

Wasim Aftab, PhD

Computational Biologist & Research Software Engineer

Munich, Germany

wasim.aftab@med.uni-muenchen.de

wasimaftab.de

linkedin.com/in/wasim-aftab

github.com/wasimaftab

PROFILE

Computational biologist with a PhD and more than 10 years of experience translating complex biomedical data into reproducible analyses and usable research software. Expertise spans multi-omics, proteomics, single-cell and spatial analysis, medical imaging, statistical methods, and scientific applications.

Biomedical data science

Multi-omics, single-cell, spatial biology, proteomics, medical imaging

Methods

Statistical inference, machine learning, deep learning, information retrieval

Research software

Python, R, Snakemake, Linux/HPC, Docker, FastAPI, Bioconductor

Professional Experience

Postdoctoral Researcher

MAY 2025–PRESENT

Institute for Stroke and Dementia Research, LMU Klinikum

- Develop computational methods integrating single-cell reference atlases, bulk proteomics, and multimodal MRI to study neurovascular disease.
- Created HASA, a hierarchy-aware method for attributing bulk proteomic signals to cell types with replicate-aware stability estimates; validated it in cerebral amyloid angiopathy data and packaged it for collaborators.
- Built SiderUNet, a reproducible 3D nnU-Net pipeline for cortical superficial siderosis segmentation, achieving an out-of-fold Dice score of 0.75.
- Designed an active-learning workflow that improved annotation Dice from 0.56 to 0.81 and reduced manual annotation by approximately three hours per case.

Postdoctoral Researcher / Bioinformatician

JAN 2022–APR 2025

Biomedical Center, LMU Munich

- Built reproducible Snakemake workflows for bulk RNA-seq, single-cell RNA-seq, and ATAC-seq, reducing processing time by approximately 50% and standardizing cross-project results.
- Partnered with chromatin-biology groups from study design and quality control through statistical analysis, visualization, and biological interpretation.
- Developed R pipelines for high-throughput proteomics, including batch correction, missing-value handling, exploratory analysis, and moderated inference.
- Created a biomedical question-answering application whose prompt-enhancement method improved information retrieval by 137.5%; the work was published in BMC Bioinformatics.

Research Assistant during PhD

SEP 2015–DEC 2021

Faculty of Medicine, LMU Munich

- Developed ImShot, open-source software integrating MALDI imaging mass spectrometry and shotgun proteomics for probabilistic in-situ protein identification.
- Designed statistical methods and R pipelines for proteomics and protein-complex discovery; built R, Shiny, and Electron tools used across more than ten collaborating laboratories.

Selected Contributions

HASA — cell-type attribution for bulk proteomics

Hierarchy-aware integration of bulk proteomic measurements with single-cell reference data, including replicate-aware stability estimates and pathway-level interpretation.

SiderUNet — multimodal MRI segmentation

Reproducible 3D deep-learning and active-learning workflows for cortical superficial siderosis, designed for GPU/HPC environments and efficient expert annotation.

ImShot — spatial proteomics software

Open-source R and Electron application for integrating MALDI imaging mass spectrometry with shotgun proteomics and visualizing probabilistic protein identifications.

Biomedical information retrieval

Python, FastAPI, and JavaScript application with a published keyword-frequency prompt enhancement method; subsequent prototypes explored Neo4j and GraphRAG for traceable evidence.

Selected Publications

Sequence and chemical specificity define the functional landscape of intrinsically disordered regions.

Nature Cell Biology, 2026. doi:10.1038/s41556-025-01867-8

Optimizing biomedical information retrieval with a keyword frequency-driven prompt enhancement strategy.

BMC Bioinformatics, 2024. doi:10.1186/s12859-024-05902-7

Improving SWATH-MS analysis by deep-learning.

Proteomics, 2023. doi:10.1002/pmic.202200179

ImShot: An Open-Source Software for Probabilistic Identification of Proteins In Situ and Visualization of Proteomics Data.

Molecular & Cellular Proteomics, 2022. doi:10.1016/j.mcpro.2022.100242

Molecular Connectivity of Mitochondrial Gene Expression and OXPHOS Biogenesis.

Molecular Cell, 2020. doi:10.1016/j.molcel.2020.07.024

Education

PhD, Bioinformatics

Computational Proteomics
LMU Munich · 2015–2021

MSc, Computer Engineering

King Abdulaziz University
2013–2015

BE, Computer Engineering

Visvesvaraya Technological University
2005–2009

Technical Toolkit

Programming

Python, R, MATLAB, C, JavaScript, shell

Data & methods

Bioconductor, PyTorch, nnU-Net, statistical inference, knowledge graphs

Engineering

Snakemake, Linux/HPC, Docker, FastAPI, Git, reproducible reporting

Applications

R Shiny, Electron, Neo4j, Nginx, scientific visualization

Earlier Experience

Senior Software Engineer

iCelero Technologies · 2012
Java, Selenium, Python, and Autolt test automation.

Scientific Programmer

JNCASR · 2009–2012
C/Fortran scientific computing, sparse eigensolvers, quantum chemistry, and protein-function prediction.

Languages

Professional

English, Hindi, Bengali

Working knowledge

German, Arabic