



Friedemann Lipphardt, PhD

Computer Science
Researcher




 friedemann.lipphardt
@gmail.com

 Tokyo, Japan

 <https://www.linkedin.com/in/friedemann-lipphardt-547759228/>

 <https://github.com/Freddi43>

 <https://orcid.org/0009-0006-9412-6844>

PROFILE

Multidisciplinary PhD graduate from the Max Planck Institute for Informatics with four years of university-level teaching experience in computer networks, distributed systems, and AI. I have taught and developed course materials for core CS curricula including Data Networks, Distributed Systems, and advanced AI safety seminars, while supervising student theses and coordinating teaching teams. My research on internet censorship and content moderation in large language models, which was published at NDSS, EuroS&P, and FOCI, directly informs my approach to teaching AI, networking, and security at the undergraduate and graduate level.

EDUCATION

Doctor of Philosophy in Computer Science,
University of the Saarland
04/2021 – 12/2025 | Saarbruecken, Germany

Master of Science in Embedded Computing, University of Turku
09/2019 – 04/2021 | Turku, Finland

Bachelor of Science in Network and Computer Engineering,
University of Applied Sciences of the Saarland
10/2016 – 07/2019 | Saarbruecken, Germany

TEACHING EXPERIENCE

Data Networks, 2022-2025, one semester per year

Core undergraduate/graduate course on Internet architecture, routing, transport protocols, and network security. Responsibilities included leading weekly exercise sessions, designing homework assignments and exam problems, grading, and holding office hours. Coordinated with the supervising professor on syllabus updates and course logistics. Class size of ~400 students per semester.

Hot Topics in Data Networks, 2022-2025, one semester per year

Advanced research seminar guiding roughly 20 students through reading, presenting, and critically analyzing recent publications in Internet measurement, network architecture, and security. Supervised student paper presentations, selected and assigned paper topics and facilitated group discussions.

Distributed Systems, 2021, one semester

Course covering theoretical foundations and practical implementation of distributed computing, including consensus algorithms, replication, and fault tolerance. Led exercise sessions and assisted with lab assignments, exam preparation and grading.

Thesis Supervision

Supervised a Master's thesis on AI safety and robustness of large language models. Provided ongoing technical mentorship, feedback on experimental design, and guidance on academic writing.

Course Administration

Contributed to course organization across multiple offerings: prepared and proofread exams, coordinated grading among teaching assistants, and supervised in-person examinations.

PROFESSIONAL EXPERIENCE

Server Engineer, Mankind Games

08/2025 – Present | Tokyo, Japan

Responsible for the infrastructure of multiple multi-tenant large-scale infrastructure projects focusing on fleet surveillance, 3D infrastructure modeling and various other projects, including full-stack development on multiple websites and video games. Also responsible for implementing and designing the company's AI strategy. The company has been granted research grants on multiple projects with my involvement by the Japanese MEXT.

PROJECTS

Contributor to Text Generation WebUI (oobabooga/textgen)

The original, tremendously popular LLM framework with over 45k stars on Github.

AI Content Moderation Classifier

Custom trained, DeBERTa based model to detect content moderation in LLMs.

INVITED TALKS

Safety in Large Language Models, IIT Bombay

11/2025 | Bombay, India

Investigating differences in content safety of Large Language Models

04/2026 | Tokyo, Japan

LANGUAGES

English	● ● ● ● ●
German	● ● ● ● ●
Japanese	● ● ● ● ●
French	● ● ● ● ●

REFERENCES

Anja Feldmann, *Director*,
Max Planck Institute for Informatics
Please contact Iris Wagner,
Secretary: iwagner@mpi-inf.mpg.de

Romain Fontugne, *Director*,
Internet Initiative Japan Research
Laboratory
romain@ij.ad.jp

Devashish Gosain,
Assistant Professor,
Indian Institute of Technology,
Bombay
dgosain@cse.iitb.ac.in

Amir Houmansadr, *Professor*,
Manning College of Information and
Computer Sciences, University of
Massachusetts, Amherst
amir@cs.umass.edu

Doctoral Researcher, Max Planck Institute for Informatics

04/2024 – 12/2025 | Saarbruecken, Germany

Doctoral researcher working under Anja Feldmann, focusing on AI safety research, AI security, internet censorship and security and traffic analysis.

Researcher, Internet Initiative Japan

07/2024 – 08/2024 | Tokyo, Japan

Working on internet security & traffic analysis projects.

Researcher, University of Turku, FitOptiVis Research Team

05/2020 – 02/2021 | Turku, Finland

Researcher on FitOptiVis ECSEL, an EU-funded research project on distributed embedded image processing.

Software Engineer, Kontron GmbH

11/2018 – 07/2019 | Saarbruecken, Germany

Embedded system developer.

PUBLICATIONS

There is No War in Ba Sing Se: A Global Analysis of Content Moderation in Large Language Models, Network and Distributed System Security (NDSS) Symposium 02/2026

The first global examination of LLM moderation across 15 models by 10 companies, 12 vantage points and 13 languages, evaluation of over 700k queries.

Dual Standards: Examining Content Moderation Disparities Between API and WebUI Interfaces in Large Language Models, Free and Open Communications on the Internet 01/2026

The first study to evaluate differences in LLM content moderation and safety when queried using different interfaces, proving the existence of interface-specific filtering algorithms depending on whether users utilize API or WebUI interfaces.

Rabid Rabbit-hole?: Safety Audit of TikTok's Recommendation System, Under Submission

The first study on TikTok political and contextual recommendation bias targeting underage users in Brazil and the US.

Can You Hear Me? A First Study of VoIP Censorship Techniques in Saudi Arabia and the UAE, IEEE European Symposium on Security and Privacy 06/2025

First study of Voice over IP censorship in the Middle East.

1-800-Censorship: Analyzing internet censorship data using the Internet Yellow Pages, Free and Open Communications on the Internet 07/2025

Integration of censorship data in popular IYP framework and showcase of initial results.

You wanna see some real speed? Comparative Analysis of global Measurement Lab and Cloudflare speed test results, AINTEC 11/2025

Speedtest comparison study across services.