

Sagun Maharjan  
Phone: (617) 233-4076  
Email: [sagunmaharjann@gmail.com](mailto:sagunmaharjann@gmail.com)  
Github: <https://github.com/sagun98>

## SUMMARY

Lead Software Engineer with **9+ years** of experience building **full-stack platforms, research tools, and scalable infrastructure** for technical and scientific users. Currently developing research and data systems at **Harvard T.H. Chan School of Public Health and Broad Institute**. M.S. in Computer Science (Interactive Intelligence), Georgia Tech, focused on human-computer interaction and human-centered design, combining strong engineering skills with empathy for users and responsible AI systems.

## EDUCATION

MS in Computer Science (Specialized in Interactive Intelligence and Machine Learning)  
Georgia Institute of Technology

Bachelor of Science in Computer Science  
University of Louisiana Monroe

## PROFESSIONAL EXPERIENCE:

Harvard T.H. Chan School of Public Health,  
Boston, MA Oct 2019 – Present

Broad Institute of MIT and Harvard, Cambridge,  
MA Oct 2019 – Present

## Lead Software Engineer and Research Associate

- Led development, maintenance, and releases of **state-of-the-art microbial community profiling tools** within the bioBakery suite, collaborat-

ing closely with **biologists, postdocs, and PhD researchers** to build scientific software and experimentation platforms.

- Distributed tools via **GitHub, PyPI, Conda, and Docker**, implementing stable packaging, automated CI/CD pipelines, versioning, and cross-platform compatibility for global research communities.
- Automated deployment of large-scale workshop VM images and instances supporting high-throughput research pipelines for **~200+ concurrent users** using Python and Docker on Guacamole in **Google Cloud**, enabling scalable hands-on training for **global workshops** where students and researchers participate in paid computational biology training from around the world.
- Designed and developed **internal infrastructure dashboards** for the Huttenhower Lab, including **storage monitoring (500TB+ across scratch, compute, permanent, and archival storage)** and **Slurm job visualization based on priority queues**, enabling researchers to track resource usage, monitor job status, and manage large-scale microbiome analysis workloads efficiently.
- Designed and developed a large-scale **Research and Data Deposition portal** using Python, GraphQL, and React deployed on **Google Cloud**, enabling researchers to **securely access open and controlled microbiome datasets at scale** for downstream analyses and collaborative scientific research.
- Actively maintaining all compute infrastructure for the Huttenhower Lab, including dedicated HPC blades, shared cluster partitions, and large-scale storage resources provided through Harvard Research Computing.
- Actively maintaining a ~50TB public access data portal (IBDMDB) that will provide an integrated resource for analyzing the gut microbial ecosystem, improving our ability to understand, diagnose, and treat IBD.
- Deployed the web-based platform Galaxy Project (version2) in docker container (on-prem VM) for computational biology and biomedical research to analyze, store, visualize, and share their genomic and proteomic data, as well as to run and visualize various bioinformatics tools and workflows.
- Migrated the existing Lab's google groups to the modern Discourse forum increasing visibility across the bioBakery user community improving response time for the user issues.
- Redesigned the lab's website using a modern CMS framework in a dockerized container adding dynamic features and robust UI.
- Migrated and currently maintaining the bioBakery repository from bitbucket/mercurial to Github, updating projects to standard template categories making it easier for lab members to implement tools, reproducible research workflows to build workflows for paper results, figures.

## Pearson PLC Co., Morrisville, NC Jun 2018 – Oct 2019

### Software Developer

- Actively participated in the development of **Pearson Developer Portal** (one-stop and shop for all internal/external APIs) implementing new features using Java (Spring boot) and Angular 6 deployed using docker containers in AWS.
- Engaged in story grooming sessions, **bug hunts** and **patches** using **JIRA** following the **Agile** development methodology.
- Expert in **TTD**(Test Driven Development) using **Junit** and **Mockito**. Additionally, automated e2e (End to End) testing of the developer portal using **Protractor**. Load tested the developer portal with 10000+ concurrent users sending an average of 30 requests/sec using Java (**Jmeter**) and **EC2** instance.
- Supported **APIGEE** API gateway, designing/developing proxies, set standard for best practices with other teams, monitoring APIs and supporting the production deployments.
- Configured two-way mutual **TLS** between internal/external clients and APIGEE generating CA and client certificates
- Automated the health check alerts in encrypted mongo servers, bastion servers, and ECS containers, using **AWS Lambda** functions with SNS email notification, additionally giving real-time insights on **cloudwatch** visualization dashboard
- Documented all the APIs of the IES (Identity Enabling Service) team in **OpenAPI 3** specification in Developer Portal deprecating their documentation from one-confluence.
- Implemented Micro Services based **Cloud Architecture** utilizing **Spring Boot** and REST Web Services
- Optimized the **Elastic search** queries giving relevant search results while searching for **APIs**, products, **forums**, and documentation in the developer portal
- Used Web services like RESTful for transmission of large blocks of XML/JSON.
- Developed the real-time dashboard integrated with using **Oracle APEX** and **D3.js** libraries generating the reports for supply chain, release management, ServiceNow and **Oracle EBS** systems.

- Deployed real-time **visualization** portal for global visibility into Key Performance Indicators (KPIs) that describes the health of operations support, risks, upcoming major events like key project releases, disaster recovery trials, environment management strategy enabling the stakeholders to tackle issues before the customers raise a ticket.
- Optimized **PL/SQL** triggers and procedures for the data visualization and calculations

## Open Learning Exchange, Cambridge, MA Jun 2017 – Jun 2018

### Software Engineer

- Collaborated with the team of 50+ members using **Git**, **GitHub**, **Maven** and **Gitter** improving the open source e-learning software (Bell Apps).
- Integrated the rating system in the dashboard with **Cassandra** and **Angular** giving scalability and high availability without compromising performance.
- Wrote automated **selenium** test cases in **Python3** troubleshooting bugs in the UI components of the Bell Apps.
- Created unit tests using **JUnit** framework in test driven development approach and developed unit and integration test cases using **JUnit** and **Mockito**.
- Tested Raspberry-Pi servers with **Vagrant** and accurately predicted the threshold traffic load of the servers.
- 

**Reviewed codes of other developers, created issues, pull requests, and branch merges in GitHub.**

## University of Louisiana at Monroe, Monroe, LA Jan 2016 – Dec 2016

### Web Developer

- Built and deployed **Docker** containers to break up monolithic app into **microservices** using **spring boot**, improving developer workflow, increasing

scalability, and optimizing speed.

- Updated the layout of the university's homepage and other subpages to support full mobile responsiveness adding session-based cookies for better user experience.
- Sanitized personnel directory of employees and students of using **Python3**, **MySQL**, and **Selenium** module.
- Involved in various Requirements and Design of the application and active involvement in many aspects of the software development lifecycle adopting **Agile** Software Development Methodology.
- Involved in designing **UML** case diagrams, class diagrams, and activity and sequence diagrams.
- Participated in semi-weekly status updates with management to discuss progress and concerns.
- Developed Front-End using HTML, **JSP**, and **JavaScript** and CSS technologies.
- Work closely with Business users/product customers to understand the business requirements.
- Involved in various Software Development Life Cycle (**SDLC**) phases of the project.
- Extensively used **Hibernate** in data access layer to access and update information in the database.
- Designed various tables required for the project in **SQL** Server database and used stored procedures and triggers in the application.

#### **Startups/Side Projects**

- **ChyaKharcha ([www.chyakharcha.com](http://www.chyakharcha.com))**: Founded and developed **Nepal's first global fundraising platform**, connecting supporters worldwide with Nepali creators and community initiatives, enabling transparent cross-border fundraising and digital payments.
- **chatFoody ([www.chatfoody.com](http://www.chatfoody.com))**: AI-powered restaurant ordering platform; B2B SaaS, voice/chat agent for Nepal restaurants
- **Bzooli ([www.bzooli.com](http://www.bzooli.com))**: Developing a **peer-to-peer marketplace for local solar energy credits**, allowing homeowners to buy and sell renewable energy credits while promoting adoption of distributed solar energy.

## Publications

- Google Scholar Link [https://scholar.google.com/citations?view\\_op=list\\_works&hl=en&user=MfoutTMAAAAJ](https://scholar.google.com/citations?view_op=list_works&hl=en&user=MfoutTMAAAAJ)
- Enhanced multi-omic viral profiling from microbial community sequencing with BAQLaVa  
<https://www.biorxiv.org/content/biorxiv/early/2026/02/12/2026.02.11.705346.full.pdf>
- Discovering and exploring the hidden diversity of human gut viruses using highly enriched virome samples  
<https://www.biorxiv.org/content/10.1101/2024.02.19.580813v1>
- Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery3  
<https://elifesciences.org/articles/65088>
- Comprehensive profile of the companion animal gut microbiome integrating reference-based and reference-free methods  
<https://academic.oup.com/ismej/article/18/1/wrae201/7819830>
- MaAsLin 3: Refining and extending generalized multivariable linear models for meta-omic association discovery  
<https://www.biorxiv.org/content/10.1101/2024.12.13.628459v1.abstract>
- A statistical model for describing and simulating microbial community profiles  
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8491899/>

## Technical skills:

- **Programming/Frameworks:** Java, Python, Javascript, J2EE, Spring MVC, Spring Boot, Hibernate, Angular, Junit, JDBC, D3.js, Django, Express-js, Bootstrap, Node.js, FastAPI, Flask, React, Next.js
- **Systems/Infrastructure/ DevOps: Linux Administration** (RHEL/CentOS, Ubuntu), macOS & Windows systems,
- **HPC infrastructure** (compute blades, shared cluster partitions, distributed storage),

- **Virtualization/Containerization/CI/CD** (Docker, Podman, Terraform, Packer, GitHub Actions, Jenkins),
- **Cloud-Infrastructure** (AWS EC2,AWS VPC,AWS S3,AWS IAM,GCP Compute),
- **Automation**(Bash, POSIX, Shell, Python),
- **MISC**(Github, PyPi, Docker Hub, Conda, Guacamole, Nginx )
- **Security/Authentication:** TLS, SAML, OAuth2
- **Data/Storage:** NFS, RAID systems, Ceph, Object storage (S3-compatible), Supabase, SQL/NoSQL

## Honors and Certifications:

- Adobe Certified Expert (ACE)
- University of Louisiana Monroe Emerging Scholar Award
  - Standalone computational physics research project