University of Alberta
CMPUT 626 Machine Learning and Practical Privacy
Section Number A2
Fall 2023

Instructor: Bailey Kacsmar
Office: See e-class version
E-mail: kacsmar@ualberta.ca

Office Hours: See e-class version
Lecture Room and Time: Tuesday/Thursday 3:30pm-4:50pm

This copy of the syllabus is for **reference only in advance of the course start date. Please see course website to ensure you are accessing the most up to date information.**

This syllabus is a guideline for the course and not a contract. As such, its terms may be altered when doing so is, in the opinion of the instructor(s), in the best interests of the class.

Course Description:

This is a seminar style course that focuses on current research in the space of machine learning and privacy. We will examine the literature as well as open questions in the space. Students will be expected to write reviews of published research papers from the field, present the paper to the class in the style of a research seminar presentation, and execute a novel research project that they will write up and present at the end of the term. *Prerequisites: This course has no prerequisites.*

Course Objectives and Expected Learning Outcomes:

By the end of this course students should be able to:

- Explain the notion of privacy within the machine learning space.
- Analyze security and privacy of machine learning protocols.
- Evaluate research on privacy of machine learning and articulate advantages and limitations.

Course Information

Email: Important course information will generally be posted to the course website, but may also be sent to your ualberta.ca email address. For personal matters, such as an illness, please email the instructors directly. We will only reply back to email from your ualberta.ca email address, following privacy rules.

Lectures: *Tuesday/Thursday 3:30-4:50pm*, content will also be posted to e-class. It is your responsibility to keep up with all course-related information posted to e-class.

e-class: `https://eclass.srv.ualberta.ca/enrol/instances.php?id=90128`

Textbooks and Readings: There is no required textbook. Additional readings will be assigned, and will appear on e-class; readings marked as mandatory contain required material for the course. You must read these mandatory readings.

Outline

Part 1: Instructional

- Privacy and Technology
  - Solove - Nothing to hide
  - Westin - Privacy categories
  - Nissenbaum - Contextual Integrity
  - Dwork - Differential Privacy
  - MPC Chapter
- Ethics, law, and policy relevant to this course
  - University policy
  - GDPR, PIPEDA, CCPA, etc.
- Basics of cryptography (e.g., Symmetric encryption, Hash functions, MAC, Public key encryption (RSA), Semantic security, etc.)
- Attacks and Adversaries

Part 2: Seminar

- A selection of papers presented by students in each class.
- See references section for the papers details
- See course overview table for which day a paper is scheduled to be presented
- If you are aware of a paper that you would really like to present, but it is not listed in the course, let the instructor know, and it will be considered for inclusion.

Grading Scheme

- 20% Seminar-style presentations as discussion lead (2-3 throughout the term)
- 5% Quality of feedback on peers
- 20% Paper reviews (2 per week)
- 10% Participation in paper discussions
- 10% Project proposal (Due October 13, 2023 at 4pm MT)
- 25% Final project report (Due December 8, 2023 at 4pm MT)
- 10% Final project presentation (To be scheduled. Nov 28, 30, Dec 5, 7)

Seminar-Style Presentations

During the weeks we have student presentations there will typically be three student presenters per class. Each presenter will create slides for a 15 minute presentation, followed by leading a 10 minute discussion section. Your presentation should highlight the research questions, methodology, results, and take-aways/impacts of the work.

Paper Reviews

During the weeks we have student presentations and the day the instructor does paper presentations you will be expected to write a review of a subset of the papers from the week. Each class there are paper presentations you are expected to read at least two of the three papers. However, you only have to write a review for one of the papers presented that day. Reviews must be submitted before the start of class (specifically thirty minutes before the start of class). For each review you will:

- Write a short paper summary (3-6 sentences)
- Identify the contributions of the paper
- Identify the research question(s) of the paper
- Identify the strong components of the paper (e.g., reasons likely associated with its acceptance, strong executions, etc). Identify 2-4 such attributes
- Identify weak components of the paper (e.g., revisions that would improve validity, aspects that could improve breadth/impact, etc). Identify 2-4 such attributes.
- A 1-2 sentence statement as to why you think this paper was included in the course/it's relationship to the content thus far.
- Identify one possible research question that could be follow up for this paper.

Peer Review

Each student will receive feedback on their seminar presentations from each course participant. The presenting students' grade is not affected by these evaluations. Rather what is graded is the quality of the feedback being given. For example, the feedback ”your presentation was really bad” is not helpful. Rather, a possible helpful comment is ”Slide n the amount of content on this slide made it difficult to know what to focus on. Consider splitting into two slides, using boxes to emphasize what to focus on, or removing any non-critical content”. Similarly, you could say ”the pacing of how you organized the introduction was really nice and made it easy to follow the flow of the motivation behind the paper”. This feedback will be submitted via e-class to each presenter before the next class.

Research Project

Projects will be done in teams of two. Exceptions (e.g., groups of one or three) are permissible only if they have acquired prior approval from the instructor before the proposal deadline. Note that a group of three would be expected to accomplish “more” than a group of two proportionately.

In writing your paper, you must become familiar with the research literature relevant to your topic. Your focus should be on academic venues, such as the USENIX Security Symposium, ACM CCS, IEEE Symposium on Security and Privacy, Privacy Enhancing Technologies Symposium (PETS), or the NDSS Symposium. You should email your topic, proposal, and paper to the instructors.

Topic approval: Your topic must be approved in advance by the instructors before you submit your proposal. Email the instructor at least one week before the proposal deadline (October5, 2023) with a brief (1-3 sentence) description of your intended topic.

Proposal: Your proposal should be one page in length and include at least 10 references, preferably including (but not limited to) papers from the aforementioned venues. It is recommended but not required that you discuss the proposal with the instructors first. Email your proposal to the instructors by Friday October 13, 2023 by 4pm MT.

Paper: Your paper should include related work (a summary of past and current work on your topic; you should provide a concise summary of work, emphasizing major accomplishments, rather than a detailed accounting of individual pieces of research activity). Email your final paper to the instructor by December 8th, 2023.

Format: Your proposal and paper paper should be formatted in the two-column ACM proceedings format, using one of the ACM SIG Proceedings Templates. Your paper should not be longer than eight pages. The ACM templates include headings for "Categories and Subject Descriptors", "General Terms", and "Keywords", which you do not need to use.

Presentation: Details of presentations will be available a few weeks into the course. It is dependent on the number of students for scheduling. It should be a presentation of your research similar to how you will have been presenting other peoples' research throughout the term.

Course Policy Information

Remarking Policy: If you have an assignment that you would like to have reappraised, please email the instructor to submit your request. Include a clear justification for your claims. The appeals deadline is **one week** after the respective graded item is first made available. If your appeal is concerned with a simple calculation error, please email me or speak with me during my office hours.

Missed or Late Assessments: Please start working on the material in advance of the deadlines. To motivate you to do so, we may require you to submit milestones for some or all of them. Late submissions for the project proposal and project proposal will be accepted only up to 72 hours after the original due date. All other graded components (paper reviews, peer feedback, and presentations must be done on time). There is no penalty for accepted late submissions. Course personnel will not normally give assistance after the original due dates.

Security Information: In this course, you will be exposed to information about security problems and vulnerabilities with computing systems and networks. To be clear, you are not to use this or any other similar information to test the security of, break into, compromise, or otherwise attack, any system or network without the express consent of the owner.

University Policy Information

Academic Integrity and Student Conduct:

The University of Alberta is committed to the highest standards of academic integrity and honesty, as well as maintaining a learning environment that fosters the safety, security, and the inherent dignity of each member of the community, ensuring students conduct themselves accordingly. Students are expected to be familiar with the standards of academic honesty and appropriate student conduct, and to uphold the policies of the University in this respect. Students are particularly urged

Week	**Tuesday**	**Thursay**
One: September 5th and 7th	Course overview, Privacy Part 1 Bailey	Privacy Part 2 Bailey
Two: September 12th and 14th	Ethics, law, and policy Bailey	Cryptography Part 1 Bailey
Three: September 19th and 21st	Cryptography Part 2 Bailey	Attacks, Adversaries Bailey
Four: September 26th and 28th	Paper Presentations Bailey [26]	Paper Presentations Student i Student ii
Five October 3rd and 5th	Paper Presentations Student i Student ii	Paper Presentations Student i Student ii
Six: October 10th and 12th	Paper Presentations Student i Student ii	Paper Presentations Student i Student ii
Seven: October 17th and 19th	Proposal Presentations Order TBA	Paper Presentations Student i Student ii
Eight: October 24th and 26th	Paper Presentations Student i Student ii	Paper Presentations Student i Student ii
Nine: October 31 and November 2nd	Paper Presentations Student i Student ii	Paper Presentations Student i Student ii
Ten: November 7th and 9th	Paper Presentations Student i Student ii	Paper Presentations Student i Student ii
READING WEEK NO CLASS	-	-
Eleven: November 21st and 23rd	Paper Presentations Student i Student ii	Paper Presentations Student i Student ii
Twelve: November 28th and 30th	ACM CCS Conference TBA	ACM CCS Conference TBA
Thirteen: December 5th and 7th	Project Presentations TBA	Project Presentations TBA

to familiarize themselves with the provisions of the Code of Student Behaviour and the Student Conduct Policy, and avoid any behaviour that could potentially result in suspicions of academic misconduct (e.g., cheating, plagiarism, misrepresentation of facts) and non-academic misconduct (e.g., discrimination, harassment, physical assault). Academic and non-academic misconduct are taken very seriously and can result in suspension or expulsion from the University.

All students are expected to consult the Academic Integrity website for clarification on the various academic offences. All forms of academic dishonesty are unacceptable at the University.

Any suspected academic offence in this course will be reported to the College of Natural and Applied Sciences. Suspected cases of non-academic misconduct will be reported to the Dean of Students. The College, the Faculty of Science, and the Dean of Students are committed to student rights and responsibilities, and adhere to due process and administrative fairness, as outlined in the Code of Student Behaviour and the Student Conduct Policy. Anyone who is found in violation is likely to receive a sanction. Typical sanctions for academic misconduct include conduct probation, a mark reduction or a mark of 0 on an assessment, a grade reduction or a grade of F in a course, a remark on the transcript, and a recommendation for suspension or expulsion. Sanctions for non-academic misconduct include conduct conditions, fines, suspension of essential or non-essential University services and resources, and suspension or expulsion from the University.

Appropriate Collaboration:

Students are not permitted to copy solutions on homework assignments. Here are some tips to avoid copying on assignments:

1. Do not write down something that you cannot explain to your instructor.

2. When you are helping other students, avoid showing them your work directly. Instead, explain your solution verbally. Students whose work is copied also receive academic sanctions.

3. If you find yourself reading another student's solution, do not write anything down. Once you understand how to solve the problem, remove the other person's work from your sight and then write up the solution to the question yourself. Looking back and forth between someone else's paper and your own paper is almost certainly copying and will result in academic sanctions for both you and your fellow student.

4. If the instructor or TA writes down part of a solution in order to help explain it to you or the class, you cannot copy it and hand it in for credit. Treat it the same way you would treat another student's work with respect to copying, that is, remove the explanation from your sight and then write up the solution yourself.

5. There is often more than one way to solve a problem. Choose the method that makes the most sense to you rather than the method that other students happen to use. If none of the ideas in your solution are your own, there is a good chance it will be flagged as copying.

Students Eligible for Accessibility-Related Accommodations:

In accordance with the University of Alberta's Discrimination, Harassment, and Duty to Accommodate policy, accommodation support is available to eligible students who encounter limitations or restrictions to their ability to perform the daily activities necessary to pursue studies at a post-secondary level due to medical conditions and/or non-medical protected grounds. Accommodations are coordinated through the Academic Success Centre, and students can learn more about eligibility on the Register for Accommodations website.

It is recommended that students apply as early as possible in order to ensure sufficient time to complete accommodation registration and coordination. Students are advised to review and adhere to published deadlines for accommodation approval and for specific accommodation requests (e.g., exam registration submission deadlines). Students who request accommodations less than a month in advance of the academic term for which they require accommodations may experience unavoidable delays or consequences in their academic programs, and may need to consider alternative academic schedules.

Academic Success Center: The Academic Success Centre (ASC) provides services to support University of Alberta students in the areas of accommodations, learning, and writing. The ASC coordinates reasonable accommodations to eligible students who encounter medical or non-medical restrictions to their ability to perform the daily activities necessary to pursue studies at a post-secondary level. To that end, they work with students to coordinate disability-related accommodation needs for participation in University programs. For more information, and to register for services, visit the Academic Accommodations webpage.

The ASC also provides peer-based and professional academic support in the areas of learning and writing. They offer individual appointments, group workshops, and online courses to students in all University of Alberta programs, and at all levels of achievement and study.

At Writing Services, undergraduate students can work with a peer tutor to get feedback on a draft of their paper. Graduate students can book an appointment with a graduate writing advisor to get feedback on their documents. For more information, please visit the Writing Services webpage.

Recording and/or Distribution of Course Materials:

Audio or video recording, digital or otherwise, of lectures, labs, seminars or any other teaching environment by students is allowed only with the prior written consent of the instructor or as a part of an approved accommodation plan. Student or instructor content, digital or otherwise, created and/or used within the context of the course is to be used solely for personal study, and is not to be used or distributed for any other purpose without prior written consent from the content author(s).

References

[1] M. Abadi, A. Chu, I. Goodfellow, H. B. McMahan, I. Mironov, K. Talwar, and L. Zhang, "Deep Learning with Differential Privacy," in *the 2016 ACM SIGSAC Conference on Computer and Communications Security*, Vienna, Austria: ACM, 2016, pp. 308–318.

[2] N. Agrawal, R. Binns, M. Van Kleek, K. Laine, and N. Shadbolt, "Exploring Design and Governance Challenges in the Development of Privacy-Preserving Computation," in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, 2021, pp. 1–13.

[3] S. Aonzo, Y. Han, A. Mantovani, and D. Balzarotti, "Humans vs. machines in malware classification," *Proc. of USENIX-23*, 2023.

[4] E. Bagdasaryan, O. Poursaeed, and V. Shmatikov, "Differential Privacy has Disparate Impact on Model Accuracy," *Advances in Neural Information Processing Systems*, vol. 32, 2019.

[5] D. G. Balash, R. A. Fainchtein, E. Korkes, M. Grant, M. Sherr, and A. J. Aviv, "Educators' Perspectives of Using (or Not Using) Online Exam Proctoring," *Proc. of USENIX-23*, 2023.

[6] F. Boenisch, C. Mühl, R. Rinberg, J. Ihrig, and A. Dziedzic, "Individualized PATE: Differentially Private Machine Learning with Individual Privacy Guarantees," *Proceedings on Privacy Enhancing Technologies*, vol. 1, pp. 158–176, 2023.

[7] K. Bonawitz, V. Ivanov, B. Kreuter, A. Marcedone, H. B. McMahan, S. Patel, D. Ramage, A. Segal, and K. Seth, "Practical Secure Aggregation for Privacy-Preserving Machine Learning," in *the 2017 ACM SIGSAC Conference on Computer and Communications Security*, Dallas, TX, USA: ACM, 2017, pp. 1175–1191.

[8] B. Bullek, S. Garboski, D. J. Mir, and E. M. Peck, "Towards understanding differential privacy: When do people trust randomized response technique?" In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 2017, pp. 3833–3837.

[9] N. Carlini, C. Liu, Ú. Erlingsson, J. Kos, and D. Song, “The Secret Sharer: Evaluating and Testing Unintended Memorization in Neural Networks,” in *28th USENIX Security Symposium (USENIX Security 19)*, 2019, pp. 267–284.

[10] K. Chaudhuri, C. Monteleoni, and A. D. Sarwate, “Differentially Private Empirical Risk Minimization,” *Journal of Machine Learning Research*, vol. 12, no. Mar, pp. 1069–1109, 2011.

[11] C. A. Choquette-Choo, F. Tramer, N. Carlini, and N. Papernot, “Label-only Membership Inference Attacks,” in *International Conference on Machine Learning*, PMLR, 2021, pp. 1964–1974.

[12] R. Cummings, G. Kaptchuk, and E. M. Redmiles, “‘I Need a Better Description’: An Investigation Into User Expectations For Differential Privacy,” in *Proceedings of the 2021 ACM SIGSAC Conference on Computer and Communications Security*, ser. CCS ’21, Virtual Event, Republic of Korea: Association for Computing Machinery, 2021, 3037–3052, ISBN: 9781450384544. DOI: 10.1145/3460120.3485252. [Online]. Available: https://doi.org/10.1145/3460120.3485252.

[13] S. Dambra, L. Bilge, P. Kotzias, Y. Shen, and J. Caballero, “One Size Does not Fit All: Quantifying the Risk of Malicious App Encounters for Different Android User Profiles,” *Proc. of USENIX-23*, 2023.

[14] A. Dionysiou, V. Vassiliades, and E. Athanasopoulos, “Exploring Model Inversion Attacks in the Black-box Setting,” *Proceedings on Privacy Enhancing Technologies*, vol. 1, pp. 190–206, 2023.

[15] A. R. Elkordy, J. Zhang, Y. H. Ezzeldin, K. Psounis, and S. Avestimehr, “How Much Privacy Does Federated Learning with Secure Aggregation Guarantee?” *Proceedings on Privacy Enhancing Technologies*, vol. 1, pp. 510–526, 2023.

[16] D. Franzen, S. Nuñez von Voigt, P. Sörries, F. Tschorsch, and C. Müller-Birn, “Am I Private and If So, how Many? Communicating Privacy Guarantees of Differential Privacy with Risk Communication Formats,” in *Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security*, New York, New York: ACM, 2022, pp. 1125–1139.

[17] M. Fredrikson, S. Jha, and T. Ristenpart, “Model inversion attacks that exploit confidence information and basic countermeasures,” in *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security*, ACM, ACM, 2015, pp. 1322–1333.

[18] C. Fung, C. J. Yoon, and I. Beschastnikh, “The Limitations of Federated Learning in Sybil Settings,” in *the 23rd International Symposium on Research in Attacks, Intrusions and Defenses*, 2020.

[19] K. Hamada, D. Ikarashi, R. Kikuchi, and K. Chida, “Efficient decision tree training with new data structure for secure multi-party computation,” *Proceedings on Privacy Enhancing Technologies*, vol. 1, pp. 343–364, 2023.

[20] J. M. Haney and W. G. Lutters, “‘It’s {Scary… It’s}{Confusing… It’s} Dull’: How Cybersecurity Advocates Overcome Negative Perceptions of Security,” in *Fourteenth Symposium on Usable Privacy and Security (SOUPS 2018)*, 2018, pp. 411–425.

[21] B. Hitaj, G. Ateniese, and F. Perez-Cruz, “Deep Models Under the GAN: Information Leakage from Collaborative Deep Learning,” in *the 2017 ACM SIGSAC Conference on Computer and Communications Security*, Dallas, TX, USA: ACM, 2017, pp. 603–618.

[22] B. Hui, Y. Yang, H. Yuan, P. Burlina, N. Z. Gong, and Y. Cao, “Practical Blind Membership Inference Attack via Differential Comparisons,” *Network and Distributed System Security Symposium (NDSS)*, 2021.

[23] T. Humphries, S. Oya, L. Tulloch, M. Rafuse, I. Goldberg, U. Hengartner, and F. Kerschbaum, “Investigating Membership Inference Attacks under Data Dependencies,” in *2023 2023 IEEE 36th Computer Security Foundations Symposium (CSF)(CSF)*, IEEE Computer Society, 2023, pp. 194–209.

[24] X. Jiang, X. Zhou, and J. Grossklags, “Comprehensive analysis of privacy leakage in vertical federated learning during prediction.,” *Proc. Priv. Enhancing Technol.*, vol. 2022, no. 2, pp. 263–281, 2022.

[25] B. Kacsmar, V. Duddu, K. Tilbury, B. Ur, and F. Kerschbaum, “Comprehension from Chaos: Towards Informed Consent for Private Computation,” Preprint, 2023.

[26] B. Kacsmar, K. Tilbury, M. Mazmudar, and F. Kerschbaum, “Caring about Sharing: User Perceptions of Multiparty Data Sharing,” in *31st USENIX Security Symposium*, USENIX Association, Boston, MA, 2022.

[27] Y. Kaya and T. Dumitras, “When Does Data Augmentation Help With Membership Inference Attacks?” In *International Conference on Machine Learning*, PMLR, 2021, pp. 5345–5355.

[28] P. G. Kelley, C. Cornejo, L. Hayes, E. S. Jin, A. Sedley, K. Thomas, Y. Yang, and A. Woodruff, “‘There will be less privacy, of course’: How and why people in 10 countries expect AI will affect privacy in the future,” in *Nineteenth Symposium on Usable Privacy and Security (SOUPS 2023)*, Anaheim, CA: USENIX Association, Aug. 2023, pp. 579–603, ISBN: 978-1-939133-36-6. [Online]. Available: https://www.usenix.org/conference/soups2023/presentation/kelley.

[29] B. Kulynych, M. Yaghini, G. Cherubin, M. Veale, and C. Troncoso, “Disparate Vulnerability to Membership Inference Attacks,” *Proceedings on Privacy Enhancing Technologies*, vol. 1, pp. 460–480, 2022.

[30] C. Liu, S. Chakraborty, and P. Mittal, “Dependence Makes You Vulnberable: Differential Privacy Under Dependent Tuples,” in *NDSS*, vol. 16, 2016, pp. 21–24.

[31] Y. Liu, R. Wen, X. He, A. Salem, Z. Zhang, M. Backes, E. De Cristofaro, M. Fritz, and Y. Zhang, “ML-Doctor: Holistic Risk Assessment of Inference Attacks Against Machine Learning Models,” in *31st USENIX Security Symposium (USENIX Security 22)*, Boston, MA: USENIX Association, Aug. 2022. [Online]. Available: https://www.usenix.org/conference/usenixsecurity22/presentation/liu-yugeng.

[32] T. Marjanov, M. Konstantinou, M. Jóźwiak, and D. Spagnuelo, “Data Security on the Ground: Investigating Technical and Legal Requirements under the GDPR,” *Proceedings on Privacy Enhancing Technologies*, vol. 3, pp. 405–417, 2023.

[33] A. M. McDonald and L. F. Cranor, “The Cost of Reading Privacy Policies,” *ISJLP*, vol. 4, p. 543, 2008.

[34] H. B. McMahan, D. Ramage, K. Talwar, and L. Zhang, “Learning Differentially Private Recurrent Language Models,” in *International Conference on Learning Representations*, 2018.

[35] H. Mozaffari, V. Shejwalkar, and A. Houmansadr, “Every Vote Counts: Ranking-Based Training of Federated Learning to Resist Poisoning Attacks.”

[36] M. Nasr, J. Hayes, T. Steinke, B. Balle, F. Tramèr, M. Jagielski, N. Carlini, and A. Terzis, “Tight Auditing of Differentially Private Machine Learning,” *Proc. of USENIX-23*, 2023.

[37] H. Nissenbaum, “Contextual Integrity Up and Down the Data Food Chain,” *Theoretical Inquiries in Law*, vol. 20, no. 1, pp. 221–256, 2019.

[38] M. Oates, Y. Ahmadullah, A. Marsh, C. Swoopes, S. Zhang, R. Balebako, and L. F. Cranor, "Turtles, Locks, and Bathrooms: Understanding Mental Models of Privacy Through Illustration," *Proceedings on Privacy Enhancing Technologies*, vol. 2018, no. 4, pp. 5–32, 2018.

[39] N. Papernot, M. Abadi, Úlfar Erlingsson, I. Goodfellow, and K. Talwar, "Semi-supervised Knowledge Transfer for Deep Learning from Private Training Data," in *the International Conference on Learning Representations*, Toulon, France, 2017.

[40] X. Qi, T. Xie, T. Wang, T. Wu, S. Mahloujifar, and P. Mittal, "Towards A Proactive ML Approach for Detecting Backdoor Poison Samples," 2023.

[41] L. Qin, A. Lapets, F. Jansen, P. Flockhart, K. D. Albab, I. Globus-Harris, S. Roberts, and M. Varia, "From Usability to Secure Computing and Back Again," in *Fifteenth Symposium on Usable Privacy and Security ({SOUPS} 2019)*, 2019, pp. 191–210.

[42] W. Qiu, D. Lie, and L. Austin, "Calpric: Inclusive and Fine-grain Labeling of Privacy Policies with Crowdsourcing and Active Learning," 2023.

[43] A. Sablayrolles, M. Douze, C. Schmid, Y. Ollivier, and H. Jégou, "White-box vs black-box: Bayes Optimal Strategies for Membership Inference," in *International Conference on Machine Learning*, PMLR, 2019, pp. 5558–5567.

[44] G. Sandoval, H. Pearce, T. Nys, R. Karri, S. Garg, and B. Dolan-Gavitt, "Lost at c: A user study on the security implications of large language model code assistants," *Proc. of USENIX-23*, 2023.

[45] R. Shokri, M. Stronati, C. Song, and V. Shmatikov, "Membership Inference Attacks Against Machine Learning Models," in *the 2017 IEEE Symposium on Security and Privacy*, IEEE, San Jose, CA, USA, 2017, pp. 3–18.

[46] M. A. Smart, D. Sood, and K. Vaccaro, "Understanding Risks of Privacy Theater with Differential Privacy," in *CSCW 2022*, New York, New York: ACM, 2022.

[47] L. Song and P. Mittal, "Systematic Evaluation of Privacy Risks of Machine Learning Models," in *30th USENIX Security Symposium (USENIX Security 21)*, 2021, pp. 2615–2632.

[48] S. Truex, N. Baracaldo, A. Anwar, T. Steinke, H. Ludwig, R. Zhang, and Y. Zhou, "A Hybrid Approach to Privacy-Preserving Federated Learning," in *the 12th ACM Workshop on Artificial Intelligence and Security*, London, England: ACM, 2019, pp. 1–11.

[49] S. Truex, L. Liu, M. E. Gursoy, L. Yu, and W. Wei, "Demystifying Membership Inference Attacks in Machine Learning as a Service," *IEEE Transactions on Services Computing*, 2019.

[50] A. Woodruff, V. Pihur, S. Consolvo, L. Brandimarte, and A. Acquisti, "Would a privacy fundamentalist sell their DNA for $1000... if nothing bad happened as a result? The Westin Categories, Behavioral Intentions, and Consequences," in *10th Symposium On Usable Privacy and Security 2014)*, 2014, pp. 1–18.

[51] J. Wu and D. Zappala, "When is a tree really a truck? Exploring Mental Models of Encryption," in *Fourteenth Symposium on Usable Privacy and Security (SOUPS 2018)*, Baltimore, MD: USENIX, 2018, pp. 395–409.

[52] Y. Wu, S. Cai, X. Xiao, G. Chen, and B. C. Ooi, "Privacy Preserving Vertical Federated Learning for Tree-based Models," *Proceedings of the VLDB Endowment*, vol. 13, no. 11,

[53] A. Xiong, T. Wang, N. Li, and S. Jha, "Towards Effective Differential Privacy Communication for Users' Data Sharing Decision and Comprehension," *arXiv preprint arXiv:2003.13922*`arXiv:2003.13922`, 2020.

[54] Y. Yang, B. Hui, H. Yuan, N. Gong, and Y. Cao, “PrivateFL: Accurate, Differentially Private Federated Learning via Personalized Data Transformation,” 2023.

[55] S. Yeom, I. Giacomelli, M. Fredrikson, and S. Jha, “Privacy Risk in Machine Learning: Analyzing the Connection to Overfitting,” in *2018 IEEE 31st computer security foundations symposium (CSF)*, IEEE, 2018, pp. 268–282.