Madelon Hulsebos

Amsterdam, Netherlands | madelon.hulsebos@gmail.com | madelonhulsebos.github.io

I am interested in **Intelligent Data Systems**. So far, my research has been on **learned table representations** and their **applications** like data preparation, search, and analysis.

EDUCATION

University of Amsterdam – Amsterdam, Netherlands
PhD Computer Science   *Sep 2020 – Sep 2023 (expected)*

Delft University of Technology – Delft, Netherlands
MSc Computer Science   *Sep 2016 – July 2018*
BSc Technology, Policy and Management   *Sep 2011 – July 2015*

ACADEMIC EXPERIENCE

University of Amsterdam – Amsterdam, Netherlands
PhD Researcher, INDE Lab   *Sep 2020 – Sep 2023 (expected)*
- Advised by Prof. Paul Groth.
- Proposed an agenda around methods and datasets for inferring table semantics and applications thereof.

Massachusetts Institute of Technology – Cambridge, United States
Visiting Collaborator, MIT Media Lab   *Aug 2018 – March 2019*
- Led a research project (Sherlock) on semantic table understanding.
- Contributed to a data visualisation training dataset and benchmarking project.

INDUSTRY EXPERIENCE

Sigma Computing – San Francisco, United States
PhD Student Researcher, previously Research Intern (Summer 2021)   *2021 – present*
- Developing systems for deploying table understanding models, and applications like data preparation and search.
- Contributing to an interactive data science tool (Decision Studio), mainly backend design and analyses.
- Publishing and presenting the outcomes of my work at scientific conferences, e.g. CIDR '22.

KPN, HEINEKEN – Rotterdam/Amsterdam, Netherlands
Data Scientist   *2019 – 2021*
- Developed and integrated an ML system for financial forecasting that produced more accurate forecasts.
- Built two ML tools for evaluating marketing campaigns based on causal inference and Bayesian modelling.
- Gave talks e.g. on data validation and transfer learning, initiated a reading group, mentored junior data scientists.

TEACHING EXPERIENCE

Delft Institute of Technology – Delft, Netherlands
Graduate Teaching Assistant, Pattern Recognition & Web Information Systems groups   *2017 – 2018*
- TA for the MSc courses: Pattern Recognition (IN4085), Web Science & Engineering (IN4252).
- Supported 250+ graduate students throughout projects and lab sessions.
- Evaluated student assignments, projects and presentations.

Aalto University – Helsinki, Finland
Research and Teaching Assistant, Machine Learning for Big Data   *2017*
- Developed assignments for the BSc course "Machine Learning: Basic Principles" (CSE3210) with 500+ students.
- Conducted experiments for semi-supervised learning over networks and presented the work at ICASSP.

ACADEMIC SERVICE

Assistant Editor, Journal of Systems Research (jsys)	*2022 - present*
Co-organizer, Table Representation Learning workshop @ NeurIPS	*2022*
Co-organizer, SemTab challenge @ ISWC	*2021, 2022*
Program Committee, PhD Workshop @ VLDB	*2023*
Program Committee, EDBT (Industry track)	*2022, 2023*
Program Committee, WWW (Industry track)	*2022, 2023*
Reviewer, NeurIPS (Datasets & Benchmarks track)	*2021, 2022*
Program Committee, AIDB @ VLDB	2022

SELECTED PUBLICATIONS

GitSchemas: A Dataset for Automating Relational Data Preparation Tasks, DBML@ICDE, 2022.
Döhmen, T., Hulsebos, M., Beecks, C., Schelter, S.

Making Table Understanding Work in Practice [abstract], CIDR, 2022.
Hulsebos, M., Gathani, S., Gale, J., Dillig, I., Groth, P., Demiralp, C.

Augmenting Decision Making via Interactive What-If Analysis, CIDR, 2022.
Gathani, S., Hulsebos, M., Gale, J., Haas, P. J., Demiralp, C.

GitTables: A Large-Scale Corpus of Relational Tables, SIGMOD (to appear), 2023.
Hulsebos, M., Demiralp, C., Groth, P.

Sato: Contextual semantic type detection in tables, PVLDB, 2020.
Zhang, D., Suhara, Y., Li, J., Hulsebos, M., Demiralp, C., Tan, W.

Sherlock: A deep learning approach to semantic data type detection, SIGKDD, 2019.
Hulsebos, M., Hu, K., Bakker, M., Zgraggen, E., Satyanarayan, A., Kraska, T., Demiralp, C., Hidalgo, C.

VizNet: Towards a large-scale visualisation learning and benchmarking repository, CHI, 2019.
Hu, K., Gaikwad, N., Hulsebos, M., Bakker, M., Zgraggen, E., Hidalgo, C., Kraska, T., Li, G., Satyanarayan, A., Demiralp, C.

INVITED TALKS

Towards "Large Table Models" for enterprise data management, KomPAKI seminar, TU Darmstadt, 2022	*June*
GitTables: a large corpus of relational tables, Database Architecture group CWI, Amsterdam, NL	*Feb 2022*
Keep The Data Flowing - Data Validation using Great Expectations, PyData, Amsterdam	*Sep 2020*

NOTABLE

VLDB Endowment Travel Award	*2022*
Honourable mention for GitTables in the SemTab challenge	*2021*

BOARD MEMBERSHIP

UniPartners Delft – Delft, Netherlands
Supervisory Board Member   *2017 – present*
- Supervise the financial and strategic positions of the company.
- Implement structured financial control, KPI oriented leadership and a supervision cycle.

Executive Board Member   *2015 – 2016*
- Controlled and optimised the quality of internal and external products and processes.
- Daily management of projects, contributing to a revenue of over €100 K.
- Moderated the CRM system (Salesforce) and trained 10+ boards.