Dan Saattrup Smart

Senior AI Specialist

Contact

+45 93 52 27 35
dan.smart@alexandra.dk
LinkedIn (saattrupdan)
GitHub (saattrupdan)
Hugging Face (saattrupdan)
Google Scholar

Key Skills

Machine Learning
Natural Language Proc
Large Language Models
Uncertainty Estimation
Python
Linux
Vue.js

Languages

English
Danish

Profile

I am a data scientist with specialised knowledge of machine learning methods, in particular Natural Language Processing (NLP). I have a PhD in Mathematics and have a strong interest within low-resource NLP. With a background in both academia and industry, I am in tune with the demands of each. I am incredibly ambitious and love what I am doing.

Experience

January 2026—Present
Principal AI Specialist • The Alexandra Institute
As my previous Senior role, this is also a mix of research and consulting, primarily within the scope of natural language processing. In this role I am in charge of most projects I'm involved in, assisting the other specialists, as well as having influence on the general direction of the focus points we have in the AI Lab. As my previous Senior role, this is also a mix of research and consulting, primarily within the scope of natural language processing. In this role I am in charge of most projects I'm involved in, assisting the other specialists, as well as having influence on the general direction of the focus points we have in the AI Lab.

June 2022—December 2025
Senior AI Specialist • The Alexandra Institute
This role is split 50/50 between a research role, in which I'm primarily working with NLP research for the Scandinavian languages, and an ML consulting role where I use this research knowledge to help Danish companies implement state-of-the-art machine learning methods to optimise and/or improve their workflows. Notable achievements:

1. Built and maintained the *de facto* standard European LLM benchmark, EuroEval (euroeval.com)
2. Lead work packages in large-scale research projects, as well as being tech lead in many commercial projects
3. Published 5 papers

March 2022—June 2022
Machine Learning Consultant • Blackbird.ai
Contracting work on improving Blackbird.ai's misinformation detection algorithms.

May 2021—June 2022
Research Associate in Machine Learning • University of Bristol
Worked on the CLARITI project, in which I used graph neural networks and multimodal machine learning methods to analyse and predict the spread of misinformation on social media channels.

June 2020—May 2021
Senior Machine Learning Consultant • Danish Business Authority
A part of the machine learning lab, in which I utilised graph algorithms and built machine learning models to detect tax fraud.

February 2020—May 2020
Machine Learning Specialist • LYS Technologies Limited
In charge of the machine learning efforts at LYS. The main task was building the LYS Light Engine, which is using data collected through thousands of LYS's proprietary light sensors to predict optimal individualised lighting at a given time.

February 2017—February 2020
Teaching Assistant • University of Bristol
In charge of weekly exercise sessions for mathematics undergraduate students, covering various mathematics courses as well as LaTeX and Python.

September 2019—January 2020
AI Researcher • Barbal Limited
Developing natural language processing models to enhance Barbal, a word processing web app made for large scale collaboration on technical documents.

February 2019—March 2019
Visiting Researcher • University of Rutgers
Visiting Professor Grigor Sargsyan to work on a joint paper in mathematical logic.

April 2018—May 2018
Assistant Lecturer • University of Bristol
I gave five lectures in the third-year logic unit MATH30100 to ~50 undergraduate students, covering basic computability theory and the incompleteness theorems.

January 2014—July 2016
Teaching Assistant • University of Copenhagen
In charge of weekly exercise sessions with ~30 undergraduate mathematics students.

January 2014—July 2016
Junior Programmer • GE Revision & Rådgivning
Automated the company's accounting process, reducing manual Excel labour from ~5 hours per account down to a minute per account. Developed cash register software which allowed day-to-day updates of cash flow from the company's clients.

Education

July 2016—March 2020
PhD in Mathematics • University of Bristol
PhD thesis: “Virtual Set Theory: Taking the Blue Pill”, advised by Philip Welch. My thesis was prize nominated for "Outstanding Excellence in a Doctoral Dissertation".

September 2014—July 2016
MSc (cand.scient) in Mathematics • University of Copenhagen
Grade: 11.7/12.0 (A)
Master's thesis: "Inner Model Theory - An Introduction", advised by Asger Törnquist.

September 2011—July 2014
BSc in Mathematics • University of Copenhagen
Grade: 10.8/12.0 (B+)
Bachelor's thesis: "Gödel's Constructible Universe", advised by Asger Törnquist.

Volunteering

April 2022—2024
Vice Chair of the Board • Danish Data Science Community
The Danish Data Science Community is an official association (Danish: forening), which has three core purposes: (a) Strengthening the relations between all the data scientists in Denmark, (b) To establish a united voice on behalf of data scientists in Denmark, and (c) To strengthen the open-source culture within Danish data science.

December 2019
Data Study Group Facilitator • The Alan Turing Institute
Led a data science team of 14 PhD students for a week. We worked with WWF to see if news data could be utilised to monitor protected sites around the world, and our resulting model correctly detects 96% of the news articles that present a potential threat to a protected site, with only 18% false positives. Final report published at https://www.turing.ac.uk/research/publications/data-study-group-final-report-wwf .

December 2019
Data Science Lead • Bristol Soup Run Trust
Starting from the Data & Community social hackathon in Bristol, I was in charge of a data science team that developed a machine learning model for a Bristol Soup Run Trust, a local charity that provides food and supplies to homeless people in Bristol who are in need. The model helps them predict how many people will show up on a given day, which both prevents a shortage of food as well as minimising food waste.

Publications

2025

MultiZebraLogic: A Multilingual Logical Reasoning Benchmark • Under review
Sofie Helene Bruun, **Dan Saattrup Smart**

ValEU: A Survey-Driven Evaluation of LLM Cultural Value Alignment • Under review
Annika Simonsen, Maximilian Müller-Eberstein, Rob van der Goot, Hafsteinn Einarsson, **Dan Saattrup Smart**

MultiWikiQA: A Reading Comprehension Benchmark in 300+ Languages • Under review
Dan Saattrup Smart

Encoder vs Decoder: Comparative Analysis of Encoder and Decoder Language Models on Multilingual NLU Tasks • NoDaLiDa ‘25
Dan Saattrup Smart, Kenneth Enevoldsen, Peter Schneider-Kamp

FoQA: A Faroese Question Answering Dataset • RESOURCEFUL ‘25
Annika Simonsen, **Dan Saattrup Smart**, Hafsteinn Einarsson

Hotter and Colder: A New Approach to Annotating Sentiment, Emotions, and Bias in Icelandic Blog Comments • NoDaLiDa ‘25
Steinunn Rut Friðriksdóttir, **Dan Saattrup Smart**, Hafsteinn Einarsson

2024

The Virtual Large Cardinal Hierarchy • Fundamenta Mathematicae
Dan Saattrup Smart, Victoria Gitman, Stamatis Dimopoulos

2023

ScandEval: A Benchmark for Scandinavian Natural Language Processing • NoDaLiDa ‘23
Dan Saattrup Smart

Addressing contingency in algorithmic (mis)information classification: Toward a responsible machine learning agenda • FAccT ’23 and Journal of Responsible Innovation
Andrés Domínguez Hernández, Richard Owen, **Dan Saattrup Smart**, Ryan McConville

Model Agnostic Explainable Selective Regression via Uncertainty Estimation
Andrea Pugnana, Carlos Mougan, **Dan Saattrup Smart**

Monitoring Model Deterioration with Explainable Uncertainty Estimation via Non-parametric Bootstrap • AAAI ‘23
Dan Saattrup Smart, Carlos Mougan

2022

MuMiN: A Large-Scale Multilingual Multimodal Fact-Checked Misinformation Social Network Dataset • SIGIR ‘22 and GLB ‘22
Dan Saattrup Smart, Ryan McConville

2019

Games and Ramsey-like Cardinals • Journal of Symbolic Logic
Dan Saattrup Smart, Philip Welch

Grants

November 2023
Horizon Europe • European Commission
Trustworthy Large Language Models (TrustLLM). This project aims to develop open, trustworthy and sustainable large language models for the Northern European languages. This pertains both to data curation, model development and model evaluation. EUR 7M awarded, EUR 600K of which the Alexandra Institute received. I am the work package lead pertaining to model evaluation.

March 2023
Grand Solution • Danish Innovation Fund
Conversational and **R**ead-**al**oud Speech Dataset (CoRal). This project aims to build and open-source a Danish ASR dataset of +1000 hours along with ASR- and speech synthesis models. A key focus is to minimise bias through intelligent selection of speakers. DKK 22.5M awarded, DKK 5.5M of which the Alexandra Institute received. I am the work package lead pertaining to the model development.

November 2017
DTP fully funded studentship • Engineering and Physical Sciences Research Council (EPSRC)
Highly competitive Doctoral Training Partnership (DTP) studentship, covering my PhD studies at the University of Bristol for 3.5 years.

April 2009
Einer Wilslev's Scholarship • Einer Wilslev's Fund
A scholarship given to extraordinary students at the Niels Brock Business College. Received it all three consecutive years I attended the school, i.e., 2009-2011.