Workshop on Privacy in Natural Language Processing (PrivateNLP at WSDM 2020)

Oluwaseyi Feyisetan*
Amazon
Seattle, Washington, USA
sey@amazon.com

Sepideh Ghanavati
University of Maine
Orono, Maine, USA
sepideh.ghanavati@maine.edu

Patricia Thaine
University of Toronto
Toronto, Ontario, Canada
pthaine@cs.toronto.edu

ABSTRACT

Privacy-preserving data analysis has become essential in Machine Learning (ML), where access to vast amounts of data can provide large gains the in accuracies of tuned models. A large proportion of user-contributed data comes from natural language e.g., text transcriptions from voice assistants. It is therefore important for curated natural language datasets to preserve the privacy of the users whose data is collected and for the models trained on sensitive data to only retain non-identifying (i.e., generalizable) information. The workshop aims to bring together researchers and practitioners from academia and industry to discuss the challenges and approaches to designing, building, verifying, and testing privacy-preserving systems in the context of Natural Language Processing (NLP).

CCS CONCEPTS

• **Security and privacy → Privacy protections**.

ACM Reference Format:
Oluwaseyi Feyisetan, Sepideh Ghanavati, and Patricia Thaine. 2020. Workshop on Privacy in Natural Language Processing (PrivateNLP at WSDM 2020). In *Proceedings of Workshop on Privacy in Natural Language Processing (PrivateNLP '20).* Houston, TX, USA, 2 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

The collection of user data has grown dramatically in recent years, raising concerns about the aggregation of sensitive data and the high risk of personally identifiable information leaks. In response, methods for privacy-preserving data analysis have been proposed to protect individual information while maintaining the utility of large quantities of aggregated data.

Preserving the privacy of training data has become essential to guaranteeing data security and to maintaining user trust for continuous access to vast amounts of data that can provide significant model performance gains. As a result, significant research has been done to provide quantifiable guarantees that a user's contribution to a system cannot be linked back to their existence within the underlying dataset. In statistical data analysis, the theoretical framework of Differential Privacy (DP) has primarily been used. While methods such as DP focus on numeric data, a large proportion of user contributions comes not in the form of statistical queries, but natural language e.g., search queries, emails, reviews, comments, or text transcriptions from the increasingly ubiquitous voice assistants.

User-generated data can be sensitive both because of the explicit and the implicit information they contain. For example, in web search systems, a user can disclose their identity or a personal preference during their query interactions either explicitly (e.g., by issuing vanity queries) or implicitly (e.g., age, gender, and nationality can be determined by the way a query is written). Explicit personally identifiable information (PII), such as an individual's PIN or SSN, can potentially be filtered out via rules or pattern matching. However, more subtle privacy attacks occur when seemingly innocuous information (combined in aggregate and in the presence of side knowledge), is used to discern the private details of an individual. A classical example can be seen from the privacy breach in the 'anonymized' AOL search logs of 2006.[1]

In addition to keeping data secure and maintaining user trust, privacy-preserving techniques now need to be more robust than ever before for companies and researchers to comply with regulations such as the General Data Protection Regulation (GDPR) and the California Consumer Privacy Act (CCPA). Failure to comply has led to a number of hefty fees. As a result, questions which were mainly of interest to the research community are getting increasingly more attention from companies: (1) how can we create and curate NLP datasets while preserving the privacy of the users whose data is collected? (2) How do we train ML models that only retain pseudonymized user data?

To address these challenges, researchers have started exploring different NLP-based methods to remove PII. These methods range from simple pattern matching techniques to advanced uses of deep learning on embedded representations. These tend to fall short because they are trained on, and operate only on *known* types of sensitive entities – signifying that more research is required. This has led to a rapidly growing research field at the intersection of NLP and Privacy.

The topic's growth has been accompanied by the creation of privacy workshop series such as PPML (Privacy Preserving Machine Learning) and TPDP (Theory and Practice of Differential Privacy) at NeurIPS and CCS. However, there are no privacy workshop focusing specifically on NLP, which has its own set of specific privacy problems.

To this end, we are holding the PrivateNLP workshop to focus on this sub-field and emerging NLP-specific challenges. The workshop aims to consolidate privacy research with NLP community. It will also help foster greater collaboration within the community and strengthen the bond between academic and industry researchers.

Our primary motivation is to advance the sub-field of privacy in text data, which is fundamental in NLP research. Additionally, the workshop can also raise awareness of privacy-related issues within NLP and can be beneficial to those not actively working in the area.

[1] A Face Is Exposed for AOL Searcher No. 4417749. https://www.nytimes.com/2006/08/09/technology/09aol.html

2 OBJECTIVES AND SCOPE

The workshop covers aspects of text-based privacy research including, but not limited to:

- Generating privacy preserving test sets
- Inference and identification attacks
- Generating differentially private derived data
- NLP, privacy and regulatory compliance
- Private Generative Adverserial Networks
- Privacy in Active Learning and Crowdsourcing
- Privacy and Federated Learning in NLP
- User perceptions on privatized personal data
- Auditing provenance in language models
- Continual learning under privacy constraints
- NLP and summarization of privacy policies
- Ethical ramifications of AI/NLP in support of usable privacy

3 ORGANIZATION AND PROGRAM

3.1 Organizers

Oluwaseyi Feyisetan (Amazon, USA). Oluwaseyi Feyisetan is an Applied Scientist at Amazon Alexa where he works on Differential Privacy and Privacy Auditing mechanisms within the context of Natural Language Processing. He holds 2 pending patents with Amazon on preserving privacy in NLP systems. He completed his PhD at the University of Southampton in the UK and has published in top tier conferences and journals on crowdsourcing, homomorphic encryption, and privacy in the context of Active Learning and NLP. He has served as a reviewer at top NLP conferences including ACL and EMNLP. He is the lead organizer of the Workshop on Privacy and Natural Language Processing (PrivateNLP) at WSDM with an upcoming event scheduled for EMNLP. Prior to working at Amazon in the US, he spent 7 years in the UK where he worked at different startups and institutions focusing on regulatory compliance, machine learning and NLP within the finance sector, most recently, at the Bank of America.

Sepideh Ghanavati (University of Maine, USA). Assistant professor in Computer Science at the University of Maine. She is also the director of Privacy Engineering - Regulatory Compliance Lab (PERC_Lab). Her research interests are in the areas of information privacy and security, software engineering, machine learning and the Internet of Things (IoT). Previously, she worked as an assistant professor at Texas Tech University, visiting assistant professor at Radboud University in Nijmegen, the Netherlands and as a visiting faculty at the Institute for Software Research at Carnegie Mellon University. She is the recipient of Google Faculty Research award in 2018. She has more than 10 years of academic and industry experience in the area of privacy and regulatory compliance especially in the healthcare domain and has published more than 25 peer-reviewed publications. She was a co-organizer of the ‘Privacy and Language Technologies’ at the 2019 AAAI Spring Symposium and has been part of the organizing committee of several workshops and conferences in the past.

Patricia Thaine (Univ. of Toronto, Canada). Patricia Thaine is a PhD Candidate at the Department of Computer Science (University of Toronto) doing research on Privacy-Preserving Natural Language Processing, with a special focus on Applied Cryptography. She also does research on computational methods for lost language decipherment. Patricia is a recipient of the NSERC Postgraduate Scholarship, the RBC Graduate Fellowship, the Beatrice ‘Trixie’ Worsley Graduate Scholarship in Computer Science, and the Ontario Graduate Scholarship. She has eight years of research and software development experience, including at the McGill Language Development Lab, the University of Toronto’s Computational Linguistics Lab, the University of Toronto’s Department of Linguistics, and the Public Health Agency of Canada. She is the Co-Founder and CEO of Private AI, the former President of the Computer Science Graduate Student Union at the University of Toronto, and a member of the Board of Directors of Equity Showcase, one of Canada’s oldest not-for-profit charitable organizations.

3.2 Program Committee

Aleksei Triastcyn (Ecole Polytechnique Federale de Lausanne), Andreas Nautsch (EURECOM), Arne Kahn (Saarland University), Avi Arampatzis (Democritus University of Thrace), Asma Eidhah Aloufi (Rochester Institute of Technology), Benjamin Zi Hao Zhao (University of New South Wales), Borja Balle (DeepMind), Claire McKay Bowen (Los Alamos National Laboratory), Congzheng Song (Cornell), Dinusha Vatsalan (Data61-CSIRO), Elette Boyle (IDC Herzliya), Fang Liu (University of Notre Dame), Isar Nejadgholi (National Research Council Canada), Jamie Hayes (University College London), Jason Xue (University of Adelaide), Julius Adebayo (MIT), Kambiz Ghazinour (State University of New York), Liwei Song (Princeton), Luca Melis (Amazon USA), Mark Dras (Macquarie University), Maximin Coavoux (University of Edinburgh), Mitra Bokaei Hosseini (St. Mary’s University), Natasha Fernandes (Macquarie University), Nedelina Teneva (Amazon USA), Olya Ohrimenko (Microsoft Research), Pauline Anthonysamy (Google), Sai Teja Peddinti (Google), Shomir Wilson (Pennsylvania State University), Tom Diethe (Amazon UK), Travis Breaux (Carnegie Mellon University)

3.3 Keynote Speaker

Tom Diethe (Amazon UK). Tom Diethe is an Applied Science Manager in Amazon Research, Cambridge UK. Tom is also an Honorary Research Fellow at the University of Bristol. Tom was formerly a Research Fellow for the “SPHERE” Interdisciplinary Research Collaboration, which is designing a platform for eHealth in a smart-home context. This platform is currently being deployed into homes throughout Bristol.

Tom specializes in probabilistic methods for machine learning, applications to digital healthcare, and privacy enhancing technologies. He has a Ph.D. in Machine Learning applied to multivariate signal processing from UCL, and was employed by Microsoft Research Cambridge where he co-authored a book titled ‘Model-Based Machine Learning.’ He also has significant industrial experience, with positions at QinetiQ and the British Medical Journal. He is a fellow of the Royal Statistical Society and a member of the IEEE Signal Processing Society.

4 ACKNOWLEDGMENTS

The organizers would like to thank the program committee for their service, as well as the the authors who submitted their research to the workshop.