ARTUUR LEEUWENBERG, Ph.D.

Email: a.m.leeuwenberg-15@umcutrecht.nl

RESEARCH TOPICS

Machine learning, (clinical) natural language processing, medical prediction modeling, temporal information extraction

ACADEMIC POSITIONS

Since 2019 **Assistant Professor – University Medical Center Utrecht, Utrecht University**

Methodological research for clinical prediction modeling and clinical natural language processing at the Julius Center for Health Sciences and Primary Care.

2015-2019 **PhD Researcher – Katholieke Universiteit Leuven**

My PhD research was on machine learning for automatic temporal language understanding, at the Language Intelligence and Informatrion Retrieval Lab, at the Department of Computer Science.

2014-2015 **Research Assistant – Universität des Saarlandes**

Conducted (published) research on automatic extraction of synonyms for better machine translation evaluation together with Mihaela Vela, Jon Dehdari, and Josef van Genabith, at the Dept. of Language Science and Technology.

Summer 2014 **Research Intern – Inria**[1]

Conducted (published) research on automatic extraction of drug-drug interactions from biomedical articles, together with Aleksey Buzmakov, Yannick Toussaint and Amedeo Napoli, at Inria Nancy-Grand-Est.

EDUCATION

2015-2019 **Ph.D. Computer Science**

Katholieke Universiteit Leuven, Belgium
Focus on Machine Learning for Event Timeline Extraction from Clinical Reports (URL).
Advisor: Prof. dr. Marie-Francine Moens

2013-2015 **M.Sc. Language Science and Technology**

Université de Lorraine, France (2014-2015) / Universität des Saarlandes, Germany (2015-2016)
With 2-year Erasmus Mundus Scholarship / Grade: 1.8 (∼ 80%)

2010-2013 **B.Sc. Cognitive Artificial Intelligence**

Universiteit Utrecht, The Netherlands

SELECTED PUBLICATIONS (PEER REVIEWED)

A complete list can be found at `http://bit.ly/PubsTuur`.

[1] Anne de Hond*, **Artuur Leeuwenberg***, Lotty Hooft, Ilse M. J. Kant, Steven W. J. Nijman, Hendrikus J. A. van Os, Jiska J. Aardoom, Thomas P. A. Debray, Ewoud Schuit, Maarten van Smeden, Johannes B. Reitsma, Ewout W. Steyerberg, Niels H. Chavannes, and Karel G. M. Moons. Guidelines and quality criteria for artificial intelligence-based prediction models in healthcare: a scoping review. *npj Digital Medicine*, 5(1):2, 2022. (*shared first authors).

[2] **Artuur Leeuwenberg**, Maarten van Smeden, Johannes A. Langendijk, Arjen van der Schaaf, Murielle E. Mauer, Karel G. M. Moons, Johannes B. Reitsma, and Ewoud Schuit. Performance of binary prediction models in high-correlation low-dimensional settings: a comparison of methods. *Diagnostic and Prognostic Research*, 6(1):1, 2022.

[3] **Artuur Leeuwenberg** and Marie-Francine Moens. Towards Extracting Absolute Event Timelines from English Clinical Reports. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 2020.

[4] **Artuur Leeuwenberg** and Marie-Francine Moens. A Survey on Temporal Reasoning for Temporal Information Extraction from Text. *The Journal of Artificial Intelligence Research (JAIR)*, 2019. (invited for presentation at IJCAI 2020, core A*).

[5] **Artuur Leeuwenberg** and Marie-Francine Moens. Temporal Information Extraction by Predicting Relative Time-lines. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Brussels, Belgium, 2018. ACL. (core A, oral presentation).

[6] **Artuur Leeuwenberg**, Mihaela Vela, Jon Dehdari, and Josef van Genabith. A Minimally Supervised Approach for Synonym Extraction with Word Embeddings. *The Prague Bulletin of Mathematical Linguistics*, 105:1–30, 2016.

1 French National Institute for Research in Computer Science and Automation

GRANTS & SCHOLARSHIPS

National Science Agenda: Subproject lead on "Responsible use of free text in medical prediction research." (NWO, €102k, 2022)

Applied Data Science Grant: "Text mining as imputation model for clinical prediction research: does it have to be perfect to be useful?" (Utrecht University, €5k, 2021)

Erasmus Mundus Scholarship (€ 23k, European Commission, 2013-2015)

TEACHING EXPERIENCE

Master Theses (daily supervisor or first examiner):

Zwierd Grotenhuis: Text mining clinical research outcomes: how accurate should it be?, 2022, UU, MSc Artificial Intelligence

Isa Spiero: Semi-supervised learning for prediction of radiotherapeutic complications, 2022, UU, Applied Data Science

Christina Heilmaier: Predicting laryngeal edema in head- and neck cancer patients, 2021, UMC Utrecht, MSc Epidemiology

Liesbeth Allein: Correction of Dutch pronouns with neural networks, 2018-2019, KU Leuven, MSc Artificial Intelligence

Lieven Kop: Semi-supervised deep learning for NLP, 2018-2019, KU Leuven, MSc Computer Science

Nathan Van Laere: ULMFit for Dutch Datasets, 2018-2019, KU Leuven, MSc Computer Science

Mariya Hendriksen: Chatbot Dialogue Breakdown Detection, 2018, KU Leuven, MSc Artificial Intelligence

Enyan Dai: Error Propagation in Relation Extraction, 2017-2018, KU Leuven, MSc Artificial Intelligence

Ellen Vissers: Sentiment Analysis in Tweets, 2017-2018, KU Leuven, MSc Computer Science

Lena Martens: Clinical Temporal Relation Extraction, 2016-2017, KU Leuven, MSc Computer Science

Vignesh Baskaran: Movie Recommendation, 2015-2016, KU Leuven, MSc Artificial Intelligence

Seppe Dijkmans: Native Language Identification, 2015-2016, KU Leuven, MSc Applied Informatics

Lecturer:

1x [MSc] Evidence Based Medicine, 2021, UMC Utrecht

2x [BSc] Clinical Scientific Research, 2021,2022, UMC Utrecht

Teaching Assistantships:

3x [BSc] Clinical Scientific Research, 2019-2021, UMC Utrecht

2x [MSc] Natural Language Processing, 2017, 2018, KU Leuven

2x [BSc] Introduction to Programming, 2015, 2016, KU Leuven

1x [BSc] Natural Language Processing: Parsing-as-Deduction, 2013, Utrecht University

1x [BSc] Introduction to Logics, 2012, Utrecht University

Assessor for 8 x MSc Artificial Intelligence theses, and 7 x MSc Computer Science theses at KU Leuven (2016-2019)

PROJECTS CONTRIBUTED TO

HTx (H2020-EU.3.1.6. 825162) "Next Generation Health Technology Assessment"

ACCUMULATE (IWT-SBO-Nr. 150056) "Acquiring Crucial Medical information Using Language Technology"

MARS (KU Leuven) "Machine Reading of Patient Records"

CALCULUS (ERC Advanced Grant 788506) "Commonsense and Anticipation enriched Learning of Continuous representations sUpporting Language UnderStanding"

DAME (KU Leuven) "A platform for Deep learning Approaches for Multilingual tExt processing"

ASSESSMENT RESPONSIBILITIES

Grant reviewer

Dutch Arthritis Society (Reuma Nederland), 2020.

PhD examination committee

Elias Moons, "Representation Learning for Automated Document Classification", KU Leuven, 2021.

Programme committee memberships

Diagnostic & Prognostic Research, The European Heart Journal, Frontiers in Epidemiology, BMC Medical Informatics and Decision Making, ACL Rolling Review (ARR), Conference on Empirical Methods in Natural Language Processing (EMNLP-IJCNLP), Annual Meeting of the Association for Computational Linguistics (ACL), North American Chapter of the Association for Computational Linguistics Conference (NAACL), International Conference on Computational Linguistics (COLING) [**outstanding reviewer mention '18**], International Workshop on Semantic Evaluation (SemEval).

EXTRACURRICULAR ACTIVITIES

Completed the "Basic course on Regulations and Organisation for Clinical Investigators" (BROK®, 2020)

Volunteer at the Conference on Empirical Methods in Natural Language Processing (EMNLP), Brussels (2018)

Board Member, Utrecht's Study Association for Cognitive Artificial Intelligence, USCKI Incognito (2012-2013)

Board Member, Dutch Study Association for Artificial Intelligence, NSVKI (2012-2013)

Student Member of the Advisory Committee for the B.Sc. Artificial Intelligence at Utrecht University (2012-2013)

TECHNICAL SKILLS

Python, R, Torch, Keras, Java, Haskell, Prolog, git, HTML, CSS, MySQL, Javascript, PHP, LaTeX, Linux

LANGUAGES

Dutch (native), English (fluent), German (limited working proficiency), French (elementary proficiency)

Last updated: October 6, 2022