Mahasweta Bhattacharya

180 Kennedy Dr | Malden, MA 02148

609-906-1583 | b.mahasweta24@gmail.com | linkedin.com/in/mahasweta-bhattacharya | github.com/mahaswetabhattacharya24

EDUCATION

University at Buffalo, State University of New York   Buffalo, NY
Doctor of Philosophy in Biomedical Engineering   *Aug. 2017 – Jan 2023*

University at Buffalo, State University of New York   Buffalo, NY
Master of Science in Electrical Engineering   *Aug. 2015 – May 2017*

West Bengal University of Technology   Kolkata, India
Bachelor of Technology in Electronics and Communication Engineering   *Aug. 2010 – May 2014*

EXPERIENCE

Senior Scientist   Sept. 2023 – Present
Sanofi   *Cambridge, MA*

- Led the design of an **agentic-AI pipeline orchestrating Claude-based agents** to autonomously aggregate and summarize multimodal biological evidence for target credentialing; delivered scalable plain-language evidence reports, reduced manual review overhead by >**50%**, and operationalized **LLM-driven reasoning in a regulated scientific workflow.**
- Designed a multi-modal foundation model integrating genetics, transcriptomics, and clinical embeddings; achieved $4\times$ **improvement in causal target recall** over genetics-only baselines and established a transferable representation space for cross-disease generalization.
- Led transcriptomic pharmacodynamics modeling to compare oral vs injectable therapies for Hidradenitis Suppurativa; identified **superior immune-pathway perturbation** for the oral candidate, enabling preclinical advancement and establishing a robust MoA modeling workflow.
- Developed a harmonized meta-analysis pipeline for public HS transcriptomes, producing a **mechanistic target-ranking framework** presented at FOCIS 2025 and forming the computational backbone for patient stratification and network modeling.
- Founded a graph-based bispecific discovery platform integrating synergy metrics, biological embeddings, and **LLM-guided evidence retrieval**; generated **5 novel bispecific target-pair candidates**.
- Built a scalable disease-mapping and indication-discovery engine scoring **232 immune indications in 3 weeks** and scaling to **17,000+ phenotypes**, enabling computational repurposing and whitespace identification.
- Developed an explainable AI-driven target-discovery engine generating **90+ hypotheses** and advancing **7 novel targets** into preclinical evaluation; incorporated causal scoring, embedding similarity, and LLM-augmented evidence synthesis.
- Co-led an automated multimodal target-credentialing platform supporting **30+ therapeutic programs** and enabling **3 preclinical nominations**; introduced modules for causal inference, uncertainty quantification, and prospective validation.

Postdoctoral Fellow, Radiation Oncology   Jan 2023 - Sept 2023
Johns Hopkins University School of Medicine   *Baltimore, MD*

- Developed a GPU-accelerated double-Gaussian proton dose engine as a Python package; achieved **0.28s** mean patient-plan computation vs **4.68s** Monte Carlo with **96% 3D-gamma agreement (2%/2mm)**, demonstrating physics-informed acceleration at ML-scale throughput.
- Architected an end-to-end beam-modeling and validation pipeline from **98 pristine Bragg-peak energies**, benchmarking against measurements, heterogeneous digital phantoms, and multi-site patient plans; built uncertainty profiles highlighting limits in highly heterogeneous regions.
- Advanced a **deep reinforcement learning framework** for VMAT machine-parameter optimization, producing deliverable prostate plans in **3.3s** and automated TPS refinement in **77s**; on a 15-patient external cohort, RL+TPS reduced mean rectum dose vs clinical plans (**17.4 vs 21.0 Gy**, p=0.024) while maintaining target coverage.
- Integrated RL-generated optimization with a commercial treatment-planning system to create a **hybrid automated + human-in-the-loop pipeline**, operationalizing ML methods into clinical workflows and demonstrating production-grade feasibility.

Research Assistant   2017 – 2023
State University of New York at Buffalo   *Buffalo, NY*

- Analyzed longitudinal two-photon calcium imaging to reconstruct session-wise functional connectomes during motor-skill learning; identified a **biphasic rewiring trajectory** (early expansion then pruning) and discovered stable L2/3 hub assemblies encoding movement.
- Integrated portable fNIRS and EEG with electric-field–informed GLMs to model cerebellar tDCS response in stroke; uncovered **0.07–0.13 Hz hemoglobin–EEG coupling signatures** predictive of responders and interpretable non-response patterns.
- Developed the **FARCI** MATLAB toolbox for fast connectome inference using OASIS spike deconvolution and partial correlations; outperformed existing spike-inference algorithms on NCC and NAOMi benchmarks in accuracy and scalability.
- Demonstrated that low-frequency coupling between frontal HbO (0.07–0.13 Hz) and EEG (1–12 Hz) tracks consciousness in acute brain injury; AMICA embeddings + k-NN achieved >**90% accuracy** distinguishing conscious vs unresponsive patients and >**99% accuracy** predicting failure to recover consciousness.

- Built a multi-sensing cerebral organoid platform integrating Vis–NIR spectroscopy (mitochondrial CCO) with MEA-derived spectral exponents; quantified maturation-linked decreases in **30–50 Hz slope** and CCO activity, enabling functional screening of metabolic–electrophysiological coupling.
- Modeled photothermal vs photobiomodulation mechanisms using Monte Carlo photon transport + bioheat modeling at 630/700/810 nm; predicted <**0.25°C scalp** and <**0.04°C cortical** heating at 810 nm, supporting safety and highlighting CCO-mediated photobiomodulation as the primary mechanism.

SKILLS

- **Machine Learning:** Supervised/unsupervised learning, representation learning, causal inference, graph neural networks, multimodal learning, VAEs, uncertainty quantification.
- **LLMs & Agentic AI:** LLM fine-tuning (LoRA/PEFT), RAG pipelines, multi-agent orchestration (Claude/GPT), structured generation, scientific summarization, grounding and consistency checking.
- **Applied ML:** End-to-end ML pipelines, feature engineering, embedding similarity search, statistical modeling, Bayesian inference, high-dimensional data analysis.
- **AI Systems:** Python (PyTorch, TensorFlow, JAX), CUDA/GPU computing, Docker, HPC (Slurm).
- **Data Engineering:** ETL pipelines, data harmonization/QC, SQL/Snowflake, graph databases (Neo4j), S3 data lakes.