# Cheng Xin

Email: xin.job2025@gmail.com | Website: jackal092927.github.io | Google Scholar: jackal092927.github.io/scholar

## Profile

Cheng Xin is a computer science Ph.D. and currently a postdoctoral researcher in the Department of Computer Science at Rutgers University, advised by Prof. Jie Gao. He received his Ph.D. in Computer Science from Purdue University in 2023 under the supervision of Prof. Tamal K. Dey; before that, he received an M.S. in Computer Science from Lehigh University and a B.Eng. in Software Engineering from Tongji University. His research spans **topological data analysis**, **topological machine learning**, **non-Euclidean representation learning**, **interpretable AI**, **AI for Science**, and **3D/video generation benchmarks**. He has published in international conferences and journals including **NeurIPS**, **ICML**, **CVPR**, **SoCG**, **TMLR**, Pattern Recognition, and Journal of Applied and Computational Topology, and has served as **Area Chair for TAG-DS Workshop 2026** and reviewer for **ICML**, **ICLR**, **NeurIPS**, and **SoCG**.

## Overseas Study and Research Experience

- Rutgers University, Department of Computer Science, Postdoctoral Researcher, Oct 2023-present. Advisor: Prof. Jie Gao. Research on interpretable graph learning, topological machine learning, non-Euclidean representation learning, hyperbolic-space algorithms, AI for Science, and 3D/video generation benchmarks.
- Purdue University, Department of Computer Science, Ph.D. student / Research Assistant, Aug 2020-Aug 2023. Advisor: Prof. Tamal K. Dey. Dissertation: Decomposition and Stability of Multiparameter Persistence Modules.
- The Ohio State University, Department of Computer Science and Engineering, Ph.D. student / Research Assistant, Jan 2017-Aug 2020. Advisor: Prof. Tamal K. Dey. Research on decomposition algorithms and stability theory for multiparameter persistent homology.
- Lehigh University, Department of Computer Science, M.S. student, Jan 2014-May 2016. Advisor: Prof. Xiaolei Huang. Thesis: Machine Learning Techniques for Cervigram Image Analysis.

## Education

- Ph.D. in Computer Science, Purdue University, United States, Aug 2020-Aug 2023
- Ph.D. student in Computer Science and Engineering, The Ohio State University, United States, Jan 2017-Aug 2020
- M.S. in Computer Science, Lehigh University, United States, Jan 2014-May 2016
- B.Eng. in Software Engineering, Tongji University, Shanghai, China, Sep 2009-Jul 2013

## Major Research Projects and Contributions

**1. TopInG: Topologically Interpretable Graph Learning (ICML 2025)**

- Project: Proposed a graph neural network interpretation framework based on persistent rationale filtration, improving both prediction performance and explanation quality on molecular graph tasks.
- Role and contribution: First author; responsible for problem formulation, method design, algorithm implementation, experimental design, and paper writing.

**2. D-GRIL: End-to-End Topological Learning (SoCG 2026, to appear)**

- Project: Integrated 2-parameter persistent homology features into machine learning models in a differentiable manner, advancing the use of multiparameter topological features in end-to-end learning.
- Role and contribution: Core contributor; participated in method design, theoretical analysis, algorithm implementation, and manuscript writing.

**3. Non-Euclidean Geometry and Hyperbolic Representation Learning Series (NeurIPS 2024, NeurIPS 2025, SoCG 2026)**

- Project: Studied non-Euclidean multidimensional scaling, Johnson-Lindenstrauss lemmas beyond Euclidean space, and locality sensitive hashing in hyperbolic space, providing theoretical and algorithmic foundations for non-Euclidean representation learning.
- Role and contribution: Core contributor / alphabetical-order author; participated in theoretical derivation, algorithm design, experimental validation, and paper writing.

**4. Decomposition and Stability Theory for Multiparameter Persistent Homology (SoCG 2018; Journal of Applied and Computational Topology 2022; Ph.D. dissertation 2023)**

- Project: Studied decomposition algorithms, bottleneck distance computation, rectangular approximation, and stability for multiparameter persistence modules, contributing foundational algorithms for topological data analysis.
- Role and contribution: Main researcher / alphabetical-order author; responsible for algorithm design, theoretical proof, complexity analysis, and paper writing.

**5. DL3DV-10K Large-Scale 3D Vision Dataset and Benchmark (CVPR 2024)**

- Project: Built a large-scale scene dataset for deep learning-based 3D vision, novel view synthesis, and video generation.
- Role and contribution: Participated in dataset and benchmark-related research, supporting experiments and publication output in 3D/video generation.

**6. Medical Image Machine Learning and Cervical Dysplasia Classification Benchmarks (MLMI 2015; Pattern Recognition 2017)**

- Project: Built a cervigram image dataset and multi-feature machine learning benchmark for cervical dysplasia classification, comparing multiple classical machine learning models for medical image classification.
- Role and contribution: Co-first author of the MLMI 2015 paper; participated in feature engineering, model evaluation, experimental analysis, and paper writing.

## Publication Record

Note: * indicates co-first author.

**1.** Chengyuan Deng, Jie Gao, Kevin Lu, Feng Luo, and **Cheng Xin**. "Locality Sensitive Hashing in Hyperbolic Space." 42nd International Symposium on Computational Geometry (**SoCG**), to appear, 2026. arXiv:2603.19724. Category: international computational geometry conference; Role: alphabetical-order author.

**2.** Soham Mukherjee, Shreyas N. Samaga, **Cheng Xin**, Steve Oudot, and Tamal K. Dey. "D-GRIL: End-to-End Topological Learning with 2-parameter Persistence." 42nd International Symposium on Computational Geometry (**SoCG**), to appear, 2026. arXiv:2406.07100. Category: international computational geometry conference; Role: core contributor.

**3.** **Cheng Xin**, Fan Xu, Xin Ding, Jie Gao, and Jiaxin Ding. "TopInG: Topologically Interpretable Graph Learning via Persistent Rationale Filtration." The 42nd International Conference on Machine Learning (**ICML**), 2025. Category: top-tier international machine learning conference; Role: first author.

**4.** Chengyuan Deng, Jie Gao, Kevin Lu, Feng Luo, and **Cheng Xin**. "Johnson-Lindenstrauss Lemma Beyond Euclidean Geometry." The 39th Advances in Neural Information Processing Systems (**NeurIPS**), 2025. Category: top-tier international machine learning conference; Role: alphabetical-order author.

**5.** Chengyuan Deng, Jie Gao, Kevin Lu, Feng Luo, Hongbin Sun, and **Cheng Xin**. "Neuc-MDS: Non-Euclidean Multidimensional Scaling Through Bilinear Forms." Advances in Neural Information Processing Systems (**NeurIPS**), Vol. 37, 2024, pp. 121539-121569. Category: top-tier international machine learning conference; Role: alphabetical-order author.

**6.** Shahrzad Haddadan, **Cheng Xin**, and Jie Gao. "Optimally Improving Cooperative Learning in a Social Setting." Proceedings of the 41st International Conference on Machine Learning (**ICML**), PMLR 235, 2024, pp. 17148-17188. Category: top-tier international machine learning conference; Role: core contributor.

**7.** Lu Ling, Yichen Sheng, Zhi Tu, Wentian Zhao, **Cheng Xin**, Kun Wan, Lantao Yu, Qianyu Guo, Zixun Yu, Yawen Lu, et al. "DL3DV-10K: A Large-Scale Scene Dataset for Deep Learning-Based 3D Vision." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (**CVPR**), 2024, pp. 22160-22169. Category: top-tier international computer vision conference; Role: contributing author.

**8.** Simon Zhang, **Cheng Xin**, and Tamal K. Dey. "Expressive Higher-Order Link Prediction through Hypergraph Symmetry Breaking." Transactions on Machine Learning Research (**TMLR**), 2024. Category: international machine learning journal; Role: core contributor.

**9.** **Cheng Xin**. "Decomposition and Stability of Multiparameter Persistence Modules." Ph.D. Thesis, Purdue University Graduate School, 2023. DOI:10.25394/PGS.23848995.v1.

**10.** **Cheng Xin**, Soham Mukherjee, Shreyas N. Samaga, and Tamal K. Dey. "GRIL: A 2-parameter Persistence Based Vectorization for Machine Learning." Proceedings of Machine Learning Research, Vol. 221, 2023, pp. 313-333. Category: machine learning / topological learning conference paper; Role: first author.

**11.** Tamal K. Dey and **Cheng Xin**. "Generalized Persistence Algorithm for Decomposing Multiparameter Persistence Modules." Journal of Applied and Computational Topology 6(3), 2022, pp. 271-322. Category: international journal; Role: alphabetical-order author.

**12.** Tamal K. Dey and **Cheng Xin**. "Rectangular Approximation and Stability of 2-parameter Persistence Modules." arXiv:2108.07429, 2021. Category: preprint; Role: alphabetical-order author.

**13.** Tamal K. Dey and **Cheng Xin**. "Computing Bottleneck Distance for 2-D Interval Decomposable Modules." 34th International Symposium on Computational Geometry (**SoCG**), LIPIcs 99, 2018, 32:1-32:15. Category: international computational geometry conference; Role: alphabetical-order author.

**14.** Tao Xu, Han Zhang, **Cheng Xin**, Edward Kim, L. Rodney Long, Zhiyun Xue, Sameer Antani, and Xiaolei Huang. "Multi-feature Based Benchmark for Cervical Dysplasia Classification Evaluation." Pattern Recognition 63, 2017, pp. 468-475. Category: SCI/SCIE journal; Role: contributing author.

**15.** Tao Xu, **Cheng Xin***, L. Rodney Long, Sameer Antani, Zhiyun Xue, Edward Kim, and Xiaolei Huang. "A New Image Data Set and Benchmark for Cervical Dysplasia Classification Evaluation." Machine Learning in Medical Imaging (**MLMI**), Springer, 2015, pp. 26-35. Category: medical image machine learning conference; Role: co-first author.

## Teaching Experience

- Lecturer, Design and Analysis of Algorithms, Rutgers University, Fall 2025, approximately 45 graduate students.
- Teaching Assistant, Data Structures and Algorithms, Purdue University, Spring 2023, approximately 200 undergraduate students.
- Teaching Assistant, Computational Geometry, Purdue University, Fall 2020, approximately 30 graduate students.

## Academic Service

- **Area Chair, TAG-DS Workshop, 2026.**
- **Reviewer: ICML, ICLR, NeurIPS, SoCG.**

## Invited Talks

- "TopInG: Topologically Interpretable Graph Learning via Persistent Rationale Filtration," Conference on Topological Data Analysis: Recent Developments and Applications, University of Missouri, Nov 2025.
- "Understanding through Shape of Data: Topological Data Analysis for Interpretable AI," Management Science and Information Systems Department Colloquium, Rutgers University, Oct 2024.
- "Exploring Representations Beyond Euclidean Geometry," John Hopcroft Center Seminar, Shanghai Jiao Tong University, Jun 2024.

- "Multiparameter Persistence and Its Applications," Theory Seminar, Department of Computer Science, Rutgers University, Nov 2023.
- "Generalized Persistence Algorithm for Decomposing Multi-parameter Persistence Modules," Applied Algebraic Topology Network Seminar, Jul 2020.

## Industry and Engineering Experience

- Machine Learning Scientist Intern, Electronic Arts, Big Data Group, May 2018-Aug 2018. Worked on large-scale machine learning models on Spark, graph learning over relational databases, attribute evaluation and selection, and dataset compression.
- Software Development Engineer Intern, Amazon AWS Infrastructure Group, May 2015-Aug 2015. Developed a data management system supporting receiving, parsing, storing, and retrieving network messages.
- Software Developer, Koal Software, Shanghai, Jul 2013-Dec 2013. Worked on backend databases, business logic, APIs, and frontend UI development.
- Developer Support Intern, Microsoft Data Intelligence Group, Shanghai, Jul 2012-Nov 2012. Handled technical support cases related to SQL Server.

## Technical Skills

Python, PyTorch, Spark, Keras, Java, C/C++, MATLAB, R; topological data analysis, graph machine learning, non-Euclidean representation learning, machine learning experimental design, and large-scale data processing.