XI HE

Assistant Professor, Cheriton School of Computer Science, University of Waterloo

xihe@uwaterloo.ca ◇ https://cs.uwaterloo.ca/~xihe/

Research Interests

My research interests span the areas of privacy and security for big-data, with a current focus on (i) designing strong and useful privacy models for data exploration and learning with theoretical understanding on the trade-offs between privacy and usability of applications and (ii) building practical systems that enable learning of useful information from the data while provably ensuring individuals' privacy.

Degrees

PhD in Computer Science, **Duke University, Durham, NC, USA** 2012-2018

BS in Mathematics with Honors & BCS with Honors, **National University of Singapore** 2008-2012

Professional Positions

Assistant Professor, Tenure Track, **Computer Science, University of Waterloo, Canada** 2019/03-Present

Faculty Member, **Vector Institute, Canada** 2022/04-Present

Research Collaborator (part-time), **Meta, Canada** 2021/08-2022/08

Postdoctoral Associate, **Computer Science, Duke University, USA** 2018/09-2018/12

Research Intern, **Microsoft Research, Redmond, WA, USA** 2016/06-2016/09

Research Intern, **AT&T Labs Research, New York, USA** 2014/06-2014/08

Research Intern, **AT&T Labs Research, New Jersey, USA** 2013/06-2013/08

Awards and Funding

H-1. **2022 Meta Research Award** for proposal on Privacy-preserving Technologies, for research on differential privacy for multi-relational databases

H-2. **2022 Processing and Analysis (DAPA) Award**, Google

H-3. **2021 John R. Evans Leader Fund (JELF)** awarded from Canada Foundation for Innovation (CFI) to support infrastructure for scalable end-to-end differentially private machine learning

H-4. **2020 Alliance Grant** awarded from Natural Sciences and Engineering Research Council of Canada (NSERC) and Waterloo Joint Innovation Lab to study privacy-preserving graph analytics engine

H-5. **2020-2022 Resource Allocation Competition Award**, Compute Canada

H-6. **2019 Discovery Grant** awarded from NSERC to support study on private data exploration

H-7. **2021 ACM SIGMOD Distinguished PC**

H-8. **2018 Outstanding Ph.D. Dissertation Award**, Duke University

H-9. **2017 Google PhD Fellowship in Privacy and Security**

H-10. **2016 International Conference on Very Large Databases Best Demo Award**

H-11. **2016 Rising Stars Workshop Participant**, Carnegie Mellon University

H-12. **Members of the U.S. Delegation to the 2nd HLF**, Heidelberg Laureate Forum Foundation, 2014

Selected Publications

CONFERENCE.

P-1. Primal Pappachan, Shufan Zhang, Xi He, Sharad Mehrotra. Don't Be a Tattle-Tale: Preventing Leakages through Data Dependencies on Access Control Protected Data. *Proc. VLDB Endow.* 2022. Accepted for publication.

P-2. Sameera Ghayyur, Dhrubajyoti Ghosh, Xi He, Sharad Mehrotra. MIDE: Accuracy Aware Minimally Invasive Data Exploration for Decision Support. *Proc. VLDB Endow.* 2022. Accepted for publication.

P-3. Shubhankar Mohapatra, Sajin Sasy, Xi He, Gautam Kamath, Om Thakkar. The Role of Adaptive Optimizers for Honest Hyperparameter Tuning. *Proc. of the AAAI Conference on Artificial Intelligence*, 36(7):7806-7813, 2022.

P-4. Priyanka Nanayakkara, Johes Bater, Xi He, Jessica Hullman, and Jennie Rogers. Visualizing Privacy-Utility Trade-Offs in Differentially Private Data Releases. *Proc. on Privacy Enhancing Technologies*, pp 601-618, 2022.

P-5. Chang Ge, Shubhankar Mohapatra, Xi He, Ihab F. Ilyas. Kamino: Constraint-Aware Differentially Private Data Synthesis". *Proc. VLDB Endow.* 14(10):1886-1899, 2021.

P-6. Helen Chen, Shubhankar Mohapatra, George Michalopoulos, Xi He, Ian McKillop. Federated Deep Learning Architecture for Personalized Healthcare. *Stud Health Technol Inform.* 281:193-197, 2021.

P-7. Johes Bater, Yongjoo Park, Xi He, Xiao Wang, and Jennie Rogers. SAQE: Practical PrivacyPreserving Approximate Query Processing for Data Federations. *Proc. VLDB Endow.* 13(12):26912705, 2020.

P-8. Zhuolun Xiang, Bolin Ding, Xi He, Jingren Zhou. Linear and Range Counting under Metric-based Local Differential Privacy. *IEEE Int. Symposium on Information Theory (ISIT)*, pp 908-913, 2020.

P-9. Yuchao Tao, Xi He, Sudeepa Roy, Ashwin Machanavajjhala. Computing Local Sensitivities of Counting Queries with Joins. In *Proc. ACM SIGMOD Int. Conf. on Management of Data*, pp 479-494, 2020.

P-10. Amrita Roy Chowdhury, Chenghong Wang, Xi He, Ashwin Machanavajjhala, Somesh Jha. Cryptϵ: Crypto-Assisted Differential Privacy on Untrusted Servers. In *Proc. ACM SIGMOD Int. Conf. on Management of Data*, pp 603-619, 2020.

P-11. Changchang Liu, Xi He, Thee Chanyaswad, Shiqiang Wang, Prateek Mitta. Investigating Statistical Privacy Frameworks from the Perspective of Hypothesis Testing. *Proc. on Privacy Enhancing Technologies*, pp 233-254, 2019.

P-12. Ios Kotsogiannis, Yuchao Tao, Xi He, Ashwin Machanavajjhala, Michael Hay and Gerome Miklau. PrivateSQL: A Differentially Private SQL Query Engine. *Proc. VLDB Endow.* 12(11):1371-1384, 2019.

P-13. Johes Bater, Xi He, William Ehrich, Ashwin Machanavajjhala, Jennie Rogers. Shrinkwrap: Differentially Private Query Processing in Private Data Federations. *Proc. VLDB Endow.* 12(3):307-320, 2019.

P-14. Chang Ge, Ihab F. Ilyas, Xi He, Ashwin Machanavajjhala. APEx: Accuracy-Aware Differentially Private Data Exploration. In *Proc. ACM SIGMOD Int. Conf. on Management of Data*, pp 177-194, 2019.

P-15. Xi He, Ashwin Machanavajjhala, Cheryl Flynn, Divesh Srivastava. Composing Differential Privacy and Secure Computation: A Case Study on Scaling Private Record Linkage. In *P*roc. ACM SIGSAC Conf. on Computer and Communications Security (CCS), pp 1389-1406, 2017.

P-16. Xi He, Graham Cormode, Ashwin Machanavajjhala, Cecilia M. Procopiuc, Divesh Srivastava. DPT: Differentially Private Trajectory Synthesis using Hierarchical Reference Systems. *Proc. VLDB Endow.* 8(11):1154-1165, 2015.

P-17. *Xi He*, Ashwin Machanavajjhala, Bolin Ding. Blowfish Privacy: Tuning Privacy-Utility Trade-offs using Policies. In *Proc. ACM SIGMOD Int. Conf. on Management of Data*, pp 1447-1458, 2014.

BOOK CHAPTER. & ARTICLES

P-18. Joe Near, Xi He. Differential Privacy for Databases. *Foundations and Trends® in Databases*: 11(2):109-225, 2021.

P-19. Sameer Wagh, Xi He, Ashwin Machanavajjhala, Prateek Mittal. DP-Cryptography: Marrying Differential Privacy and Cryptography in Emerging Applications. *Commun. ACM* 64(2):84-93, 2021.

P-20. Ashwin Machanavajjhala, Xi He. Analyzing Your Location Data with Provable Privacy Guarantees. In: *Gkoulalas-Divanis, A., Bettini, C. (eds) Handbook of Mobile Data Privacy. Springer, Cham.* 2018.

WORKSHOP & DEMO.

W-21. Shufan Zhang, Runchao Jiang, Xi He. DProvSQL: Privacy Provenance Framework for Differentially Private SQL Engine. In *Theory and Practice of Differential Privacy (TPDP)*, part of *ICML* 2022.

W-22. Shubhankar Mohapatra, Florian Kerschbaum, Xi He. Differentially Private Data Generation with Missing Data. In *TPDP*, part of *ICML* 2022.

W-23. Jiaxiang Liu, Karl Knopf, Yiqing Tan, Bolin Ding, Xi He. Catch a Blowfish Alive: A Demonstration of Policy-AwareDifferential Privacy for Interactive Data Exploration. *Proc. VLDB Endow.* 14(12):2859-2862, 2021.

W-24. Siyuan Xia, Beizhen Chang, Karl Knopf, Yihan He, Yuchao Tao, Xi He. DPGraph: A Benchmark Platform for Differentially PrivateGraph Analysis. In *Proc. ACM SIGMOD Int. Conf. on Management of Data*, pp 2808-2812, 2021.

W-25. Miti Mazmudar, Thomas Humphries, and Matthew Rafuse, and Xi He. Cache Me If You Can: Accuracy-Aware Inference Engine for Differentially Private Data Exploration. In *TPDP*, part of *CCS*, 2020.

W-26. Harry Sivasubramaniam, Haonan Li, and Xi He. Differentially Private Sublinear Average Degree Approximation. In *TPDP*, part of *CCS*, 2020.

W-28. Xi He, Nisarg Raval, Ashwin Machanavajjhala. *"A Demonstration of VisDPT: Visual Exploration of Differentially Private Trajectories". Proc. VLDB Endow.* 9(13):1489-1492, 2016.

Selected Talks

T-1. Tutorial on "Practical Security and Privacy for Database Systems." In *Proc. ACM SIGMOD Int. Conf. on Management of Data*, pp 2839-2845, 2021.

T-2. "Privacy is Not an Afterthought". *UC Berkeley Data System and Foundation Lab*, Fall 2020; *Boston-area DP seminar*, Spring 2021; *Google DAPA seminar*, Spring 2022; *Joint Speaker Series on Privacy and Security in AI and Big Data at the Hong Kong Polytechnic University*, Spring 2022.

T-3. "A Tour of Differential Privacy". *Privacy Engineering Section Forum at the International Association of Privacy Professions (IAPP)*, US, Winter 2019.

T-4. "Private Exploration Primitives for Data Cleaning". *Mathematical Foundations of Data Privacy Workshop*, Banff, Canada, Spring 2018.

T-5. "Trading off Accuracy Privacy and Efficiency in Secure Joins using Differential Privacy". *Differential Privacy Meets Multi-Party Computation (DPMPC) Workshop*, US, Fall 2018.

T-6. "Privacy with Constraints: Opportunities & Challenges". Theory and Practice of Differential Privacy *TPDP*, part of *CCS*, Fall 2017.

T-7. "Differential Privacy in the Wild: A tutorial on current practices & open challenges". *Proc. VLDB Endow.* 9(13):16111614, 2016. In *Proc. ACM SIGMOD Int. Conf. on Management of Data*, pp 1727-1730, 2017.

Professional Services and Leadership

Program Committee. Co-chair for the Student Research Competition, ACM SIGMOD 2020-2021; Tutorial chair for ICDE 2023; Review committee for VDLB 2020-2023, SIGMOD 2020-2023, ICDE 2020-2022, CCS 2020-2022, EDBT 2020-2022, PETS 2020-2023, TPDP 2019-2021.

Professional Affiliations. Women in Computer Science(WiCS) Grad Group Sub-Committee at Univ. of Waterloo (2022-); NUCC Adjudication Committee Member 2022; Member of Waterloo CS Graduate Recruitment Committee (2019-); Committee member of ACM-W Duke University Chapter (president in 2015, vice president in 2014,2016,2017).

Graduate Supervisions

S-1. **Shubhankar Mohapatra** (MA → PhD)	2019/05-2020/08, 2021/09-
S-2. **Karl Knopf** (PhD)	2019/09-
S-3. **Shufan Zhang** (MA → PhD)	2020/09-
S-4. **Christian Covington** (MA)	2020/09-2021/06
S-5. **Harry Sivasubramaniam** (MA)	2019/05-2021/12

Courses

C-1. **Building Privacy-Aware Database Systems** (CS848), Seminar	2021/01
C-2. **Privacy and Fairness in Data Science** (CS848), Seminar	2019/09, 2018/09
C-3. **Introduction to Database Management** (CS348), Undergraduate Course	2019/03, 2020/01, 2021/05, 2022/09