

AMAN SHARMA

Email ◇ GitHub

Twitter ◇ LinkedIn ◇ Portfolio

EDUCATION

KTH Royal Institute of Technology, Stockholm, Sweden

February 2023 - January 2028

Doctoral Degree

Computer Science

Indian Institute of Technology Roorkee (IITR)

July 2017 - June 2021

Bachelor in Technology (B.Tech.)

Metallurgical and Materials Engineering

Amity International School, Noida, India

April 2015 - March 2017

Senior Secondary

Central Board of Secondary Education (CBSE)

PUBLICATIONS

- [1] A. Sharma, B. Baudry, and M. Monperrus, ‘Canonicalization for Unreproducible Builds in Java’, Apr. 30, 2025, arXiv: arXiv:2504.21679. doi: 10.48550/arXiv.2504.21679.
- [2] F. Bono, F. Reyes, A. Sharma, B. Baudry, and M. Monperrus, ‘Java-Class-Hijack: Software Supply Chain Attack for Java based on Maven Dependency Resolution and Java Classloading’, Jul. 26, 2024, arXiv: arXiv:2407.18760. Accessed: Jul. 29, 2024. [Online]. Available: <http://arxiv.org/abs/2407.18760>
- [3] A. Sharma, M. Wittlinger, B. Baudry, and M. Monperrus, ‘SBOM.EXE: Countering Dynamic Code Injection based on Software Bill of Materials in Java’, Jun. 28, 2024, arXiv: arXiv:2407.00246. doi: 10.48550/arXiv.2407.00246.
- [4] M. Balliu et al., ‘Challenges of Producing Software Bill Of Materials for Java’, IEEE Secur. Privacy, vol. 21, no. 6, pp. 12–23, Nov. 2023, doi: 10.1109/MSEC.2023.3302956.
- [5] K. Etemadi, A. Sharma, F. Madeiral, and M. Monperrus, ‘Augmenting Diffs With Runtime Information’, IEEE Transactions on Software Engineering, pp. 1–20, 2023, doi: 10.1109/TSE.2023.3324258.

PROFESSIONAL EXPERIENCE

Google, New York City, USA

August 2025 - Present

Student Researcher

Working as an intern on OSS-Rebuild project.

KTH Royal Institute of Technology, Stockholm, Sweden

November 2021 - January 2023

Research Engineer

Worked full-time as research engineer in the EECS school. My role is to support the research work in ASSERT, lead by Dr Martin Monperrus and Benoit Baudry, by building and maintaining scientific tools.

Google Summer of Code 2020

June 2020 - August 2020

Accord Project

Worked as a remote intern for Accord Project to develop an **MS Word Add-in** which enabled lawyers to draft **Smart Legal Contracts** directly on the Word document itself. To preserve the grammar (text structure) of the smart contract, I also wrote a parser to convert the CiceroMark (JSON) to OOXML (XML based file to represent Word documents) and vice-versa which automated checking of

consistency of the draft with the original contract template.

[Source Code]

Iterative.ai

October 2019 - November 2019

Data Version Control

Contributed to Data Version Control (DVC) - a command line tool used to **manage and version huge data science projects**. I implemented **dvc version** command from scratch with the corresponding test suites. Apart from this feature, I solved a number of issues to make the user experience smooth and the application robust. In addition to this tool, I also contributed to the documentation of this application by keeping it up to date with the additions in DVC and introducing features to make the documentation more interactive and user-friendly.

[Source Code]

Google Summer of Code 2019

May 2019 - August 2019

The Vega Project at the University of Washington

Developed a back-end service for **Vega Editor** - a tool which creates data visualizations. This service's purpose is to allow users to publish, store and, share visualizations directly from the editor. Apart from the service, I added many enhancements to the editor to make it feature-rich and robust.

[Source Code]

PROJECTS

by-the-pool

Goal to identify causes of unreproducible builds in Java artifacts and suggest ways to mitigate them.

Paper: <https://arxiv.org/abs/2504.21679>

Dataset: <https://github.com/chains-project/reproducible-central>

sbom.exe

Tool with a goal to provide runtime integrity for Java applications.

Paper: <https://arxiv.org/abs/2407.00246>

Tool: <https://github.com/chains-project/sbom.exe>

collector-sahab

It augments the static line-based diff with runtime information about variable value changes in two executions of a Java program.

Paper: <https://arxiv.org/abs/2212.11077v2>

Tool: <https://github.com/ASSERT-KTH/collector-sahab/>

sorald

It automatically repairs static analysis warnings reported by SonarQube.

Paper: <https://arxiv.org/abs/2103.12033>

Tool: <https://github.com/ASSERT-KTH/sorald/>

sbom-2023

A dataset of 156 SBOMs that were studied for the quality of dependency list they produce.

Paper: <https://arxiv.org/abs/2303.11102>

Tool: <https://github.com/chains-project/SBOM-2023>

Content Management System

Information Management Group

CMS is a multi-level role based application that allows specific users in the institute to manage the IITR's official website. It publishes digital content on the website using a WYSIWYG interface. Thus, organizing the content on the web page without writing any code.

[[IITR Website](#)]

IMG Website

Information Management Group

Developed the IMG website which is a portfolio for IMG, IITR which displays information about the IMG members, the blogs they have written, and the projects they have worked on or contributed to.

[[Website](#)]

POSITIONS OF RESPONSIBILITY

Google Summer of Code 2021 Mentor

May 2021 - Present

Google Summer of Code

I am responsible for mentoring Kushal Kumar on Cicero Word add-in - the same project which I worked on during Google Summer of Code 2020. We collaborate to introduce features for transformations between OOXML and CiceroMark (JSON).

Chief Technical Coordinator

April 2020 - May 2021

Information Management Group (IMG)

IMG is a group of students that constantly strive to lead in the innovation, development, and maintenance of advanced information technologies including computer systems, software, and database systems at IITR. As the chief technical co-ordinator, I was responsible for **maintaining the technological infrastructure** of applications and making sure all services are up and running at all times, especially the IITR Website. I also conducted the interviews for selection of a new team and eventually mentored them by **organising regular lectures** and **reviewing their GitHub pull requests**.

COURSES

Mathematics - I [[Syllabus](#)]

Computer Programming [[Syllabus](#)]

Mathematics - II [[Syllabus](#)]

Basic Calculus 1 and 2 [[Syllabus](#)][[Certificate](#)]

Design & Analysis of Algorithms [[Syllabus](#)][[Certificate](#)]

Discrete Mathematics [[Syllabus](#)][[Certificate](#)]

Database Management System [[Syllabus](#)][[Certificate](#)]

Introduction to Automata, Languages and Computation [[Syllabus](#)][[Certificate](#)]

Computational Complexity [[Syllabus](#)][[Certificate](#)]