David Vos

+31 6 40127944 | d.j.a.vos@uva.nl | davidvos.dev | linkedin.com/in/davidjavos | github.com/davidvos

Research interests

In my research, I study machine learning methods for information retrieval, recommender systems, and data management. My interests include text representation learning, in-context learning for ranking, and generative modeling for personalized recommendations. Through my collaboration with DPG Media, I investigate retrieval challenges in the news domain, where rapidly evolving content and short item lifetimes pose unique modeling challenges. In the long term, I aim to develop retrieval systems that account for their societal implications, supporting information ecosystems that reflect diverse user needs.

Publications

David Vos, Till Döhmen, Sebastian Schelter.
"Towards Parameter-Efficient Automation of Data Wrangling Tasks with Prefix-Tuning."
NeurIPS 2022 First Table Representation Learning Workshop. Best paper runner-up award.

Yuyue Zhao, Jin Huang, **David Vos**, Maarten de Rijke.
"Revisiting Language Models in Neural News Recommender Systems."
European Conference on Information Retrieval (ECIR), 2025.

David Vos, Jin Huang, Maarten de Rijke.
"The Role of LLMs in Democratic News Recommender Systems."
Report on the Search Futures Workshop at ECIR 2024, *ACM SIGIR Forum*, Volume 58, Issue 1.

David Vos, Andrew Yates, Maarten de Rijke.
"Reliable and Controllable Learned Sparse News Recommendations for Short-History Users." Under review.

David Vos, Jin Huang, Maarten de Rijke.
"A Study of Normative Diversity Metrics in News Recommendations." Under review.

Education

PhD Student in Information Retrieval Sep. 2023 - Present
Information Retrieval Lab - University of Amsterdam
- Supervised by prof. dr. Maarten de Rijke & dr. Harrie Oosterhuis, in collaboration with DPG Media.
- Research topic: Responsible retrieval and ranking for personalized recommendations.

MSc. Artificial Intelligence (cum laude) Sep. 2019 - Sep. 2022
University of Amsterdam, GPA: 8.2/10
- Thesis: *Language control prefixes: Conditional prefix-tuning for efficient multilingual data-to-text generation in low-resource languages.*
- Supervisors: prof. dr. Sebastian Schelter & dr. Marlies van der Wees.

Erasmus Exchange Sep. 2018 - Dec. 2018
University of Edinburgh
- Elective courses in computer science and anthropology.

BSc. Artificial Intelligence (cum laude) Sep. 2016 - Jun. 2019
University of Amsterdam, GPA: 8.3/10
- Thesis: *Automating a weekly newsletter for news article recommendations at De Volkskrant.*
- Supervisors: dr. Harrie Oosterhuis, dr. Anne Schuth & Lucas de Haas.

Honours Programme, BSc. Artificial Intelligence Sep. 2016 - Jun. 2019
University of Amsterdam
- Art-inspired fashion: Siamese networks and GANs for clothing design generation.
- Automated ESG benchmarking: NLP-based sustainability assessment of large multinationals.

Training programmes & qualifications

- Oxford Machine Learning School (OxML), 2025.
- European Summer School on Information Retrieval (ESSIR), 2025.
- Seeds for the Future, 2019. Summer internship at Huawei in Shenzhen and Beijing on network technologies.
- Cambridge English Certificate of Proficiency (C2 level).
- The Next Web T500, 2018-2022. Awarded a spot on TNW's T500 promising tech-talent under 25 shortlist.

Academic activities

Talks & lectures

University of Amsterdam and external venues

- University lectures on text representation learning, sequential recommender systems, and transformers.
- Search Futures Workshop (ECIR 2024): *The Role of LLMs in Democratic News Recommender Systems.*
- NeurIPS 2022 Table Representation Learning Workshop: *Prefix-tuning for data wrangling.*
- DSDSD 2023: *Towards parameter-efficient automation of data wrangling tasks with prefix-tuning.*
- Talk for the ING Data Science team on large language models for data management.

Teaching & supervision

University of Amsterdam

- Supervised MSc. thesis: *Efficiently optimizing tree-structured identifiers for generative recommendations.*
- Supervised MSc. thesis: *Overcoming feedback bias in IKEA recommender systems.*
- Supervised BSc. thesis: *Uncovering gender bias after finetuning large language models.*
- Taught tutorials and supervised students in MSc. courses on Recommender Systems and Information Retrieval.

Service & organisation

IRLab Amsterdam

- Research lab representative at the PhD council of the university's Institute of Informatics.
- Co-organised the European Summer School for Information Retrieval (ESSIR) 2025.
- Chair of weekly research group meetings and reading groups on neural information retrieval, recommender systems and generative retrieval.

Experience

Visiting researcher, Intelligent Data Engineering (INDE) Lab Jun. 2022 - Dec. 2022

University of Amsterdam

- Worked on parameter-efficient optimisation of large language models for data management tasks.
- Supervised by prof. dr. Sebastian Schelter & dr. Marlies van der Wees.

Software engineer Dec. 2022 - Apr. 2023

Vaccinatieregister

- Developed integrations aligned with medical standards at a European provider of travel vaccination software.

Machine learning engineer intern Feb. 2019 - Jun. 2019 & Nov. 2021 - Jun. 2022

DPG Media

- Intern in the Recommendation & Search team at a large European media company.

Software engineer Aug. 2020 - Sep. 2021

Finly

- Developed software for the customer support pipeline of large insurance and telecom companies.

Data engineer Apr. 2019 - Jun. 2020

A-INSIGHTS

- Developed and maintained data pipelines for a market insights scale-up.

Teaching assistant Aug. 2017 - Nov. 2021

University of Amsterdam

- Assisted in bachelor courses in mathematics, programming, and logic. Led tutorials and graded assignments.

Co-founder & developer Feb. 2017 - Jul. 2019

JUNO Amsterdam

- Co-founded a digital consulting agency delivering web and data solutions for small businesses.

Software engineer Apr. 2017 - Apr. 2019

Syncable Music

- Developed a music licensing platform for online video as part of a startup team.