

# Anders E. Nielsen

DATA DRIVEN PHARMA · STATISTICS · DATA INFRASTRUCTURE · AWS

✉ andellegaard@gmail.com | 🏠 anhosu.com | 📺 AnHosu | 📄 anders-ellegaard

I am an ambitious data professional with 6 years of experience applying modern data technology to solve a diverse set of problems within drug development, supply chain, and pharmaceutical manufacturing. I am motivated by complex problems and am constantly looking to develop my skills.

## Experience

---

### Gubra

*Hørsholm, Denmark*

#### RESEARCH SCIENTIST, COMPUTATIONAL BIOLOGY

*2021-05 to present*

- I am contributing to preclinical research and drug discovery at Gubra, where I am responsible for the data and app infrastructure in our drug discovery platform.
- Digitalisation of internal discovery processes and customer facing services, through Shiny applications and robust data-informed processes.
- Building a robust and scalable backend for the peptide drug discovery platform and making data available to users and scientists through web apps and APIs.
- Building a pipeline for potency data analysis. I ensured automatic and uniform treatment so results were comparable across hundreds of experiments and made results easily available for downstream processes.
- Building a pipeline for summarising and analysing mass spectrometry data for use in drug discovery.
- Applying Large Language Models (LLMs) to peptide design.

### Novo Nordisk

*Hillerød, Denmark*

#### DATA SCIENTIST

*2017-09 to 2021-04*

- Increasing overall equipment efficiency of an assembly line by applying a combination of Design of Experiments and Gaussian Processes Regression and Bayesian optimisation. Skills applied: ML, A/B testing, Python, Bayesian statistics.
- Predictive maintenance on brownfield manufacturing equipment. Skills used: AWS Greengrass, AWS IoT, CI/CD with Azure DevOps
- Digital transformation of 3 factories through fast-paced prototyping and development of data based products. One of these factories is recognised as a Digital Lighthouse Factory by the World Economic Forum. Skills applied: Agile/Scrum, data visualisation, optimisation.
- Creating dashboards for planning in factories. Skills applied: Docker, Tableau, Streamlit, InfluxDB, visualisation, Python.
- We built a regulatory framework for applying machine learning to quality-critical pharma processes. Skills applied: GxP, models in production, model validation, model monitoring.
- Involved in hiring 15+ data professionals, personally leading 100+ interviews.
- I started out in Supply Chain where I applied data and occasionally machine learning to simplify processes and understand trends while maintaining SQL databases.
- From 2019-02, I helped establish a Manufacturing Intelligence department. We applied modern data technology to manufacturing problems while advocating data science and machine learning within the organisation. My responsibilities were quite diverse ranging from data engineering tasks and coming up with machine learning cases to teaching colleagues and hiring.

### Innovation Centre Denmark, Embassy of Denmark in South Korea

*Seoul, South Korea*

#### INTERN

*2016-08 to 2017-01*

- Collected market intelligence, facilitated Korean-Danish technology collaboration, arranged and executed large events, and had various diplomatic tasks.

## Education

---

### Technical University of Denmark (DTU)

Lyngby, Denmark

M.Sc. BIOINFORMATICS AND SYSTEMS BIOLOGY ENGINEERING

2015-08 to 2017-08

- Thesis on applying Deep Learning techniques to identify potential enzyme sequences in large metagenomic datasets.
- Courses in advanced bioinformatics, protein structure & computational biology, systems biology, metagenomics, cell factory engineering, machine learning, etc.

### Technical University of Denmark (DTU)

Lyngby, Denmark

B.Sc. BIOTECHNOLOGY ENGINEERING

2012-08 to 2015-07

- Thesis on analyzing terabyte-scale NGS metagenomic data to discover ancient pathogens.
- Semester abroad at Seoul National University.

## Skills

---

### Programming & Scripting

I have been writing **Python** for 6+ years and I am comfortable with several common data manipulation libraries. I also have 2+ years of experience writing **R**, using **Shiny** to build apps and **Plumber** to build **APIs**. I am **Unix**-savvy and used to work in command-line environments.

### Machine Learning / AI

I have a solid theoretical understanding of ML and deep learning methods, as well as 5+ years of experience applying them to various domains. I have applied ML to solve complex problems, often with sparse or little data. Toolbox: **Stan**, **LLMs**, **SciKit**, **Pytorch**, etc.

### Databases & SQL

I have been developing and using databases for 5+ years. I am comfortable with multiple dialects of SQL, though mostly **MSSQL** and **Postgres**. I have designed and built databases for **production** from scratch on multiple occasions. I have NoSQL experience too, mostly DynamoDB.

### Statistics

I have a strong theoretical understanding of classical statistics and experience applying **Bayesian** and multivariate statistics, including ML methods like Gaussian Processes. I have experience implementing statistical workflows in R and Stan.

### Computational Biology

Metagenomics, Peptide Design, NGS

### Developer Tools

**Docker**, Git, CI/CD with Azure DevOps, **Agile/Scrum**

### AWS

Certified **SAA**. Experience with IoT, S3, EC2, ECS, Kinesis, DynamoDB etc.

### Languages

Fluent in Danish, English, and Korean