

SOHOM DATTA

W MediaWiki Gerrit •  GitHub •  LinkedIn

EDUCATION

Manipal Institute of Technology, Manipal

July 2019 – Present

B.Tech in Computer Science and Engineering

ACHIEVEMENTS

- Awarded 3133.7 USD for discovering authentication bypasses in the dart:core URI parsing module in Dart-lang by the Google Vulnerability Rewards Program in 2022. (CVE-2022-3095, GHSA-m9pm-2598-57rj)
- Awarded 3133.7 USD for discovering a URL validation bypass in Google's Clojure library by the Google Vulnerability Rewards Program in 2021.
- Found and reported security/privacy issues in Google Chrome and Firefox's implementation of the ResourceTiming API . (CVE-2022-1146, CVE-2022-29915)
- Found and reported a high severity Denial-of-service attack against the popular jpeg-js javascript library to snyk.io. (CVE-2022-25851, SNYK-JS-JPEGJS-2859218)

TECHNICAL EXPERIENCE

North Carolina State University

February 15 2023 - (Ongoing)

Research intern - Wolfpack Privacy and Security Lab

Implemented large scale architectural changes to the VisibleV8 crawling automation pipeline that were able to provide significant performance gains in terms of crawling speed

Google Summer of Code

June 2022 – November 2022

Student Software Developer, Chromium

- Worked on aligning Chromium's implementation of the Performance API with the W3C specifications.
- Discovered and fixed bugs related to incorrect First Contentful Paint and Largest Contentful Paint entries in the process of refactoring the way the PaintTiming API marked contentful images in Chromium.
- Collaborated with the W3C Web Performance Working Group to rectify issues in the way LCP timing entries were being reported at the Painting layer.
- Discovered and helped patch systemic security issues in the implementation of the ResourceTiming API in the major browser engines, Webkit, Blink and Gecko (Firefox). (CVE-2023-1232)

Manipal Institute of Technology

December 2021 – Present

Research Assistant

- Conducted research into inter-process communication mechanisms and their application as a data-exfiltration vector in cross-process Spectre attacks under Professor Nisha P Shetty.
- Worked on building a novel obfuscation technique aimed at reducing the efficacy of linkage attacks against ego-centric networks.

Google Summer of Code

March 2021 – Sep 2021

Mentor for Wikimedia Foundation org

- Guided and helped student developers build and integrate an image zooming and panning interface into the Wikisource Page Editor.
- Assisted in writing automated tests to detect accidental/malicious changes in minified javascript blobs imported as part of dependencies.

Student developer for Wikimedia

March 2020 – Sep 2020

- Developed a software feature that made it easier for volunteers to work with "pagelist"s, a custom syntax used to store page number information for multi-page files.
- Developed heuristics that allowed users to create and edit "pagelist" without interacting with custom XML tag-based syntax.

Wikimedia Foundation

Oct 2019– Present

Volunteer developer

- Authored over 80+ major patches across various projects under Wikimedia.
- Helped in maintaining the infrastructure for the ProofreadPage extension, which adds proofreading capabilities to MediaWiki and is deployed in over 70 production wikis managed by Wikimedia.

LEADERSHIP POSITIONS

Cryptonite Manipal (cybersecurity student project)

April 2021 – August 2022

Subsystem Head

- Led a cybersecurity team of 22 engineering students and participated in Capture The Flag (CTF) competitions.
- The team placed among the top 15 in India in CSAW CTF '21 hosted by NYU (one of the oldest capture the flag competitions for academic teams), secured 2nd place in India in ASIS CTF Finals 2021 and was ranked among the top 12 teams in India on CTFTime in 2021-2022.
- Created challenges and took an active role in managing the security infrastructure for niteCTF, an international cybersecurity capture the flag event that attracted over 1200+ participants from 43 countries.

Wikimedia Foundation

October 2021 – Present

Maintainer, ProofreadPage

- Improved developer documentation related to the extension, providing guides and detailed walkthroughs regarding its setup and use to help ease the onboarding process for new developers.
- Introduced Selenium integration tests across the ProofreadPage codebase to better validate critical frontend changes.
- Took initiative in overhauling the preload and caching mechanism provided by ProofreadPage to decrease load times for editor-facing components.
- Built EditInSequence, a community requested feature that allows users to edit multiple pages via a fast and easy to use interface.

Association of Computing Machinery, Manipal

August 2021 – August 2022

Web-development Head

- Led and mentored a team of 10 students who were actively involved in building and maintaining websites for events conducted by the club.
- Took a lead in developing and maintaining the infrastructure for Scavenger Hunt, a inter-college event with over 700+ participants.
- Conducted workshops and one-on-one sessions to bring the freshers up to speed with modern web-development standards.

Research Society Manipal

August 2021 - August 2022

Senior Cybersecurity Mentor

- Conducted workshops on control flow integrity, binary exploitation and format strings exploits detailing the state-of-the-art research in the area, including mitigation techniques such as ASLR, stack cookies and fuzzy testing.
- Mentored new recruits and provided resources and guidance on getting started on the basics of cybersecurity.

PROJECTS

- **Fuzzing sudo** - Performed a fuzzing campaign on sudo, and found latent use-after-frees, out-of-bound reads and integer overflows. (sudo-project/sudo Issue#198, PR#196, PR#218, PR#227)
- **Pagelist Widget (GSoC project)** - Developed a Javascript (ES5) based custom widget that allowed users to interact with a representation of the "pagelist syntax" (a syntax based on XML tags) graphically. Additionally, implemented heuristic logic that translated user inputs back into the "pagelist syntax".
- **Dijkstra's pathfinding algorithm in Python** - Developed a distributed system of nodes that dynamically computed the best path inside a maze using the publisher-subscriber model provided by the Robotic Operating System (ROS)
- **Software sandbox using seccomp-bpf** - Developed a toy process sandbox using the kernel seccomp API that enabled users to selectively allow and deny specific sequences of syscall usages that were considered malicious according to a set of heuristics rules.
- **Dis, a stack based language** - Developed a compiler in Typescript for a toy stack-based language that used polish notation to perform arithmetic operations and compiled into x86 assembly.

TECHNICAL STRENGTHS

Programming Languages	Javascript, C, C++, Python, HTML, CSS
Other Software/Frameworks	jQuery, NodeJS, Express, React, Make, Tensorflow, IDA, Ghidra, Blink