

Christoph Möhl
Dorfstraße 58 A, 53125 Bonn

+49 (0)228 38767742
info@moehl-data-services.de
moehl-data-services.de



Dr. Christoph Möhl

Data Scientist - Data Engineer - Python Developer

Your partner for **Machine Learning**, **Artificial Intelligence**, **Data Science**, **Data Engineering** and **Python**.

As an experienced data specialist in scientific and corporate environments, I offer design and implementation of machine learning models, data transformation and data modeling, data analysis service as well as consulting in the field of data science, big data and image processing.

Experience

I have **more than 10 years of practical experience as a Data Scientist, AI expert, image analysis specialist and Python developer**. I use AI methods pragmatically and can critically assess their benefits.

As a long-time **leader of a data analysis service department** in a scientific environment, I have gained experience in **project management and personnel recruitment** and actively participated in the development of new technologies.

As a Data Scientist in the field of wholesale and distribution, **I know the requirements of medium-sized companies** and can develop customized solutions in the field of **AI, Big Data and Data Engineering** with a sense of proportion for the respective budget framework.

Among others, I have worked on the following use cases:

- Prediction of **sales figures** (wholesale and retail)
 - Processing and analysis of **sensor data** (biotechnology)
 - Automated **object recognition** in medical **image data** (Biomedical Research).
 - **Data Science Consulting** and **Data Engineering** and **Big Data** handling for the development of new therapies (drug discovery, pharma)
 - Backend and frontend development of **web based BI dashboards** (investment banking).
-

Employments

- since 2022: **Freelance Data Scientist**, *Möhl Data Services*
 - 2021-2022: **Data Scientist and Data Engineer**, *IPS Press Services GmbH, Meckenheim, Germany*
Development of AI based sales prediction models.
 - 2009-2021: **Leader of Image and Data Analysis Facility**, *German Center of Neurodegenerative Diseases (DZNE), Bonn, Germany*
Lead Data Scientist of a workgroup offering inhouse software development and data consulting services for research groups in the field of biomedical research.
 - 2009-2011: **Research associate (PostDoc)**, *EMBL (European Molecular Biology Lab), Biophysics Unit, Gilmour Group, Heidelberg, Germany*
Development of quantitative fluorescence microscopy assays in the field of zebrafish development.
 - 2006-2009: **Research associate (PhD candidate)**, *Research Centre Jülich, Biomechanics Unit, Merkel Group, Jülich, Germany*
Development of quantitative fluorescence microscopy assays in the field of cell mechanics and cell migration.
-

Education

- 2006-2009: **PhD Biophysics**, *Research Center Jülich, Germany*
Quantitative live cell imaging studies on focal adhesion dynamics in migrating skin cells.
 - 1999-2005: **Diploma Biology/Computer Science**, *University of Bonn, Germany*
Thesis: Mathematical Modelling and Measurement of Cell Migration Dynamics (Group of Wolfgang Alt)
-

Certificates

- Azure Data Scientist Associate



More info: https://www.credly.com/badges/3a17db83-d4c1-494b-9ba7-0401b5f3a5fa/public_url

- Azure Fundamentals



More info: https://www.credly.com/badges/f25a53fb-ee2e-4d9a-8237-d6dfb172f375/public_url

- Azure Data Fundamentals



More info: https://www.credly.com/badges/4e918531-b8ae-414d-98f3-60f39e18f787/public_url

- Azure AI Fundamentals



More info: https://www.credly.com/badges/402014eb-2742-4fc0-a218-4a232a0fcc03/public_url

Projects

Python Developer, Frontend and Backend

Time: 11/2023 to 12/2023

Client: University Cologne (CECAD Imaging Facility)

Development of an exporter for OMERO, a database for biomedical image data

Tasks:

- Extension of the open source command line tool omero-cli-transfer to transfer research data from the OMERO database to ARC repositories.
- Development of a mapping specification for transferring OMERO projects to ARC repositories.
- Extension of the OMERO web frontend to display ARC metadata and control data export.
- Documentation

Tools: Python, OMERO, Postgres, Django, Git, Github Actions, Docker, Pytest,

Python developer, ML Engineer, AI Expert (image analysis, computer vision)

Time: 09/2023 to 10/2023

Automated detection of neurons in microscopic images with artificial intelligence.

Tasks:

- Development of a command line tool for Deep Learning based recognition of cell objects in large image files (>10GB)
- Development of a workflow for users to train an individual Deep Learning model and for quality control purposes
- Consulting for optimization of data measurement (confocal microscopy).
- Documentation

Tools: Python, Tensorflow, Scikit-Image, aicsimageio, cellpose,

Data Scientist

Time: 09/2023 to 10/2023

Client: Industry

Feature ranking analysis for an industrial manufacturing facility.

Tasks:

- Data import
- Data cleansing
- Explorative data analysis
- Feature analysis
- Training of supervised machine learning models
- Feature ranking analysis with SHAP
- Documentation

Tools: Python, Pandas, Scikit-Learn, Machine Learning,

Python Developer / Data Engineer

Time: 03/2023 to 07/2023

Client: DekaBank

Frontend and backend development of a web application for portfolio management.

Tasks:

- Frontend and Backend development with Python and Javascript
- Design of the class-based software architecture in the backend
- Data model design (combination of relational and json based model), implementation of test, staging and production database
- Documentation
- Setup of continuous integration pipeline (package installation, unit testing, pep8 checks, automated builds of documentation)
- Work within an interdisciplinary team of financial experts, software developers and analysts

Tools: Python, Plotly Dash, Pydantic, Mypy, Sphinx, Gitlab, Pandas, Mssql, Pytest, Javascript,

Python Developer / Data Engineer

Time: 09/2022 to 12/2022

Client: DekaBank

Frontend and backend development of a business intelligence web application.

Tasks:

- Frontend and Backend development
- Development of data buffering solutions for fast data provision of fragmented data sources.
- Handling and effective provisioning of big data tables.
- Implementation of a business intelligence web app
- Refactoring prototype scripts to production code (unit tests, continuous integration...)
- Work within an interdisciplinary team of financial experts, software developers and analysts

Tools: Python, Pydantic, Mypy, Plotly Dash, Flask, Gitlab, Pandas, Mssql, Pytest, MLflow, Parquet, Javascript,

Data Scientist and Data Engineer

Time: 02/2021 to 06/2022

Client: Medium-sized Trading and Logistics Company

Development of AI based sales prediction models.

Design and implementation of an AI-supported prediction model for sales of print media in magazine distribution, prototyping, implementation and deployment of the productive system.

Tasks:

- Conception of a deep learning model for the prediction of sales figures
- Development of a data model and ETL processes for processing raw data
- Implementation of the automated prediction service based on an AI model

Tools: Tensorflow, Postgres, MS SQL Server, Python, Ssqlalchemy, Alembic, Docker, docker-compose, git,

Data Engineer / Project Lead

Time: 04/2020 to 12/2020

Client: Research Institute

Design and implementation of a Postgres database for data management of an automation system.

An automation plant produces sensor data of various types, which are fed into mathematical prediction models together with various metadata. Measurement data and metadata are to be stored centrally in an SQL database. The project requires close interaction with employees who operate the automation system and evaluate the data. The interdisciplinary team includes biologists, chemists, technicians, data analysts and software developers.

Tasks:

- Project management
- Development of the data model in numerous workshops.
- Development of import specifications in close cooperation with future users.
- Implementation of the model in Python/Ssqlalchemy
- Setting up a Postgres test database with docker-compose and Gitlab-CI
- Implementation of importer tools in Python.

Tools: Postgres, Python, Ssqlalchemy, Docker, docker-compose, Gitlab-CI,

Data Scientist / Python Developer

Time: 09/2019 to 03/2020

Development of a machine learning application for the analysis of sensor data.

Highly noisy time series data are recorded via a sensor within a measuring apparatus. In these time series, sporadically occurring events are to be automatically detected and characterized. For this purpose, a machine learning software with MVC architecture was implemented in Python. The software includes the following functionality: Filter and normalize the raw data, calculate robust metrics based on the pre-processed data, provide a graphical user interface to display data and interactively create training data sets, automatic detection of events using a supervised machine learning method, export of the result data.

Tools: Python, Tkinter, Scikit-Learn, git,

Python Developer / Project Lead

Time: 03/2016 to 12/2020

Development of YAPiC, an open source software for analyzing biomedical image data using deep learning.

Tasks:

- Project management
- Conception and algorithm development
- Presentation of the software at international conferences
- Deployment, development of CI/CD pipelines
- Management of the further development of the tool by the open source community

Tools: Python, TensorFlow, Travis-CI, GitHub,

Data Scientist / Image Processing Specialist

Time: 02/2017 to 07/2017

Automated characterization of tissue samples using deep learning

As part of a medical research project, tissue samples were photographed with an automated microscope. Based on Python and Tensorflow, software was developed to identify and classify specific cell types in the tissue. In this way, different cell types could be counted automatically for user-defined tissue regions.

Tools: Python, TensorFlow, git,

Python Developer, DevOps

Time: 10/2018 to 06/2019

Development of a parallelized image analysis pipeline for processing massive image data of an automated microscope.

An automation system within a pharmaceutical laboratory produces terabytes of image data daily. Based on CellProfiler software, object recognition and feature calculation was implemented to extract structured data from the raw image data. For robust deployment on an in-house CPU cluster, the application was containerized with Docker and orchestrated with SLURM.

Tasks:

- Definition of the specifications in collaboration with the domain experts
- Planning and acquisition of necessary hardware
- Conception and implementation
- Big Data handling

Tools: Python, CellProfiler, Docker, SLURM,

Data Scientist Drug Discovery

Time: 01/2014 to 03/2015

Drug Screening Analysis including Feature Engineering, Clustering and Ranking Analysis

Tasks:

- Setup and execution of image analysis pipeline for automated object detection of image based drug screening data.
- Feature processing and selection
- Clustering analyses to identify drug candidate groups
- Development of ranking algorithms to identify drugs with high potential and low toxicity
- Big Data processing

Tools: Python, CellProfiler, Pandas, Numpy, Scipy, ScikitLearn, Apache Spark,

Created: 2023-11-02