



NEMANIA BOROVITS

Data Scientist

📍 Uilenburg 11, 5211 EV, Den Bosch

☎ +31640604480

✉ nborovits@gmail.com

in [linkedin.com/in/nemania-borovits-8306b812a/](https://www.linkedin.com/in/nemania-borovits-8306b812a/)

🌐 nborovits.com/

Professional skills

Data Analysis:

Programming language: Python.
Data engineering: Apache Spark, Apache Kafka, Elastic Stack.
Data Visualization: Plotly Dash, Matplotlib, Seaborn, Kibana.
Machine learning: Scikit-learn, NumPy, Pandas, Snorkel.
Deep learning: TensorFlow, Keras, PyTorch.
NLP: NLTK, Hugging face Transformers, spaCy, gensim, t-SNE, Snorkel.
Web Crawling: BeautifulSoup.
Privacy Frameworks: Flower, Opacus, SEAL.

Cloud Technologies:

Google Cloud
Heroku
Dataiku
Microsoft Azure

Back-end Development:

Data modeling: Microsoft SQL Server, MongoDB
ORM: nHibernate, Entity Framework.
Business Logic: asp .net MVC, asp .net (C#).
API: asp .net Web API 2 (C#), Flask.
Web Services: SOAP, RESTful (C#, Python).

Front-end development:

Summary

A passionate Data Scientist from Thessaloniki, Greece located in 's-Hertogenbosch, the Netherlands. I have worked for 5 years as a full-stack developer and data engineer in Greece. Then, I decided to specialize my skill set in the domain of Data Science. After graduating with a Master's in Data Science & Entrepreneurship, I started the journey towards the Ph.D. degree. The scientific domain I specialize in is Data Engineering and AI for Privacy. My Ph.D. is co-financed by the Jheronimus Academy of Data Science and the Dutch Telco provider KPN.

Professional Experience

15/10/2020 - Data Scientist

KPN

Member of the Customer Contact Analytics department and the KPN Data Science Lab.

Involved in Natural Language Processing projects analyzing call center text data in the Dutch Language. Collaborating with subject matter experts and business stakeholders to set, assess and ideate over business KPIs. Collaborating with Data and MLOps engineers to optimize the development, testing, deployment, inference and maintenance of the generated models. Following the SCRUM framework for day-to-day project management.

Projects involving unstructured call center text data in the Dutch language:

1. Rule-based and NER-based data anonymization.
2. Dependency parsing-based and NER-based business-relevant ontology creation.
3. Labeling automation through Weak Supervision for text classification.
4. Privacy-preserving deep learning for data anonymization and business-relevant ontology creation.

Projects involving tabular TV viewership data:

1. Development of a framework for validation and evaluation of synthetic data.

Over the past year, I have contributed to KPN's shift towards a decentralized data governance model, namely data mesh, through my direct collaboration with the Data Governance team. My contributions include:

1. Have been leading the initiative on researching, developing, and planning the adoption of data contracts for internal data products. This involved conducting a structured multivocal academic literature review to establish a robust foundation on the components of data contracts. I defined both functional and non-functional requirements, with a particular emphasis on privacy engineering.
2. Collaborating with various domain units within the organization to engineer privacy into their data products, adhering to the principles of Privacy by Design.

15/10/2019 - Data Scientist / Research Assistant

31/03/2020 Tilburg University

Performed academic research in the domain of Infrastructure as Code (IaC).

The practical part of the work included:

1. Data collection.
2. Data verification.
3. Data cleaning and preparation.
4. Data pre-processing.
5. Deep learning model development.
6. Model evaluation and tuning.
7. Model deployment.

Web skills: HTML5, Javascript, jQuery, AJAX, JSON, CSS, Bootstrap.
JS Frameworks: React js, Knockoutjs, VueJs.

Source control tools:



GitHub, BitBucket, Mercurial

The result of my work is an academic paper that was accepted for the Workshop on Machine Learning Techniques for Software Quality Evaluation (MaLTeSQuE). The carried out research applies Deep Learning (CNN) in order to detect name-based bugs in Infrastructure as Code (IaC).

This work is supported by the European Commission grant no. 825480 (SODALITE H2020) and no. 825040 (RADON H2020).

01/10/2016 - Data Engineer

17/08/2018 [Mouzenidis Group](#)

Designed, developed and maintained scalable data pipelines for member-companies of Mouzenidis Group holding.

Main duties included:

- 1) Design, develop and optimize architectures for data acquisition and storage.
- 2) Develop scripts for database maintenance, monitoring, and performance tuning.
- 3) Perform data integrations through Extract Transform Load (ETL) operations.
- 4) Deployment of machine learning models.
- 5) Collaboration with Analytics team in order to meet business goals.

01/10/2014 - Full Stack Developer

30/09/2016 [Mouzenidis Group](#)

Developed web applications for member-companies of Mouzenidis Group holding.

Main duties included:

- 1) Design and development of server-side (back-end) and client-side (front-end) architecture.
- 2) Management and development of well-functioning databases.
- 3) Security and data protection settings implementation.
- 4) Service-Oriented Architecture design and implementation through web services and APIs.
- 5) Web server management and maintenance.
- 6) Ideation of Web solutions in collaboration with marketing and sales department.

10/04/2013 - Junior Web Developer

30/09/2014 [Mouzenidis Travel](#)

Supported development of Mouzenidis Travel's websites.

Main duties included:

- 1) Responsive web page redesigning.
- 2) Javascript and CSS optimization.
- 3) Page loading and SEO optimization.
- 4) Modules development for the Mouzenidis Group CMS. (<https://cms.mouzenidis.com/adminb>).
- 6) Trained personnel from other departments on how to use the Mouzenidis Group CMS.

Academic publications

Published articles: [My profile on Google Scholar](#).

Articles already accepted and scheduled for publication in 2024:

Borovits, N., Kumara, I., Tamburri, D. A., & van den Heuvel, W.-J. (2024). *Privacy Engineering in the Data Mesh: Towards a Decentralized Data Privacy Governance Framework*. In Service-Oriented Computing – ICSOC 2023 Workshops. Springer.

Education

15/10/2020 PhD in Data Engineering and AI for Privacy
[Jheronimus Academy of Data Science](#)

27/08/2018 - MSc Data Science & Entrepreneurship
10/07/2020 [Jheronimus Academy of Data Science](#)

01/10/2006 - BSc in Applied Informatics
18/07/2012 [University of Macedonia](#)

Languages

Mother tongues:

Greek, Serbian.

Other languages:

English: IELTS, Proficiency in English (University of Michigan).

Dutch, Russian: Basic communication skills / working knowledge.

Organizational Competencies

Creating and keeping deadlines

Project management

Teamwork

Goal setting and meeting goals

Creative thinking

Strategic thinking

Other interests

Artificial intelligence ethics enthusiast.

Jogging, playing football and basketball, cooking and travelling.