Myrthe Reuver

Computational Linguist, Data Scientist, PhD Candidate

I am a PhD (started: Sept 2020) at the Free University of Amsterdam (VU), advisors: prof. dr. Antske Fokkens (VU) and dr. Suzan Verberne (Leiden University). My research is on NLP to support democracy in news recommendation. My interest is in **societally and scientifically responsible NLP in complex domains**, and unpacking complex concepts (such as democratic values and diversity) computationally.

CV last edited: January 12, 2022

🌐 https://myrthereuver.github.io/   ✉ myrthe.reuver[at]vu.nl

🐙 myrthereuver   Twitter: @myrthereuver

Research Experience

PhD candidate: Diverse News Recommendation   **Amsterdam**
Free University of Amsterdam (VU)   *Sept 2020 – present*
- PhD project on diversity in news recommender systems. Work on language technology (language models) to automatically detect diversity in news texts, and enable a "democracy supporting" news recommender. Supervised by prof. dr. Antske Fokkens of the CLTL research group at the Free University of Amsterdam and dr. Suzan Verberne at Leiden University

Graduation Projects Clinical Natural Language Processing   **Deventer**
Topicus Healthcare   *Dec 2019 – Sept 2020*
- MSc thesis on classification of complaint category in emergency GP triage (huisartsenpost)
- ResMA thesis on detection of smoking status from EMRs with weak supervision & transfer learning

ResMA Research Internship: *Classifying Urban Legends*   **Amsterdam**
Meertens Institute   *Feb 2019 – July 2019*
- Results: a popular science article in *Neerlandistiek*, a simple interface demo for the model, a conference presentation at the DHBenelux conference in Liège (September 2019), and a poster presentation at CLIN30 (Computational Linguistics in the Netherlands) in January 2020.

Education

Master of Science (MSc) Cognitive Science & Artificial Intelligence   **Tilburg**
Tilburg University   *2018 – 2020*
- GPA: 7.9 / 10 (with distinction)
- Thesis: "What Kind of Emergency? Complaint Labels for Dutch Short Clinical Texts" (Grade: 8/10)
- Courses (selection): Machine Learning, Deep Learning, Advanced Data Processing with Python

Research MA Linguistics, track: Language & Speech Technology   **Nijmegen**
Radboud University   *2017 – 2020*
- GPA: 8.3 / 10 (cum laude)
- Thesis: "Finding the Smoke Signal: Smoking Status Extraction & Classification" (Grade: 8 / 10)
- Internship: Classifying Urban Legends at the Meertens Institute (Grade: 9 / 10)
- **Certificate**: Interdisciplinary Honours Program for Master's Students "Thinking Machines in Healthcare"
- Courses (selection): Text & Multimedia Mining, Language & Speech Technology, Advanced Statistics in R

BA English Language & Culture, specialization: Linguistics   **Nijmegen**
Radboud University   *2014 – 2017*
- GPA: 8.6 / 10 (cum laude)
- **certificate** BA Honours Program: 2-year Text Mining Project + Summerschool Aarhus University, Denmark.

Bachelor of Journalism, first-year diploma   **Ede**
Christelijke Hogeschool Ede (CHE)   *2012 – 2013*
Radio Presenter at "Radio Almelo" (AAFM)

HAVO (higher general secondary education), Culture & Society profile Almelo
OPDC 't Nijrees Almelo *2007 – 2012*

Teaching

Teacher "Programming in Python for Text" **Vrije Universiteit Amsterdam**
prof. dr. Antske Fokkens *Sept 2021 – Nov 2021*
- A first programming course in Python for both BA and MA students in Text Mining, with a focus on basic analysis with (data with) text. Tasks included: updating material from previous years, lecturing on more advanced topics: errors, nested data structures, and XML.

Teaching Assistant BA course "Information Science" **Radboud University, Nijmegen**
prof. dr. Martha Larson *April 2018 – June 2020*
- A first-year bachelor's course about the basics of (digital) information science, data, privacy, algorithms and machine translation. Tasks: leading practical sessions, lecturing segments of lectures + designing, preparing and grading exercises, assignments, and exams.

Events

Winning team Hackathon "Diverse Recommendation" **Virtual**
NPO and Media Perspectives *October 2020*
- A 24-hour hackathon on designing a topic, viewpoint, and content diverse news & media recommender for Dutch public broadcaster NPO. 5 teams of Data Scientists and researchers competing.
- Won the hackathon with an interdisciplinary team of philosophers, communication scientists, and computationals (Marijn Sax, Felicia Loecherbacher, Sanne Vrijenhoek and Nicolas Mattis)

Information Technology (ICT) with Industry research week **Leiden**
Lorentz Centre and NWO *January 2019*
- Full week of research teams working with an industry partner on their technology-related research challenge. Participated in developing a prototype of an automatic keyword classifier for the National Library (Koninklijke Bibliotheek, KB).
- Results: a demo prototype, a report, and an NWO Science vlog about my experiences.

Leadership and Academic Service

On-site student volunteer **Amsterdam & online (hybrid)**
RecSys2021 *27 Sept 2021 – 1 Oct 2021*
- Tasks: guiding attendees, monitoring the live stream, welcoming + registration of attendees

Founder & Editor-in-Chief **Nijmegen**
RU:ts Student Linguistics Journal *Sept 2018 – Sept 2020*
- Together with fellow master's students Iris Faber & Sander Nederveen, started a journal where students are authors, editors, and peer reviewers.

Skills

- Machine Learning, Natural Language Processing
- (Academic) Writing
- Public Speaking and Presenting

Languages

- **Python** including sklearn & pandas, & **R** – proficient
- **html & css**, familiar

- **Dutch**, mothertongue, **(Academic) English**, C2 (near-Native)
- **German**, B2 (decent), **Latin**, B1 (basic)

Communicating Research Outcomes

Peer Reviewed Abstracts / Posters (No Paper)

- **Myrthe Reuver**, Suzan Verberne, Roser Morante and Antske Fokkens. 9 July 2021. "Viewpoints in the news: claim detection for diverse news recommendation." **Poster** at Computational Linguistics in the Netherlandsa (CLIN) 31 (Online), 9 July 2021. Link.
- **Myrthe Reuver**, Iris Hendrickx, Jeroen Kuijpers. "Smoking Status Detection with a Pre-Trained Transformer and Weak Supervision in Dutch Electronic Patient Files." Computational Linguistics in the Netherlandsa (CLIN) 31 (Online), 9 July 2021. Link.
- **Myrthe Reuver**. *"How Similar are Poodles in the Microwave? Classification of Urban Legend Types"*. Poster at Computational Linguistics in the Netherlands (CLIN) 30, Utrecht, January 2020. Link.
- **Myrthe Reuver**. *"Making the Computer Understand Urban Legend Types (And Making Humans Understand the Computer)"*. Moving Humanities Graduate Conference, Nijmegen: 24 October 2019. Link.
- **Myrthe Reuver**. *"Amputation or Accident? Classifying Dutch Urban Legends into Story Types with a Hierarchical Classifier"*. Digital Humanities BeNeLuX, Liège: 13 Sept 2019. Link.
- **Myrthe Reuver**. *"Whose Gender is It? Possessive Pronoun Errors by Portuguese L2 speakers of English."* 13th TWIST Student Conference in Linguistics, Leiden. 13 April 2019. Link.
- **Myrthe Reuver**. *"First-Person Visibility: Ælfric's Pronouns in his Prefaces."*, StuTs Student Linguistics Conference, Nijmegen. 24 November 2017. Link.

Invited Talks

- **Myrthe Reuver**. *"Hackathons, Shared Tasks, and Papers: Collaboration and Interdisciplinarity with Data on Complex Problems"*. Free University of Amsterdam, University Library Data Conversations/ Online. 7 April 2021. Video

Popular Science

- **Myrthe Reuver**. (2021) "5 Things I Learned During the First Year of my PhD." Blog post. Link
- **Myrthe Reuver**, Nicolas Mattis. (2020) "Hackers of Science: Lessons from interdisciplinary science in a Hackathon" Socializing Science, blog post. Link
- **Myrthe Reuver**. (2019). "Hoe een computer broodjeaapverhalen leert categoriseren." Vertelcultuur Online. Link
- **Myrthe Reuver**. Vlog for research week "ICT with Industry". Link

Reports (Not Peer Reviewed)

- **Myrthe Reuver**, Nicolas Mattis. (2021). "Implementing Evaluation Metrics Based on Theories of Democracy in News Comment Recommendation (Hackathon Report)" In: Proceedings of the EACL Hackashop on News Media Content Analysis and Automated Report Generation (co-located at EACL 2021, online). Association of Computational Linguistics, p. 134–139. Link
- Anthony Burton, Elena Aversa, Alessandra Facchin, Ivana Emily Škoro, Henri Mütschele, **Myrthe Reuver**, Shenglang Qing. (2019). "Streams of the Deep Web: Rebel Media, YouTube, and the algorithmic shaping of media ecosystems." Digital Methods Wiki, UvA Summerschool. Link
- Brandsen, A., Kleppe, M., Veldhoen, S., Zijdeman, R., Huurman, H., Vos, H. De, Goes, K., Huang, L., Kim, A., Mesbah, S., **Reuver, M.**, Wang, S., Hendrickx, I. (2019). "Brinkeys.." KB Lab:The Hague, the Netherlands. Link

Publications

Myrthe Reuver, Antske Fokkens, and Suzan Verberne. No NLP task should be an island: Multi-disciplinarity for diversity in news recommender systems. In *Proceedings of the EACL Hackashop on News Media Content Analysis and Automated Report Generation*, pages 45–55. Association for Computational Linguistics, 2021.

Myrthe Reuver, Nicolas Mattis, Marijn Sax, Suzan Verberne, Nava Tintarev, Natali Helberger, Judith Moeller, Sanne Vrijenhoek, Antske Fokkens, and Wouter van Atteveldt. Are we human, or are we users? The role of natural language processing in human-centric news recommenders that nudge users to diverse content. In *Proceedings of the 1st Workshop on NLP for Positive Impact, co-located at ACL-IJCNLP 2021*, pages 47–59. Association for Computational Linguistics, 2021.

Myrthe Reuver, Suzan Verberne, Morante Roser, and Antske Fokkens. Is stance detection topic-independent and cross-topic generalizable?– A reproduction study. In *8th Workshop on Argument Mining, co-located with EMNLP 2021*, page 46–56, Punta Cana, Dominican Republic, 2021. Association for Computational Linguistics.