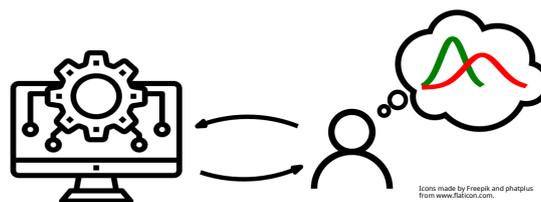


Algorithm Auditing via Statistical Hypothesis Testing

Background. In principle, the market of digital services has the lowest entry threshold one could possibly imagine. Whether you are a billion-dollar enterprise or a private individual; anyone can, at any time, offer their services by merely putting a website online. Same rights for everyone, equal opportunity for participation, the perfect democracy? That is, of course, not the whole story.

Big tech monopolists such as Alphabet/Google, Meta/Facebook, or Amazon effectively control the attention of most users. Their algorithms decide which content you see and how and where you consume it. If the information they provide is biased, this can boost or close businesses and even change election outcomes. If their automated decision-making (ADM) systems operate unfairly, the subjects of those decisions are discriminated. If the platforms they offer are insecure, huge amounts of private data are exposed. If they crawl third-party content and present it as part of their service to eliminate any need for a user to ever leave their sites, those third parties are put out of business.

Of course, all monopolists claim to do no wrong; they simply care for their users' needs and want to provide the best service they possibly can. But even if one believes that and additionally puts aside the fact that good intentions do not necessarily lead to good outcomes, it still holds: With great power comes great need for oversight.¹



Challenge. Usually, no outsider has full access to the internal systems of any service provider. And even if they had, a complete audit or the reverse engineering of the employed algorithms would be practically infeasible. Some services provide an API endpoint, however even the field of explainable AI is struggling to come up with satisfying solutions. Thus if a service provider claims to implement a particular property how would you test that claim on a black box?

¹ See for example a GI report on the topic: <https://gi.de/themen/beitrag/studie-regulierung-von-algorithmischen-entscheidungssystemen>

We aim to approach the problem from a different angle and focus on an approach that gained recent traction in the algorithm theory community. This approach utilises efficient algorithms for learning and testing a statistical hypothesis. That is, given a black-box algorithm, decide whether the outputs fulfill the statistical hypothesis or divert from the claimed output distribution – be it maliciously or unintentionally.

Vision. We want to develop a software system to detect aberrations to quantitative or qualitative claims of platform operators. This system will be based on prior work in the field of statistical hypothesis testing, but should also extend on current work. As auditing commercial platforms is the final step, we will initiate the bachelor project with a game of cops and robbers. Half of the project team will develop a very simple platform API, corresponding claims as well as hidden discrepancies. The other half will implement a test system to confirm or refute those claims and detect the discrepancies. From this game the project will then mature to choose and take on a real-world platform to audit. This includes coming up with hypotheses as well as the means to gather and analyze test data.

Industry partner contribution. We team up with Corint Media, which legally enforces digital rights to ensure the funding of independent and free media. They just successfully opened up an antitrust case against Alphabet/Google at the German Federal Cartel Office (Bundeskartellamt) and are currently also negotiating with Meta/Facebook. They will advise us on what algorithmic properties of the ADM systems of these platforms are of particular interest for our studies.

Our contribution. The research group will work closely with you on the project and will provide expertise in algorithm engineering, software development, and data science. We will help you to develop hypotheses and theoretical models to investigate the collected data and evaluate the quality of the developed solutions. As in previous bachelor and master projects of our group, we plan to write a joint scientific publication about our research findings. For that and your bachelor theses, we will offer a workshop on scientific writing in March 2023. We also offer a round table discussion on the project before the voting takes place on August 1st, 2022 at 2pm (just contact Stefan Neubert).

Your contribution. In your work as a team, you will review related work on statistical hypotheses testing, carry out an experiment to test this work in practice, and develop a software system and algorithms for solving the problem at hand. Besides team experience, you will learn about data collection, efficient data processing, and algorithmic fairness as well as the practical use of probability theory, statistics and other related fields.

Supervisors.



Prof. Dr. Tobias Friedrich
Dr. Andreas Göbel
Stefan Neubert
Marcus Pappik