

VISHAL SOOD

(+41) 79 942 0489 · vishal.chandra.sood@protonmail.com · Bussigny, CH-1030
Swiss Permit C · github.com/visood · linkedin.com/in/physicistvisha

SENIOR RESEARCH ENGINEER

Designing platforms that abstract computational complexity while preserving analytical power

A Senior Research Engineer with a PhD in Physics and extensive experience building robust, scalable computational tools that accelerate scientific discovery. Proven ability to translate complex research requirements—from neuroscience to genomics—into production-grade software platforms. Combines a first-principles approach to algorithm development with deep expertise in high-performance computing (HPC), scientific workflow automation, and principled data engineering. Passionate about architecting the foundational software that bridges the gap between complex data and scientific insight.

CORE EXPERTISE

Analytical Thinking | Research Design | Scientific Writing | Problem Solving | Project Management | Data Interpretation | Process Optimization | Critical Reasoning | Cross-Disciplinary Collaboration | Experimental Design | Knowledge Management | Technical Documentation | Workflow Development | Strategic Planning | Quality Assurance

TECHNICAL SKILLS

- Languages: Python (Expert), C++ (Expert), SQL, Bash, Julia, Haskell, Scala
- Scientific Computing: Algorithm Development, HPC (SLURM), Data Pipelines, NumPy, Pandas, HDF5, Dask, Performance Optimization, Statistical Modeling
- Software Architecture: Scientific Workflow Automation, API Design, Plugin-based Frameworks, Distributed Systems, Declarative & Compositional APIs
- Data Engineering & DevOps: Data Modelling, CI/CD (Jenkins, GitHub Actions), Docker, Build Systems (Waf, CMake), Database Systems (MariaDB, PostgreSQL, PostGIS), Data Provenance, FAIR
- Research & Validation: Principled Validation Frameworks, Reproducibility by Design, Statistical Analysis, Hypothesis Testing, Monte Carlo Methods.

PROFESSIONAL EXPERIENCE

SENIOR SCIENTIFIC SOFTWARE DEVELOPER 2017 – 2024

Blue Brain Project, EPFL

Architected and built the core computational frameworks that enabled large-scale analysis and validation of the world's most detailed brain circuit models, directly empowering a large team of neuroscientists.

- Translated complex neuroscientific inquiry into robust, automated computational workflows; architected configuration-driven systems (Python, HPC/Slurm) that allowed scientists to execute and manage petabyte-scale analyses reproducibly.
- Led the design and development of a novel, plugin-based scientific validation framework (Python), creating an extensible ecosystem for rigorously comparing computational models against diverse experimental data.
- Engineered user-centric Python APIs that provided an intuitive, high-level interface to complex, terabyte-scale hierarchical data stores (HDF5), significantly lowering the barrier for exploratory data analysis.

SENIOR PRODUCT DEVELOPER 2015 – 2017

Saphetor SA

Applied software engineering rigour to build a high-performance clinical genomics platform from the ground up, rapidly mastering a new scientific domain to deliver production-critical systems.

- Engineered the high-performance C++ backend for a real-time genomic variant annotation service, translating complex clinical and bioinformatics requirements into a scalable, low-latency system.
- Developed the end-to-end data processing pipeline for integrating and querying multiple large-scale genomic databases, creating the core knowledge base for variant interpretation.
- Designed and implemented a robust Python REST API to serve complex clinical queries, bridging the gap between the high-performance backend and user-facing applications.

DATA SCIENTIST 2015- 2015

Citiziv Sàrl, EPFL Innovation Park, Switzerland

MAÎTRE ASSISTANT 2011- 2014

Lab of Statistical Biophysics, EPFL, Switzerland

POST-DOCTORAL FELLOW 2008- 2010

Niels Bohr Institute, Copenhagen, Denmark

RESEARCH FELLOW 2006- 2008

Physics Dept., University of Calgary, AB, Canada

ACADEMIC FOUNDATION

Doctor of Philosophy in Physics, Boston University, Boston, MA, USA 2006

Thesis: Interacting Particle Systems on Graphs. Developed theoretical and computational models for stochastic processes, providing a deep, first-principles foundation for modelling complex systems and developing data-driven algorithms.

Bachelor of Technology in Engr. Physics, Indian Institute of Technology Bombay, India 2000

KEY RESEARCH ENGINEERING PROJECTS

My work has focused on building foundational platforms that operationalize research, transforming complex computational tasks into reproducible, scalable, and user centric workflows.

Automating Large-Scale Neuroscience Analysis on HPC

Architected a workflow engine (connsense-TAP) to manage petabyte-scale analyses of brain circuit models on distributed HPC systems.

- Scientific Enablement: Abstracted away the complexity of parallel execution (Slurm), allowing scientists to launch massive computational experiments via simple, declarative configuration files.
- Integrated Data System: Engineered a unified HDF5 data store with a lazy-loading Python API, transforming terabytes of results into a single, queryable scientific asset.

Architecting a Framework for Collaborative Scientific Validation

Conceived and built a novel software ecosystem (DMT) for rigorous, automated, and collaborative model-data comparison.

- Interoperable by Design: Used the Adapter Pattern and formal API contracts to create a plugin-based system where any community model could be tested against any experimental dataset.
- Automated Scientific Narrative: Engineered a system to programmatically generate rich validation reports, integrating text, data, and figures into a complete, reproducible scientific argument.

Architecting a Knowledge Portal for a Digital Brain Reconstruction

Architected the end-to-end data and analysis pipelines that power the SSCx Portal, a platform making one of the world's most detailed brain reconstructions accessible to the global neuroscience community.

- Data Processing Engine: Designed the core engine (connsense-TAP) to transform raw, petabyte-scale simulation outputs from HPC into a structured, queryable HDF5 knowledge base.

- Scientific Content Generation: Engineered the analysis framework (factology) to programmatically generate the portal's content hierarchical "factsheets" with integrated data and figures.

Building a Production-Grade Platform for Clinical Genomics

Engineered core components for a real-time genomic variant annotation and interpretation service (Saphetor).

- Research to Production: Translated complex bioinformatics requirements and clinical guidelines (ACMG) into a robust, high-performance C++ backend and Python REST API.
- Data Integration: Designed pipelines to ingest and integrate multiple large-scale genomic databases, creating a comprehensive knowledge base for clinical decision support.