C1

Subhashini Sivagnanam Advisor(s): Salvador Dura-Bernal Ph.D.

Co-author(s): William Lytton

Enhancing the Integrity and Reproducibility of Computational Neuroscience Models through Provenance Tracking and Similarity Analysis

The reproducibility and validation of computational neuroscience simulations remain challenging due to the complexity of maintaining comprehensive records of datasets, parameters, and execution processes. This poster describes a pipeline for tracking and verifying the provenance of computational neuroscience experiments through three specific aims.First, we develop an automated pipeline that captures and stores cryptographic hashes of datasets alongside their associated metadata in an immutable distributed ledger, enabling independent validation of research artifacts.The system automatically extracts execution parameters, generates integrity checksums, and creates verifiable records of computational experiments. Second, we demonstrate this provenance tracking system using a neuroscience connectivity analysis workflow. Our implementation captures standardized metadata at each stage of the analysis pipeline, including data pre-processing, signal coupling computation, and topological analysis. The system records critical experimental parameters such as epoch duration, coupling measures, and execution environment details while maintaining cryptographic verification of input and output files. This comprehensive tracking ensures that researchers can reproduce experiments with identical parameters and verify the integrity of their results. Third, we enhance the system with similarity detection capabilities using Locality Sensitive Hashing techniques, enabling researchers to identify related computational models based on their parameters and structures. This feature facilitates the discovery of similar experiments and promotes knowledge sharing within the research community. Our implementation leverages existing computational infrastructure while establishing a pipeline for verifiable research practices.This approach promotes transparent, verifiable computational neuroscience that enhances reproducibility, enables independent validation, and accelerates scientific progress.