

## Curriculum Vitae – Jan Šnajder

---

CONTACT INFORMATION	University of Zagreb Faculty of Electrical Engineering and Computing Dept. of Electronics, Microelectronics, Computer and Intelligent Systems HR-10000 Zagreb, Croatia	<i>Phone:</i> +385-1-6129-935 <i>Fax:</i> +385-1-6129-653 <i>E-mail:</i> jan.snajder@fer.hr <i>Web:</i> www.zemris.fer.hr/~jan
EDUCATION	<b>PhD in Computer Science</b> , June 2010 University of Zagreb, Faculty of Electrical Engineering and Computing, Croatia  <b>M.Sc. in Computer Science</b> , October 2006 University of Zagreb, Faculty of Electrical Engineering and Computing, Croatia  <b>Diploma in Computer Engineering</b> , July 2002 University of Zagreb, Faculty of Electrical Engineering and Computing, Croatia	
RESEARCH INTERESTS	<ul style="list-style-type: none"><li>• Natural language processing (lexical semantics, opinion mining, user profiling)</li><li>• Text analytics</li><li>• Machine learning</li></ul>	
WORK EXPERIENCE	<b>Associate Professor</b> , Sep 2016 – present University of Zagreb, Faculty of Electrical Engineering and Computing, Croatia  <b>Assistant Professor</b> , April 2011 – Sep 2016 University of Zagreb, Faculty of Electrical Engineering and Computing, Croatia  <b>Visiting Researcher</b> , June 2016 – September 2016 Dept. of Computing and Information Systems, University of Melbourne, Australia  <b>Visiting Researcher</b> , February 2015 – March 2015 National Institute of Information and Communications Technology, Kyoto, Japan  <b>Visiting Researcher</b> , September 2014, September 2015 Institute for Natural Language Processing, University of Stuttgart, Germany  <b>Visiting Researcher</b> , October 2012 – September 2013 Department of Computational Linguistics, Heidelberg University, Germany  <b>Postdoctoral Researcher</b> , June 2010 – March 2011 University of Zagreb, Faculty of Electrical Engineering and Computing, Croatia  <b>Research Assistant</b> , September 2002 – June 2010 University of Zagreb, Faculty of Electrical Engineering and Computing, Croatia	
RESEARCH EXPERIENCE	<b>Project leader</b> <ul style="list-style-type: none"><li>• Establishment of an integrated system for managing the official documentation of the Republic of Croatia – R&amp;D project funded by the European Social Fund (project leader at the partner institution) (2017 – present)</li><li>• SenseHive: Dynamic Crowdsourcing Models for Incremental Construction of Lexico-Semantic Resources – Research project funded by the Croatian Science Foundation (2015 – 2018)</li></ul>	

- CATA CX: Cog-Affective Social Media Text Analytics for Customer Experience Analysis – Proof of Concept project funded by the Croatian Agency for SMEs, Innovations and Investments (HAMAG-BICRO) (2016 – 2017)
- Derivational Semantic Models for Information Retrieval – Research project funded by the Croatian Science Foundation (2012 – 2013)

#### Participation in projects

- EVerBEST: REtrieval Based on semantically Enriched Structures for interactive user Tasks (EVERBEST) – Unity Through Knowledge Fund (2015 – 2017)
- PARSEME (Parsing and Multi-Word Expressions – Towards linguistic precision and computational efficiency in natural language processing) – EU COST Action IC1207 (2013 – 2017)
- Intelligent Text Analyzer – Funded by the Croatian Institute of Technology (2011 – 2012)
- MUMIA (Multilingual and Multifaceted Interactive Information Access) – EU COST Action IC1002 (2011 – 2015)
- Knowledge Discovery in Textual Data – Research project funded by the Croatian Ministry of Science, Education and Sports (2008 – 2014)
- Computer Aided Document Indexing for Accessing Legislation (CADIAL) – Joint Flemish-Croatian research project (2007 – 2009)
- Knowledge Discovery and Visualization for Textual Data – Croatian-French research project, COGITO programme (2007)
- Automating Indexing Using Eurovoc Descriptors (AIDE) – Joint project with Croatian Information Documentation Referral Agency HIDRA (2004 – 2007)
- Automatic Indexing System – IT application project funded by the Croatian Ministry of Science, Education and Sports (2003)
- Multiagent Systems for Interpretation of Dynamic Scenes – Research project funded by the Croatian Ministry of Science, Education and Sports (2002 – 2006)
- Agent Systems on the Web – IT application project funded by the Croatian Ministry of Science, Education and Sports (2001 – 2002)
- Intelligent systems – Educational project funded by the Croatian Academic and Research network CARNET (2001)

#### SELECTED PUBLICATIONS

#### Book chapter

Snajder, J., Dalbelo Basic, B., Tadic, M. (2009) Lexicon-Based Morphological Normalisation and its Application to Croatian Language. *Technologies for the Processing and Retrieval of Semi-Structured Documents: Experience from the CADIAL Project*, 23–80.

#### International journals

Brooke, J., Snajder, J., Baldwin, T. (2017). Unsupervised Acquisition of Comprehensive Multiword Lexicons using Competition in an n-gram Lattice. *Transactions of the Association for Computational Linguistics*, 5: 455–470.

Karan, M., Snajder, J. (2018). Paraphrase-Focused Learning to Rank for Domain-Specific Frequently Asked Questions Retrieval. *Expert Systems with Applications*, 91: 418–433.

- Glavas, G., Snajder, J. (2015) Construction and Evaluation of Event Graphs. *Natural Language Engineering*, 21 (4), 607–652.
- Glavas, G., Snajder, J. (2014) Event Graphs for Information Retrieval and Multi-Document Summarization. *Expert Systems with Applications*, 41 (15), 6904–6916.
- Snajder, J. (2013) Models for Predicting the Inflectional Paradigm of Unknown Croatian Words. *Slovenščina 2.0*, 1 (2), 1–34.
- Karan, M., Glavas, G., Saric, F., Snajder, J., Mijic, J., Silic, A., Dalbelo Basic, B. (2013) CroNER: Recognizing Named Entities in Croatian Using Conditional Random Fields. *Informatica*, 37, 165–172.
- Petrovic, S., Snajder, J., Dalbelo Basic, B. (2010) Extending Lexical Association Measures for Collocation Extraction. *Computer Speech and Language*, 24 (2), 383–394.
- Snajder, J., Dalbelo Basic, B., Tadic, M. (2008) Automatic Acquisition of Inflectional Lexica for Morphological Normalisation. *Information Processing and Management*, 44 (5), 1720–1731.
- Malenica, M., Smuc, T., Snajder, J., Dalbelo Basic, B. (2008) Language Morphology Offset: Text Classification on a Croatian-English Parallel Corpus. *Information Processing and Management*, 44 (1), 325–339.
- Petrovic, S., Snajder, J., Dalbelo Basic, B., Kolar, M. (2006) Comparison of Collocation Extraction Measures for Document Indexing. *Journal of Computing and Information Technology*, 14 (4), 321–327.
- Kolar, M., Vukmirovic, I., Dalbelo Basic, B., Snajder, J. (2005) Computer-Aided Document Indexing System. *Journal of Computing and Information Technology*, 13 (4), 299–305.

### Conference proceedings

- Tutek, D., Snajder, J. (2018). Iterative Recursive Attention Model for Interpretable Sequence Classification. Proceedings of Analyzing and interpreting neural networks for NLP (BlackBoxNLP 2018), Brussels, (in press).
- Alagic, D., Snajder, J., Padó, S. (2018). Leveraging Lexical Substitutes for Unsupervised Word Sense Induction. Proceedings of Thirty-Second AAAI Conference on Artificial Intelligence (AAAI-18), New Orleans, 5004–5011.
- Boltuzic, G., Snajder, J. (2017). Toward Stance Classification Based on Claim Microstructures. Proceedings of the 8th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis (WASSA 2017), Copenhagen, 74–80.
- Medic, Z., Snajder, J., Padó, S. (2017). Does Free Word Order Hurt? Assessing the Practical Lexical Function Model for Croatian. Proceedings of the 6th Joint Conference on Lexical and Computational Semantics (\*SEM 2017), Vancouver, 115–120.
- Boltuzic, F., Snajder, J. (2016). Fill the Gap! Analyzing Implicit Premises between Claims from Online Debates. Proceedings of the 3rd Workshop on Argumentation Mining (ArgMining 2016), ACL 2016, Berlin, 124–133.
- Alagic, D., Snajder, J. (2016). Cro36WSD: A Lexical Sample for Croatian Word Sense Disambiguation. Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016), Portorož, 1689–1694.

Boltuzic, F., Snajder, J. (2015) Identifying Prominent Arguments in Online Debates Using Semantic Textual Similarity. In Proceedings of the 2nd Workshop on Argumentation Mining, (ArgMining 2015), Denver, 110–115.

Kisselew, M., Padó, S., Palmer, A., Šnajder, J. (2015) Obtaining a Better Understanding of Distributional Models of German Derivational Morphology. Proceedings of IWCS 2015, London, 58–63.

Padó, S., Zeller, B., Snajder, J. (2014) Towards Semantic Validation of a Derivational Lexicon. Proceedings of the 25th International Conference on Computational Linguistics (COLING 2014), Dublin, 1728–1739.

Snajder, J. (2014) DerivBase.hr: A High-Coverage Derivational Morphology Resource for Croatian. Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC 2014), Reykjavik, 3371–3377.

Boltuzic, F., Snajder, J. (2014) Back up your Stance: Recognizing Arguments in Online Discussions. Proceedings of the First Workshop on Argumentation Mining (ArgMining 2014), Baltimore, 49–58.

Glavas, G., Snajder, J. (2013) Event-Centered Information Retrieval Using Kernels on Event Graphs. Proceedings of the 8th Workshop on Graph-Based Methods in Natural Language Processing (TextGraphs-8), 1–5.

Snajder, J., Pado, S., Agic, Z. (2013) Building and Evaluating a Distributional Memory for Croatian. Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (ACL 2013), Sofia, 784–789.

Zeller, B., Snajder, J., Pado, S. (2013) DERIVBASE: Inducing and Evaluating a Derivational Morphology Resource for German. Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (ACL 2013), Sofia, 1201–1211.

Pado, S., Snajder, J., Zeller, B. (2013) Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (ACL 2013), Sofia, 731–735.

Glavas, G., Snajder, J. (2013) Recognizing Identical Events with Graph Kernels. Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (ACL 2013), Sofia, 797–803.

Saric F., Glavas, G., Snajder, J., Dalbelo Basic, B. (2012) TakeLab: Systems for Measuring Semantic Text Similarity. Proceedings of the First Joint Conference on Lexical and Computational Semantics (\*SEM), Montreal, 441–448.

Saratlija, J., Snajder, J., Dalbelo Basic, B. (2011) Unsupervised Topic-Oriented Keyphrase Extraction and its Application to Croatian. *Lecture Notes in Artificial Intelligence (Text, Speech and Dialogue, 14th International Conference, TSD 2011)*. 6836, 340–347.

Snajder, J., Dalbelo Basic, B. (2009) String Distance-Based Stemming of the Highly Inflected Croatian Language. Proceedings of Recent Advances in Natural Language Processing (RANLP-2009), Borovets, 411–415.

Snajder, J. Dalbelo Bašić, B. (2008) Higher-order Functional Representation of Croatian Inflectional Morphology. Proceedings of the Sixth International Conference on Formal Approaches to South Slavic and Balkan Languages (FASSBL 2008), Dubrovnik, 121–130.

Snajder, J., Dalbelo Basic, B., Petrovic, S., Sikiric, I. (2008) Evolving New Lexical Association Measures Using Genetic Programming. Proceedings of the 46th Annual Meeting of the Association for Computational Linguistics (ACL 2008), Columbus, 181–184.

A complete list of publications is available from: <http://zemris.fer.hr/~jan/>

TEACHING  
EXPERIENCE

**Thesis supervision**

Supervised three PhD theses, 47 master theses, and 41 bachelor theses. Assisted in supervision of 26 bachelor and master theses.

**Lecturer**

Heidelberg University, Department of Computational Linguistics

- Computational Morphology Seminar (Spring 2013)

**Lecturer**

University of Zagreb, Faculty of Electrical Engineering and Computing, Croatia

- Artificial Intelligence (Spring 2011, Spring 2012, Spring 2014, Spring 2015, Spring 2016, Spring 2017, Spring 2018)
- Machine Learning (Fall 2010, Fall 2011, Fall 2012, Fall 2013, Fall 2014, Fall 2015, Fall 2016, Fall 2017)
- Programming in Haskell (Fall 2011, Fall 2013, Fall 2014, Fall 2015, Fall 2016, Fall 2017)
- Text Analysis and Retrieval (Spring 2014, Spring 2015, Spring 2016, Spring 2017, Spring 2018)

**Teaching assistant**

University of Zagreb, Faculty of Electrical Engineering and Computing, Croatia

- Artificial Intelligence (Spring 2008, Spring 2009, and Spring 2010)
- Intelligent Systems (Fall 2005, Fall 2006, and Fall 2007)
- Machine Learning (Fall 2002, Fall 2003, Fall 2004, and Fall 2006)
- Fuzzy, Evolutionary, and Neuro-computing (Spring 2002, Spring 2004, Spring 2005, and Spring 2006)
- Digital Logic (Fall 2008, Fall 2009, and Fall 2010)
- Digital System Design (Spring 2003, Spring 2004, Spring 2005, and Spring 2006)
- Computer Architecture and Organization (Fall 2002, Fall 2003, Fall 2004, Fall 2005, and Fall 2006)

PROFESSIONAL  
SERVICE

- Associate Editor of the Journal of Computing and Information Technology (2014–present)
- Secretary of the Association for Computational Linguistics Special Interest Group for Slavic Natural Language Processing (ACL SIGSLAV) (2014 – 2017)
- Vice-chair of the IEEE Croatia Section Computer Chapter (2013 – 2015)
- Secretary of the Croatian Language Technologies Society (2012 – present)
- UNIZG FER Departmental ECTS coordinator (2012 – 2016)
- Reviewer for international journals *Automatika*, *DMKD*, *EAAI*, *ESWA*, *IP&M*, *JAI*, *ACMTOIT*, and *SLO 2.0*

- Program Committee member of international scientific conferences: EMNLP 2018, COLING 2018, ACL 2018, SocialNLP2018, ACL 2017, BSNLP 2017, EACL 2017, EMNLP 2017, NAACL 2017, SocialNLP 2017, TextGraphs 2017, CIKM 2016, ECIR 2016, COLING 2016, EMNLP 2016, SocialNLP 2016, NAACL 2016, IJCAI 2016, AIST 2016, LREC 2016, BSNLP 2015, LSDSem 2015, EMNLP 2015, RANLP 2015, ECIR 2015, ISJT 2014, CLIB 2014, EMNLP 2014, PoITAL 2014, COLING 2014, MWE 2014, EACL 2014, MIPRO 2014, ACL 2013, RANLP 2013, MIPRO 2013, KONVENS 2012, MIPRO 2012, FedCSIS 2012, IRFC 2012, ECBS 2011, MIPRO 2011, SNDP 2010, MIPRO 2009, and KES 2008
- Member of the BSNLP 2017, ECIR 2015, and BSNLP 2015 organizing committees
- Reviewer for National IT Application Projects (2004)

PROFESSIONAL MEMBERSHIPS      ACL, ACM, IEEE, IEEE Computer Society, FoLLI, Croatian Language Technologies Society, AMAC-FER

- AWARDS
- Endeavour Research Fellowship of the Australian Government, December 2015
  - Fellowship of the Japanese Society for the Promotion of Science, August 2014
  - Croatian Science Foundation Postdoc Scholarship Award, March 2012
  - Silver plaque *Josip Lončar* for outstanding PhD thesis and successful scientific work, November 2010
  - Member of team awarded *Prime Minister Web Award*, 2009
  - Member of team awarded *VIDI e-Innovation Award*, 2009 and 2007
  - University of Zagreb Rector Award (best seminar work in Computer Science), May 2002

TECHNICAL SKILLS      Programming languages: C/C++, Pascal, Python, SML, Haskell, Prolog, UNIX shell scripting, SQL, VHDL,  $\text{\LaTeX}$ , and Wolfram Mathematica

Operating systems: Linux and Microsoft Windows

LANGUAGES      Mother tongue: Croatian

Fluent: English (TOEFL score 114/120), German (TestDaF score 5.0/5.0)