

Marjan Van de Kauter

Contact details

LT3 Language and Translation Technology Team
Department of Translation, Interpreting and Communication
Faculty of Arts and Philosophy
Ghent University
Groot-Brittanniëlaan 45
B-9000 Ghent
Belgium

E-mail: Marjan.VandeKauter@UGent.be

Web page: <http://www.lt3.ugent.be/people/marjan-van-de-kauter>

SCIENTIFIC EXPERIENCE

PhD researcher, Ghent University January 2012 - present
I'm currently a PhD researcher in the LT3 Language and Translation Technology Team. The goal of my project SentiFM (Sentiment mining for Financial Markets) is to create a system for (feature-based) sentiment analysis in Dutch and English financial news articles. My PhD is supervised by Véronique Hoste of the LT3 team (Ghent University) and co-supervised by Diane Breesch of the Department of Business (BUSI) - Accounting and Auditing (Vrije Universiteit Brussel). I'm funded by a PhD grant for Strategic Basic Research of the Agency for Innovation by Science and Technology in Flanders (IWT).

Visiting researcher, University of Mannheim February - April 2015
I was a visiting researcher in the Data and Web Science Group under the supervision of Prof. Dr. Simone Paolo Ponzetto (head of the Natural Language Processing focus group). I investigated the use of Linked Open Data and web data for sentiment analysis in the financial domain. My research stay was funded by an FWO travel grant.

Research assistant, Ghent University College January - December 2011
As a research assistant in the LT3 team, I worked on the IWT-funded TExSIS project (Terminology Extraction for Semantic Interoperability and Standardization). The goal of this project was to develop a system for the automatic extraction of mono- and multilingual terminology from domain-specific Dutch, English, French and German texts. My work focused on creating automatic preprocessing tools (for sentence splitting, tokenization, Part-of-Speech tagging, lemmatization, Named Entity Recognition and chunking) for all four languages. I was also involved in the development of the modules for sentence alignment and terminology extraction.

ACADEMIC DEGREES

Master in Translation, Ghent University College 2010
Dutch, German, Czech
Graduated magna cum laude
Master's thesis: "Die Verwendung von CAT-Tools in Sprachendiensten des öffentlichen Sektors in Belgien: eine Marktforschung"

Bachelor in Applied Linguistics, Ghent University College 2009
Dutch, German, Czech
Graduated summa cum laude
Erasmus semester at Charles University in Prague, Czech Republic

ADDITIONAL TRAINING

Doctoral Training Programme, Ghent University

2012-2015

- Specialist courses:
 - Behavioural Finance and Equity Investment Strategies, London Financial Studies
 - Using Bayesian methods to model cognition, LOT Summer School Utrecht
 - Reference and Coherence in Discourse, LOT Summer School Utrecht
 - Statistics: Introductory Statistics - Basics of Statistical Inference, Ghent University IPVW
- Transferable skills courses:
 - Statistics: Introduction to R, Ghent University IPVW
 - Advanced Academic English: Conference Skills, Ghent University doctoral schools
 - Project management, Ghent University doctoral schools

Machine Learning, Coursera

2014

Online course taught by Andrew Ng of Stanford University

TermITes and CATs: terminology and translation technology, TermCoord - European Parliament

2011

Products and services for Flemish companies, Flemish Innovation Network and IWT

2011

Bachelor in Applied Informatics, Ghent University College

2010-2012

Credit contract for the following courses:

- Object-oriented programming (Java), first-year course
- Object-oriented design, first-year course
- Databases and data modelling, first-year course
- Front-end programming (Java), second-year course

TEACHING AND SUPERVISION EXPERIENCE

- Introduction to Language Technology (hands-on session about terminology extraction), Postgraduate Dutch and Translation, Faculty of Translation, Interpreting and Communication - Ghent University, fall semester 2014
- Introduction to Language Technology (hands-on sessions about terminology extraction), 3rd year Applied Linguistics, Faculty of Translation, Interpreting and Communication - Ghent University, spring semester 2014
- Thesis (co-)supervision:
 - 2014-2015: Emilie Van Kelf - Master's thesis on sentiment analysis in tweets about companies (supervised by Lieve Macken), Laura Werbrouck - Bachelor's thesis on event detection in financial news reports
 - 2013-2014: Mieke Van Malderen - Bachelor's thesis on sentiment analysis in product reviews
- Recruitment and supervision of student workers for the manual annotation of preprocessing information, domain-specific terminology and sentiment

SCIENTIFIC OUTPUT

1. Organisation of workshops/conferences

- Atila 2014 workshop, Ghent, Belgium, November 20th-21st 2014 (together with Orphée De Clercq and Els Lefever)

2. Publications

In press

- C1: De Clercq, O., Van de Kauter, M., Lefever, E., & Hoste, V. (in press). LT3: Applying Hybrid Terminology Extraction to Aspect-Based Sentiment Analysis. *Proceedings of the 9th International Workshop on Semantic Evaluation (SemEval 2015)*. Denver, Colorado, USA.

2015

- A1: Van de Kauter, M., Breesch, D., & Hoste, V. (2015). Fine-grained analysis of explicit and implicit sentiment in financial news articles. *Expert Systems with Applications*, 42(11), 4999-5010.
- A1: Van de Kauter, M., Desmet, B., & Hoste, V. (2015). The good, the bad and the implicit: a comprehensive approach to annotating explicit and implicit sentiment. *Language Resources and Evaluation*. doi:10.1007/s10579-015-9297-4.

2014

- A1: Lefever, E., Van de Kauter, M., & Hoste, V. (2014). HypoTerm: Detection of hypernym relations between domain-specific terms in Dutch and English. *Terminology*, 20(2): Special issue: Lexical-semantic approaches to terminology, 250-278.
- B2: Desmet, B., De Clercq, O., Van de Kauter, M., Van Hee, C., & Hoste, V. (2014). Taaltechnologie 2.0: sentimentanalyse en normalisatie. *Beschouwingen uit een talenhuis*, 157-161. Academia Press, Ghent, Belgium.
- B2: Vanopstal, K., Macken, L., Lefever, E., Van de Kauter, M., Buysschaert, J., & Hoste, V. (2014). Terminologie: op het snijvlak van ambacht en technologie. *Beschouwingen uit een talenhuis*, 179-189. Academia Press, Ghent, Belgium.
- C1: Lefever, E., Van de Kauter, M., & Hoste, V. (2014). Evaluation of Automatic Hypernym Extraction from Technical Corpora in English and Dutch. *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC 2014)*, 490-497. Reykjavik, Iceland.
- C1: Van Hee, C., Van de Kauter, M., De Clercq, O., Lefever, E., & Hoste, V. (2014). LT3: Sentiment Classification in User-Generated Content Using a Rich Feature Set. *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)*, 406-410. Dublin, Ireland.
- V: Van de Kauter, M., Desmet, B., & Hoste, V. (2014). Guidelines for the Fine-Grained Analysis of Polar Expressions, version 2.0. LT3 Technical Report - LT3 14.02.

2013

- A2: Van de Kauter, M., Coorman, G., Lefever, E., Desmet, B., Macken, L., & Hoste, V. (2013). LeTs Preprocess: The multilingual LT3 linguistic preprocessing toolkit. *Computational Linguistics in the Netherlands Journal*, 3, 103-120.

3. Lectures and workshop/conference contributions

- accepted paper, presentation: “LT3: Applying Hybrid Terminology Extraction to Aspect-Based Sentiment Analysis” (with Orphée De Clercq, Els Lefever and Véronique Hoste), SemEval-2015 at NAACL HLT 2015, Denver, Colorado, USA, June 4th-5th 2015.
- presentation: “SentiFM: Sentiment mining for Financial Markets” (with Véronique Hoste), Data and Web Science Group colloquium, Mannheim, Germany, February 20th 2015.
- accepted abstract, presentation: “Applying Terminology Extraction to Aspect-Based Sentiment Analysis” (with Orphée De Clercq, Els Lefever and Véronique Hoste), CLIN25, Antwerp, Belgium, February 6th 2015.
- presentation: “LeTs Preprocess: the Multilingual LT3 Linguistic Preprocessing Toolkit” (with Lieve Macken), GLIMS, Ghent, Belgium, October 21st 2014.
- accepted paper, poster presentation: “LT3: Sentiment Classification in User-Generated Content Using a Rich Feature Set” (with Cynthia Van Hee, Orphée De Clercq, Els Lefever and Véronique Hoste), SemEval-2014 at COLING 2014, Dublin, Ireland, August 24th 2014.
- accepted paper, presentation: “Evaluation of Automatic Hypernym Extraction from Technical Corpora in English and Dutch” (with Els Lefever and Véronique Hoste), LREC 2014, Reykjavik, Iceland, May 28th 2014.
- accepted abstract, poster presentation: “Evaluation of Automatic Hypernym Extraction from Technical Corpora in English and Dutch” (with Els Lefever and Véronique Hoste), CLIN24, Leiden, the Netherlands, January 17 2014.
- accepted abstract, poster presentation: “Fine-grained Analysis of Explicit and Implicit Sentiment in Financial News” (with Diane Breesch and Véronique Hoste), CLIN24, Leiden, the Netherlands, January 17 2014.
- presentation: “Fine-grained Analysis of Explicit and Implicit Sentiment in Financial News” (with Véronique Hoste), ATILA 2013, Corsendonck, Belgium, September 27th 2013.
- accepted abstract, poster presentation: “The Good, the Bad and the Implicit: Annotating Polarity” (with Bart Desmet and Véronique Hoste), CLIN23, Enschede, the Netherlands, January 18 2013.
- presentation: “The Good, the Bad and the Implicit: Annotating Polarity” (with Bart Desmet and Véronique Hoste), LTCI 2012, Ghent, Belgium, December 14th 2012.
- presentation: “The Good, the Bad and the Implicit: Annotating Polarity” (with Bart Desmet and Véronique Hoste), ATILA 2012, Nijmegen, the Netherlands, November 23rd 2012.
- accepted abstract, poster presentation: “LeTs Preprocess: the Multilingual LT3 Linguistic Preprocessing Toolkit” (with Geert Coorman, Els Lefever, Bart Desmet, Lieve Macken and Véronique Hoste), CLIN22, Tilburg, the Netherlands, January 20th 2012.
- presentation: “Automatic terminology extraction for machine translation” (with Lieve Macken, Els Lefever and Véronique Hoste), TiNT day, Ghent, Belgium, December 9th 2011.
- poster presentation: “Terminology Extraction for Semantic Interoperability and Standardization” (with Lieve Macken, Els Lefever and Véronique Hoste), PhD day Ghent University College, Ghent, Belgium, March 10th 2011.