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 QUANTITATIVE RESEARCH ENGINEER

Designing platforms that abstract computational complexity while preserving analytical power

A first-principles thinker with a PhD in Statistical Physics and over a decade of experience architecting high-performance computational ecosystems. Proven ability to translate the complex stochastic systems underlying financial derivatives into robust, low-latency C++ applications and scalable Python validation pipelines. Specialized in building foundational quantitative libraries and validation frameworks that ensure model accuracy, reproducibility, and scalability.

PROFESSIONAL GOAL

To leverage a deep background in statistical physics to model complex stochastic systems in finance, and to translate those models into high-performance C++ and Python ecosystems for systematic strategy research, validation, and execution.

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

- Programming: Python (Expert), C++ (11/14/17/20) (Expert), SQL, Bash, Julia, Haskell, Scala, Object Oriented and Functional, Metaprogramming

- Scientific Computing: Algorithm Development, HPC (SLURM), Data Pipelines, NumPy, Pandas, HDF5, Parquet, Dask, Joblib, Performance Optimization, Statistical Modeling, Low-Latency Design, Memory & I/O Optimization

- Quantitative Modeling: Stochastic Processes, Monte Carlo Methods, Time-Series Analysis, Numerical Methods & Optimization

- Software Architecture: Scientific Workflow Automation, API Design (REST, Declarative), Plugin-based Frameworks, Distributed Systems, Declarative &Compositional APIs, Quantitative Library Design, Reproducible & Auditable Pipelines, Validataoin & QA Frameworks

- Data Engineering & DevOps: Data Modelling, CI/CD (Jenkins, GitHub Actions), Docker, Build Systems (Waf, CMake), Database Systems (MariaDB, PostgreSQL, PostGIS), Data Provenance, FAIR, Singularity, Data Pipelines (ETL)

- 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 orchestrated the core computational ecosystem for one of the world's largest-scale neuroscience simulation initiatives.

- Designed and implemented massively parallel data processing workflows (Python, Dask, Joblib) on a 100+ node SLURM cluster to analyze multi-terabyte HDF5 datasets, mastering HPC I/O optimization, memory management, and parallelization principles.

- Spearheaded the design of a plugin-based validation framework (Python), enabling systematic, component-based testing of complex models against empirical data—a pattern directly applicable to validating trading strategies and risk components.
- Developed novel algorithms grounded in statistical physics and computational geometry to model complex, multi-scale stochastic systems.
- Built and maintained the flagship data construction and analysis pipelines, reducing manual setup for researchers by over 90% and eliminating entire classes of configuration errors through principled architectural design.
- Drove significant improvements in operational efficiency by automating complex pipeline orchestration and dependency management, accelerating the research-to-results cycle in a production-scale scientific environment.

SENIOR PRODUCT DEVELOPER 2015 − 2017

Saphetor SA

Proved domain-agnostic capabilities by rapidly mastering clinical genomics to engineer the high-performance C++ backend for a precision medicine platform.

- Engineered a high-performance C++ backend and Python REST API for a real-time, low-latency genomic data service, capable of annotating thousands of variants per second.
- Authored a high-performance computational parser in C++ for a complex, domain-specific data format (HGVS), demonstrating expertise in efficiently processing structured, high-throughput data streams analogous to financial protocols.
- Integrated multiple large-scale public and proprietary databases, designing algorithms for evidence classification and pattern detection in noisy, high-dimensional data.
- Implemented rigorous quality control and data validation mechanisms within the real-time pipeline, ensuring data integrity and system resilience for a mission-critical clinical application.
- Led the migration and modernization of the core C++ build system (Make to Waf), significantly improving dependency management, build times, and developer productivity.

DATA SCIENTIST 2015- 2015

Citiviz Sárl, 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

SPECIALIZED EXPERTISE

Quantitative Modeling & First-Principles Thinking

- Leverages a PhD in Statistical Physics to develop models of complex stochastic systems from the ground up—the same mathematical foundation used to describe asset dynamics and value derivatives.
- Implements numerical methods, including Monte Carlo and optimization algorithms, in C++ and Python to solve complex simulation and parameter estimation problems.

High-Performance C++ for Quantitative Systems

- Engineers and optimizes low-latency C++ applications for real-time data processing and complex computation,

sharing core architectural patterns with a valuation engine.

- Masters modern C++ (11/17/20), memory management, and I/O optimization to build robust, scalable, and numerically intensive systems.

Principled Validation & QA Architecture

- Architects formal validation frameworks for the rigorous, systematic benchmarking of competing predictive models against empirical data.
- Enforces reproducibility and auditability by design, creating systems where every result is traceable to its inputs and assumptions—a foundational principle for robust risk management.

Scalable Infrastructure for Systematic Simulation

- Designs automated, configuration-driven pipelines to orchestrate massively parallel simulations and parameter sweeps on large datasets.
- Engineers integrated data systems that transform terabyte-scale raw data into queryable, interactive assets for model development and analysis.

KEY ARCHITECTURAL EXPERIENCE

My work has focused on building domain-agnostic frameworks that operationalize quantitative research at scale. This philosophy is crystallized in three key architectural patterns:

1. A Framework for Principled Model Validation & Benchmarking Conceived and built a novel software ecosystem (**DMT**) for rigorous model-data comparison. Key features:

   - *Interoperable, Plugin-Based Core:* Used the Adapter Pattern to create a formal contract allowing any pricing model to be tested against any market dataset, ensuring systematic benchmarking of competing models.

   - *Automated Reporting Engine:* Engineered a "scientific narrative" engine that programmatically generates rich validation reports from pipeline results.

     This framework provides a direct blueprint for validating and comparing different pricing models and numerical methods within a quantitative library, ensuring accuracy and robustness.

2. High-Performance C++ for Low-Latency Systems Engineered the core C++ backend for a real-time, high-throughput data processing platform (**Saphetor**).

   - *Low-Latency Service Design:* Built a C++ backend and REST API for a service processing thousands of complex data transactions per second.

   - *Modern C++ Expertise:* Leveraged C++11/14/17 features and advanced build systems (Waf, CMake) to deliver robust, maintainable, and high-performance production code.

     This system shares core architectural patterns with the high-performance components of a financial valuation engine.

3. A Framework for Scalable, Reproducible Simulation Architected a workflow engine (**connsense-TAP**) to manage petabyte-scale analyses on HPC clusters.

   - *Decoupling Intent from Execution:* Rigorously separating a researcher's high-level goal (defined declaratively) from the complexity of parallel execution.

   - *Configuration as an Auditable Record:* Ensuring any result is perfectly reproducible—a critical requirement for backtesting, stress testing, and regulatory compliance.

     This architecture provides a model for building auditable, scalable backtesting and simulation platforms.

4. End-to-End Data and Analysis Pipelines for a Scientific Webportal 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

   - *Scientific Content Generation*: Engineered the analysis framework