Danni **Dromi**

SENIOR DATA SCIENTIST · COMPUTER SCIENCE MSc.

Persillevej 44 - 2300 Copenhagen

☐ +45 61 78 75 84 | ✉ danni@finechariot.com | ☐ dromi | in danni-dromi

Senior Data Scientist specializing in AI-powered chatbots and large language models. Experienced in developing innovative solutions based on AI and Large Language Models, with expertise in RAG applications and Azure OpenAI. Skilled in Python, machine learning, and stakeholder communication. Proven track record of delivering impactful AI projects across various industries, combining technical prowess with strong collaboration abilities.

Past Project Experience

RockGPT – Internal Chatbot System

CLIENT: ROCKWOOL

ROLE: SENIOR MACHINE LEARNING ENGINEER (GENAI)

2024 - 2025

- Helped take Rockwool's internal chatbot, RockGPT, from prototype to a production-ready solution
- Matured and stabilized the existing backend application to meet production standards
- Optimized prompting strategies for large language models to improve chatbot performance
- Performed full database migration from MongoDB to PostgreSQL
- Redesigned and rebuilt the entire frontend client for improved usability and maintainability
- Collaborated closely with designers to enhance the overall user experience
- Ensured the system met reliability and performance requirements for internal use at scale

METHODS:

Generative AI, Large Language Models, Prompt Engineering, Database Migration, Stakeholder Collaboration, Software Development

TECHNOLOGIES:

Python, Microsoft Azure, FastAPI, PostgreSQL, MongoDB, JavaScript, Vue.js, OpenAI

Semantic Search on Biology Research Documents

CLIENT: NOVO NORDISK

ROLE: SENIOR DATA SCIENTIST

2023 - 2024

- Team at Novo (BraiNN) wanted to investigate if AI would benefit their product
- Investigate use case for using AI in biology research literature search
- Discuss with stakeholders to determine and user needs and how AI can facilitate these needs
- Preform several internal presentations for the team and stakeholders on my methods and process for researching and developing AI solutions
- Provide continuous advisory to stakeholders on methods and practices for working with and integrating AI
- Analyze user tracking data to deeper understand user behavior
- Determine best models and approaches. Present and discuss these with stakeholders and the team.
- Build prototype AI application for an internal data platform
- Train and deploy several prototypes for internal and external user testing

METHODS:

Semantic Search, Data Search, Large Language Models, NLP, Machine Learning, Data Modeling, Stakeholder Management, Client Presentation

TECHNOLOGIES:

Amazon Web Services, Sentence Transformers, Open Search, BERT, Jupyter Notebooks, Azure Platform, Python, Polars, Streamlit, Neo4J

Course Planing Data Analysis

CLIENT: TEKNOLOGISK INSTITUT

2024 - 2024

ROLE: SENIOR DATA ANALYST

- Client wanted a thorough understanding of the data for the last 5 years course planing
- Understand which courses perform well, and at what time and location
- Identify patterns in course planning
- Continuous correspondence with client to understand their needs and steer focus of analysis
- Analyze data and create visualizations of key patterns and trends at different levels of granularity
- Present results multiple times to different client departments
- Propose solutions for how they can resolve undesired patterns in existing course planing
- Propose several solutions for how these findings can be operationalised using AI

METHODS:

Data Analysis, Data Modelling, Stakeholder Management & Advisory, Client Presentations, Knowledge sharing

TECHNOLOGIES:

PowerBI, Python, Pandas, Polars

SkoleGPT

CLIENT: KØBENHAVNS PROFESSIONSHØJSKOLE

2023 - 2023

ROLE: SENIOR DATA SCIENTIST

- Prototype project to build a ChatGPT style chatbot for use in danish school system
- Focus on GDPR, Open Source, Compliance, without logging or tracking
- Worked with client to sketch and understand an MVP solution
- Researching Open Source Large Language Models to find best candidate for project
- Deployed hosted model with scaling
- Implemented front-end solution based on existing Open Source framework
- Further development is planned
- Project available at https://skolegpt.dk/

METHODS:

Chatbots, Large Language Models, GDPR, Open Source, NLP, Machine Learning, Data Modeling, Prototyping, Full Stack

TECHNOLOGIES:

Azure Platform, Hugging Face, Llama 2, Python, Pandas, NextJS

Automated Fact Checking Study

CLIENT: AMESTO NEXTBRIDGE

2022 - 2023

ROLE: SENIOR DATA SCIENTIST

- A preliminary study on how current AI models & tools can support fact checking organisations
- Collaborated with Amesto NextBridge on a delivery for their client Faktisk.no (Norwegian factchecking medium)
- Study and understand current processes in fact checking
- Analyze and understand cutting edge methods for using AI in fact checking from both the academic and professional world
- Map out all related technologies
- Develop and measure POC models & their performance
- Supported client (Amesto NextBridge) on building slidedecks & handover material
- Collaborated with client (Amesto NextBridge) on presenting the outcome of the project to their client (Faktisk.no)
- Currently working on publishing a research paper with client on project findings

METHODS:

Fact Checking, Claim Detection, Claim Validation, NLP, Machine Learning, Exploratory Data Analysis, Academic Study, Client Management, Data Modeling

TECHNOLOGIES:

Azure Platform, Python, SQL, Pandas, OpenAI, BERT, PowerPoint

Food to Nutrition Matching

CLIENT: UNIKK.ME

2022 - 2022

ROLE: SENIOR DATA SCIENTIST

- Develop a product for matching grocery store items to a nutritional database
- Analyze input data to understand and detect patterns and types of food items
- Map out and investigated multiple AI based models and hypotheses
- Continuously report on results and findings to clientside stakeholders
- Delivered a functional model that could easily be integrated in their existing Azure framework
- Demonstrate further methods of improvement to the client

METHODS:

Data Analysis, Data Modeling, Data Visualization, NLP, Machine Learning, Stakeholder Management, Clustering

TECHNOLOGIES:

Azure Platform, Python, SQL, Pandas, Jupyter Notebooks, SpaCy, GIT

Modelling Data Architecture & Flow

CLIENT: UNIKK.ME

2022 - 2022

ROLE: SENIOR DATA SCIENTIST

- Map out and specify the data model for client product under development
- Create a clear overview of how data flows through product stages
- Continuous dialog with client stakeholders to extract requirements for data
- Specify data columns and requirements at different product stages
- Create multiple data architecture diagrams and data structure tables based on client requirements
- Present findings to client over a series of workshop-style meetings

METHODS:

Data Analysis, Data Modeling, Data Architecture, Stakeholder Management, Data Presentation

TECHNOLOGIES:

Python, SQL, Pandas, Jupyter Notebooks

AI Lead Scoring - Multiple deliveries

CLIENT: GROUPM

2020 - 2022

ROLE: DATA SCIENTIST

- Predict users likelihood to purchase from client on-site behavior
- Create custom ID matching and pull visitor behavior data from client website using Google Tag Manager
- Create ETL pipelines for continuously pulling data to GCP
- Perform data exploration and data modeling on gathered data to analyze and detect patterns and signals
- Build predictive models using BigQueryML
- Report to relevant stakeholders on progress & performance of solution
- Deliver and handover audiences on predefined advertisement platforms (Google DV360, GoogleAds, Facebook, etc.)
- Implemented on-site behavior based customer targeting strategies using Machine Learning.

METHODS:

Machine Learning, Stakeholder Management, ETL Pipelines, Cloud Engineering, Exploratory Data Analysis, Data Modeling, Data Visualization

TECHNOLOGIES:

Google Cloud (BigQuery(ML), AutoML, Cloud Functions, Pub/Sub, Firestore, Data Studio, Stackdriver), Azure DevOps, Python, SQL, Pandas, NodeJS, GIT

ETL marketing attribution pipeline - Multiple Deliveries

CLIENT: GROUPM

2021 - 2022

ROLE: DATA ENGINEER

- ETL integration for a marketing attribution dashboard - Multiple Deliveries
- Build a flexible ETL framework for pulling data from multiple sources
- Engineered for robustness and ease of use
- Fast deployment and easy to customize
- Running correspondence with project stakeholders regarding supporting additional platforms
- Data delivered in cleaned state in BigQuery in GCP

METHODS:

Data Engineering, Cloud Engineering, Stakeholder Management, Cloud Architecture

TECHNOLOGIES:

Google Cloud (BigQuery, Cloud Functions, Pub/Sub, Firestore, Stackdriver), Python, SQL, Git

Sellout Data Pipeline

CLIENT: GROUPM

ROLE: DATA ENGINEER

2021 - 2022

- Worked for a client of GroupM as part of a team to build a unified pipeline for ingesting sellout data
- Data arrived from multiple sources using different methods (email, FTP-server, self-service platforms, etc)
- Built ETL pipeline for ingesting massive batch datasets
- Contributed on ETL Framework for flexibly ingesting data from various email sources
- Helped client migrate an existing solution to a cloud hosted environment
- Partook in workshops and presentations to handover solutions and knowledge obtained during the implementation phase to client

METHODS:

Data Engineering, Client Management, Client Mediation, ETL Pipelines, Cloud Architecture

TECHNOLOGIES:

Google Cloud (Pyspark, BigQuery, Cloud Functions, Compute (VM's), Stackdriver), Azure DevOps, Python, Pandas, SQL, Git

Activation Framework

CLIENT: GROUPM

ROLE: DATA ENGINEER

2020 - 2021

- Worked for a client of GroupM as part of a team to build pipelines for uploading datasets to activation platforms
- Partook in client process meetings (Daily standup, sprint planning, scoping, etc.)
- Collaborated with client-side engineers
- Extended upon existing client framework and client methodologies
- Performance tested and reported on built solutions

METHODS:

SCRUM, Data Engineering, ETL Pipeline, Client Management

TECHNOLOGIES:

Microsoft Azure (Databricks, DataLake Storage, Cosmos DB), Pyspark, Apache Spark, Python, SQL, Git

Video Analysis Framework

CLIENT: GROUPM

ROLE: DATA SCIENTIST

2020 - 2020

- Worked in collaboration with two MIT students to build a pipeline for analyzing a commercial video
- Split a commercial into separate cuts by Machine Learning model
- Analyze each cut in various steps based on still image, audio & detected text
- Each analysis done using different pretrained Machine Learning model
- Partial results & final results written to GCP
- Project used experimentally for various use cases (eg. analyzing an episode of the DR2 show "Deadline")

METHODS:

Machine Learning Engineering, Cloud Engineering, Data Science

TECHNOLOGIES:

Google Cloud (Storage (buckets), BigQuery, AutoML, Stackdriver, Compute (VMs), Python, Shell Script, Keras, Git

Docunet

CLIENT: PAPERFLOW

ROLE: MACHINE LEARNING ENGINEER

2019 - 2020

- Worked on a Machine Learning model for digitally flattening creased documents
- Method based on published academic research
- Built code for creating and synthesizing training set
- Developed model architecture based on described research

METHODS:

Machine Learning Engineering, Data Science

TECHNOLOGIES:

Tensorflow, Python, Numpy, Pandas, SQL, Scipy, Convolutional Neural Networks (CNN), Deep Neural Networks (DNN)

Catalog Classifier

2018 - 2019

Role: Machine Learning Engineer

- Classify and match invoices by creditor to a customer provided catalog
- Build infrastructure for gathering data from multiple data sources
- Clean-up and structure data
- Experiment with model architectures
- Implement system of Machine Learning models
- Dynamically predict and retrain models using GCP
- Report and correspond with customers and stakeholders on performance of solution and ad-hoc problems

Methods:

Machine Learning Engineering, Data Science, Stakeholder Management, Cloud Engineering

Technologies:

Google Cloud (AI Engine, Storage), Python, Keras, Tensorflow, Numpy, Pandas, SQL, Scipy, Convolutional Neural Networks (CNN), Deep Neural Networks (DNN), Transfer Learning

Catalog Classifier

Client: Paperflow

2018 - 2019

Role: Machine Learning Engineer

- Classify and match invoices by creditor to a customer provided catalog
- Build infrastructure for gathering data from multiple data sources
- Clean-up and structure data

Workshops & Courses

PNO
Half day workshop

AGENTIC AI
Mar 2025

Follow-up workshop on agentic AI held with Philip Petersen from Prehype. Included an introductory talk on the underlying technology and mechanics of agentic AI, followed by hands-on sessions where participants explored practical applications. Served as technical expert and facilitator, assisting participants with implementation questions throughout the workshop.

Herning Bibliotekerne
2 day workshop

AI & PROMPTING FOR LIBRARIANS
Sep 2024

Included an internal course and an open workshop for libraries across Jutland, covering prompting, image and video generation, advanced AI tools, songwriting with AI, chatbot building, and responsible use of the technology.

Nordisk Ministerråd
2 hour mini-workshop

AI IN THE NORDIC SETTING
Aug 2024

Mini workshop on modern AI technologies and their applications in communication and policy across the Nordic region. Covered how language models like ChatGPT work, their use in public administration, and the representation of Nordic languages in large language models.

Lemvigh-Müller
90 minute segment with workshop

AI & PROMPT ENGINEERING
Jun 2024

A talk and workshop for top management covering the transformative nature of modern AI, fundamentals of prompt engineering, hands-on activities to invoke curiosity, discussions on AI's impact on work, and ethical considerations.

Pharma4Ever
Full Day Workshop

AI & CHATGPT WORKSHOP
May 2024

Focused on generative AI in the pharmaceutical industry, including fundamentals, applications of ChatGPT, hands-on assignments with advanced AI tools, discussions on potential and ethical concerns.

Danmarks Medie og Journalisthøjskole (DMJX)
Full day Course

GENERATIVE AI FOR JOURNALISM
March 2024, May 2024 & May 2025

Courses on using generative AI for journalism, including text and image generation with ChatGPT and DALL-E 3, hands-on assignments, and discussions on ethical considerations.

Danske Spil
90 minute intensive workshop

GENERATIVE AI @ DANSKE SPIL
2024

Part of a kickoff conference, includes intro to prompt engineering, case study, group assignments on applying Danske Spil products to new user segments, and result discussions.

PROSA
3 hour after work workshop

GET STARTED ON GENERATIVE IMAGES/VIDEO/AUDIO
May 2024 & September 2023

Focus on visual/audial generative AI tools, open workshop format, emphasis on collaboration and result sharing.

Dansk IT
2 day and 1 day course

CHATGPT FOR IT-PROFESSIONALS
November 2023, March 2024

Workshop on ChatGPT application for IT professionals, including hands-on assignments, discussions on AI landscape and ethical considerations.

Transmission Learning
Full day workshop

GENERATIVE AI WORKSHOP
2023

Conducted six times, focus on generative AI for text and image generation, hands-on assignments, discussions on practical applications and ethical considerations.

Talks

ComputerWorld Cloud & AI Festival
Copenhagen

INNOVATION IN PRACTICE: ROCKWOOL BUILDS ITS OWN GEN AI TOOL
Sep 2025

A talk together with ROCKWOOL on RockGPT their internal AI Chatbot and my contributions to the project.

Driving AI (IDA)

IDA Conference

ASSESING THE CLIMATE IMPACT OF AI

May 2025

On behalf of Danish Data Science Community (DDSCA), I gave a talk together Kasper Groes regarding our latest publication on AI's climate impact.

Socialdemokratiet

Christiansborg Palace

POLITICAL AI

Sep 2024

Presentation on the role of AI in political and organizational contexts, covering AI's place in the broader digitalization trend, practical applications in administrative work, and future challenges in responsible adoption. Included interactive discussions on efficiency, ethics, and the societal impact of AI.

ComputerWorld Summit

Dansk Industri

THE AI REVOLUTION: TRENDS YOU HAVE TO KNOW

May 2024

A talk on the current trends and hottest topics in AI.

Driving AI (IDA)

IDA Conference

SKOLEGPT

May 2024

A talk together with Martin Exner of Københavns Professionshøjskole on the process and future perspective of the SkoleGPT project I helped create.

Nutanix

ComputerWorlds studio

(AI) MASTERCLASS

May 2024

I cohosted an online Masterclass on AI together with Thomas Jensen from Nutanix.

Dansk Erhverv/Danmarks Rejsebureau Forening (DRF)

Aller Medier

AI IN THE TRAVEL INDUSTRY

Apr. 2024

A talk on AI and its influence and use in the travel industry, originally planned to be hosted at Børsen but moved due to fire.

Beauty 2024

Bella Centret

AI IN THE BEAUTY INDUSTRY

Mar. 2024

A talk on how AI is used within the beauty industry.

VL Group 53

Company Onsite

GENERATIVE AI & THE FUTURE OF WORK

Jan. 2024

A talk on the last 6 months development within AI and how these trends can affect the future of work.

COMPUTERWORLD - AI Business Excellence Day

Virtual

SAFE APPLICATIONS OF AI ON YOUR OWN DATA

Dec. 2023

I gave a talk on how businesses can apply AI LLMS on their own data without compromising security. I also showcases uses for this technology.

Driving IT

IDA - Copenhagen

WILDER GENERATIVE AI

Nov. 2023

I gave a talk on generative AI and what progress we will see in the near and distant future.

ComputerWorld Cloud Festival

Copenhagen

3 IMPRESSIVE CASES OF APPLIED AI

Sep. 2023

I gave a talk together with Søren Christian Søndergaard Poulsen showcasing some of the most impressive usecases of AI both in Denmark and internationally.

Internal Presentation

UFST (Skat) - Copenhagen

SKOLEGPT PRESENTATION

Sep. 2023

I gave a talk at UFST on my work in developing SkoleGPT as well as the experiences I've gained in working with cutting edge Large Language Models.

Driving IT

IDA - Copenhagen

DEEP GENERATIVE MODELS

Nov. 2022

I gave a talk on Generative AI, its history and progress (right before ChatGPT came out).

PyData Copenhagen *Alm Brand - Copenhagen*

8 YEARS OF DEEP GENERATIVE MACHINE LEARNING *Sep. 2022*

I gave a talk on the exciting and rapidly progressing history of Deep Generative ML and also showcased some of my own creative work.

Education

UCPH (University of Copenhagen)

MSC. IN COMPUTER SCIENCE *Nov. 2014*

- Elective Project: Hybridization of mesh-free and mesh based simulation methods
  developed using MATLAB, grade: 10
- Studied a semester abroad (aug. - dec. 2013) at the *University of British Columbia*, Vancouver - Canada
- Master's Thesis: The Use of Moving Meshes for Segmentation and Registration of 2D Medical Imaging
  developed using C++, grade: 12

UBC (University of British Columbia)

EXCHANGE STUDENT *Aug. 2013 - Dec. 2013*

- I studied a semester abroad, at UBC in Vancouver - Canada

UCPH (University of Copenhagen)

BSC. IN COMPUTER SCIENCE *Jun. 2012*

- BSc. Project: Using Compounds of Convex Shapes in Interactive Rigid Body Simulation
  developed using C++, grade: 12

Frederiksberg Gymnasium

MATHEMATICAL STUDENT *Jul. 2007*

Additional

Languages   Danish (mother tongue), English (high professional level), German (still learning)

Skill Matrix

Skill Name/Category	Level	Experience	Last Used
Industry			
Pharmaceutical	★★★☆☆	1 year	2024
Education	★★★☆☆	1 year	2024
Marketing	★★★★☆	2 years	2022
FinTech	★★★★☆	4 years	2020
Artificial Ingelligence & Data Science			
Generative AI	★★★★★	4 years	2025
RAG Applications	★★★★★	2 year	2025
Azure OpenAI	★★★★☆	2 years	2025
Data Modelling	★★★★★	5 years	2024
Python	★★★★★	10 years	2025
Jupyter Notebooks	★★★★★	5 years	2024
Large Language Models	★★★★★	5 years	2025
Natural Lanauge Processing	★★★★★	5 years	2024
Semantic Search	★★★★☆	1 years	2024
Conversational Language Models (Chatbots)	★★★★☆	3 years	2025
Data Analysis	★★★★☆	4 years	2024
Machine Learning	★★★★★	5 years	2024
Pandas + Polars	★★★★★	5 years	2025
Data Visualization	★★★★☆	3 years	2023
Communication			
Presentation	★★★★★	4 years	2025
Workshop Facilitation	★★★★☆	3 years	2025
Knowledge Sharing	★★★★★	3 years	2025
Stakeholder Management	★★★★☆	3 years	2025
Client Communication	★★★★☆	4 years	2025
SCRUM/Agile	★★★★★	7 years	2024
Data			
SQL	★★★★★	10 years	2024
ETL Pipelines	★★★★☆	4 years	2023
Database Migration	★★★☆☆	1 years	2025
Neo4J	★★★☆☆	1 years	2024
PostgreSQL	★★★★☆	5 years	2020
MongoDB	★★★☆☆	1 years	2025
Redis	★★★☆☆	2 years	2020
Platform			
Microsoft Azure	★★★★☆	2 year	2025
Google Cloud Platform (GCP)	★★★★☆	4 years	2024
Amazon Web services (AWS)	★★★☆☆	3 years	2024
Linux	★★★★★	12 years	2024
Development Language & Technology			
Python	★★★★★	7 years	2024
Node.js/Javascript	★★★★★	4 years	2023
Vue.js	★★★★☆	1 years	2025
Git	★★★★★	10 years	2024
Docker	★★★☆☆	3 years	2024
Continuous Integration/Continuous Deployment (CI/CD)	★★★★☆	5 years	2022
Estimation	★★★★☆	4 years	2024