

Eugene Bagdasaryan

✉ eugene@cs.cornell.edu
🌐 cs.cornell.edu/~eugene

Research Interests

I study security and privacy in emerging machine learning technologies under real-life conditions and attacks. My goal is to build ML systems that are ethical, safe, and private by design.

Education

Cornell University, New York, NY, USA

2016 – 2023* PhD in Computer Science. Advised by Vitaly Shmatikov and Deborah Estrin
Aug expected

2016 – 2019 MSc in Computer Science
Aug Dec

Bauman Moscow State Technical University, Moscow, Russia

2009 – 2016 Engineer's degree in Computer Science, *summa cum laude*
Sep June

2009 – 2013 BS in Computer Science, *summa cum laude*
Sep June

Awards and Honors

- 2021 Apple Scholars in AI/ML PhD Fellowship
- 2019 Digital Life Initiative Doctoral Fellowship
- 2017 Bloomberg Data For Good Exchange Award
- 2017 Computer Science Dept TA Excellence Award
- 2011,'12,'13 Potanin Foundation Scholarship
- 2011,'12 Bauman University Academic Excellence Fellowship

Full-time Experience

2014 – 2016 **Software Engineer II**, *Cisco Innovation Center*, Moscow, Russia
Sep Jul
Worked on DevOps tools for large-scale OpenStack cloud deployments.

Internships

- 2021 – 2021 **Research Intern**, *Apple*, Cupertino, CA, USA
May Aug
Conducted research on federated learning and language models.
- 2020 – 2020 **Research Intern**, *Google Research*, New York, NY, USA
May Aug
Researched local differential privacy and secure aggregation for federated analytics.
- 2018 – 2018 **Applied Scientist Intern**, *Amazon*, Seattle, WA, USA
May Aug
Worked on a novel multi-service recommendations engine for Alexa.

2013 Aug	–	2014 July	Software Engineering Intern , <i>Cisco Systems</i> , Boston, MA, USA Developed front-end and back-end for the SocialMiner data analytics web application.
2012 Dec	–	2013 Apr	Intern , <i>Deloitte</i> , Moscow, Russia Performed data analytics tasks for the audit department.

Publications

Conference Publications

Eugene Bagdasaryan and Vitaly Shmatikov. Spinning language models: Risks of propaganda-as-a-service and countermeasures. In *SEIP*, 2022.

Eugene Bagdasaryan and Vitaly Shmatikov. Blind backdoors in deep learning models. In *USENIX Security*, 2021.

Eugene Bagdasaryan, Andreas Veit, Yiqing Hua, Deborah Estrin, and Vitaly Shmatikov. How to backdoor federated learning. In *AISTATS*, 2020.

Eugene Bagdasaryan, Omid Poursaeed, and Vitaly Shmatikov. Differential privacy has disparate impact on model accuracy. In *NeurIPS*, 2019.

Zhiming Shen, Zhen Sun, Gur-Eyal Sela, **Eugene Bagdasaryan**, Christina Delimitrou, Robbert Van Renesse, and Hakim Weatherspoon. X-containers: Breaking down barriers to improve performance and isolation of cloud-native containers. In *ASPLOS*, 2019.

Longqi Yang, **Eugene Bagdasaryan**, Joshua Gruenstein, Cheng-Kang Hsieh, and Deborah Estrin. Openrec: A modular framework for extensible and adaptable recommendation algorithms. In *WSDM*, 2018.

Longqi Yang, **Eugene Bagdasaryan**, and Hongyi Wen. Modularizing deep neural network-inspired recommendation algorithms. In *RecSys*, 2018.

Journal Publications

Eugene Bagdasaryan, Peter Kairouz, Stefan Mellem, Adrià Gascón, Kallista Bonawitz, Deborah Estrin, and Marco Gruteser. Towards sparse federated analytics: Location heatmaps under distributed differential privacy with secure aggregation. In *PETS*, 2022.

Workshop Papers and Preprints

Eugene Bagdasaryan and Vitaly Shmatikov. Hyperparameter search is all you need for training-agnostic backdoor robustness. *Preprint*, 2023.

Eugene Bagdasaryan, Congzheng Song, Rogier van Dalen, Matt Seigel, and Áine Cahill. Training a tokenizer for free with private federated learning. In *FL4NLP at ACL*, 2022.

Kleomenis Katevas, **Eugene Bagdasaryan**, Jason Waterman, Mohamad Mounir Safadiéh, Eleanor Birrell, Hamed Haddadi, and Deborah Estrin. Policy-based federated learning. *Preprint*, 2020.

Tao Yu, **Eugene Bagdasaryan**, and Vitaly Shmatikov. Salvaging federated learning by local adaptation. *Preprint*, 2020.

Eugene Bagdasaryan, Griffin Berstein, Jason Waterman, Eleanor Birrell, Nate Foster, Fred B Schneider, and Deborah Estrin. Ancile: Enhancing privacy for ubiquitous computing with use-based privacy. In *WPES at CCS*, 2019.

Jonathan Behrens, Ken Birman, Sagar Jha, Matthew Milano, Edward Tremel, **Eugene Bagdasaryan**, Theo Gkountouvas, Weijia Song, and Robbert Van Renesse. Derecho: Group communication at the speed of light. Technical report, Cornell University, 2016.

Tutorials

Dec 2018 **RecSys Tutorial**, *Modularizing Deep Neural Network-Inspired Recommendation Algorithms*
In collaboration with Longqi Yang and Hongyi Wen

Media Coverage

Apr 2023 **The Economist**, “It doesn’t take much to make machine-learning algorithms go awry”
Oct 2022 **Pluralistic: Cory Doctorow**, “Backdooring a summarizerbot to shape opinion”
Oct 2022 **Schneier on Security**, “Adversarial ML Attack that Secretly Gives a Language Model a Point of View”
Dec 2021 **VentureBeat**, “Propaganda-as-a-service may be on the horizon if large language models are abused”
Aug 2021 **ZDNet**, “Cornell University researchers discover ‘code-poisoning’ attack”
Jun 2020 **Cornell Chronicle**, “Platform empowers users to control their personal data”

Mentoring Experience

Undergraduate Students

2019–2020 Mohamad Mounir Safadiéh, Vassar College → Apple
2017–2019 Griffin Berstein, Vassar College → Cornell PhD program

Masters Students

2021–2022 Anastasia Sorokina, Cornell Tech
2020 Pargol Gheissari, Cornell Tech → Palantir
2020 Kuan-Ting Liu, Cornell Tech → Facebook

2020 Calvin Li, Cornell Tech → Evernorth
2020 Chinmay Bhat, Cornell Tech → Shift Technology
2020 Surya Omesh, Cornell Tech → Bloomberg
2020 Saloni Gandhi, Cornell Tech → Twitter
2020 Devansh Gosalia, Cornell Tech → OppFi
2019 Rony Krell, Cornell Tech → UnitedHealth Group

Teaching Experience

Spring 2022 CS 5436/INFO 5303: Privacy in the Digital Age, Part-time
Spring 2021 CS 5436/INFO 5303: Privacy in the Digital Age, Part-time
Spring 2020 CS 5450: Networked and Distributed Systems, Part-time
Fall 2018 CS 5450: Networked and Distributed Systems, Part-time
Spring 2017 CS 5450: Networked and Distributed Systems, Full-time, **Excellence Award**
Fall 2016 CS 4320: Introduction to Database Systems, Full-time

Professional Activity

Conference Reviewing

ICLR'22, ICML'22, NeurIPS'21

Journal Reviewing

TMLR'22, IEEE T-IFS'22

Workshop Reviewing

FL4NLP@ACL'22, AdvML@ICML'22, MAISP@MobiSys'21

Department Service

2018–2019 Co-lead of the PhD Student at Cornell Tech (PACT) organization

Invited Talks

Apr 2023 **Michigan CS**, *Research Seminar*
Untrustworthy Machine Learning: How to Balance Security, Accuracy, and Privacy?
Apr 2023 **Columbia CS**, *Research Seminar*
Untrustworthy Machine Learning: How to Balance Security, Accuracy, and Privacy?
Apr 2023 **BU CDS**, *Research Seminar*
Untrustworthy Machine Learning: How to Balance Security, Accuracy, and Privacy?
Mar 2023 **UW Allen School CSE**, *Research Seminar*
Untrustworthy Machine Learning: How to Balance Security, Accuracy, and Privacy?
Mar 2023 **McGill**, *Research Seminar*
Untrustworthy Machine Learning: How to Balance Security, Accuracy, and Privacy?

- Feb 2023 **CISPA**, *Research Seminar*
Untrustworthy Machine Learning: How to Balance Security, Accuracy, and Privacy?
- Feb 2023 **UMass CS**, *Research Seminar*
Untrustworthy Machine Learning: How to Balance Security, Accuracy, and Privacy?
- Jan 2023 **UCLA CS**, *Research Seminar*
Untrustworthy Machine Learning: How to Balance Security, Accuracy, and Privacy?
- Sep 2022 **Brave Software**, *Research Seminar*
Sparse federated analytics: location heatmaps and language tokenizations.
- Jul 2022 **Google Research**, *Google Federated Talks*
Sparse federated analytics: location heatmaps and language tokenizations.
- Mar 2022 **University of Chicago**, *The SAND Lab Talks*
Spinning Language Models: Propaganda-As-A-Service and Countermeasures.
- Jan 2022 **University of Cagliari**, *Machine Learning Security Seminar Series*
Spinning Language Models: Propaganda-As-A-Service and Countermeasures.
- Jan 2022 **Samsung AI Center Cambridge**, *Invited Talk Series*
Evaluating privacy preserving techniques in machine learning.
- Dec 2021 **University College London**, *Privacy and Security in ML Interest Group*
Blind Backdoors in Deep Learning Models.
- Nov 2021 **University of Cambridge**, *Computer Laboratory Security Seminar*
Blind Backdoors in Deep Learning Models.
- Sep 2021 **Telefonica Research**, *Research Seminar*
Evaluating privacy preserving techniques in machine learning.
- Jan 2021 **Microsoft**, *Applied Research Invited Talk Series*
Evaluating privacy preserving techniques in machine learning.
- Jun 2020 **Google Research**, *Google Federated Talks*
Salvaging federated learning with local adaptation.
- Feb 2020 **Cornell Tech**, *Digital Live Initiative*
Evaluating privacy preserving techniques in machine learning.