

# Vivek Appadurai, Copenhagen, DK 1799

appadurai.vivek@gmail.com | [Google Scholar](#) | GITHUB: [VAQM2](#) | Permanent resident in Denmark. | +45 28749660

## Professional Summary

I'm a bioinformatics professional with over 12 years of experience in pharmaceutical research, clinical science and academic settings, specializing in the development, modernization, and support of scientific software solutions and informatics processes. I possess proven expertise in Python, SQL, automating workflows, using high performance, cloud computing platforms, and contributing to AI/ML-enabled drug discovery initiatives. I'm adept at bridging scientific research needs with technical execution, driving data automation, and optimizing underlying databases. In addition, I have demonstrated strong problem-solving, collaboration and communication skills, learning agility, along with a commitment to delivering user-friendly, research-centric solutions within dynamic software ecosystems.

## Skills

**Programming & Scripting:** Python, R, Perl, SQL, Bash, Git, RESTful APIs, Docker.

**Software Development & DevOps:** Microservice Architecture, Azure pipelines for CI/CD, Agile Methodologies, Software System Maintenance & Enhancement, Technical Documentation.

**AI/ML:** Hypothesis Testing, Regression Modeling, Classification, Clustering, Imputation, Dimensionality Reduction, Conceptual & Logical Data Modeling, Ontologies.

**Data Management:** Relational Databases (Optimization & Restructuring), Data Registration, Metadata, Master and Reference Data Management, Data Lineage, Data Governance, Information Security, Data Complexity Management.

**Cloud & HPC:** AWS (S3, EC2), Sun Grid Engine, MOAB, LSF.

**Data Visualization:** Streamlit, Shiny, Dash, Tibco Spotfire.

**Domain Expertise:** Bioinformatics, RDKit, Omics Data Integration, Pharmaceutical Research & Development workflows.

**Languages:** English (Native speaker), Danish (Intermediate, passed PD3)

## Experience

### Senior Technical Data Steward, Novo Nordisk – Maaloev, Denmark

Sep 2023 – Present

- Led the development of a microservice application, leveraging natural language processing, to enhance experimental metadata capture during experimental data registration in research portfolio projects, significantly improving on feature sets delivered to AI engineers and contributing to AI/ML-enabled solutions for automating data registration within the drug discovery workflow.
- Architected and implemented conceptual and logical data models for Novo Nordisk's Compound Database (NNCD), a key initiative for developing a semantic layer and knowledge graphs, directly supporting restructuring and optimization of underlying databases for future AI/ML-enabled solutions.
- Authored architectural decision records as part of the NNCD modernization task force, translating knowledge on existing systems into documentation and actionable system improvements to expand the R&ED entity registration application landscape to accommodate novel drug modalities.
- Established standard operating procedures for onboarding, identity and access management, and robust data governance principles for omics datasets, ensuring information security and managing data complexity in external partnerships.
- Provided hands-on bioinformatics, data visualization, and machine learning support to research portfolio projects, enabling data-driven decision making and acting as a bridge between scientific research needs and technical execution.
- Mentored new hires and junior colleagues, fostering sustainable team growth while aligning to standard operating procedures that are conducive to maintaining high-quality delivery standards.

### Postdoctoral Researcher, Institute of Biological Psychiatry – Roskilde, Denmark

Apr 2021 – Sep 2023

- Designed and implemented a comprehensive bioinformatics data infrastructure for iPSYCH, an international consortium harmonizing genetics, metabolomics, socio-demographic, and clinical data to build a large biobank, establishing harmonized and streamlined end-to-end informatics processes for complex scientific data.
- Led a research study applying machine learning methods on genetic and transcriptomic data to identify novel genetic features, directly contributing to AI/ML-enabled solutions in psychiatric disease research.
- Engaged with clinical researchers, epidemiologists and bioinformaticians to identify critical pain points and

design user-friendly, research-centric solutions for data analysis and interpretation, demonstrating ability to translate research needs of diverse stakeholders into technical solutions and iterate based on feedback.

**Doctoral Intern in Bioinformatics, Lundbeck A/S – Valby, DK      Aug 2020 – Oct 2020**

- Provided real world evidence during the target discovery phase, by calculating the prevalence of loss-of-function mutations in potential targets for severe migraine, demonstrating direct experience within biopharma scientific software platforms.

**PhD Candidate, Institute of Biological Psychiatry – Roskilde, DK      Jan 2018 – Apr 2021**

- Spearheaded bioinformatics analyses for multiple research studies, employing regression modeling and causal inference methodologies, and empirically estimating the impact of missing data imputation algorithms on genetic risk scores, showcasing strong problem-solving and analytical skills.

**Lead Bioinformatician, Institute of Biological Psychiatry – Roskilde, DK      July 2016 – Dec 2018**

- Led bioinformatics efforts to identify and enhance the quality of large exome sequencing datasets, addressing data degradation and contributing to highly-cited studies, demonstrating expertise in data analysis and data quality improvement.

**Bioinformatician, Human Genome Sequencing Center – Houston, TX USA      Mar 2012 – Jun 2016**

- Developed software and automated workflows aimed at identifying candidate causal mutations in pediatric-onset mitochondrial diseases, integrating evidence from open source clinical variant repositories, and machine learning based variant effect prediction models to identify candidate causal mutations, which upon molecular follow-up led to a 26% increment in the resolution of unsolved cases.

**Scientific Programmer, The McDowell Genome Institute – St. Louis, USA      Jan 2010 – Dec 2011**

- Developed software to enhance the scaffolding quality of de novo genome assemblies delivered to the UCSC genome browser, demonstrating hands-on technical solutions development and contributions to open-source scientific data repositories.

**Awards**

---

**Lundbeck Foundation** – Postdoctoral Fellowship      **Apr 2022**

**Education**

---

**University of Copenhagen** – PhD in Biostatistics and Bioinformatics      **Apr 2021**

**Missouri University of Science and Technology** – MS in Electrical and Computer Engineering      **Dec 2009**

**Anna University** – BE in Electronics Engineering      **May 2006**

**Extracurricular**

---

I'm an avid runner (PRs 10k: 46:01, 5k: 22:30) and I enjoy reading neo-western, classical American, Russian literature when I'm not spending time with my soon to be one-year old daughter. Prior to becoming a dad, I volunteered with the coding pirates branch in Østerbro, teaching programming skills in Python and Javascript to teenagers.