万之瑜 Zhiyu Wan

邮箱： zhiyu.wan.1@vumc.org ♦ 电话： +1(615) 839-8083

学习经历 (EDUCATION)

- **范德堡大学 Vanderbilt University** 美国 田纳西州 纳什维尔市
  - 计算机科学专业 哲学博士学位 2012 年 8 月 – 2020 年 12 月

  博士毕业论文：Privacy-Preserving Sharing of High-Dimensional Data based on Computational Game Theory（基于计算博弈论的高维数据的保密分享）

  毕业论文委员会：Bradley Malin, Daniel Fabbri, Douglas H. Fisher, Yevgeniy Vorobeychik, Murat Kantarcioglu
  - 计算机科学专业 理学硕士学位 2017 年 9 月
- **西安交通大学** 中国 陕西省 西安市
  - 控制理论与控制工程专业 攻读硕士学位 2011 年 9 月 – 2012 年 6 月
  - 自动化专业 工学学士学位 2007 年 9 月 – 2011 年 6 月
  - 工商管理专业 管理学学士学位（辅修） 2007 年 9 月 – 2011 年 6 月
  - 少年班 2005 年 9 月 – 2007 年 6 月

工作经历 (RESEARCH EXPERIENCE)

- **范德堡大学医学中心 (VUMC) 公众基因隐私及身份识别研究中心 (Center for Genetic Privacy and Identity in Community Settings)** 2021 年 1 月至今

  - ***博士后研究员***，设计基因及健康数据的重识别风险评估模型；应用该模型至范德堡大学的 SD 数据库、波士顿儿童医院领衔的精准医疗项目、及 All of Us 精准医疗项目；撰写相关文章；使用 Python 编写基因汇总数据博弈求解器；指导暑期实习本科生。（导师: Ellen Wright Clayton, Bradley Malin）
- **范德堡大学 (Vanderbilt University) 健康信息隐私实验室 (HIPLab)** 2013 年 6 月 – 2020 年 12 月

  - ***研究助理***，建立了用于健康及基因数据保密分享的博弈论模型；编写了 4 个 Python 程序（例如多阶段博弈求解器）；指导了 3 位暑期实习本科生；发表了 5 篇期刊论文、1 篇书籍章节及 5 篇会议论文。（导师: Bradley A. Malin）
- **西安交通大学信息工程科学研究中心** 2011 年 1 月 – 2012 年 6 月

  - ***研究助理***，比较了估计算法的多种性能评估指标；完成了学士学位的本科毕业论文。（导师: 李晓榕，段战胜）
- **西安交通大学智能网络与网络安全教育部重点实验室** 2009 年 6 月 – 2010 年 12 月

  - ***信息新蕾计划实习生***, 提出了一个基于蜜罐并优化了黑名单算法的协同防御框架；支持并测试了 Botwarden（一个僵尸网络检测及追踪系统）；发表了一篇会议论文。（导师: 管晓宏，郑庆华）

已出版的经过同行评审的刊物 (PEER-REVIEWED PUBLICATIONS)

- **书本章节 (Book Chapter)**

1. **Zhiyu Wan**, Yevgeniy Vorobeychik, Ellen Wright Clayton, Murat Kantarcioglu, and Bradley A. Malin. Game theory for privacy-preserving sharing of genomic data. In *Responsible Genomic Data Sharing: Challenges and Approaches, Xiaoqian Jiang and Haixu Tang, Eds.*, pp. 135–160, Academic Press, 2020.

- **期刊文章 (Journal)**

2. Abinitha Gourabathina*, **Zhiyu Wan***, J. Thomas Brown, Chao Yan, and Bradley A. Malin. PanDa Game: Optimized Privacy-Preserving Publishing of Individual-Level Pandemic Data Based on a Game Theoretic Model. *IEEE Transactions on NanoBioscience*, 22(4): 808–817, 2023. [影响因子：3.9]

3. Rajagopal Venkatesaramani, **Zhiyu Wan**, Bradley A. Malin, and Yevgeniy Vorobeychik. Defending against membership inference attacks on Beacon services. *ACM Transactions on Privacy and Security*, 26(3): 42, 2023. [影响因子：2.3]

4. Rajagopal Venkatesaramani, **Zhiyu Wan**, Bradley A. Malin, and Yevgeniy Vorobeychik. Enabling Trade-offs in Privacy and Utility in Genomic Data Beacons and Summary Statistics. *Genome Research*, 33(7): 1113–1123, 2023. [影响因子：9.438]

5. Jia Guo, Ellen Wright Clayton, Murat Kantarcioglu, Yevgeniy Vorobeychik, Myrna Wooders, **Zhiyu Wan**, Zhijun Yin, and Bradley Malin. A Game Theoretic Approach to Balance Privacy Risks and Familial Benefits. *Scientific Reports*, 13(1): 6932, 2023. [影响因子：4.996]

6. Weiyi Xia, Melissa Basford, Robert Carroll, Ellen Wright Clayton, Paul Harris, Murat Kantacioglu, Yongtai Liu, Steve Nyemba, Yevgeniy Vorobeychik, **Zhiyu Wan**, and Bradley A. Malin. Managing Re-identification Risks While Providing Access to the All of Us Research Program. *Journal of the American Medical Informatics Association*, ocad021, 2023. [影响因子：7.942]

7. Yongtai Liu, Zhijun Yin, Congning Ni, Chao Yan, **Zhiyu Wan**, and Bradley A. Malin. Examining Rural and Urban Sentiment Difference in COVID-19 Related Topics on Twitter: Word Embedding–Based Retrospective Study. *Journal of Medical Internet Research*, 23(1): e42985, 2023. [影响因子：7.076]

8. Chao Yan*, Yao Yan*, **Zhiyu Wan***, Ziqi Zhang, Larsson Omberg, Justin Guinney, Sean D. Mooney, and Bradley A. Malin. A Multifaceted Benchmarking of Synthetic Electronic Health Record Generation Models. *Nature Communications*, 13(1): 7609, 2022. [影响因子：17.694]

9. Yongtai Liu, Zhijun Yin, **Zhiyu Wan**, Chao Yan, Weiyi Xia, Congning Ni, Ellen Wright Clayton, Yevgeniy Vorobeychik, Murat Kantarcioglu, and Bradley A. Malin. Implicit Incentives Among Reddit Users to Prioritize Attention Over Privacy and Reveal Their Faces When Discussing Direct-to-Consumer Genetic Test Results: Topic and Attention Analysis. *JMIR Infodemiology*, 2(2): e35702, 2022.

10. **Zhiyu Wan**, James W. Hazel, Ellen Wright Clayton, Yevgeniy Vorobeychik, Murat Kantarcioglu, Bradley Malin. Sociotechnical safeguards for genomic data privacy. *Nature Reviews Genetics*, 23(7): 429–445, 2022. [影响因子：59.581]

11. J. Thomas Brown, Chao Yan, Weiyi Xia, Zhijun Yin, **Zhiyu Wan**, Aris Gkoulalas-Divanis, Murat Kantarcioglu, and Bradley A. Malin. Dynamically adjusting case-reporting policy to maximize privacy and utility in the face of a pandemic. *Journal of the American Medical Informatics Association*, 29(5): 853–863, 2022. [影响因子：7.942]

12. Congning Ni, **Zhiyu Wan**, Chao Yan, Yongtai Liu, Ellen Wright Clayton, Bradley Malin, and Zhijun Yin. The Public perception of the #GeneEditedBabies event across multiple social media platforms: Observational study. *Journal of Medical Internet Research*, 24(3): e31687, 2022. [影响因子：7.076]

13. **Zhiyu Wan**, Yevgeniy Vorobeychik, Weiyi Xia, Yongtai Liu, Myrna Wooders, Jia Guo, Zhijun Yin, Ellen Wright Clayton, Murat Kantarcioglu, and Bradley A. Malin. Using game theory to thwart multistage privacy intrusions when sharing data. *Science Advances*, 7(50): eabe9986, 2021. [影响因子：14.136]

14. Weiyi Xia, Yongtai Liu, **Zhiyu Wan**, Yevgeniy Vorobeychik, Murat Kantacioglu, Steve Nyemba, Ellen Wright Clayton, and Bradley A. Malin. Enabling realistic health data re-identification risk assessment through adversarial modeling. *Journal of the American Medical Informatics Association*, 28(4): 744–752, 2021. [影响因子：4.497]

15. Weiyi Xia, **Zhiyu Wan**, Zhijun Yin, James Gaupp, Yongtai Liu, Ellen Wright Clayton, Murat Kantarcioglu, Yevgeniy Vorobeychik, and Bradley A. Malin. It's all in the timing: calibrating temporal penalties for biomedical data sharing. *Journal of the American Medical Informatics Association*, 25(1): 25–31, 2018. [影响因子：4.270]

16. **Zhiyu Wan**, Yevgeniy Vorobeychik, Murat Kantarcioglu, and Bradley A. Malin. Controlling the signal: Practical privacy protection of genomic data sharing through Beacon services. *BMC Medical Genomics*, 10(Suppl 2): 39, 2017. [影响因子：2.848]

17. **Zhiyu Wan**, Yevgeniy Vorobeychik, Weiyi Xia, Ellen Wright Clayton, Murat Kantarcioglu, and Bradley A. Malin. Expanding access to large-scale genomic data while promoting privacy: a game theoretic approach. *The American Journal of Human Genetics*, 100(2): 316–322, 2017. ***被国际医学信息学协会评为年度最佳论文 Best papers in IMIA Yearbook of Medical Informatics***. [影响因子：10.794]

18. **Zhiyu Wan**, Yevgeniy Vorobeychik, Weiyi Xia, Ellen Wright Clayton, Murat Kantarcioglu, Ranjit Ganta, Raymond Heatherly, and Bradley A. Malin. A game theoretic framework for analyzing re-identification risk. *PLoS*

ONE, 10(3): e0120592, 2015. [影响因子：3.234]

• 会议文章 (Conference)

19. Rajagopal Venkatesaramani, **Zhiyu Wan**, Bradley A. Malin, and Yevgeniy Vorobeychik. Enabling Trade-offs in Privacy and Utility in Genomic Data Beacons and Summary Statistics. *Proceedings of the 27th Annual International Conference on Research in Computational Molecular Biology (RECOMB)*, 2023. [已接受]

20. Abinitha Gourabathina*, **Zhiyu Wan***, James Brown, Chao Yan, and Bradley Malin. Privacy-Preserving Publishing of Individual-Level Pandemic Data Based on a Game Theoretic Model. *Proceedings of the 2022 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, Las Vegas, NV, USA, 2022, pp. 961–968.

21. Chao Yan, Yao Yan, Ziqi Zhang, **Zhiyu Wan**, Justin Guinney, Sean Mooney, and Bradley Malin. Assessing Machine Learning Based Generators for Synthetic Electronic Health Records: A Benchmarking. *Proceedings of the 2022 American Medical Informatics Association Annual Fall Symposium (AMIA)*, Washington, DC, USA, 2022.

22. J. Thomas Brown, **Zhiyu Wan**, Aris Gkoulalas-Divanis, Murat Kantarcioglu, and Bradley Malin. Supporting COVID-19 Disparity Investigations with Dynamically Adjusting Case Reporting Policies. *Proceedings of the 2022 American Medical Informatics Association Annual Fall Symposium (AMIA)*, Washington, DC, USA, 2022.

23. Weiyi Xia, Yongtai Liu, **Zhiyu Wan**, Yevgeniy Vorobeychik, Murat Kantacioglu, Ellen W. Clayton, and Bradley A. Malin. A Scalable Tool for Realistic Health Data Re-identification Risk Assessment. *Proceedings of the 2022 American Medical Informatics Association Annual Fall Symposium (AMIA)*, Washington, DC, USA, 2022.

24. Xinmeng Zhang, **Zhiyu Wan**, Chao Yan, Thomas Brown, Weiyi Xia, Aris Gkoulalas-Divanis, Murat Kantarcioglu, and Bradley Malin. How Adversarial Assumptions Influence Re-identification Risk Measures: A COVID-19 Case Study. In *Privacy in Statistical Databases. PSD 2022. Lecture Notes in Computer Science, vol 13463, Domingo-Ferrer, J., Laurent, M.., Eds.*, pp. 361–374, Springer, 2022.

25. Yongtai Liu, Douglas Downey, **Zhiyu Wan**, Murat Kantarcioglu, Yevgeniy Vorobeychik, Bradley Malin. De-identifying socioeconomic data at the census tract level for medical research through constraint-based clustering. In *AMIA 2019 Annual Symposium Proceedings*, pp. 793–802, American Medical Informatics Association, 2021.

26. Yongtai Liu, Chao Yan, Zhijun Yin, **Zhiyu Wan**, Weiyi Xia, Murat Kantarcioglu, Yevgeniy Vorobeychik, Ellen Wright Clayton, and Bradley Malin. Biomedical research cohort membership disclosure on social media. In *AMIA 2019 Annual Symposium Proceedings*, pp. 607–616, American Medical Informatics Association, 2019. ***获最佳论文奖***

27. Yongtai Liu, **Zhiyu Wan**, Weiyi Xia, Murat Kantarcioglu, Yevgeniy Vorobeychik, Ellen Wright Clayton, Abel Kho, David Carrell, and Bradley Malin. Detecting the presence of an individual in phenotypic summary data. In *AMIA 2018 Annual Symposium Proceedings*, pp. 760–769, American Medical Informatics Association, 2018.

28. Fabian Prasser, James Gaupp, **Zhiyu Wan**, Weiyi Xia, Yevgeniy Vorobeychik, Murat Kantarcioglu, Klaus Kuhn, and Bradley Malin. An open source tool for game theoretic health data de-identification. In *AMIA 2017 Annual Symposium Proceedings*, pp. 1430–1439, American Medical Informatics Association, 2017.

29. Weiyi Xia, Murat Kantarcioglu, **Zhiyu Wan**, Raymond Heatherly, Yevgeniy Vorobeychik, and Bradley Malin. Process-driven data privacy. In *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management*, pp. 1021–1030, ACM, 2015.

30. Xiaobo Ma, Jiahong Zhu, **Zhiyu Wan**, Jing Tao, Xiaohong Guan, and Qinghua Zheng. Honeynet-based collaborative defense using improved highly predictive blacklisting algorithm. In *2010 8th World Congress on Intelligent Control and Automation*, pp. 1283–1288, IEEE, 2010.

*：共同第一作者

预出版刊物 (PREPRINT)

31. Chris Clifton, Bradley Malin, Anna Oganian, Ramesh Raskar, Vivek Sharma, Weiyi Xia, Jeremy Seeman, **Zhiyu Wan**, and Abhishek Singh. A Roadmap for Greater Public Use of Privacy-Sensitive Government Data: Workshop Report. *arXiv*:2208.01636, 2022.

受邀讲座 (INVITED TALKS)

- **Zhiyu Wan**. Strategic Privacy-Preserving Sharing of High Dimensional Biomedical Data Enabled by Artificial Intelligence. *UTHealth School of Biomedical Informatics Spring 2023 Research Seminar*. Virtual. (February 1, 2023).
- **Zhiyu Wan**. Game Theory For Shielding Personal Data Against Privacy Attacks On Anonymity. *2022 INFORMS Annual Meeting*. Indianapolis, IN, USA. (October 16, 2022).

公开演讲或学术海报 (PRESENTATIONS/POSTERS)

- **Zhiyu Wan**. Privacy-preserving collection and sharing of unbiased human voice data for automatic assessment of voice disorders. *17th Annual Vanderbilt Postdoctoral Association Symposium*. Nashville, TN, USA (November 17, 2023).
- **Zhiyu Wan**. Privacy-preserving collection and sharing of unbiased human voice data for automatic assessment of voice disorders. *Inaugural Bridge2AI Voice Symposium*. Washington, DC, USA (April 19, 2023). ***获最佳海报奖***
- **Zhiyu Wan**. A Game Theoretic Model for Sharing Biomedical Data While Preventing Multistage Privacy Intrusions. *16th Annual Vanderbilt Postdoctoral Association Symposium*. Nashville, TN, USA. (October 27, 2022).
- **Zhiyu Wan**. Privacy-preserving collection and sharing of unbiased human voice data for automatic assessment of voice disorders and respiratory diseases: A pilot study. *Informatic Con 22*, Nashville, TN, US (September 28, 2022).
- **Zhiyu Wan**. Game Theory for Genomic Data Privacy. *2nd event of the "Speak Easy" - A Postdoc Talk Series*. Vanderbilt University, Nashville, TN. July 21, 2022.
- **Zhiyu Wan**. A Game Theoretic Model for Privacy-Preserving Genomic Data Sharing. *9th Annual Oak Ridge Postdoctoral Association Research Symposium*. Virtual. July 29, 2021.
- **Zhiyu Wan**. Game Theory Can Expand Access to Genomic Data While Promoting Privacy. *AAAI 2017 Spring Symposium on AI for Social Good*. Stanford University. March 28, 2017.
- **Zhiyu Wan**. Practical Protection of Genomic Data Sharing Through Beacon Services. *iDASH Privacy and Security Workshop*, co-located with AMIA 2016 Annual Symposium. Chicago, IL. November 11, 2016.

教学经历 (TEACHING EXPERIENCE)

- **范德堡大学 (Vanderbilt University) 教学中心 (Center for Teaching)**
  本科教学证, 完成了本科教学的研习班、实践课及专业练习 2019 年 8 月 – 2020 年 5 月
- **范德堡大学 (Vanderbilt University) 电子工程及计算机科学系 (Department of EECS)**
  客座讲师, 每学期教一节研究生课程：生物医学中的数据隐私 (CS 8396) 2018/’19/’20/’22 年春
  教学助理, 作业考试评分答疑：离散数学 (CS 2212) 及编程入门 (CS 1103) 2012 年秋及 2013 年春

荣誉及奖励 (HONORS AND AWARDS)

• 首届 Bridge2AI Voice 研讨会学术海报竞赛	**第一名**	2023 年 4 月
• 2019 美国医学信息学协会年度研讨会	**最佳论文**	2019 年 11 月
• Russell G. Hamilton 研究生领导力项目旅费资助金	**$1,000 美元**	2019 年 11 月
• 2018 国际医学信息学协会年鉴	**年度最佳论文**	2018 年 8 月
• 范德堡大学学生旅费资助金	**$500 美元**	2017 年 3 月
• 2016 iDASH 隐私安全竞赛	**第一名**	2016 年 11 月
• iDASH 旅费奖励金	**$1,250 美元**	2016 年 11 月
• 西安交通大学优秀毕业生	**最佳 2%**	2010 年 12 月
• 海航慈航奖学金（全国 60 名）	**~$3,000 美元**	2010 年 9 月
• COMAP 国际数学建模竞赛 (MCM)	**二等奖**	2010 年 2 月

学术服务 (ACADEMIC SERVICE)

- **会议环节主席 (Session Chair)**:
  - 电子电气工程师学会生物信息学与生物医学国际峰会 (BIBM ’22/’21)

- **评审委员会成员 (Committee Member)**:
  - 美国医学信息学协会临床信息学峰会 (AMIA CIC ’23)
  - 国际医学信息学协会 2022 版年鉴 (IMIAYB ’22)
  - 社会公益人工智能工作交流会 (AI4SG Workshop ’21)
  - 电子电气工程师学会生物信息学与生物医学国际峰会 (BIBM ’22/’21)
  - 人工智能工具国际峰会 (ICTAI ’22/’21/’20)
- **审稿人 (Reviewer)**:
  - 自然 - 生物技术 (Nature Biotechnology)
  - 美国计算机协会 – 隐私与安全汇刊 (ACM Transactions on Privacy and Security)
  - 应用数学建模 (Applied Mathematical Modelling)
  - 生物信息学 (Bioinformatics)
  - 生物医学中心 - 医学信息学及决策制订 (BMC Medical Informatics and Decision Making)
  - 生物医学中心-医学研究方法论 (BMC Medical Research Methodology)
  - 公共卫生前沿 (Frontiers in Public Health)
  - 电气电子工程师学会 - 可信与安全计算汇刊 (IEEE Transactions on Dependable and Secure Computing)
  - 美国医学信息学协会期刊 (Journal of the American Medical Informatics Association)
  - 公共科学图书馆杂刊 (PLoS ONE)
  - 医疗决策 – 政策与实践 (MDM Policy & Practice)
  - 医学信息方法 (Methods of Information in Medicine)
  - 数据隐私汇刊 (Transactions on Data Privacy)
  - 美国人工智能促进协会 - 人工智能峰会 (AAAI ’19/’18/’17/’16)
  - 电气电子工程师学会 - 人工智能与知识工程峰会 (AIKE ’19)
  - 美国医学信息学协会年度研讨会 (AMIA ’22/’21/’20/’19/’18/’17)
  - 电气电子工程师学会 - 大数据国际峰会 (BigData ’16/’15)
  - 计算机协会 - 数据与应用安全隐私峰会 (CODASPY ’23/’22/’21/’20/’19)
  - 数据及应用安全隐私峰会 (DBSec ’20)
  - 电气电子工程师学会 - 高可靠性系统工程国际研讨会 (HASE ’15)
  - 电气电子工程师学会 - 数据挖掘国际峰会 (ICDM ’19/’18/’17)
  - 学习表征国际峰会 (ICLR ’22)
  - 电子电气工程师学会机器学习国际峰会 (ICML ’21)
  - 人工智能国际联合峰会 (IJCAI ’19)
  - 神经信息处理系统峰会 (NeurIPS ’21)
  - 生物计算太平洋研讨会 (PSB ’21)
  - 统计数据库隐私会议 (PSD ’20)
  - 跨学科人工智能会议 (TransAI ’20) …

领导力及会员 (LEADERSHIP AND MEMBERSHIPS)

• 范德堡大学博士后联合会 (VPA)	**副主席**	2021 年 8 月 – 2023 年 5 月
• 范德堡大学中国学生学者联合会 (VUCSSA)	**副主席**	2015 年 9 月 – 2018 年 5 月
• 人工智能促进协会 (AAAI)	**学生会员**	2017 年 3 月– 2020 年 12 月
• 美国人类基因学协会 (ASHG)	**学生会员**	2016 年 9 月– 2020 年 12 月
• 美国医学信息学协会 (AMIA)	**会员**	2016 年 11 月至今
• 电气电子工程师学会 (IEEE)	**会员**	2015 年 1 月至今
• 计算机协会 (ACM)	**会员**	2014 年 1 月至今
• 运筹学及管理科学学会(INFORMS)	**会员**	2022 年 9 月– 2022 年 12 月

技术能力 (TECHNICAL SKILLS)

- 编程语言： Python, MATLAB, Java, C/C++, Delphi, UML, BASIC, MASM
- 机器学习： Pytorch, TensorFlow, Keras, Pandas, SciPy, Matplotlib, Numpy
- 大数据： Mahout, Scala, Spark, Hive, Apache Cassandra
- 云计算： Hadoop, Hbase, Amazon Web Services
- 模拟平台： GME, Formula, MATLAB Simulink, LabVIEW, OMNeT++, Deterlab
- 网络安全工具： VMware, GnuPG, OpenSSL, Snort, Netcat, Nessus, Wireshark, Sniffer
- 网站开发： HTML, PHP, ASP, JavaScript, SQL Server

课程项目 (COURSE PROJECTS)

- 大数据 (A) **基于 Mahout 的自由职业者推荐系统** 2014 年春
- 机器学习(A) **社交网络中的社群检测** 2014 年春
- 数据隐私 (A-) **对准标识符匿名性的更精确的估计** 2013 年春
- 计算机网络 (A) **带多媒体信息的即时通讯软件** 2013 年春
- 分布式系统 (A) **健康数据流中的安全与隐私保护** 2012 年秋
- 网络安全 (A) **对于信息物理系统的僵尸网络攻击** 2012 年秋
- 模型集成计算(A+) **蜜网协同防御系统的元建模及其用于防御僵尸网络的演示** 2012 年秋

其他相关课程 (OTHER RELEVANT COURSES)

- **范德堡大学**: 高级人工智能 (B+), 云计算 (A), 算法 (B+).
- **西安交通大学**: 网络安全 (A+), 计算机网络 (B+), 数据库 (B+), 操作系统 (A), 数据结构与算法 (B+), 面向对象程序设计 (A).