# Zheguang Zhao

Research Fellow in Secure Systems
The University of Melbourne
700 Swanston Street
Parkville, VIC 3010
Australia

| | |
|---|---|
| Phone: | +61-433-440-582 |
| Email: | zheguang.zhao@alumni.brown.edu |
| Homepage: | zheguang.github.io |
| LinkedIn: | www.linkedin.com/in/zheguang |
| D Google Scholar: | goo.gl/DR8pSa |

## Education

Ph.D. in Computer Science, Brown University, Rhode Island, 2021
Advisors: Stan Zdonik, Seny Kamara
Readers: George Kollios, Moti Yung
Thesis: *Building a Structurally-Encrypted Relational Database*
Software: KafeDB

M.Sc. in Computer Science, Brown University, Rhode Island, 2016
Advisor: Stan Zdonik
Readers: Tim Kraska, Ugur Cetintemel
Thesis: *Approximate Data Structures for Visualization*

B.Sc. in Computer Science, University of Wisconsin at Madison, Wisconsin, 2012
Advisor: Jignesh Patel, Ben Liblit

## Certification & Technologies

Deep Learning Specialization, *Coursera / deeplearning.ai* [link]

ML/Stats: Tensorflow, Keras, Julia, MATLAB

Systems: Spark, MapReduce, Linux Kernel, Network Security, Databases

PL: C/++, Haskell, Julia, Java, Python, SQL

Infra: Google Cloud, Jenkins, Git, Jira, Docker, Chef, Puppet, Kubernetes

## Experiences

The University of Melbourne, *Research Fellow in Secure Systems* (2023-now)

> Security in adversarial ML and human-AI collaboraiton. Hosted by Prof. Olya Ohrimenko and Prof. Ben Rubinstein

Technical University of Darmstadt, *Postdoc Researcher* (2021-2022)

> Secure federated learning and databases. Hosted by Prof. Carsten Binnig in Data Management Group

Sifr Systems (Acquired by MongoDB), *Database Scientist* (2017-2021)

> KafeDB, an end-to-end encrypted relational database based on SparkSQL

Los Alamos National Laboratory, *Research Intern* (2019)

ML model reconstruction of mixed dynamics in cyber-physical systems, with application to network verification and security

Microsoft AI & Research, *Research Intern* (2017)

Constraint learning. Hosted by DMX Group

Intel Labs, *Research Intern* (2015)

Efficiency of machine learning algorithms in Apache Spark

In-memory transactional processing using non-volatile memory

Hadapt (Acquired by Teradata), *Software Engineer* (2013-2014)

Enterprise SQL-on-Hadoop system including query execution, storage engine, high availability and analytics

Kosmix (Acquired by @WalmartLabs), *Software Engineer* (2012-2013)

In-memory distributed queue system for in-house stream processing

Great Lakes Bioenergy Research Center, *Software Engineer* (2009-2012)

Scientific database for biological enzyme research

## Open Source

KafeDB: End-to-End Structurally-Encrypted Relational Database [link]

ML framework for Cypber-physical Systems [link]

Encrypted Searchable Signal [link]

Macau: statistical hypothesis testing based on resampling [link]

Machine learning algorithms in Spark [link]

Consistency control for machine learning algorithms [link]

R-tree in Rust [link]

Spark performance analysis tool [link]

VoltDB on non-volatile memory [link]

## Preprints

*An Optimal Relational Database Encryption Scheme* [link]
Cryptology ePrint Archive: Report 2020/274

*Learning of Cyber-Physical Systems*
Advanced Network Science Initiative, Los Alamos National Laboratory, 2019

*Behavior of Large Random Graph.* [link]
Randomized Algorithms for Counting, Integration and Optimization, Brown University, 2017

*Signal Search.*[link]
Brown University, 2017

# Publications

*Towards Decentralized Parameter Servers for Secure Federated Learning* DATA 2022

*ACID-V: towards a New Class of DBMSs for Data Sharing*
Polystores Workshop at VLDB, 2021

*Encrypted Databases: from Theories to Systems*
CIDR, January 2021

*Dynamic Query Refinement for Interactive Data Exploration*
EDBT/ICDT Joint Conference, March 2020

*Investigating the Effect of the Multiple Comparisons Problem in Visual Analysis*
CHI Conference, April 2018

*Controlling False Discoveries During Interactive Data Exploration*
SIGMOD Conference, May 2017

*Safe Visual Data Exploration*
SIGMOD Conference, Demo, May 2017

*Bridging the Gap between HPC and Big Data frameworks*
VLDB Journal, 2017

*Towards Sustainable Insights*
CIDR Conference, January 2017

*Towards a Benchmark for Interactive Data Exploration*
IEEE Data Engineering Bulletin, 2016.

*Larger-than-memory Data Management on Modern Storage Hardware for In-memory OLTP Database Systems*
SIGMOD DaMoN Workshop, June 2016

*VisTrees: Fast Indexes for Interactive Data Exploration*
SIGMOD HILDA Workshop, June 2016

*Data Tiering in Heterogeneous Memory Systems*
EuroSys Conference, April 2016

# Teaching

Security Analytics, The University of Melbourne, 2023.

Data Management Labs, TU Darmstadt, 2021.

Introduction to Database Systems, Brown University 2015, 2018

Advanced Topics in Database Systems, Brown University 2015

## Supervision

Benedikt Völker, *Client-side Validation for Detecting the Model Poisoning Attack in Federated Learning*
M.Sc.'21, TU Darmstadt

Philipp Imporatori, *Building an Oblivious Relational Database*
M.Sc.'22, TU Darmstadt

Shan Li, *Gradient-based Attacks on Federated Learning*
M.Sc.'22, TU Darmstadt

## Services

Reviewer: OSDI'16, TVCG'21, TODS'21, IEEE S&P (Journal)'24, IEEE S&P (Oakland)'24

Program committee: VLDB'22 (Demo Track)

Member: IACR, ACM

## Honors

Graduate Fellowship, Brown University, 2014-2019

Dean's Honor List, University of Wisconsin at Madison, 2007, 2008, 2009

Eta Kappa Nu, 2009

Upsilon Pi Epsilon, 2010

Golden Key International Honour Society, 2010

Dissertation Fellowship, Brown University, 2018-2019

## References

Available upon request.