissar, Bulgaria, 5-11 September 2015.
13. Weakly Supervised Definition Extraction. L. Espinosa, **H. Saggion**, F. Ronzano. In Proceedings of the Recent Advances in Natural Language Processing (RANLP), Hissar, Bulgaria, 5-11 September 2015.
14. (\*) Stimulating and Simulating Creativity with Dr Inventor. D. O'Donoghue, Y. Abgaz, D. Hurley, F. Ronzano, **H. Saggion**. In Proceedings of the Sixth International Conference on Computational Creativity (ICCC 2015).
15. (\*) Hypernym Extraction: Combining Machine-Learning and Dependency Grammar. L. Espinosa-Anke, F. Ronzano, **H. Saggion**. In Proceedings of CICLing, 2015, pp372-383.
16. (\*) *Dr. Inventor Framework: Extracting Structured Information from Scientific Publications.* F. Ronzano, **H. Saggion**. In Discovery Science 2015, pp209-220.
17. (\*) *Do We Criticise (and Laugh) in the Same Way? Automatic Detection of Multi-Lingual Satirical News in Twitter.* F. Barbieri, F. Ronzano, **H. Saggion**. In Proceedings of IJCAI 2015, pp1215-1221.

18. (\*) Porting a methodology for syntactic simplification from English to Spanish. M. Marimon, **H. Saggion** and D. Fereés. Workshop on Replicability and Reproducibility in Natural Language Processing: adaptive methods, resources and software. IJCAI 2015.
19. (\*) Towards Dr Inventor: A Tool for Promoting Scientific Creativity. D. P. O’Donoghue, **H. Saggion**, D. Hurley, Y. Abgaz, F. Dong, X. Zheng, O. Corcho, J.J. Zhang, J-M Careil, B. Mahdian and X. Zhao. In Proceedings of the International Conference on Computational Creativity (ICCC). 2014.
20. On the Automated Generation of Scholarly Publishing Linked Datasets: The Case of CEUR-WS Proceedings. F. Ronzano, B. Fisas, G. Casamayor, **H. Saggion**. Semantic Web Evaluation Challenges 2015. Volume 548 of the series Communications in Computer and Information Science. p 177-188.
21. Applying Dependency Relations to Definition Extraction. L. Espinosa, **H. Saggion**. 2014. 19th International Conference on Application of Natural Language to Information Systems.
22. Italian Irony Detection in Twitter: a First Approach. F. Barbieri, F. Ronzano, **H. Saggion**. In Proceedings of The First Italian Conference on Computational Linguistics 2014.
23. Automatic Detection of Irony and Humour in Twitter. F. Barbieri, **H. Saggion**. In Proceedings of the International Conference on Computational Creativity (ICCC). 2014.
24. (\*) Semantify CEUR-WS Proceedings: Towards the Automatic Generation of Highly Descriptive Scholarly Publishing Linked Datasets. F. Ronzano, G. Casamayor, **H. Saggion**. SemWebEval@ESWC 2014, pp83-88.
25. (\*) Creating Summarization Systems with SUMMA (DEMO). **Saggion, H.** In Proceedings of the Language Resources and Evaluation Conference (LREC). 2014.
26. *Modelling Irony in Twitter: Feature Analysis and Evaluation*. Barbieri, F. and **Saggion, H.** In Proceedings of the Language Resources and Evaluation Conference (LREC). 2014.
27. (\*) Can Numerical Expressions Be Simpler? Implementation and Demonstration of a Numerical Simplification System for Spanish (DEMO). Bautista, S. and **Saggion, H.** In Proceedings of the Language Resources and Evaluation Conference (LREC). 2014.
28. Keyword Highlighting Improves Comprehension for People with Dyslexia. Rello, L, and **Saggion, H.** and Baeza-Yates, R. In Proceedings of the 3rd Workshop on Predicting and Improving Text Readability for Target Reader Populations (an EACL workshop). 2014.
29. One Step Closer to Automatic Evaluation of Text Simplification Systems. Stajner, S. and Mitkov, R. and **Saggion, H.** In Proceedings of the 3rd Workshop on Predicting and Improving Text Readability for Target Reader Populations (an EACL workshop). 2014.
30. (\*) *Modelling Irony in Twitter*. Barbieri, F. and **Saggion, H.** In Proceedings of the 4th Conference of the European Chapter of the Association for Computational Linguistics (EACL). 2014.
31. (\*) Readability Indices for Automatic Evaluation of Text Simplification Systems: A Feasibility Study for Spanish. Stajner, S. & **Saggion, H.** In Proceedings of the International Joint Conference on Natural Language Processing. Nagoya, Japan, 14-18 October 2013.
32. (\*) *Comparing Resources for Spanish Lexical Simplification*. **Saggion, H.** & Bott, S. & Rello, L. In Lecture Notes in Artificial Intelligence 7978. 2013: pps236-247.
33. Frequent Words Improve Readability and Short Words Improve Understandability for People with Dyslexia. Rello, L & Baeza-Yates, R.A. & Dempere-Marco, L. & **Saggion, H.** INTERACT (4) 2013, pps203-219.



34. One Half or 50%? An Eye-Tracking Study of Number Representation Readability. Rello, L & Bautista, S. & Baeza-Yates, R.A. & Gervás, P. & Hérvas, R. & **Saggion, H.** INTERACT (4) 2013: 229-245
35. Simplify or help?: text simplification strategies for people with dyslexia. Luz Rello & Ricardo A. Baeza-Yates & Stefan Bott & Horacio Saggion. In Proceedings of Web for All. 2013. **Best paper award.**
36. (\*) Unsupervised Learning Summarization Templates from Concise Summaries. **Saggion, H.** In Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Atlanta, USA. 9th-14th June 2013.
37. (\*) *Automatic Text Simplification in Spanish: A Comparative Evaluation of Complementing Components.* Drndarevic, B. & Stajner, S. & Bott, S. & Bautista, S. & **Saggion, H.** In Proceedings of the 12th International Conference on Intelligent Text Processing and Computational Linguistics. Lecture Notes in Computer Science. Samos, Greece, 24-30 March. 2013.
38. The Impact of Lexical Simplification by Verbal Paraphrases for People with and without Dyslexia. Luz, R. & Baeza-Yates, R. & **Saggion, H.** In Proceedings of the 12th International Conference on Intelligent Text Processing and Computational Linguistics. Lecture Notes in Computer Science. Samos, Greece, 24-30 March. 2013.
39. (\*) *Can Spanish Be Simpler? LexSiS: Lexical Simplification for Spanish.* Bott, S. & Rello, L. & Drndarevic, B. & **Saggion, H.** Proceedings of the 24th International Conference on Computational Linguistics, Mumbai, India, 8-15 December. 2012.
40. Graphical Schemes May Improve Readability but not Understandability for People with Dyslexia. Rello, L. & Baeza-Yates, R. & **Saggion, H.** & Graells, E. In Proceedings of the Workshop Predicting and improving text readability for target reader populations. Montreal, Canada, 7 June, p. 25-32. 2012.
41. A First Approach to the Creation of a Spanish Corpus of Dyslexic Texts. Rello, L. & Baeza-Yates, R. & **Saggion, H.** & Jennifer Pedler. In Proceedings if the 1st Workshop Natural Language Processing for Improving Textual Accessibility. Istanbul, Turkey, 27 May, p. 22-26. 2012.
42. From Ontology to NL: Generation of Multilingual User-Oriented Environmental Reports. Bouayad-Agha, N. & Casamayor, G. & Mille, S. & Rospocher, M. & **Saggion, H.** & Serafini, L. & Wanner, L. 2012. 17th International conference on Applications of Natural Language Processing to Information Systems (NLDB 2012), 26-28 June 2012.
43. (\*) Can Text Summaries help Predict Ratings? A Case Study of Movie Reviews. **Saggion, H.** & Lloret, E. & Palomar, M. 2012. 17th International conference on Applications of Natural Language Processing to Information Systems (NLDB 2012), 26-28 June 2012.
44. (\*) Towards Automatic Lexical Simplification in Spanish: An Empirical Study. Drndarevic, B. & **Saggion, H.** 2012. Workshop Predicting and improving text readability for target reader populations (PITR2012). NAACL 2012.
45. (\*) Unsupervised Content Discovery from Concise Summaries. Saggion H. 2012. Joint Workshop on Automatic Knowledge Base Construction and Web-scale Knowledge Extraction (AKBC-WEKEX 2012). NAACL 2012.
46. (\*) Automatic Simplification of Spanish Text for e-Accessibility. Bott, S. & **Saggion, H.** 2012. Proceedings of the 13th International Conference on Computers Helping People with Special Needs. pp54-56.
47. (\*) The CONCUSUS Corpus of Event Summaries. **Saggion, H.** & Szasz, S. 2012. In Proceedings of the Language Resources and Evaluation Conference LREC 2012.

48. (\*) A Hybrid System for Spanish Text Simplification. Bott, S. & **Saggion, H.** & Figueroa, D. 2012. Third Workshop on Speech and Language Processing for Assistive Technologies (SLPAT). NAACL 2012.
49. (\*) A Text Simplification Tool for Spanish. Bott, S. & **Saggion, H.** & Mille, S. Proceedings of the Language Resources and Evaluation Conference (LREC). 2012.
50. (\*) An Empirical Comparison of Predicates Used in Document Abstracting. **Saggion, H.** III Workshop RST e os Estudos do Texto. 2011.
51. (\*) Multi-domain Cross-lingual Information Extraction from Clean and Noisy Texts. **Saggion, H.** & Szasz S. 8th Brazilian Symposium in Information and Human Language Technology. 2011.
52. Building an Environmental Information System for Personalized Content Delivery. Wanner, L. & Vrochidis. S. & Tonelli, S. & Mossgraber, J. & Bosch, H. & Karppinen, A. & Myllynen, M. & Rospocher, M. & Bouayad-Agha, N. & Bügel, U. & **Saggion, H.** 9th IFIP WG 5.11 International Symposium on Environmental Software Systems - Frameworks of eEnvironment (ISESS 2011). 2011.
53. (\*) Learning Predicate Insertion Rules for Document Abstracting. **Saggion, H.** 2011. In Proceedings of the 12th Conference on Intelligent Text Processing and Computational Linguistics. Springer. LNCS. pp301-312.
54. (\*) Experiments on Summary-based Opinion Classification. Lloret, E. & **Saggion, H.** & Palomar, M. Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text. pp107-115. 2010.
55. (\*) Human Language Technology for Text-based Analysis of Psychotherapy Sessions in the Spanish Language. **Saggion, H.** & Stein-Sparvieri, E. & Maldavsky, D. & Szasz, S. Proceedings of the NAACL HLT 2010 Young Investigators Workshop on Computational Approaches to Languages of the Americas. pp32-39. 2010.
56. (\*) *Multilingual Summarization Evaluation without Human Models.* **Saggion, H.** & Torres Moreno, JM. & da Cunha, I. & SanJuan, E. & Velazquez-Morales, P. In Proceedings of the 23rd International Conference in Computational Linguistics (COLING). 2010
57. (\*) Évaluation automatique de résumés avec et sans références. Torres Moreno, JM. & **Saggion, H.** & da Cunha, I. & Velazquez-Morales, P. & SanJuan, E. Proceedings of the Conference Traitement Automatique des Langues Naturelles. 2010.
58. (\*) Interpreting SentiWordNet for Opinion Classification. Saggion H, Funk A. In Proceedings of the Language Resources and Evaluation Conference (LREC). 2010.
59. (\*) NLP Resources for the Analysis of Patient/Therapist Interviews. Saggion H, Stein-Sparvieri E, Maldavsky D, Szasz S. In Proceedings of the Language Resources and Evaluation Conference (LREC). 2010.
60. (\*) Using Text Summaries for Predicting Rating Scales. **Saggion, H.** & Lloret, E. & Palomar, M. Proceedings of the 1st Workshop on Computational Approaches to Subjectivity and Sentiment Analysis (WASSA). pp:44-51. 2010.
61. (\*) A classification algorithm for predicting the structure of summaries. **Saggion, H.** In Proceedings of the 2009 Workshop on Language Generation and Summarisation (ACL). pp:31-38. 2010.
62. (\*) Experiments on Metadata Generation of Web Business Charts. **Saggion, H.** Proceedings of II International Workshop on Web and Text Intelligence (WTI). 2009.

63. A Framework for Identity Resolution and Merging for Multi-source Information Extraction. Yankova, M. & **Saggion, H.**, Cunningham, H. In Proceedings of LREC 2008. Marrakech, Morocco, 27-29 May 2008.
64. Ontology-Driven Human Language Technology for Semantic-Based Business Intelligence. Declerck, T., Krieger, H.-U., Spies, M., **Saggion, H.** European Conference on Artificial Intelligence. Patras, Greece, 2008.
65. Human Language and Semantic Web Technologies for Business Intelligence Applications (Poster). Declerck, T., Krieger, H.-U., Spies, M., **Saggion, H.** Proceedings of LangTech 2008. Rome, Italy, 2008.
66. (\*) *Experiments on Semantic-based Clustering for Cross-document Coreference.* Saggion, H. International Joint Conference on Natural language Processing. Hyderabad, India, 2008.
67. (\*) Semantic Tagging and Summarization for Searching People on the Web. **Saggion, H.** In Proceedings of the 2nd Workshop on Text Mining and Applications. Guimarães, Portugal, 2007.
68. (\*) Ontology-based Information Extraction for Business Applications. **H. Saggion**, A. Funk, D. Maynard, K. Bontcheva. In Proceedings of the 6th International Semantic Web Conference (ISWC 2007), Busan, Korea, November, 2007.
69. Ontological Integration of Information Extracted from Multiple Sources. A. Funk, D. Maynard, **H. Saggion** and K. Bontcheva. In the Multi-source Multilingual Information Extraction and Summarization (MMIES) workshop at Recent Advances in Natural Language Processing (RANLP07), Borovets, Bulgaria, September, 2007.
70. (\*) Natural Language Technology for Information Integration in Business Intelligence. D. Maynard, **H. Saggion**, M. Yankova, K. Bontcheva and W. Peters. 10th International Conference on Business Information Systems, Poland, 25-27 April, 2007.
71. (\*) Experiments on Passage Selection and Answer Identification for Question Answering. **Saggion, H.** & Gaizauskas, R. Lecture Notes in Artificial Intelligence. Vol. 4139. 2006.
72. (\*) Language Resources for Background Gathering. Saggion, H & Gaizauskas, R. Proceedings of LREC 2006.
73. (\*) Multilingual Multidocument Summarization Tools and Evaluation. **Saggion, H.** Proceedings of LREC 2006.
74. (\*) Experiments on Statistical and Pattern-based Biographical Summarization. **Saggion, H.** & Gaizauskas, R. Text Mining and Applications (TeMA) Workshop, 2005. Lecture Notes in Artificial Intelligence. Springer-Verlag.
75. (\*) Integrating NLP Tools to Support Information Access to News Archives. **Saggion, H.** & Barker, E. & Gaizauskas, R. & Foster, J. Recent Advances in Natural Language Processing, 2005.
76. SUPPLE: A Practical Parser for Natural Language Engineering Applications. Robert Gaizauskas, Mark Hepple, Horacio Saggion, Mark A. Greenwood and Kevin Humphreys. 9th International Workshop on Parsing Technologies (IWPT'05). 2005.
77. (\*) Exploring the Performance of Boolean Retrieval Strategies for Open Domain Question Answering. **Saggion, H.** & Gaizauskas, R. & Hepple, M. & Roberts, I. & Greenwood, M. Proceedings of the Workshop on Information Retrieval for Question Answering. SIGIR 2004.
78. (\*) *Mining on-line sources for definition knowledge.* Saggion, H. & Gaizauskas, R. Proceedings of the 17th International FLAIRS Conference. 2004.

79. (\*) A Pattern Based Approach to Factoid and Definition Question Answering. Greenwood, M. & **Saggion, H.** Proceedings of RIAO 2004.
80. (\*) Identifying Definitions in Text Collections for Question Answering. **Saggion, H.** Proceedings of LREC 2004.
81. Contribution of NLP to the Content Indexing of Multimedia Documents. Declerck, T. & Kuper, J. & **Saggion, H.** & Samiotou, A. & Wittenburg, P. & Contreras, J. Proceedings of CIVR 2004.
82. MEAD - a platform for multidocument multilingual summarization. LREC 2004.
83. (\*) *Evaluation Challenges in Large Scale Document Summarization*. Radev, D. & Teufel, S. & **Saggion, H.** & Lam, W. & Blitzer, J. & Qi, H. & Çelebi, A. & Liu, D. & Drabek, E. Proceeding of the Association for Computational Linguistics, ACL-2003.
84. (\*) *Intelligent Multimedia Indexing and Retrieval through Multi-source Information Extraction and Merging*. Kuper, J. & **Saggion, H.** & Declerck, T. & Reidsma, D. & Cunningham, H. & Wittenburg, P. & DeJong, F. & Wilks, Y. International Joint Conference on Artificial Intelligence, IJCAI-2003.
85. MUMIS – Advanced information extraction for multimedia indexing and searching. Declerck, T. & Cunningham, H. & Saggion, H. & Kuper, J. & Reidsma, D. & Wittenburg, P. Proceedings of WIAMIS, 2003.
86. Event-coreference across Multiple Multi-lingual Sources in the MUMIS project. Kuper, J. & **Saggion, H.** & Cunningham, H. & Declerck, T. & Hoenkamp, E. & Puts, M. & de Jong, F.M.G. & Wilks, Y. & Wittenburg, P. BNAIC 2003, T. Heskes, P. Lucas, L. Vuurpijl and W. Wiegerinck (eds), Nijmegen, pp. 425-426, 2003.
87. Cross document annotation for multimedia retrieval. Reidsma, D. & Kuper, J. & Declerck, T. & **Saggion, H.** & Cunningham, H. EACL Workshop Language Technology and the Semantic Web (NLPXML), Budapest, 2003
88. Cross document ontology based information extraction for multimedia retrieval. Reidsma, D. & Kuper, J. & Declerck, T. & **Saggion, H.** & Cunningham, H. Supplementary proceedings of the ICCS03, Dresden, 2003.
89. (\*) Colouring Summaries Bleu. Pastra, K. & **Saggion, H.** EACL-2003 Workshop on Evaluation Initiatives. Hungary April 12-17, 2003.
90. (\*) Event-coreference across Multiple, Multi-lingual Sources in the Mumis Project. **Saggion, H.** & Kuper, J. & Cunningham, H. & Declerck, T. & Wittenburg, P. & Puts, M. & Hoenkamp, E. & DeJong, F. & Wilks, Y. 10th Conference of the European Chapter of the Association for Computational Linguistics. EACL-2003. Hungary April 12-17, 2003.
91. (\*) NLP for Indexing and Retrieval of Captioned Photographs. Pastra, K. & **Saggion, H.** & Wilks, Y. 10th Conference of the European Chapter of the Association for Computational Linguistics. EACL-2003. Hungary, April 12-17, 2003.
92. (\*) Robust Generic and Query-based Summarization. **Saggion, H.** & Bontcheva, K. & Cunningham, H. 10th Conference of the European Chapter of the Association for Computational Linguistics. EACL-2003. Hungary, April 12-17, 2003.
93. (\*) Extracting Relational Facts for Indexing and Retrieval of Crime-Scene Photographs. Pastra, K. & **Saggion, H.** & Wilks, Y. ES2002. 22nd SGAI International Conference on Knowledge-based Systems and Applied Artificial Intelligence. Cambridge, UK, December 2002.

94. (\*) Solving Coreference under Event Semantic Constraints. **Saggion, H.** Discourse Anaphora and Anaphora Resolution Conference. DAARC 2002. Lisbon, Portugal, September 2002.
95. Developing Reusable and Robust Language Processing Components for Information Systems using GATE. Bontcheva, K. & Cunningham, H. & Tablan, V. & Maynard, D. & **Saggion, H.**. 3rd International Workshop on Natural Language and Information Systems (NLIS'2002).
96. (\*) *Meta-evaluation of Summaries in a Cross-lingual Environment using Content-based Metrics.* **Saggion, H.** & Radev, D. & Teufel, S. & Lam, W. International Conference on Computational Linguistics. COLING 2002. Taipei, 2002.
97. Using a Text Engineering Framework to Build an Extendable and Portable IE-based Summarisation System. Maynard, D. & Bontcheva, K. & **Saggion, H.** & Cunningham, H. & Hamza, O. Proceedings of the ACL Workshop on Text Summarisation. 2002.
98. (\*) Access to Multimedia Information through Multisource and Multilanguage Information Extraction. 7th International Workshop on Applications of natural language to Information Systems (NLDB'2002). 27-28 June 2002.
99. (\*) Extracting Information for Automatic Indexing of Multimedia Material. Saggion, H. & Cunningham, H. and Bontcheva, K. & Maynard, D. & Hamza, O. & Cursu, C. & Wilks, Y. & Language Resources and Evaluation Conference. LREC 2002. Las Palmas, Spain, 29-31 May 2002.
100. (\*) Developing Infrastructure for the Evaluation of Single and Multi-document Summarization Systems in a Cross-lingual Environment. **Saggion, H.** & Radev, D. & Teufel, S. & Lam, W. & Strassel, S.M. Language Resources and Evaluation Conference. LREC 2002. 29-31 May 2002.
101. (\*) Shallow-based Robust Summarization. **Saggion, H.** ATALA Workshop "Automatic Summarization: Solutions & Perspectives", Paris, France, December 2002.
102. (\*) Using the GATE architecture for named entity recognition in the football domain. Saggion & al. 2nd International Workshop on Spanish Language Processing and Language technologies (SPLPT2). September 2001.
103. (\*) Summary Generation and Evaluation in SumUM. **Saggion, H.** & Lapalme, G. IBERAMIA/SBIA'2000. Atibaia, São Paulo, Brasil, 22-24 November 2000.
104. (\*) Selective Analysis for the Automatic Generation of Summaries. **Saggion, H.** & Lapalme, S. 6th International Conference for Knowledge Organization ISKO, University of Toronto, Toronto, Canada, July 10 to 13, 2000.
105. (\*) *Concept Identification and Presentation in the Context of Technical Text Summarization.* **Saggion, H.** & Lapalme, H. Workshop on Automatic Abstracting, NAACL-ANLP2000, Association for Computational Linguistics, Seattle, USA, April 30, 2000.
106. (\*) Selective Analysis for Automatic Abstracting: Evaluating Indicativeness and Acceptability. **Saggion, H.** & Lapalme, H. Content-Based Multimedia Information Access, RIAO'2000, Paris, France, April 12-14, 2000.
107. (\*) Using Linguistic Knowledge in Automatic Abstracting. Saggion, H. ACL'99: the 37th Annual Meeting of the Association for Computational Linguistics. University of Maryland, June 20 - 26 1999. Proceedings.
108. (\*) Where does Information come from? Corpus Analysis for Automatic Abstracting. **Saggion, H.** & Lapalme, G. RIFRA'98 Conference, Sfax, Tunes, November 11-14 1998. Proceedings.

109. (\*) The Generation of Abstracts by Selective Analysis. AAAI'98 Spring Symposium Series, Intelligent Text Summarization Symposium, March 23-25 1998, Stanford, CA, USA. AAAI Technical Report (SS-98-06).
110. (\*) Definite Anaphora in Portuguese Abstracts. DAAR'96 Conference, Lancaster University, July 1996, England.
111. (\*) Calculo de la Estructura de un Texto en un Sistema de Procesamiento de Lenguaje Natural. **Saggion, H.** & Carvalho, A. XI Congress of Natural and Formal Languages, September 25-29 1995, Tortosa, Spain.
112. (\*) Text Structure Aiming at Machine Translation. **Saggion, H.** & Carvalho, A. XII Artificial Intelligence National Meeting - RNIA'95, September 20-23 1995, Cuernavaca, Morelos, Mexico.
113. (\*) Análise Textual Visando a Tradução Automática. XV Congress of the Brazilian Computer Science Society -SEMISH'95 and XXI Latin American Conference of Computer Science - PANEL'95, July 31 -August 4 1995, Canela - RS, Brazil. Proceedings.
114. (\*) Anaphora Resolution in a Machine Translation System. **Saggion, H.** & Carvalho, A. International Conference on Machine Translation: Ten Years On, November 1994. Cranfield University, Bedford, England.
115. (\*) Un Sistema de Texto a Habla para el Español. Aspectos Teóricos e implementación. **Saggion, H.** & Schijman, P. V Argentine Congress on Computer Science, October 1992, Buenos Aires, Argentina.

## Papers Describing Participation in Evaluation Programmes

1. UPF-taln: SemEval 2015 Tasks 10 and 11 Sentiment Analysis of Literal and Figurative Language in Twitter. F Barbieri, F Ronzano, H Saggion. In SemEval 2015. **2nd Best Approach in Figurative Language Detection.**
2. Matching Texts with SUMMA. Saggion H. Défi Fouille de Texte (DEFT 2011). Proceedings of TALN 2011. 2011. **Best Approach in Matching with Confidence Scores.**
3. Using SUMMA for Language Independent Summarization at TAC 2011. **Saggion, H.** Text Analysis Conference. 2011.
4. SHEF: Semantic Tagging and Summarization Techniques Applied to Cross-document Coreference. Saggion, Horacio. In Proceedings of Semeval 2007, Association for Computational Linguistics, June, 2007.
5. Multi-document summarization by sentence cluster relevance and redundancy removal. **Saggion, H.** & Gaizauskas, R. Proceedings of DUC 2004. **2nd Best Approach in Generic Multidocument Summarization.**
6. The University of Sheffield's TREC 2004 Q&A Experiments. Gaizauskas, R. & Greenwood, M. & Hepple, M. & Roberts, I. & **Saggion, H.** Proceedings of TREC 2004.
7. The University of Sheffield's TREC 2003 Q&A Experiments. Gaizauskas, R. & Greenwood, M. & Hepple, M. & Roberts, I. & **Saggion, H.** & Sargaison, M. Proceedings of TREC 2003. **4th Best Definition Extraction System. In Top-10 Best Systems for QA.**
8. Summaries with SumUM a text summarization system and its Expansion for Document Understanding Conference. Farzindar, A. & Lapalme, G. & **Saggion, H.** Proceedings of DUC 2002.

## Books, Chapters, and Volumes

NOTES: With asterisk (\*) I indicate the publications for which I am one of the lead authors. With *italics*, I indicate the most significant works.

1. (\*) *Automatic Text Simplification*. Synthesis Lectures on Human Language Technologies, Morgan & Claypool Publishers 2017
2. (\*) *Technologies for Persons with Dyslexia*. **H. Saggion**, L. Rello. In The SAGE Encyclopedia of Educational Technology. SAGE Publications, Inc. pp710-712.
3. (\*) *Multi-source, Multilingual Information Extraction and Summarization*. Poibeau, T. & **Saggion, H.** & Piskorski, J. & Yangarber, R. (Eds.). Springer Series *Theory and Applications of Natural Language Processing*. 2013.
4. (\*) *A Study of the Effect of Document Representations in Clustering-based Cross-document Coreference Resolution*. **Saggion, H.** In Multi-source, Multilingual Information Extraction and Summarization. Springer. 2013.
5. (\*) Automatic Text Summarization: Past, Present and Future. **Saggion, H.** & Poibeau, T. In Multi-source, Multilingual Information Extraction and Summarization. Springer. 2013.
6. (\*) *Second Workshop on Natural Language Processing for Improving Textual Accessibility - NLP4ITA*, (LREC Workshop), Baeza-Yates, R. & Rello, L. & **Saggion, H.** (Eds.), 2013.
7. (\*) *First Workshop on Natural Language Processing for Improving Textual Accessibility - NLP4ITA*, (LREC Workshop), Rello, L. & **Saggion, H.** (Eds.), 2012.
8. (\*) Résumé automatique par filtrage sémantique. **Saggion, H.** in *Filtrage sémantique de l'annotation la navigation textuelle*, Jean-Luc Minel (Ed). Éditions Hermès. 2009.
9. (\*) *Second Workshop on Multisource, Multilingual Information Extraction and Summarization*, (COLING Workshop), Sivaji Bandyopadhyay and Thierry Poibeau and Horacio Saggion and Roman Yangarber (Eds.), 2008.
10. (\*) *Information Access and Natural Language Processing: An Stimulating Dialogue*. Gaizauskas, R., **Saggion, H.**, Barker, E. In Words and Intelligence II - Essays in Honor of Yorick Wilks. Ahmad, K., Brewster, C. and Stevenson, M. (Eds.) Springer, 2007.
11. (\*) *First Workshop on Multisource, Multilingual Information Extraction and Summarization* (RANLP Workshop), Thierry Poibeau and Horacio Saggion (Eds.), 2007.
12. (\*) Natural Language Processing for Mining Profiles and Definitions, **Saggion, H.** in Emerging Technologies of text Mining: Techniques and Applications. H.A. do Prado and E. Ferneda (Eds.). Idea Group Publishing. 2007.
13. (\*) *Crossing Barriers in Text Summarization Research*, (RANLP Workshop), Horacio Saggion and Jean-Luc Minel (Eds.), 2005.
14. Content-based Indexing and Searching of Multimedia Documents in “Annotation for the Semantic Web”, S. Handschuh and S. Staad (Eds.), 2003.
15. (\*) SumUM in “Filtrage sémantique. Du résumé automatique à la fouille de textes” by Jean-Luc Minel. Éditions Hermès. 2002.

## Book Reviews

1. (\*) Review of the Book “Indexing and Abstracting in Theory and Practice”, F.W. Lancaster. University of Illinois. 2003. Journal of the American Society for Information Science and Technology. In Press.
2. (\*) Review of the Book “Automatic Summarization”, I. Mani. John Benjamins Publishing Company. Journal of Document Design, Issue: 5(1), April 2004.
3. (\*) Review of the Book “Advances in Automatic Text Summarization”, I. Mani and M.T. Maybury (Eds). The MIT Press. Journal of the Society for Knowledge Organization, Issue: January 2001.

## Lectures at Summer Schools / Tutorials

1. Data-Driven Text Simplification. Tutorial. Stajner, S. & **Saggion, H.** COLING 2018, Santa Fe, USA, August 2018.
2. Scholarly Data Mining: Making Sense of Scientific Literature. Tutorial. **Saggion, H. & Ronzano, F.** ACM/IEEE Joint Conference on Digital Libraries, JCDL 2017, Toronto, ON, Canada.
3. Natural Language Processing for Intelligent Access to Scientific Information. Tutorial. **Saggion, H. & Ronzano, F.** COLING 2016, Osaka, Japan
4. Introduction to Automatic Text Simplification. **Saggion, H.** Tutorial. Recent Advances Natural Language Processing. Hissar, Bulgaria, September 2015.
5. Text Simplification with a Purpose. **Saggion, H.** Tutorial. International Joint Conference in Natural Language Processing. Nagoya, Japan, 2013.
6. Introduction to Automatic Text Simplification (in Spanish). **Saggion, H.** Short Course. Altamira Summer School. Alicante, Spain, 2012.
7. Hands-on Natural Language Processing Applications for Information Access. **Saggion, H.** ECI 2008. Course. Universidad de Buenos Aires, Argentina. July-August 2008.
8. Introduction to Text Summarization and Information Access: Resources and Evaluation. Tutorial. **Saggion, H.** Language Resources and Evaluation Conference (LREC). May 2008, Marrakech, Morocco.
9. Introduction to Text Summarization and Other Information Access Technologies. Tutorial. **Saggion, H.** International Joint Conference on Natural Language Processing. January 2008. Hyderabad, India.
10. Hands-on Natural Language Processing Applications for Information Access. **Saggion, H.** ESSLLI 2007. Course. University College Dublin. Dublin, Ireland. August 2007.
11. Text Summarization: Resources and Evaluation. **Saggion, H.** LREC 2006. Tutorial. Genoa, Italy, May 2006.
12. Automatic Text Summarization. **Saggion, H.** ESSLLI 2005. Course. Heriot-Wat University, Edinburgh, Scotland, August 2005.
13. Natural Language Processing and Automatic Text Summarization. **Saggion, H.** Rio 2004, Summer School. Universidad Nacional de Rio Cuarto, Cordoba, Argentina. Course Notes. February 2004.



## Teaching Experience

- Universitat Pompeu Fabra, DTIC, Ramón y Cajal Fellow (2010-2015), Associate Professor (2010-present). 60hs per year from 2010 until 2014. 100hs per year from 2015.
  - Algorithms and Data Structures II (2017, 2018) Course coordinator since 2017. Course responsibilities: material preparation, administration, teaching assistants coordination. Teaching responsibilities: theory.
  - Data Bases (2010, 2011, 2012, 2013, 2014, 2015, 2016). Course coordinator since 2012. Course responsibilities: material preparation, administration, teaching assistants coordination. Teaching responsibilities: theory, laboratory, problem solving, marking.
  - Intelligent Web Applications. 2011, 2012, 2013, 2014, 2015, 2016, 2017. Course coordinator. Course responsibilities: material preparation, administration, teaching assistants coordination. Teaching responsibilities: theory, laboratory, problem solving, marking.
- Université de Montréal, DIRO, 1998-1999, Teaching Assistant (1998, 1999)
  - Artificial Intelligence (4 hs per week). Teaching responsibilities: problem solving, marking.
- Universidad de Buenos Aires, Computer Science Department, Assistant Professor (1995-1996), Teaching Assistant (1986-1987, 1991-1993).
  - Algorithms and Data Structures I (full time / 8 hs per week). Course responsibilities: teaching assistants coordination, material preparation, administration. Teaching responsibilities: demonstration, problem solving, marking.
  - Algorithms and Data Structures II (part time / 4 hs per week). Teaching responsibilities: demonstration, problem solving, marking.
  - Logic for Computer Science (part time / 4 hs per week). Teaching responsibilities: demonstration, problem solving, marking.
  - Introduction to Natural Language Processing (2 hs per week). Teaching responsibilities: demonstration.
- Universidad Nacional de Quilmes, Secretaría de Ciencia y Técnica, Assistant Professor (1992-1993).
  - Informatics I and II (part time / 6 hs per week). Teaching responsibilities: demonstration, problem solving, marking.

## Lectures and Invited Talks

- Keynote Speech “Current Challenges and Opportunities in Mining Multilingual Scientific Text Collections ” at WOSP 2018 Workshop, LREC 2018, Miyazaky, Japan, May 2018.
- Invited tutorial at RANLP 2015. Introduction to Automatic Text Simplification. Hissar, Bulgaria. September 2015.
- Simplifying Texts in Spanish. Keynote Speech. XXIX Conference of the Spanish Association for Natural Language Processing. Madrid, Spain. 18-20 September, 2013.
- Automatic Text Simplification: What for? Keynote Speech at the Recent Advances in Natural Language Conference, RANLP 2013, Hissar, Bulgaria. 7-13 September 2013.
- Text Simplification in Spanish. Seminar. LIMSI Laboratoire, Paris, France. 6th December, 2012.
- Spanish Text Simplification in Simplext: Making Texts more Accessible. Seminar. XEROX Research Centre, Grenoble, France. 12th April, 2012.

- Easy Access to Textual Information: The Role of Natural Language Processing. Keynote Speech. 16th International Conference on Application of Natural Language to Information Systems (NLDB2011). Alicante, Spain. June 28-30, 2011.
- Journée Résumé Automatique Multimédia. Paris, France. Keynote Speech. 17th March 2011.
- Applying Human Language Technology for Business Intelligence. Invited Seminar. Aston University, UK. 11th March 2009.
- Information Extraction and Summarization for Information Access. Invited Seminar. Universidad Pompeu Fabra, Barcelona, Spain. 11th December 2008.
- Summarization and Information Extraction for Cross-document Coreference Resolution. Invited Seminar. Laboratoire d'Informatique Appliquée d'Avignon, Université d'Avignon. 4th December 2008.
- Applying Human Language Technology for Business Intelligence: The Musing Project. Computational Linguistics Research Seminar, University of Wolverhampton. 8th October 2008.
- Natural Language Processing Applications for Information Access and Business Intelligence. Natural Language Processing Seminar. Facultad de Matemáticas, Astronomía, y Física, Universidad Nacional de Córdoba, Córdoba, Argentina. 4th August 2008.
- Experiments on Cross-document Coreference: The Role of Semantic Information and Summarization. KTH Seminar - Stockholm University. 12 June 2007.
- Natural Language Processing for Information Access Applications (Course). Escuela de Ciencias Informáticas (ECI 2008). 28 July - 1st August 2008. Universidad de Buenos Aires, Buenos Aires, Argentina.
- Introduction to Text Summarization and Other Information Access Technologies: Resources and Evaluation (Tutorial). Language Resources and Evaluation Conference LREC 2008. 26th May 2008. Marrakesh. Morocco.
- Introduction to Text Summarization and Other Information Access Technologies (Tutorial). International Joint Conference on Natural Language Processing. Hyderabad, India. 7th January 2008.
- Hands-on NLP Applications for Information Access. 19th European Summer School of Logic, Language, and Information (ESSLLI 2007), University College Dublin, Dublin. 13-17 August 2007.
- Applications in Human Language Technology. Departamento de Ciencias de la Computación, Universidad de Buenos Aires, Argentina, 13th April 2007.
- Mining the Web to Improve Semantic-based Multimedia Search and Digital Libraries. IST 2006. Helsinki, Finland, 21st November 2006.
- An Overview of GATE. IST 2006. Helsinki, Finland, 22nd November 2006.
- Finding Definitions in Text Collections for Question Answering. University of Montreal/EBSI. 23rd March 2005.
- Automatic Text Summarization 5-days course, 17th European Summer School of Logic, Language, and Information (ESSLLI 2005), Heriot-Watt University, Scotland. 8-12 August 2005.
- Electronic Cubreporter: New Approaches for Information Access to News Archives. ICOSS Seminar, University of Sheffield. 15th March 2005.
- Finding Definitions in Text Collections for Question Answering. Departamento de Ciencias de la Computación, Universidad de Buenos Aires, Argentina, 11th March 2004.

- Natural Language Processing and Automatic Text Summarization. 5-day course. Rio 2004 Summer School in Computer Science. Rio Cuarto, Cordoba, Argentina. 16-20 of February 2004.
- Answering Definitions by Mining On-line Resources. NLP Seminar, University of Sheffield. 2003.
- Multimedia Indexing and Searching Environment: The MUMIS Project. NLP Seminar, University of Sheffield, 2003.
- Shallow-based Robust Summarization. ATALA Seminar on “Automatic Summarization Solutions and Perspectives”, Paris, France, 14 December 2002.
- Using Natural Language Processing for Indexing Multimedia Material. ILASH Seminar on Accessing Multimedia Information. 26 November 2002. Sheffield, UK.
- The SOCIS Project. NLP Seminar, University of Sheffield, December 2000.
- Analysing the Language at the Scene of the Crime. Forensic Linguistics Seminar, Department of Linguistics, University of Birmingham, UK. 23rd March 2001.
- La génération de résumés par analyse sélective. 67ème Congrès de l’ACFAS, University of Ottawa, Canada, May 10-11, 1999.
- La génération de résumés par analyse sélective. University of Nancy II, Workshop on Automatic Abstracting and Information Extraction. Nancy, France, 6 March 1998.
- Generación automática de resúmenes. Instituto de Computación, Universidad de la República, Uruguay. 28 December 1998.
- La génération de résumés par analyse sélective. University of Paris IV-Sorbonne, CAMS Seminars. Paris, France, 16 March 1998.
- La génération de résumés. University of Paris VII, TALANA seminars. Paris, France, 16 January 1998.

## Academic Awards

- Award (first finalist) in Knowledge Transfer from Consell Social Universitat Pompeu Fabra 2018 for project ”SemaMed: Semantic Medical Record Understanding”.
- 3rd Prize Second NLP Hackathon 2018 for application “re.facil.es”.
- Best Paper Award 3rd Workshop on Noisy User-generated Text 2017
- 1st Prize (with Francesco Ronzano) First NLP Hackathon 2017 for application “TextDigester”.
- Best Communication Award Conference Web for All 2017.
- Best Paper Award. SEPLN 2015.
- Most Innovative Approach. ESWC / Semantic Web Publishing Challenge 2014.
- Cátedra Telefónica Universidad de Alicante Award, best system demonstrating the application of NLP to social inclusion or disability. 2013.
- Web for All Best paper award. 2013.
- Vodafone Innovation Award (Simplext Project). 2011.

- BDigital Awards finalist (Simplext Project). 2012.
- Ramón y Cajal Fellowship. 2009.
- University of Sheffield Travel Grant (ESSLLI). 2005.
- ACL Student Travel Award ANLP/NAACL 2000.
- Fundación Antorchas Award for PhD Research Project 1999-2000.
- French/Quebec Award, Stage at University of Paris VII, January - April 1998.
- Argentine Education Ministry's Award (PhD fellowship) 1997-1999.
- Canadian Agency of International Development (CIDA) Award 1997-2000.
- Bourse d'excellence de l'Université de Montréal 1997-2000.
- Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) Award (MsC research project - University of Campinas - UNICAMP), São Paulo, Brazil, 1994.
- CNPq Travel Award IX Escola de Computação, 1994.
- Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) Award (MsC fellowship - University of Campinas - UNICAMP), São Paulo, Brazil, 1993.
- IBM Argentina Internship Award (top ranked undergraduate student), 1986.
- Universidad de Buenos Aires Ciencia y Técnica (UBACyT) Award (Research project), 1987.
- Award from Programa Argentino Brasileiro de Informatica (top ranked student; best 1987 student) - 1986, 1987, 1988.

## News Coverage

- Several appearances in TV3 (Catalunya Broadcasting Corporation) and RTVE (Spanish National Broadcasting Corporation)
- Press coverage for our work on Social Media analysis  
ABC: [https://www.abc.es/sociedad/abci-dime-emoticonos-utilizas-y-dire-cual-lucha-201809090132\\_noticia.html](https://www.abc.es/sociedad/abci-dime-emoticonos-utilizas-y-dire-cual-lucha-201809090132_noticia.html)
- Press release book publication  
UPF News: [https://www.upf.edu/es/web/media/enoticies/-/asset\\_publisher/wdGAWZ7EMj53/content/id/8512490/maximized#.W6XoJvlRc2w](https://www.upf.edu/es/web/media/enoticies/-/asset_publisher/wdGAWZ7EMj53/content/id/8512490/maximized#.W6XoJvlRc2w)
- Press coverage on our text analysis technology for social media.  
Diario Ara [http://www.ara.cat/societat/possible-detecti-satira-amagada-llenguatge\\_0\\_1448855120.html](http://www.ara.cat/societat/possible-detecti-satira-amagada-llenguatge_0_1448855120.html)  
La Vanguardia <http://www.lavanguardia.com/20151017/54438173485/tot-al-peu-de-la-lletra-quim-monzo.html#.ViIQmvjur8Y.twitter>
- Press coverage for our simplification technology  
EL Faro de Vigo: <https://www.farodevigo.es/sociedad/2017/04/07/crean-prototipo-correo-electronico-personas/1656453.html>  
El Mundo: <http://www.elmundo.es/cataluna/2017/04/06/58e66283e5fdea25318b45f8.html>

- Press coverage for our project KOLUMBA  
ABC: <http://agencias.abc.es/agencias/noticia.asp?noticia=2659236>
- Best paper Award at SEPLN 2015.  
UPF News <http://www.upf.edu/enoticies/es/1516/0941.html#.Vkbbyrb-vbGm>.
- Most Innovative Approach Award  
UPF News <http://www.upf.edu/enoticies/1314/0606.html#.Vkb257-vbGk>
- Text Summarization  
El Periodico <http://www.elperiodico.com/es/noticias/sociedad/busca-del-robot-escritor-3244295>
- Best computer system applied to human language technologies in social inclusion and disability.  
UPF News <http://www.upf.edu/dtic/en/actualitat/0501.html#.VkbYJ7-vbGk>
- My PhD Student obtains Best Young European Research Award.  
20 minutos <http://www.20minutos.es/noticia/1854132/0/entrevista-mejor-joven/investigadora-europea/luz-reello-upf/>
- Book Publication  
UPF News <http://www.upf.edu/enoticies/1112/0226.html#.Vkb2kb-vbGk>

## Other Activities

- Academic Selection Committees  
Member of the selection committee for a Natural Language Processing Lecturer Vacancy, University of Avignon, France, 2012
- Research Project Reviewer:  
Programa de Proyectos de Investigación y Desarrollo, Universidad de la República (Uruguay);  
Agence National de la Recherche (France);  
Programa UBA-CyT, Universidad de Buenos Aires (Argentina);  
STIC-AmSud Programme (France & South America).
- Student Supervision: PhD Students  
Pablo Accuosto (PhD Candidate, Universitat Pompeu Fabra, Spain). "Scientific Argument Mining". Since October 2016.  
Ahmed Aburaéd (PhD Candidate, Universitat Pompeu Fabra, Spain). "Text Simplification / Text Summarization". Since October 2015.  
**Francesco Barbieri (PhD Candidate, Universitat Pompeu Fabra, Spain), "Machine Learning Methods for Understanding Social Media Communication: Modeling Irony and Emojis". PhD Awarded in January 2018.**  
**Luis Espinoza (PhD Candidate, Universitat Pompeu Fabra, Spain), "Knowledge Acquisition in the information Age: The interplay between Lexicography and Natural Language Processing". PhD Awarded in July 2017.**  
**Luz Rello (Universitat Pompeu Fabra, Spain), "DysWebxia: NLP Model to Improve Accessibility of the Textual Web for Dyslexic Users". Thesis defended and approved on 27th June 2014 with qualification "Cum Laude".**

**Sanja Stajner (University of Wolverhampton, UK), Multilingual Text Simplification. PhD Awarded in March 2015.**

Fabio Bif Goularte (visiting PhD Candidate, Federal University of Santa Catarina), summarization. 2017 (4 months).

Elena Lloret (visiting PhD Candidate, University of Sheffield, UK), Summarization and Opinion Mining. 2010 (3 months).

Susana Bautista (visiting PhD Candidate, Universitat Pompeu Fabra, Spain), Numerical Text Simplification. 2012 (6 months).

**Milena Yankova (University of Sheffield), “Identity Resolution Framework”, 2006-2010. Graduated in 2010.**

- Student Supervision: Undergraduate Students

Enric Martos (Undergraduate Student, Universitat Pompeu Fabra, Spain). Opinion Mining of Amazon Reviews. 2017-2018.

Francesc Xavier Julià (Undergraduate Student, Universitat Pompeu Fabra, Spain). Análisis de Eventos Musicales en Twitter. 2017-2018.

Adrià Jaumot García (Undergraduate Student, Universitat Pompeu Fabra, Spain). Enriquecimiento de Publicaciones Científicas. 2016-2017.

Andrea Planas Cabani (Undergraduate Student, Universitat Pompeu Fabra, Spain). Análisis de opiniones por aspecto. 2015-2016.

Marc Sanmiquel (Undergraduate Student, Universitat Pompeu Fabra, Spain). Best Soccer Tweets. 2014-2015.

Lia Diaz (Undergraduate Student, Universitat Pompeu Fabra, Spain). Resúmenes Hablados. 2014-2015.

Martí Rella (Undergraduate Student, Universitat Pompeu Fabra, Spain). Seguimiento de emociones en partidos de fútbol. 2014-2015.

Jaume Rella (Undergraduate Student, Universitat Pompeu Fabra, Spain). TwIrony: Identificación de la ironía en tweets en catalán. 2014-2015.

Carles Guell (Undergraduate Student, Universitat Pompeu Fabra, Spain). Summarizing Tweets in Catalan. 2013-2014.

Lola Lage (Undergraduate Student, Universitat Pompeu Fabra, Spain). Monitoring Positive and Negative News in Spanish Media. 2013-2014.

Raul Oñate (Undergraduate Student, Universitat Pompeu Fabra, Spain). Opinion Mining in Tweets. 2012-2013.

Jordi Vallverdu (Undergraduate Student, Universitat Pompeu Fabra, Spain). Opinion Mining. 2010-2011.

Siti Salwa Hasbullah (MSc student from University of Malaysia), “Semantic Role Labelling”, January - March 2008;

- PhD Panel member

Marco Del Tredici, Universitat Pompeu Fabra. “A distributional approach to the modeling of metaphor and polysemy” 2015.

Ahmet Aker, University of Sheffield, “Multilingual Multi-Document Summarization for Geo-Referenced Images”. 2008-2010.

Silvia Vázquez (Universitat Pompeu Fabra, Spain). “Context cues to Automatic Extraction of Opinion Words”. 2011.

Simon Mille (Universitat Pompeu Fabra, Spain). "On Deep Stochastic Sentence Generation: Resources and Experiments". 2010.

- Senior Researcher Evaluation Panel member

Dimplôme d'habilitation à diriger des recherches, Nuria Gala (Université d'Aix en Provence, France) "Approches multidisciplinaires pour l'étude du lexique et la construction de ressources lexicales nouvelles". 2015.

- PhD jury:

Ali Naderi (Universitat Politècnica de Catalunya, Spain). "Unsupervised Entity Linking using Graph-based Semantic Similarity". 2016.

Dominic Paul Rout (University of Sheffield, UK). "A Ranking Approach to Summarising Twitter Home Timeline". 2015.

Reda Siblani (Université Concordia, Montréal, Canada). "Recognizing Textual Entailment using Description Logic and Semantic Relatedness". 2014.

Jesus Antonio Motta (Université Laval, Quebec, Canada). "VENCE : un modele performant d'extraction de resumes base sur une approche d'apprentissage automatique renforcee par de la connaissance ontologique". 2014.

Hector Martinez Alonso (University of Copenhagen / Universitat Pompeu Fabra). "Annotation of regular polysemy. An empirical assessment of the underspecified sense". 2013.

Anne-Lyse Minard (Université Paris-Sud, France). "Extraction de relations en domaine de spécialité". 2012.

Gabriela Ferraro (Universitat Pompeu Fabra, Spain). "Towards Deep Content Extraction from Specialized Discourse: The Case of Verbal Relations in Patent Claims". 2012.

Rémi Lavallay (Université d'Avignon, France). "Extraction automatique de segments textuels, détection de rôles, de sujets et de polarités". 2012.

Aiala Rosá (Université de Paris Ouest - Nanterre, France and Universidad de la República, Uruguay). "Identificación de opiniones de diferentes fuentes en textos en español". 2011.

Rogelio Nazar (Universitat Pompeu Fabra, Spain). "A Quantitative Approach to Concept Analysis". 2010.

Rodrigo Alarcón Martínez (Universitat Pompeu Fabra, Spain). "Descripción y evaluación de un sistema basado en reglas para la extracción automática de contextos definitorios". 2009.

Ester Boldrini (Universidad de Alicante, Spain). "EMOTIBLOG: A model to learn subjective information detection in the new textual genres of the Web 2.0 - a multilingual and multi-genre approach". 2012.

Elena Lloret (Universidad de Alicante, Spain). "Text Summarisation based on Human Language Technologies and its Applications". 2011.

Laura Plaza (Universidad Complutense de Madrid, Spain). "The Use of Semantic Graphs in Automatic Summarization: Comparative Case Studies in Biomedicine, Journalism and Tourism". 2010.

Seemab Latif (University of Manchester, UK). "Automatic summarisation as pre-processing for document clustering". 2010.

Silvia Fernandez (Université de Nancy 2, France). "Modèles de mémoires associative basés en physique statistique et leurs applications en TAL". 2009.

Florian Boudin (Université d'Avignon, France). "Exploration d'approches statistiques pour le résumé automatique de texte". 2008.

Iria da Cunha (PhD Candidate 2008, University Pompeu Fabra), "Hacia un modelo lingüístico de resumen automático de artículos médicos en español";

Martin Hassel (PhD Candidate 2007, University of Stockholm), “Resource Lean Automatic Text Summarization“

- MsC Jury  
Emilia Garcia Casemont (Universitat Pompeu Fabra, Spain). “Embodied Cognitive Semantics and Grammar for Gradable Adjectives”. 2012.
- PhD evaluator:  
PhD thesis Universidad Politecnica de Catalunya, 2007 (Maria Fuertes Fort), “A Flexible Eclectic Multitask Summarizer for Documents in Different Language, Domain and Media”;  
PhD thesis Universidad de Alicante, 2008 (Paloma Moreda) “Los Roles Semánticos en la Tecnología del Lenguaje Humano: Anotación y Aplicación”;  
PhD dissertation awards (Conferencia Lationamericana en Informatica CLEI 2000, 2002)
- Member of the following conference program committees (or reviewer): LREC (2006, 2008, 2010, 2012, 2014, 2016); ACL (2005, 2006, 2007, 2008, 2016); NAACL (2016, 2013); IJCNLP 2008; TIL (2007, 2009, 2011); MMIU 2008; ATALA 2006; EMNLP 2006; PROPOR 2006; IJCAI 2005; COLING (2004, 2010, 2012); CICLING (2011, 2013); Text Summarization Workshop/ACL 2004; EsTAL 2004; ECAI 2004; ESSLLI (2004, 2006); IR4QA/SIGIR (2004, 2008); Workshop on Evaluation Initiatives/EACL 2003; ATEM-2003; EuroLan’03; CLEI 2003; Workshop on Spanish Language Resources/SEPLN 2001; CLEI 2001
- Reviewer of the following journals: Information Processing & Management; Journal of Documentation; Journal of Natural Language Engineering; Journal of Language Resources and Evaluation; Computational Linguistics; Traitement Automatique des Langues; AI in Medicine; IEEE Intelligent Systems; IPSI Journal; TALIP; Data & Knowledge Management; Journal of Information Science; Expert Systems

## Conference Organization

- Co-Chair of the Workshop & Shared Task on Quality Assessment for Text Simplification at LREC 2016.
- Co-Chair of the Workshop on Improving Social Inclusion using NLP: Tools and resources at LREC 2016.
- Scientific Chair, 30th Conference of the Spanish Society for Natural Language Processing. 2014.
- Scientific Summarization Track Chair, 24th International Conference on Computational Linguistics. 2012
- Co-Chair of the 14th European Workshop on Natural Language Generation (ENLG). 2013.
- Co-Chair of the 2st Workshop on Natural Language Processing for Improving Text Accessibility. 2013.
- Scientific Summarization Chair, 24th International Conference on Computational Linguistics. 2012
- Co-Chair of the 1st Workshop on Natural Language Processing for Improving Text Accessibility. 2012.
- Scientific Program Chair 7th Brazilian Symposium in Information and Human Language Technology. 2009.
- Co-Chair of the 2nd Workshop on Multisource, Multilingual Information Extraction and Summarization at COLING 2008.



- Co-Chair of the 1st Workshop on Multisource, Multilingual Information Extraction and Summarization at RANLP 2007.
- Organizer of the Workshop on Text Summarization for RANLP 2005.
- Natural Language Processing Seminar's coordinator from 2000 to 2002 and Summarisation reading Group Conveyor from 2002 until 2004.
- Assistant to the organisation of COLING/ACL 1998.
- Member of the organization of 10th ECI (Escuela de Computación, Departamento de Computación, Universidad de Buenos Aires, Argentina) 1996.
- Assistant to the organisation of SBIA (Brazilian Symposium on Artificial Intelligence) 1995.

## Contributions to Natural Language Processing

- Dr Inventor Scientific Text Mining Framework free for research purposes  
<http://taln.upf.edu/pages/dri.upf/>
- Dr Inventor Multi-layer Scientific Corpus  
<http://sempub.taln.upf.edu/dricorpus#download>
- SUMMA toolkit, a text summarization library free for research purposes  
<http://taln.upf.edu/pages/summa.upf/>
- CONCUSUS, a comparable corpus in English and Spanish of summaries free for research purposes  
<http://www.taln.upf.edu/pages/concusus/>
- SummBank, a bilingual aligned corpus in English and Chinese. Distributed by the LDC.  
<https://catalog.ldc.upenn.edu/LDC2003T16>
- Contribution of resources to the GATE system  
<http://gate.ac.uk>

## Professional Societies

- Association for Computer Machinery (ACM) since 2004
- Sociedad Argentina de Infomatica e Investigacion Operativa (SADIO) since 2005
- Institute of Electrical and Electronics Engineers (IEEE) since 2007
- British Computer Society (BCS) 2006-2010.
- Association for Computational Linguistics (ACL) since 1998

## Complementary Studies in Computer Science

- I Escuela Brasilerio-Argentina de Informática (EBAI), Summer 1986, Universidad de Campinas, Brazil.
- II Escuela Brasilerio-Argentina de Informática (EBAI), Summer 1987, Universidad del Centro de la Provincia de Buenos Aires, Argentina.
- III Escuela Brasilerio-Argentina de Informática (EBAI), Summer 1988, Univesidade de Curitiba, Brazil.
- I Escuela de Ciencias Informáticas (ECI), 1987, Universidad de Buenos Aires, Argentina.
- III Escuela de Ciencias Informáticas (ECI), 1989, Universidad de Buenos Aires, Argentina.
- V Escuela de Ciencias Informáticas (ECI), 1991, Universidad de Buenos Aires, Argentina.
- IX Escuela de Computação, 1994, Universidade de Pernambuco, Recife, Brazil.
- ESSLLI European Summer School in Language Logic and Information. 1997 Univertité d'Aix en Provençe, France.

## Software

- Platforms: Unix, Linux, Windows-based Systems, PICK/MICRODATA, PICK/ULTIMATE, VM/SP, Digital-based systems.
- Programming languages: Java, Prolog, Perl, C, DATA/BASIC, PICK/BASIC, PICK/ASSEMBLER, PICK/PROC, PICK/RPL, TCL, CLIPPER, LISP, PICK/ACCESS, PASCAL, COBOL, FORTRAN
- Document preparation: Latex, Word
- Other tools: Office, ACCESS, PICK/DATABASE, DBASE, IBM/VM-AS, IBM/GDDM

## Industrial Experience

- 1991-1993: Software analyst and developer (IBM DOS and PICK SYSTEM) for a number of companies in Buenos Aires, Argentina.
- 1988-1991: System programmer (PICK SYSTEM) - SADE SACCIFIM, Buenos Aires, Argentina. Technical support.
- 1987-1988: Internship SADE SACCIFIM, Buenos Aires, Argentina. Technical support.
- 1986: Internship IBM Argentina, Martinez, Argentina. Technical support.

## Languages

- English: excellent communication skills (oral, written)
- Spanish: excellent communication skills (oral, written)
- French: excellent communication skills (oral, written). Alliance Française Diploma, Buenos Aires, Argentina, 6-years Course
- Portuguese: good communication skills (oral, written)
- Italian: attended introductory courses
- Catalan: level I course