SANCHIT SAHAY

New York University, New York, NY ◇ ss19723@nyu.edu ◇ stupendoussuperpowers.github.io

RESEARCH INTERESTS

Software Supply Chain Security, Operating Systems, Virtualization, File Systems

EDUCATION

New York University, Tandon School of Engineering, New York, NY — 2024-2026
Master of Science, Computer Science — GPA: 3.889/4.0

Manipal Institute of Technology, Manipal, India — 2018-2022
Bachelor of Technology - Information Technology — Cumulative GPA: 9.10/10
Minor: Big Data Analysis

RESEARCH AND WORK EXPERIENCE

Secure Systems Lab, NYU, New York, NY — June 2025 – Present
Researcher, Advisor: Prof. Justin Cappos — *Rust, Go, C, eBPF, Linux Kernel & FS*

SBOMit (OpenSSF) — https://github.com/SBOMit

- Contributed to SBOMit: an OpenSSF initative to augment Software Bill of Materials (SBOMs) with in-toto attestations.
- Extended in-toto with eBPF-based network attestations to record package sources and detect suspicious network calls during builds, exposing attack vectors missed by manifest-based SBOM tools.
- Integrated attestations into Syft to demonstrate blindspots in existing SBOM generators.
- Presented implementation as a talk at **KubeCon + CloudNativeCon Atlanta** (Nov 2025)

Lind-Wasm (Secure Systems Lab) — github.com/Lind-Project/lind-wasm

- Contributed to Lind, a WebAssembly-based sandbox that securely executes POSIX-applications through a minimal kernel microvisor.
- Designed APIs for writing syscall monitors that intercept and override syscalls in sandboxed processes
- Built the in-memory filesystem for TriSeal, enabling trusted I/O for secure C compilation on Intel SGX enclaves.

Commvault Systems, Bangalore, India — January 2022 – August 2024
Engineer, Virtual Server Agent Team — *Python, .NET, VMware, Huawei Cloud*

- Built data-protection software for private-cloud deployments running on VMware and Huawei hypervisors, supporting environments used by 500+ enterprises and government organizations.
- Extended platform coverage for new vendor features and expanded automation across internal test suites.
- Refactored Commvault's vCloud Python SDK to significantly improve reliability of end-to-end testing.
- Developed Commvault's VMware Cloud Director plugin, streamlining data-protection workflows and reducing operational overhead for customers.

LegalAI, Bangalore, India (Remote) — April 2021 – December 2021
Full-Stack & DevOps Intern — *Node.js, React, Google Cloud Platform*

- Built an end-to-end claims-processing system composed of multiple GCP-hosted microservices and React portals for collecting claim details from clients and reviewing the generated legal drafts.
- Built the CI/CD pipeline that included a local App Engine–like runtime to emulate GCP behavior, enabling consistency between local and cloud machines.

OPEN SOURCE AND COURSE PROJECTS

HFS+ Port For FreeBSD: C FreeBSD Status Report

- Ported Apple's open-source HFS+ implementation to FreeBSD 14 by adapting its VFS-layer operations to modern FreeBSD interfaces
- Developed userland tools for mounting and management of HFS+ volumes.

Cargo: Rust github.com/stupendoussuperpowers/cargo

- Open source contributor to Cargo, Rust's package manager.

Improving Learned Bloom Filters github.com/stupendoussuperpowers/wise-bloom-filters

- Compared and benchmarked techniques to improve Learned Bloom Filters through Projection Hashing, Caching, and Low-Rank Approximation (LoRA).

Talk2Data: Python, Google Cloud Platform (GCP) github.com/Sitanshuk/Talk2Doc

- Developed a centralized AI platform designed to assist college students.
- Built scalable pipelines to extract and organize emails and Notion data for job applications, course materials and upcoming deadlines.
- Leveraged RAG with personalized LLMs to power queryable tables and chatbots.
- Designed load-efficient mechanisms for hosting the platform on Google Cloud Platform (GCP).

MTA Ridership Prediction: Python github.com/stupendoussuperpowers/mta-ridership

- Used machine learning models to predict NYC subway ridership from temporal and fare-class features
- Applied K-Shape clustering to analyze neighborhood-level ridership patterns.

RELEVANT COURSEWORK

New York University
Cloud Computing, Software Supply Chain Security, Programming Languages, Algorithmic Machine Learning and Data Science, Computer Networking

Manipal Institute of Technology
Operating Systems, Database Management Systems, Distributed Systems, Software Reliability, Cloud Computing

TECHNICAL SKILLS

Languages: Rust, Python, Node.js, C/C++, Java, C#, GoLang, Kotlin
Frameworks: eBPF, React, Next.js, .NET, Android SDK
Databases: MongoDB, Postgres, MSSQL
Cloud: Google Cloud Platform, Amazon Web Services, VMware vCenter and Cloud Director, Huawei FusionCompute
ML/Big Data: Pandas, Numpy, Keras, Spark, Hadoop, neo4j

POSITIONS OF RESPONSIBLITY

Project Head, IECSE, *Official Computer Science club, Manipal Insitute of Technology* 2020-2021

Teaching Mentor, *Problem Solving Using Computers, Manipal Institute of Technology* 2019-2020

President, LDQ, *Literary, Debate and Quizzing club, Manipal Institute of Technology* 2020-2021