

CONTACT
INFORMATION

[Computational Lexicology & Terminology Lab \(CLTL\)](#)
[Faculty of Humanities](#)
[The Network Institute](#)
[Vrije Universiteit Amsterdam](#)
De Boelelaan 1105
1081 HV Amsterdam
The Netherlands
Mob: +31 6 499 099 59
E-mail: marieke.van.erp@vu.nl

ABOUT

Marieke van Erp is a researcher and lecturer at the Computational Lexicology and Terminology Lab at the Faculty of Humanities of Vrije Universiteit Amsterdam. She has worked on the intersection of language technology, digital humanities, and cultural heritage since 2005. Since 2009, she has added experience and expertise in Semantic Web to her portfolio.

She is currently working in the interdisciplinary [CLARIAH Project](#) where she is coordinating the VU contributions to the linguistics work package and developing tools to discover and link entities in text.

Her research is focused on making text collections more accessible to researchers and information professionals through applying and developing state-of-the-art natural language processing techniques. Her research is driven by the combination of statistical and knowledge-based methods. She is particularly interested in the role entities and events play in digital humanities and in creating domain-specific knowledge bases to facilitate quantitative methods in traditionally qualitative fields.

EXPERIENCE

[Computational Lexicology and Terminology Lab, Vrije Universiteit Amsterdam](#)

- **November 2016 - present: researcher Robot Revolution**
- **January 2016 - present: researcher [CLARIAH](#)**
- **March 2013 - January 2016: researcher [NewsReader](#)**

- Sept. 2014 - July 2015 Project coordinator “Mining causal graphs from patient records”
- Sept. 2014 - July 2015 Project coordinator “Polemics Visualized - Experiments in Syriac text comparison”

Web and Media Group, Vrije Universiteit Amsterdam

- Oct. 2009 - March 2013: researcher [Agora](#)
- Jan. 2012 - August 2012 Project coordinator “Networking for Business Success”

EDUCATION AND CERTIFICATES

Vrije Universiteit Amsterdam

University Teaching Qualification (BKO), completed Summer 2015

English Language Proficiency Assessment: Level C2, February 2016

Tilburg University

Ph.D., Sept. 2005 - Sept. 2009

- Thesis:
 - Title: Accessing Natural History: Discoveries in Data Cleaning, Structuring, and Retrieval
 - Promotor: [Prof. dr. Antal van den Bosch](#)
 - The goal of the [MITCH project](#) (Mining for Information in Texts from the Cultural Heritage) was to increase accessibility to textual data from the cultural heritage. In this project I developed automatic approaches for data cleaning and I investigated how domain knowledge can aid data retrieval on the log-books and databases from the Netherlands Centre for Biodiversity ([Naturalis](#))
- Internship [USC/ISI](#) Sept. 2007 - Apr. 2008:
 - Supervisor: [Dr. E. Hovy](#)
 - During this internship, I explored manual and automatic methods for knowledge representation, in particular through ontologies. I developed an ontology for the MITCH project based on CIDOC-CRM guidelines to organise the data within MITCH and further the transition from a flat database to a knowledge base.

M.A. Language and Artificial Intelligence, Jan. 2005 - Sept. 2005 (2.1 HONS),

- Thesis:
 - Title: Time for Data Categories
 - Supervisor: [Prof. dr. H.C. Bunt](#)

B.A. Language and Artificial Intelligence, Sept. 2001 - Sept. 2005 (2.1 HONS)

- Thesis:
 - Title: Improving Automatic Text Classification by means of Multiword Expressions: a preparatory study
 - Supervisor: [Prof. dr. H.C. Bunt](#)
- Internship, [Knowledge Concepts](#), Dec. 2003 - Apr. 2004

TEACHING

- Feb-March 2017: Text Mining, Coordinator and instructor
- Nov-Dec 2016: Taal, Media & Communicatie, co-instructor
- Feb-March 2016: Text Mining, coordinator and co-instructor
- December 2015: Taal, Media & Communicatie, co-instructor
- Feb-March 2015: Text Mining, coordinator and co-instructor

- Feb-March 2014: Text Mining, Co-instructor
- November 2013: Computational Linguistics Tutorial 2, Research Master Linguistics. Co-Instructor
- September 2013: Computational Linguistics Tutorial 1, Research Master Linguistics. Co-Instructor
- Feb-March 2013: The Social Web, Master Information Sciences, 2012. Co-Instructor & co-developer
- Nov-Dec 2012: Knowledge & Media (mIS), Co-Instructor
- Feb-March 2012: The Social Web, Master Information Sciences, 2012. Co-Instructor & co-developer
- Knowledge & Media, Master Information Sciences, 2011, Coordinator and Instructor
- Information Retrieval Course, Bachelor Information, Multimedia, and Management, 2011, 2012. Guest Lecturer.
- *Text Mining for Natural History* Richness of the World Module, EBC / Fundamentals of Evolution, Biodiversity and Conservation Course, Master Evolution, Biodiversity and Conservation, Leiden University, 2009-2012. Guest lecturer.
- Session SIKS Advanced Course The Semantic Web, 2011, post-graduate course. Co-instructor & co-developer hands-on session
- *Text Mining for Cultural Heritage* University of Washington Exploration Seminar. 2010, 2011. Guest lecturer.
- Seminar 2011: Search Engines, Bachelor Information, Multimedia, and Management. Track Chair
- Human-Computer Interaction, Bachelor Information, Multimedia, and Management, 2011. Assistant instructor.
- Knowledge & Media, Master Information Sciences, 2011. Assistant instructor

SUPERVISION

- PhD Student: Filip Ilievski: A Quantum Model of Text Understanding (since March 2015)
- MSc Students:
 - 2017:
 - Niels Dekker: Out with the Old and in with the Novel? Evaluating Social Network Extraction for Classic and Modern Fiction. MSc Information Sciences, Supervisor. Ongoing
 - 2016:
 - Sanne Vrijenhoek: Illuminating Dark Entities. MSc Artificial Intelligence, Supervisor. Cum Laude
 - 2015:
 - Filip Ilievski: NED with two-stage coherence optimization. MSc Artificial Intelligence, Supervisor. Cum Laude
 - 2012:
 - Guido van Oorschot: Analysing Football Matches through Twitter. Supervisor. Cum Laude
 - Ardjan van Nuland: Building a Community around Historical Events and Narratives. Supervisor.
 - 2011:
 - Robert Hensel: Improving geographical data in animal specimen datasets. Supervisor.
 - Youssef Ibrahim: Toward Automated Document Classification and Text Segmentation. Daily supervisor.
 - 2010:
 - Vincent Osinga: Fighting Pirates With Words. Daily supervisor.
- BSc Students:
 - 2016:
 - Marvin Kennis: Multi-channel discourse data as an indicator for Bitcoin price movements, Supervisor. Ongoing
 - Niels Dekker: It's About Time; Using Temporal Data to Improve Sentiment Classification on Tweets. A Volkswagen Use-Case. BSc Informatie, Multimedia &

Management, supervisor

- 2015:
Tim Loyen Concept extraction in the medical domain. BSc Informatie, Multimedia & Management, Supervisor
- 2014:
Gleb Satyukov: Semi-Informed Data Sampling. BSc Lifestyle Informatics, Supervisor
- 2012:
Jeffrey Bruijntjes: Creating a mobile interface for the Agora project. Supervisor.
- 2011:
Niek Siekman (2011) Using knowledge of maritime piracy- events to find relevant newspaper articles. Second reader.
- [Information Sciences Project](#), Bachelor Information, Multimedia, and Management , 2009-2012, Steering Committee
- 2009:
Daniel van Norren: Cultural Differences in Enterprise Resources Planning Systems. Thesis committee.
Pieter Kanters: Automatic Mood Classification for Music. Thesis committee.
- 2007:
Niels Bosman: User Centred Interface Design. Een onderzoek naar het ontwikkelen van een User Interface voor de error corrector voor MITCH. Co-supervisor.

INVITED TALKS INTERNATIONAL

- NewsReader: Automatically extracting Events, Entities and Perspectives from Newspapers. CLARIN-Plus workshop. Leuven, September 2016
- NewsReader: Making large news streams manageable. The University of Sydney, School of Information Technologies, Sydney, Australia, October, 2013
- Agora: putting museum objects into their art-historic context. EURECOM: The Multimedia Group Seminar Series. EURECOM: The Multimedia Group Seminar Series. Sophia Antipolis, France. July 2012

INVITED TALKS NATIONAL

- Information Extraction from Semi-structured Text. CREATE Salon. Universiteit van Amsterdam. 15 November 2016
- The domain as unifier, how focusing on social history can bring technical fields together. CEDAR Final Event. International Institute for Social History. Amsterdam. June 2016
- From Events to Stories. Different ways of structuring the same bag of events over time. eHumanities workshop. Soeterbeeck. June 2013

SERVICE

- Programme Committee
 - [Language, Data and Knowledge 2017 \(LDK 2017\)](#)
 - [ESWC Research Track 2016, 2017](#)
 - [European Chapter of the Association for Computational Linguistics \(EACL 2017\) Student Research Workshop](#)
 - [Digital Humanities 2012, 2013, 2014, 2015, 2016, 2017](#)
 - [European Conference on Information Retrieval \(ECIR\) 2014, 2015, 2016, 2017](#)
 - [ACM SIGAPP Symposium On Applied Computing 2017](#)
 - [International Conference on Knowledge Engineering and Knowledge Management \(EKAW\) 2016 Research Track](#)
 - [Language Resources and Evaluation Conference \(LREC\) 2010, 2012, 2014, 2016](#)
 - [International Semantic Web Conference \(ISWC\) 2016 Research Track](#)
 - [International Semantic Web Conference \(ISWC\) Poster & Demo track 2014, 2015, 2016](#)
 - [Empirical Methods in Natural Language Processing \(EMNLP\) 2015 short papers](#)

- [Linked Data in Linguistics 2013, 2014, 2015](#)
- [Natural Language Processing for Informal Text \(NLPIT\) 2015](#)
- [Replicability and Reusability in Natural Language Processing \(RRNLP\) 2015](#)
- [Italian Conference on Computational Linguistics \(CLIC\) 2014](#)
- [Extended Semantic Web Conference\(ESWC\) 2014 Poster & Demo track](#)
- [Machine Learning and Data Mining for Sports Analytics \(MLSA\) 2013](#)
- [Web of Linked Entities \(WoLE\) 2013](#)
- [Dutch-Belgian Information Retrieval Workshop \(DIR\) 2011, 2012, 2013](#)
- [Intelligent User Interfaces \(IUI\) 2013](#)
- [WebScience 2012](#)
- [Benelearn 2012](#)
- [ESSLLI 2011 Student Session](#)

Journal reviewing

- [Journal of Web Semantics](#) 2011, 2012, 2016 (member of the editorial board since October 2016)
- [Data Mining and Knowledge Discovery](#), 2016
- [Semantic Web Journal](#), 2015
- [Computers in Industry](#), 2014
- [International Journal On Semantic Web and Information Systems](#), 2013
- [Journal on Data Semantics](#), 2013
- [Transactions in GIS](#), 2013
- [Computational Linguistics in the Netherlands](#), 2009, 2010

Organising Committee

- [The 3rd Workshop on Noisy User-generated Text \(W-NUT\) Shared task #2: Novel and Emerging Entity Recognition](#)
- [Events and Stories in the News 2017](#)
- [NewsStory: Computing News StoryLines: 2015, 2016,](#)
- [Named Entity rEcognition and Linking \(NEEL\) Challenge 2016](#)
- [NLP&DBpedia: 2014, 2015, 2016](#)
- [Linked Science: 2013, 2014, 2015](#)
- [LaTeCH 2015: 9th Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities](#)
- [Detection, Representation and Exploitation of Events in the Semantic Web \(DeRiVE\): 2011, 2012, 2013, 2015](#)
- [Semantic Web and Information Extraction \(SWAIE\): 2012, 2013, 2014](#)
- [New Models of Interaction: GLAMS, Linked Open Data & User participation, 2011](#)
- [Belgian-Dutch Conference on Machine Learning 2009 \(Benelearn 09\), Tilburg, May 2009](#)
- [Eighth International Conference on Computational Semantics \(IWCS-8\), Tilburg, January 2008](#)

Other

- Sept. 2016: Co-organiser of [CLARIAH LOD workshop](#) September 2016
- 2015: Organisation of NewsReader Hackathon & Evaluation Events, [January](#) and [November](#) 2015
- Sep. 2016 - present: Network Institute Academy Assistants Coordinator
- Feb. 2015 - present: Founder and co-organiser Language, Literature and Communication Lunch Seminar
- Dec. 2013 - Dec.2015: [ACL Special Interest Group on Language Technologies for the Socio-Economic Sciences and Humanities \(SIGHUM\) Liason Representative](#),
- Oct. 2010 - May 2012: Member of VU CS Outreach group (defining strategies for collaboration with secondary schools etc.).
- June 2011: Coordinated and presented for "Informatie, Multimedia & Management" degree programme at CS Last Minute day for prospective students.

- Mar. 2011, 2012: 'Dagje studeren': Lecture for prospective "Informatie, Multimedia & Management" students.
- Nov. 2010, Feb. 2011, Nov. 2011: Promoted "Informatie, Multimedia & Management" degree programme at VU Bachelor day for prospective students.

LANGUAGES (CEFR LEVELS)

- Dutch: native speaker
- English: C2
- French: B1
- German: A2
- Spanish: A2

PROGRAMMING & OTHER

- Python, Perl
- SPARQL, MySQL
- Linux, Mac OS X

SOFTWARE

- <https://github.com/cltl/multilingual-finegrained-entity-typing> Module to detect entity types for Dutch and Spanish based on Wikipedia, DBpedia and word embeddings
- <https://github.com/MvanErp/entity-typing> Modules to detect entity types based on distributional semantics and DBpedia
- <https://github.com/cltl/entity-link-postprocess> Postfix for entities detected by the NewsReader pipeline
- <https://github.com/cltl/entity-relink-pipeline> Postfix for entities detected by the NewsReader pipeline, pipeline version
- <https://github.com/MvanErp/tm2016-staffscrape> Scripts to download VU staff homepages for Text Mining students to perform text analysis on
- <https://github.com/newsreader/vua-nedtype-reranker> Postfix for entity links assigned by the NewsReader pipeline
- <https://github.com/newsreader/vua-srl-nl> Reimplementation of Dutch semantic role labeller as described in Orphée De Clercq and Veronique Hoste and Paola Monachesi (2012) Evaluating automatic cross-domain semantic role annotation. In: Proceedings of the 8th International Conference on Language Resources and Evaluation Conference (LREC-2012). Istanbul, Turkey. pp 88-93
- <https://github.com/MvanErp/Factuality-Classifier> First version of factuality detector to automatically classify whether statements in a text are factual (have happened or haven't happened), probable, possible or underspecified using FactBank.

MEDIA

- [My presentations on SlideShare](#)
- [Interview with NPO Radio 1 on the National Library's Hackalod](#), November 2016
- [Audience Prize at National Library's Hackalod](#)

- [Video of my presentation at CLARIN-Plus workshop "Working with Digital Collections of Newspapers"](#), September 2016
- January 2016: [Visualisation of NewsReader deliverables evolution](#)
- February 2015: [NewsReader Hackathon promo video](#)
- November 2009: Interview for "CATCH: the Future" promo video
- January 27, 2009: Interview for Noorderlicht Radio on text mining and the MITCH project

PUBLICATIONS

JOURNAL

- Giuseppe Rizzo, Bianca Pereira, Andrea Varga, Marieke van Erp, Amparo Elizabeth Cano Basave (2017) Lessons Learnt from the Named Entity rEcognition and Linking (NEEL) Challenge Series. *To appear in: Semantic Web Journal*
- Piek Vossen, Rodrigo Agerri, Itziar Aldabe, Agata Cybulska, Marieke Van Erp, Antske Fokkens, Egoitz Laparra, Anne-Lyse Minar, Alessio Palmero Aprosio, German Rigau, Marco Rospocher (2016) NewsReader: using knowledge resources in a cross-lingual reading machine to generate more knowledge from massive streams of news. *Knowledge-Based Systems*.
- Marco Rospocher, Marieke van Erp, Piek Vossen, Antske Fokkens, Itziar Aldabe, German Rigau, Aitor Soroa, Thomas Ploeger, and Tessel Bogaard (2016) Building Event-Centric Knowledge Graphs from News. *Journal of Web Semantics*. Volumes 37-38. Pages. 132-151
- L. Derczynski, D. Maynard, G. Rizzo, Marieke van Erp, G. Gorrell, R. Troncy, J. Petrak, and K. Bontcheva (2015) "Analysis of named entity recognition and linking for tweets" *Information processing & management*, vol. 51, iss. 2, pp. 32-49.
- Albert Meroño-Peñuela, Ashkan Ashkpour, Marieke van Erp, Kees Mandemakers, Leen Breure, Andrea Scharnhorst, Stefan Schlobach, Frank van Harmelen (2015) Semantic Technologies for Historical Research: A Survey. *Semantic Web journal*
- Marieke van Erp, Robert Hensel, Davide Ceolin and Marian van der Meij (2014) Georeferencing Animal Specimen Datasets. *Transactions in GIS*.
- Willem Robert van Hage, Marieke van Erp and Véronique Malaisé (2012) Linked Open Piracy: A story about e-Science, Linked Data, and statistics, *Journal of Data Semantics*
- Antal van den Bosch, Piroska Lendvai, Marieke van Erp, Steve Hunt, Marian van der Meij and René Dekker (2009) Weaving a New Fabric of Natural History, *Interdisciplinary Science Reviews*, Vol. 34 No. 2, June, 2009, 210-227
- Antal van den Bosch, Marieke van Erp, and Caroline Sporleder (2009) Making a Clean Sweep of Cultural Heritage. *IEEE Intelligent Systems, Special Issue on AI and Cultural Heritage* March/April 2009 (vol. 25 no. 2), pp. 54-63

CONFERENCE

- Filip Ilievski, Wouter Beek, Marieke van Erp, Laurens Rietveld, Stefan Schlobach (2016). LOTUS: Adaptive Text Search for Big Linked Data. *Proceedings of The Extended Semantic Web Conference (EWSC 2016)*. 29 May-2 June. Crete, Greece. Pages: 470-485

- Marieke van Erp, Pablo Mendes, Heiko Paulheim, Filip Ilievski, Julien Plu, Giuseppe Rizzo and Jörg Waitelonis (2016) Evaluating Entity Linking: An Analysis of Current Benchmark Datasets and a Roadmap for Doing a Better Job. *Proceedings of LREC 2016*.
- Giuseppe Rizzo, Filip Ilievski, Marieke van Erp, Julien Plu and Raphaël Troncy (2016) Context-enhanced Adaptive Entity Linking. *Proceedings of LREC 2016*.
- Anne-Lyse Minard, Manuela Speranza, Ruben Urizar, Begoña Altuna, Marieke van Erp, Anneleen Schoen and Chantal van Son (2016) MEANTIME, the NewsReader Multilingual Event and Time Corpus. *Proceedings of LREC 2016*.
- Giuseppe Rizzo, Marieke van Erp and Raphaël Troncy (2014) Benchmarking the Extraction and Disambiguation of Named Entities on the Semantic Web. *Proceedings of LREC 2014*. Reykjavik, Iceland. May 26 - 31
- Chantal van Son, Marieke van Erp, Antske Fokkens and Piek Vossen (2014) Hope and Fear: Interpreting Perspectives by Integrating Sentiment and Event Factuality. *Proceedings of LREC 2014*. Reykjavik, Iceland. May 26 - 31
- Marieke van Erp, Gleb Satyukov, Piek Vossen and Marit Nijssen (2014) Discovering and visualising stories in news. *Proceedings of LREC 2014*. Reykjavik, Iceland. May 26 - 31
- Jordi Atserias, Marieke van Erp, Isa Maks, German Rigau, J. Fernando Sánchez-Rada (2014) EuroLoveMap: Confronting feelings from News. *Proceedings of OpeNER Hackathon*. Reykjavik, Iceland, 26 May 2014
- Antske Fokkens, Marieke van Erp, Marten Postma, Ted Pedersen, Piek Vossen and Nuno Freire (2013) Offspring from Reproduction Problems: What Replication Failure Teaches Us. *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics*, pages 1691-1701, Sofia, Bulgaria, August 4-9 2013.
Nominated for Best Paper Award
- Chiel van den Akker, Marieke van Erp, Lora Aroyo, Ardjan van Nuland, Lourens van der Meij, Guus Schreiber and Susan Legêne (2013) From Information Delivery to Interpretation Support: Evaluating Cultural Heritage Access on the Web. *Proceedings of WebSci'13*. Paris, France, May 2013.
- Johan Oomen and Marieke van Erp and Lotte Belice Baltussen (2012) Sharing cultural heritage the linked open data way - Why you should sign up. *Museums and the Web 2012*, San Diego, CA, USA, April 11-14 2012.
- Chiel Van Den Akker, Susan Legêne, Marieke van Erp, Lora Aroyo, Roxane Segers, Lourens Van Der Meij, Jacco Van Ossenbruggen, Guus Schreiber, Bob Wielinga, Johan Oomen and Geertje Jacobs (2011) Agora Digital Hermeneutics: Online Understanding of Cultural Heritage. *Proceedings of the 3rd International Conference on Web Science (WebSci'11)*. Koblenz, Germany 14 - 17 June 2011
Nominated for Best Paper Award
- Marieke van Erp, Johan Oomen, Roxane Segers, Chiel van den Akker, Lora Aroyo, Geertje Jacobs, Susan Legêne, Lourens van der Meij, Jacco van Ossenbruggen and Guus Schreiber (2011) *Automatic Heritage Metadata Enrichment with Historic Events*. *Museums and the Web 2011*. Philadelphia, PA, USA, April 6-9 2011
- Marieke van Erp (2006) Bootstrapping Multilingual Geographical Gazetteers from Corpora. *Proceedings of the 11th ESSLLI Student Session*, pp. 192-202, Malaga, Spain, 31 July - 11 August 2006.
- Caroline Sporleder, Marieke van Erp, Tijn Porcelijn, Antal van den Bosch and Pim Arntzen (2006) Identifying Named Entities in Text Databases from the Natural History Domain. *Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC-06)*, pp. 1742-1745, Genoa, Italy, 2006.

THESIS

- Marieke van Erp (2010) *Accessing Natural History: Discoveries in Data Cleaning, Structuring, and Retrieval*. Ph.D. Thesis, Tilburg University, 2010. ISBN: 978-90-8559-027-9. 197 pages

BOOK CHAPTER

- Marieke van Erp, Antal van den Bosch, Steve Hunt, Marian van der Meij, René Dekker and Piroska Lendvai(2011) *Natural Selection - Finding Specimens in a Natural History Collection*. *Changing Diversity in a Changing Environment*, ISBN 978-953-307-796-3

WORKSHOP

- Marieke van Erp and Piek Vossen (2016) *Entity Typing using Distributional Semantics and DBpedia*. *NLP&DBpedia 2016*. Kobe, Japa. 18 October 2016.
- Giuseppe Rizzo, Marieke van Erp, Julien Plu, and Raphaë Troncy (2016). *Making Sense of Microposts (#Microposts2016) Named Entity rEcognition and Linking (NEEL) Challenge*. *Proceedings of the 6th Workshop on 'Making Sense of Microposts'* (pp. 50-59). Montreal, Canada. Filip Ilievski, Wouter Beek, Marieke van Erp, Laurens Rietveld, and Stefan Schlobach (2015) *LOTUS: Linked Open Text UnleaShed*. *Proceedings of COLD 2015*.
- Marieke van Erp, Filip Ilievski, Marco Rospocher and Piek Vossen (2015) *Missing Mr. Brown and buying an Abraham Lincoln - Dark Entities and DBpedia*. *Proceedings of NLP&DBpedia 2015*
- Tommaso Caselli, Piek Vossen, Marieke van Erp, Antske Fokkens, Filip Ilievski, Ruben Izquierdo Bevia, Minh Lê, Roser Morante, and Marten Postma (2015) *When it's all piling up: investigating error propagation in an NLP pipeline*. *NLP Applications: completing the puzzle*. Passau, Germany. 16 June 2015
- Anne-Lyse Minard, Manuela Speranza, Eneko Agirre, Itziar Aldabe, Marieke van Erp, Bernardo Magnini, German Rigau, Rub'On Urizar (2015) *SemEval-2015 Task 4: TimeLine: Cross-Document Event Ordering*. *Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015)* Denver, Colorado, USA. 4 June 2015.
- Giuseppe Rizzo, Marieke van Erp and Raphaël Troncy (2014) *Inductive Entity Typing Alignment*. *Linked Data for Information Extraction 2014 (LD4IE 2014)*. Riva del Garda, Italy, October 2014
- Marieke van Erp, Antske Fokkens and Piek Vossen (2014) *Finding Stories in 1,784,532 Events: Scaling up computational models of narrative*. *Workshop on Computational Models of Narrative (CMN'14)*. Quebec City, Canada, July 31 – August 2.
- Antske Fokkens, Marieke van Erp, Piek Vossen, Sara Tonelli, Willem van Hage, Luciano Serafini, Rachele Sprugnoli and Jesper Hoeksema (2013) *GAF: A Grounded Annotation Framework for Events*. *The 1st Workshop on Events: Definition, Detection, Coreference, and Representation*. Workshop in conjunction with NAACL-HLT 2013, Atlanta, GA, USA, June 2013.
- Marieke van Erp, Giuseppe Rizzo and Raphaël Troncy (2013) *Learning with the Web: Spotting Named Entities on the intersection of NERD and Machine Learning*. *#MSM2013 Concept Extraction Challenge*. Rio de Janeiro, Brazil, May 2013. **System ranked 2nd in Challenge**

- Guido van Oorschot, Marieke van Erp, and Chris Dijkshoorn (2012) Automatic Extraction of Soccer Game Events from Twitter. *Proceedings of the Workshop on Detection, Representation, and Exploitation of Events in the Semantic Web (DeRiVE 2012)*, Boston, MA, USA, November 12, 2012.
- Chiel van den Akker, Marieke van Erp, Lora Aroyo, Roxane Segers, Lourens van der Meij, Susan Legêne, and Guus Schreiber (2012) Understanding Objects in Online Museum Collections by Means of Narratives. *Computational Models of Narrative*. Istanbul, Turkey, May 26-27, 2012.
- Marieke van Erp (2012) Reusing Linguistic Resources: Tasks and Goals for a Linked Data Approach. *Linked Data in Linguistics Workshop 2012*, Frankfurt/Main, Germany, March 7-9 2012.
- Willem Robert van Hage, Véronique Malaisé and Marieke van Erp (2011) Linked Open Piracy. *Proceedings of Detection, Representation and Exploitation of Events in the Semantic Web (DeRiVE 2011)*. Bonn, Germany 23 October 2011.
- Chiel van den Akker, Lora Aroyo, Agata Cybulska, Marieke van Erp, Peter Gorgels, Laura Hollink, Cathy Jager, Susan Legene, Lourens van der Meij, Johan Oomen, Jacco van Ossenbruggen, Guus Schreiber, Roxane Segers, Piek Vossen and Bob Wielinga (2010) Historical Event-based Access to Museum Collections. *Proceedings of the First International Workshop on Recognising and Tracking Events on the Web and in Real Life (EVENTS2010)*, Athens, Greece, May, 2010.
- Marieke van Erp and Steve Hunt (2010) Knowledge-driven Information Retrieval for Natural History. *Proceedings of the 10th Dutch-Belgian Information Retrieval Workshop (DIR2010)*, pp 31-38, Nijmegen, The Netherlands, January 25, 2010.
- Marieke van Erp, Antal van den Bosch, Sander Wubben and Steve Hunt (2009) Instance-Driven Discovery of Ontological Relation Labels. *Proceedings of EACL 2009 Workshop on Language Technology and Resources for Cultural Heritage, Social Sciences, Humanities, and Education (LaTeCH - SHELTeR 2009)*, Athens, Greece, March 30, 2009
- Marieke van Erp, Piroska Lendvai, and Antal van den Bosch (2009) Comparing Alternative Data-Driven Ontological Vistas of Natural History. *Proceedings of IWCS-8*, Tilburg, The Netherlands, 2009.
- Marieke van Erp, Piroska Lendvai, Antal van den Bosch and Steve Hunt (2008) Automatic Ontology Construction for Improved Access to Taxonomic Databases. *International Workshop on Distributed Sensing and Collective Intelligence in Biodiversity Monitoring (BioDivGrid'08)*. Amsterdam, The Netherlands, 2008.
- Marieke van Erp (2007) Retrieving Lost Information from Textual Databases: Re-discovering Expeditions from an Animal Specimen Database. *Proceedings of the ACL 2007 Workshop on Language Technology for Cultural Heritage Data (LaTeCH 2007)*, Prague, Czech Republic, June 28, 2007.
- Caroline Sporleder, Marieke van Erp, Tijn Porcelijn and Antal van den Bosch (2006) Correcting 'Wrong-Column' Errors in Text Databases. *Proceedings of the Annual Machine Learning Conference of Belgium and The Netherlands (Benelearn-06)*, pp. 49-56, Ghent, Belgium, 2006.
- Caroline Sporleder, Marieke van Erp, Tijn Porcelijn and Antal van den Bosch (2006) Spotting the 'Odd-one-out': Data-Driven Error Detection and Correction in Textual Databases. *Proceedings of the EACL 2006 Workshop on Adaptive Text Extraction and Mining (ATEM-06)*, pp. 41-48, Trento, Italy, 2006.
- Marieke van Erp and Piek Vossen (2016) Deep Reading for Disasters. Presented at RIMMA 2016, 27 - 28 June 2016. Berlin, Germany

- Leon van Wissen, Marieke van Erp, Ben Peperkamp (2016) Discovering an Encyclopaedic Novel: a case study in automatically analysing Harry Mulisch's *The Discovery of Heaven* (1992). Presented at DHBenelux 2016, 9-10 June 2016. Belval, Luxembourg
- Isa Maks, Marieke van Erp, Piek Vossen, Rinke Hoekstra, Nicoline van der Sijs (2016) Integrating Diachronous Conceptual Lexicons through Linked Open Data. Presented at DHBenelux 2016, 9-10 June 2016. Belval, Luxembourg
- Chantal van Son, Marieke van Erp, Antske Fokkens, Paul Huygen, Ruben Izquierdo Bevia and Piek Vossen (2015) Details from a distance? A Dutch pipeline for event detection. Presented at CLIN 26, 18 December 2015. Amsterdam, The Netherlands
- Hannes Vlaardingerbroek, Marieke Van Erp, Wido van Peursen (2015) Polemics Visualized - morphological analysis for Syriac. Presented at DHBenelyx 2015, Antwerp, Belgium
- Hannes Vlaardingerbroek, Marieke Van Erp, Wido van Peursen (2015) Polemics Visualised: experiments in Syriac text comparison, Presented at: CLIN 25, February 2015, Antwerp, Belgium
- Filip Ilievski, Marieke van Erp, Piek Vossen, Wouter Beek, Stefan Schlobach (2015) Named Entity Disambiguation with two-stage coherence optimization, Presented at: CLIN 25, February 2015, Antwerp, Belgium
- Marieke Van Erp, Antske Fokkens, Piek Vossen, Sara Tonelli, Willem Robert Van Hage, Luciano Serafini, Rachele Sprugnoli and Jesper Hoeksema (2013) Denoting Data in the Grounded Annotation Framework. ISWC 2013 Posters and Demos. Sydney Australia, 21-25 October 2013
- Cristina-Iulia Bucur, Katja Jonsas, Marieke van Erp and Heidi Dahles (2012) Network Capitalism: Chinese business linkages traced through website analysis. *WebSci12*, Evanston, Ill, USA, June 22-24, 2012.
- Peter Boot, Marieke van Erp, Lora Aroyo and Guus Schreiber (2012) The changing face of the book review. *WebSci12*, Evanston, Ill, USA, June 22-24, 2012.
- Roxane Segers, Marieke van Erp, Lourens van der Meij, Lora Aroyo, Guus Schreiber, Bob Wielinga, Jacco van Ossenbruggen, Johan Oomen, Geertje Jacobs (2011) Hacking History via Event Extraction. Poster presented at The Sixth International Conference on Knowledge Capture (K-CAP 2011), Banff, Canada, 25 - 29 June 2011
- Willem R. van Hage, Véronique Malaisé, Marieke van Erp, Guus Schreiber (2011) Linked Open Piracy. Poster presented at The Sixth International Conference on Knowledge Capture (K-CAP 2011), Banff, Canada, 25 - 29 June 2011
- Marieke van Erp, Véronique Malaisé, Willem Robert van Hage, Vincent Osinga and Juan-Manuel Coletto (2011) Parse and Tag Somali Pirates. Abstract presented at Computational Linguistics in the Netherlands. Ghent, Belgium, February 11 2011.
- Roxane Segers, Marieke van Erp and Lourens van der Meij (2011) Extraction of Historical Events from Unstructured Texts. Abstract presented at Computational Linguistics in the Netherlands. Ghent, Belgium, February 11 2011.
- Marieke van Erp (2009) Constructing a Domain Ontology from a Domain Database and the Web. Abstract presented at The 19th Meeting of Computational Linguistics in the Netherlands (CLIN-09)

- Marieke van Erp, Antal van den Bosch, Piroska Lendvai and Steve Hunt (2008) Feature selection techniques for database cleansing: knowledge-driven vs. greedy search. Abstract presented at and in Proceedings of the annual machine learning conference of Belgium and The Netherlands (Benelearn 2008) Spa, Belgium, May 20-21 2008.
- Marieke van Erp (2008) Text Mining in Natural History. Poster presented at SIREN 2008, Amsterdam, Sept. 29, 2008.
- Antal van den Bosch, Caroline Sporleder, Marieke van Erp, and Stephen Hunt (2007) Automatic techniques for generating and correcting cultural heritage collection metadata. Poster presented and abstract in Proceedings of Digital Humanities 2007, Urbana-Champaign, June 2-8, 2007
- Marieke van Erp (2007) What do objects 11551 and 15576 have in common? Finding hidden variables in a textual database. CLIN 17 - Computational Linguistics in the Netherlands, Leuven, Belgium, January 12, 2007
- Marieke van Erp (2006) Bootstrapping Multilingual Geographical Gazetteers from Corpora. Poster presented at the Annual Machine Learning Conference of Belgium and The Netherlands (Benelearn-06), Ghent, Belgium, May 10-11 2006.
- Caroline Sporleder, Marieke van Erp, Tijn Porcelijn, Antal van den Bosch and Pim Arntzen (2006) Identifying Named Entities in Text Databases from the Natural History Domain. Poster presented at the Fifth International Conference on Language Resources and Evaluation (LREC-06), Genoa, Italy, May 24-26, 2006.
- Caroline Sporleder, Marieke van Erp, Tijn Porcelijn, Antal van den Bosch and Pim Arntzen Cleaning and Enriching Textual Databases from the Natural History Domain. Poster presented at the 28th Annual Meeting of the DGfS, Poster Session SIG Computational Linguistics, Bielefeld, Germany, February 22-24, 2006, and at SIREN 2006, Utrecht, 12 October 2006.
- Marieke van Erp, Caroline Sporleder, Tijn Porcelijn and Antal van den Bosch (2005) Correcting Taxonomic Errors in a Specimen Database. Abstract presented at The 16th Meeting of Computational Linguistics in the Netherlands (CLIN-05), Amsterdam, December 16, 2005.
- Caroline Sporleder, Marieke van Erp, Tijn Porcelijn, Antal van den Bosch and Pim Arntzen (2005) Cleaning and Enriching Data on Reptiles and Amphibians. Poster presented at SIREN 2005, Eindhoven, 6 October 2005
- Anneleen Schoen, Chantal van Son, Marieke van Erp and Hennie van der Vliet. NewsReader Document-Level Annotation Guidelines – Dutch. NWR-2014-08. VU University Amsterdam.
- Anne-Lyse Minard, Manuela Speranza, Bernardo Magnini, Marieke van Erp, Itziar Aldabe, Rubén Urizar, Eneko Agirre and German Rigau (2014) TimeLine: Cross-Document Event Ordering. SemEval 2015 – Task 4. NWR-2014-10. Fondazione Bruno Kessler.
- Anne-Lyse Minard, Alessandro Marchetti, Manuela Speranza, Bernardo Magnini, Marieke van Erp, Itziar Aldabe, Rubén Urizar, Eneko Agirre and German Rigau (2015) TimeLine: Cross- Document Event Ordering. SemEval 2015 – Task 4. Annotation Guidelines. NWR-2014- 11. Fondazione Bruno Kessler.
- Marieke van Erp and Lourens Van der Meij (2013) Reusable Research? A Case Study in Named Entity Recognition. CLTL 2013-01, Computational Lexicology & Terminology Lab, VU University Amsterdam.

PROJECT
DELIVERABLES

- Caroline Sporleder, Marieke van Erp, Tijn Porcelijn, Antal van den Bosch, Pim Arntzen and Erik van Nieukerken (2006) Cleaning and Enriching Research Data on Reptiles and Amphibians. The MITCH Pilot Project and “nulmeting”. Technical Report, ILK 06-01, Tilburg University, 2006.
- Rodrigo Agerri, Itziar Aldabe, Zuhaitz Beloki, Egoitz Laparra, German Rigau, Aitor Soroa, Marieke van Erp, Antske Fokkens, Filip Ilievski, Ruben Izquierdo, Roser Morante, Chantal van Son, Piek Vossen, Anne-Lyse Minard (2015) Event Detection v3. NewsReader Deliverable D4.2.3.
- Rodrigo Agerri, Itziar Aldabe, Zuhaitz Beloki, Egoitz Laparra, Maddalen Lopez de Lacalle, German Rigau, Aitor Soroa, Antske Fokkens, Ruben Izquierdo, Marieke van Erp, Piek Vossen, Christian Girardi, Anne-Lyse Minard (2014) Deliverable D4.2.2 Event Detection, version 2.
- Marieke van Erp, Roxane Segers, Filip Ilievski, Antske Fokkens, Rodrigo Agerri, Marco Rospocher, Itziar Aldabe, Egoitz Laparra, German Rigau, Sanne Vrijenhoek (2015) Domain model for financial and economic events, version 3. NewsReader Deliverable D.5.2.2
- Marieke van Erp, Roxane Segers, Piek Vossen, Antske Fokkens, Filip Ilievski, Marco Rospocher, Rodrigo Agerri, Egoitz Laparra, German Rigau (2014) Deliverable D5.2.1 Domain model for financial and economic events, version 1.
- Pim Stouten, Rutger Kortleven, Tessel Bogaard, Thomas Ploeger, Marieke van Erp (2015) Evaluation by decision makers NewsReader Deliverable D8.2.3.
- Ian Hopkinson , Francis Irving , Aidan McGuire, Piek Vossen, Marieke van Erp, Antske Fokkens, Pim Stouten, Rutger Kortleven, Marco Rospocher, Mohammed Qwaider, Anne-Lyse Minard, Itziar Aldabe, German Rigau, Tessel Bogaard (2015) NewsReader Exploitation Plan v2. NewsReader Deliverable D9.2
- Ian Hopkinson, Aidan McGuire, Piek Vossen, Marieke van Erp, Pim Stouten, Rutger Kortleven, Marco Rospocher, Itziar Aldabe (2015) Deliverable 9.1 NewsReader Exploitation Plan v1 .
- Anne-Lyse Minard, Manuela Speranza, Marieke van Erp, Antske Fokkens, Marten Postma, Piek Vossen, Eneko Agirre, Itziar Aldabe, German Rigau, Ruben Urizar (2015) Evaluation tasks in open competitions. NewsReader Deliverable D10.4
- Marieke van Erp, Selene Kolman, Rutger Kortleven (2015) Project Workshop. NewsReader Deliverable D10.5.
- Marieke van Erp, Piek Vossen (2015) Project Showcase. NewsReader Deliverable D10.6.
- Marieke van Erp, Piek Vossen, Rodrigo Agerri, Anne-Lyse Minard, Manuela Speranza, Ruben Urizar, Egoitz Laparra, Itziar Aldabe, German Rigau (2015) Deliverable D3.3.2 Annotated Data, version 2