



10

YEARS OF
EXPERIENCE

Introduction

Johannes is an experienced and gifted developer with a strong problem-solving mindset. He has extended experience and a special interest in data science/data engineering. He prefers interactive processes, involving stakeholders and a relevant team early on in the project. That way he can ensure involvement in the requirement process and create the baseline for communication in the project. His greatest technical strengths are in python and data management. Johannes has a strong sense of responsibility and has great interpersonal skills, besides his technical palate, Johannes is a warm, professional person with a helpful attitude which makes him a great Project Manager or part of a team.

Github: <https://github.com/johannesoccean>

Medium: <https://medium.com/@johannes.ocean>

Personality



Ambitious



Nature lover



Analytical

Skills

Python	AI	SQL	JavaScript	HTML	CSS	Matlab	Git	Github	Gitlab	Subversion	Agile
Docker	Windows	Linux	Microsoft Office	QGIS	PyCharm	VSCode	Jupyter Notebook	Conda			
SQLite	draw.io	PostgreSQL	PGvector	PostGIS	Apache	Azure	Snowflake Cortex	OpenAI			
Machine Learning	PyTorch	Scikit-learn	Numpy	Pydantic	Pandas	Geopandas	Shapely	Fiona			
Rasterio	Seaborn	Matplotlib	Basemap	Streamlit	Bokeh	Dash	Plotly	Folium	PyQGIS		
Scipy	Statsmodels	Satpy	Tkinter	Kivy	Flask	FastAPI	Alembic	Pytest	Unittest	YAML	
JSON	netCDF										

Work Experience

Software Engineer Level 3 / Senior Data scientist / Volvo Cars Corporation 2024 Q3 - Ongoing

As a Python developer / Senior Data Scientist at Volvo Cars Safety Centre, I apply advanced data analytics to understand accident causation, driving safety enhancements in car development.

My primary role involves:

- developing and maintaining advanced dashboards using Streamlit.
- working extensively with complex data sets, conducting data processing using mainly Pandas, and Numpy.
- development of a Chatbot-RAG solution using Azure OpenAI, Langchain, PGVector and ChromaDB.
- work with Unittest employing Pytest.
- development and maintaining of a machine learning project that implements a pre-trained resnet model from NASA.

Python

Streamlit

Unittest

Docker

Azure

Snowflake Cortex

Pandas

Numpy

PGVector

ChromaDB

AI

RAG

LLM Output Parsing

Langchain

English

Chatbot-POC, Developer / Data Scientist / Svensk Byggtjänst 2023 Q4

The work largely entailed breaking down the AMA library into manageable chunks and creating embeddings from these data sets. Utilization of an Azure OpenAI model (inside EU borders) and vector similarity search in Postgres using pgvector, ultimately provided optimized embedding matches and responses.

Backend

Python

FastAPI

SQLAlchemy

PostgreSQL

PGvector

Alembic

Docker

Git

OpenAI

Embeddings

Agile

Chatbot-POC, Developer / Data Scientist / Erik Olsson 2023 Q2

Implemented few-shot learning in a chatbot designed to generate real estate advertisements for houses and apartments.

Python

Azure OpenAi Studio

Prompt engineering

Agile



Integration-API, Developer / Belid lighting group

2023 Q2-Q4

The work involved developing an API that is responsible for shipping information from the Sanity CMS to a Centra system using webhooks from Sanity and GraphQL queries and mutations to Centra. Another key responsibility of the API was to update the inventory balance in a Monitor database. This involved regularly checking for updates and synchronizing inventory information between the various systems.

Backend

Python

FastAPI

Pydantic

Sanity CMS

GraphQL

Git

Agile



Chatbot-AI-API, Developer / Data Scientist / Fortum

2023 Q2

As part of a team, Johannes assumed the primary responsibility for setting up the main structure of the chatbot API, utilizing the FastAPI framework. Within the project we employed OpenAI's LLM and GPT models, embeddings, and langchain for enhanced language understanding and implemented precise prompts to improve user experience. Additionally, Johannes managed code using Git, collaborated effectively in an agile international environment.

Backend

Python

FastAPI

Pandas

Streamlit

OpenAI

AI

LLM

GPT

Embeddings

Langchain

Prompt engineering

Git

Agile

English



Machine Learning-API project, Developer / QueensLab

Nov 2022 - 2023 Q1

Without revealing too much, this is an inhouse project with focus on setting up the following:

- An API using FastAPI
- Database CRUD-handling with SQLAlchemy
- Database migrations using Alembic
- Text recognition using Azure Text Analytics
- Azure Blob Storage
- OpenAI integrations (embeddings)

Backend

Python

FastAPI

SQLAlchemy

PostgreSQL

Alembic

Azure

Docker

Git

OpenAI

Agile



Integrating microservices, Developer / SMHI

2022 Q1 - Oct 2022

Within SMHI there is a need to get away from manual file changes. Johannes has taken the initiative to utilize, develop and implement microservices (architecture) into some of the daily workflow inside the data center at SMHI. With a web app (https://github.com/sharksmhi/flask_station_app) a list of monitoring stations is retrieved from a microservice (https://github.com/shark-microservices/microservice_station), manage, visualize and connect them to "The national station register" via an API. Using tools like this eliminates some of the human errors that otherwise occur from time to time.

Fullstack

Python

Folium

Flask

Pandas

HTML

CSS

Apache

Agile

Github

Git



Weather Station Website, Developer / Data Engineer / Hobby project

2022 Q1

Data from a weather station is collected and sent via a Raspberry Pi to an API hosted on a VPS. The data runs through quality control and is then stored in a database. The data is presented on a dash/plotly website (<https://utmaderna.se/>). From the open API at SMHI I retrieve forecast data as a complement to the observations.

Fullstack

Python

Flask

Pandas

SQLite

Dash

CSS

Apache

Ubuntu Server

ETL

Meteorology



Developing Microservices, Developer / SMHI

2021 Q3 - 2022 Q3

There was a need to enhance the digital cooperation within the Swedish marine community, and between government authorities. In 2021, SMHI received a development assignment from SwAM (The Swedish Agency for Marine and Water Management) to develop microservices that can be used for various forms of data validation and code translations. The idea is that the work put forward in this assignment should be used as an example of how to use microservices, not only at SMHI, but on a national level. Here, Johannes got the chance to expand his skill set as a developer with technologies such as the Apache software and OpenAPI-specifications.

Backend

Python

Flask

OpenAPI

FastAPI

Pandas

Geopandas

Apache

Agile

Github

Git



GIS-system - Satellite Algae Surveillance, Developer / Data Scientist / SMHI

2019 - Oct 2022

During summer time SMHI operates the "Algae situation". In 2019, with the aim of replacing the old system, Johannes started to develop a new python plugin for QGIS (Quantum Geographic Information System), which allows the user to manually adjust algae maps based on satellite data. The plugin is used daily by SMHI personal for creation of algae maps. The main data formats are shape- and raster files. The work also contained alot of data aggregations, -analysis, -visualization. Some final products can be seen in this annual report

(<https://www.smhi.se/publikationer/publikationer/cyanobakterier-i-ostersjon-sommaren-2022-1.191705>)

QGIS

Python

PyQGIS

Fiona

Rasterio

Shapely

Geopandas

Matplotlib

Seaborn

Numpy

Pandas

Agile

ETL

Oceanography



Svea - Dataflows, Developer / Data Scientist / SMHI

2021 Q3 - Oct 2022

R/V Svea is the top notch research vessel aimed for marine monitoring over Swedish waters. A large team at SMHI (oceanographers, marine biologists, scientists, technicians

and developers) was given the means by the Swedish government to work with setting

up a sustainable data flow (from the ocean to story telling products). Johannes work revolved

around setting up data structures, quality control routines and visualize data aggregations.

Python

Pandas

Numpy

Matplotlib

Seaborn

Bokeh

JavaScript

SQLite

ETL

Github

Git

Oceanography



SHARKvalidator, Developer / SMHI

2021-2022

In the work against bad data the Oceanographic data center at SMHI needed open and transparent validation of data deliveries. Johannes pushed for a new code structure, published at <https://github.com/sharksmhi/sharkvalidator>. The standard validation routines checks: mandatory data types (metadata, data), mandatory data fields, data formats.

Python

Pandas

Github

Git

Oceanography



Climate scenarios, Developer / Data Scientist / SMHI

2021

In the beginning of 2022 SMHI updated the ocean related climate scenarios service. During the months prior to publication, Johannes responsibility was to convert model data (netCDF) into shapefiles. It might sound quite straight forward, but the different data transformation steps include a few pitfalls to look out for (eg. bowtie geometries). Beside the data transformation, Johannes also participate in quality assurance in regards to the oceanographic model data (NEMO, RCO-SCOBI).

Python

Geopandas

Shapely

Rasterio

netCDF

QGIS

Agile

Gitlab

Git

Oceanography



SHARKweb 2.0, Stakeholder / SMHI

2020

SHARKweb (<https://sharkweb.smhi.se>) needed a face lift. The result is a new re-designed web based user interface for searching and downloading of marine environmental monitoring data. As an expert on oceanographic data within SHARK, Johannes specified the work tasks for a team of front-end developers. It also included communication with product owner at SMHI, the end-users at eg. universities / coastal zone county government and the co-financier SwAM.

Agile



Data Management Team FO SHARK, Developer / Data Engineer / Oceanographer / SMHI

2015 - Oct 2022

SMHI is responsible for the largest marine environmental monitoring project in Sweden, which generates a significant amount of data from lab analyses and in situ sensors. A dedicated team is in charge of supporting the various departments at SMHI that rely on oceanographic data. The team's responsibilities may include developing new routines for quality control or data visualization, as well as engaging in agile communication with end-users. In recent years, a major focus has been on creating new data flows for instruments on the research vessel, Svea, which has led to a slight shift in the team's priorities since 2020.

Python

Tkinter

Postgresql

Pandas

Numpy

Matplotlib

Seaborn

Bokeh

Agile

ETL

Github

Git

Oceanography



2015 - Oct 2022

SMHI is, on behalf of SwAM, the national data host for oceanographic and marine biological data. The “data treasure” dates back to the year 1893, enabling scientists, students and authorities to explore time series of e.g. temperature, oxygen conditions or algal blooms. Data is accessible via SHARKweb and the SHARKdata API. The data host's mission is to quality control, store and make data available with an open license, in a transparent manner. Within the framework of the assignment, Johannes used his expertise in physical oceanography and passion for python programming to develop tools that reduce manual work and facilitate: handling, quality control and visualisation of data. As a data host, a big part of the daily work involves interacting with data providers to ensure the delivery of high-quality data and metadata. Effective communication with these data providers is essential to achieve this goal. Furthermore, Johannes was responsible for the annual international reporting of physical and chemical oceanographic data to ICES (International Council for the Exploration of the Sea).

Python

Bokeh

Matplotlib

Seaborn

Numpy

Pandas

Geopandas

Folium

Subversion

Agile

ETL

Oceanography

Education



M.Ed., Physical Oceanography / University of Gothenburg, Gothenburg

2013–2015



B.S., Oceanography / University of Gothenburg, Gothenburg

2010–2013



Natural Science Base Year / University of Gothenburg, Gothenburg

2009–2010



Military service, Medic / T2 Regiment, Skövde

2006-2007



Technology - Data / Upper Secondary School - Lindälv, Kungsbacka

2003-2006