Igor Shilov

London, UK | hello@igorshilov.com | +447523506334 | igorshilov.com | ffuuugor

Summary

AI Researcher focused on ML Privacy with a strong software engineering background. Currently pursuing a PhD in AI Privacy.

Experience

Meta, Staff Software Engineer — 2017 - 2023, *London, UK*

- **Lead developer on PyTorch framework for Differential Privacy.** I was a lead developer and maintainer of Opacus, a PyTorch library to train ML models with Differential Privacy. With over 1k stars on github, the tool is helping to advance the state-of-the-art in privacy preserving ML both internally and for the wider community of researchers.
- **Research Engineer on Privacy Preserving ML team.** I supported various product applications of Differential Privacy and Federated Learning, as well as contributed to the original research on Differential Privacy and Privacy Attacks.
- **Lead Engineer on StopNCII.org.** I have lead the team developing a privacy-preserving platform helping combat non-consensual intimate image sharing, a joint effort between Meta and a UK-based NGO 'Revenge Porn Helpline'. The platform takes advantage of on-device perceptual hashing to protect privacy.
- **Lead engineer on Safety.** I've led the team of 4 engineers working on building ML models to detect and remove harmful content.

Zvooq, Senior Software Engineer — 2014 - 2017, *Moscow, Russia*

- **Lead ML Engineer.** I was a team lead for 4 software engineers building music recommendation engine.
- **Analytics engine development.** I've developed company's internal analytics system to handle copyright holders reporting.

Mail.ru, Software Engineer — 2013 - 2014, *Moscow, Russia*

- **Search ranking algorithm.** I've worked on the ranking algorithm for the company's search engine (10% of market share in Russia).

Côc Côc, Software Engineer — 2011 - 2012, *Moscow, Russia* / *Hanoi, Vietnam*

- **Web crawler development.** I've worked on building a large-scale distributed web-crawler and improving query intent classification.

Education

PhD student — **Imperial College London**, Computing Research — 2023 - Present, *London, UK*

Specialist degree (5 years) — **Lomonosov Moscow State University**, Applied Mathematics and Computer Science — 2008 - 2013, *Moscow, Russia*

Selected Publications

Copyright Traps for Large Language Models — 2024

Matthieu Meeus*, ***Igor Shilov****, Manuel Faysse, Yves-Alexandre Montjoye

In Forty-first International Conference on Machine Learning, 2024

Opacus: User-Friendly Differential Privacy Library in PyTorch 2021

Ashkan Yousefpour, ***Igor Shilov***, Alexandre Sablayrolles, Davide Testuggine, Karthik Prasad, Mani Malek, John Nguyen, Sayan Ghosh, Akash Bharadwaj, Jessica Zhao, Graham Cormode, Ilya Mironov

In NeurIPS Workshop on Privacy in Machine Learning, 2021

Antipodes of label differential privacy: PATE and ALIBI 2021

Mani Malek Esmaeili, Ilya Mironov, Karthik Prasad, ***Igor Shilov***, Florian Tramer

In Advances in Neural Information Processing Systems, 2021

Complete Publication List

Free Record-Level Privacy Risk Evaluation Through Artifact-Based Methods 2024

Joseph Pollock*, ***Igor Shilov***, Euodia Dodd, Yves-Alexandre Montjoye

Under review, arXiv preprint 2411.05743

Certification for Differentially Private Prediction in Gradient-Based Training 2024

Matthew Wicker, Philip Sosnin, ***Igor Shilov***, Adrianna Janik, Mark N. Müller, Yves-Alexandre Montjoye, Adrian Weller, Calvin Tsay

Under review, arXiv preprint 2406.13433

SoK: Membership Inference Attacks on LLMs are Rushing Nowhere (and How to Fix It) 2024

Matthieu Meeus, ***Igor Shilov***, Shubham Jain, Manuel Faysse, Marek Rei, Yves-Alexandre Montjoye

Under review, arXiv preprint 2406.17975

Mosaic Memory: Fuzzy Duplication in Copyright Traps for Large Language Models 2024

Igor Shilov, Matthieu Meeus*, Yves-Alexandre Montjoye

arXiv preprint 2405.15523

Defending against Reconstruction Attacks with Rényi Differential Privacy 2022

Pierre Stock, ***Igor Shilov***, Ilya Mironov, Alexandre Sablayrolles

arXiv preprint 2202.07623