Daniel R. Jiang

✉ danielrjiang@gmail.com
⌂ https://danielrjiang.github.io
G Google Scholar
in LinkedIn

Professional Experience

Meta

2023 –   Research Scientist, Applied Reinforcement Learning
2019 – 2023   Research Scientist, Adaptive Experimentation (Central Applied Science)

University of Pittsburgh

2021 –   Adjunct Professor, Industrial Engineering
2016 – 2021   Assistant Professor (tenure-track), Industrial Engineering
　　　　On leave-of-absence from Jan. 2019 – Sept. 2021.

Education

Princeton University

2011 – 2016   Ph.D. Operations Research & Financial Engineering
　　　　Advisor: Warren B. Powell
　　　　Research topics: Reinforcement learning, approximate dynamic programming

Purdue University

2007 – 2011   B.S. Electrical and Computer Engineering, *with highest distinction (summa cum laude)*
2007 – 2011   B.S. Mathematics, *with highest distinction (summa cum laude)*

Research Interests

Keywords: *Reinforcement learning, LLM post-training, RLHF with performance & user signals, multi-turn RL, resolving epistemic uncertainty, adaptive experimentation, Bayesian optimization.*

I am interested in sequential decision-making and reinforcement learning (RL). In particular, I want to understand how AI agents can *effectively plan over long horizons* and *strategically acquire information to resolve epistemic uncertainty* in real-world environments. At Meta, my work has resulted in:

- The first real-world deployment of an RL-trained LLM for generative AI ads [1] using performance signals, leading to millions in revenue and +6.7% in advertiser-level click-through-rate;

- The first RL-based conversational agents trained using online multi-turn RLHF on user signals [2];

- Multi-step lookahead methods for adaptive experimentation, and co-creation of BoTorch [14].

Publications

§: direct mentoring, *: equal contribution, (α-β): alphabetical

Working Papers

1. *Improving Generative Ad Text on Facebook Using Reinforcement Learning.*
   Daniel R. Jiang*, Alex Nikulkov*, Yu-Chia Chen, Yang Bai, & Zheqing Zhu
   Preprint, 2026.
   Media: featured in *Jack Clark's Import AI Newsletter,* cited by *The New York Times.*

2. *Aligning LLMs Toward Multi-Turn Conversational Outcomes Using Iterative GRPO.*
   Daniel R. Jiang, Jalaj Bhandari, Alexis Yang, Rémi Munos, & Tyler Lu
   Preprint, 2026.

3. *Self-Improvement Through Deliberation: Post-Training Language Models on Multi-Agent Debate.*
   A. Samanta, A. Magesh, Y. Yu, R. Wu, A. Jain,
   D. R. Jiang, B. Vidolov, P. Sajda, Y. Efroni*, & K. Hassani*
   Preprint, 2026.

4. *Structure Enables Effective Self-Localization of Errors in LLMs*
   A. Samanta, A. Magesh, A. Jain, K. Asadi, Y. Yu,
   D. R. Jiang, B. Vidolov, K. Hassani, P. Sajda, J. Bhandari, & Y. Efroni
   Preprint, 2026.

5. *Optimization-Driven Adaptive Experimentation.*
   (α-β) Ethan Che, Daniel R. Jiang, Hongseok Namkoong, & Jimmy Wang
   Preprint, 2024.

Peer-reviewed Publications

1. *Faster Reinforcement Learning by Freezing Slow States.*
   Yijia Wang[§] & Daniel R. Jiang
   Management Science, **2025**.

2. *Carbon Aware Transformers Through Joint Model-Hardware Optimization.*
   B. Acun, N. Ardalani, M. Elhoushi, S. Hsia, D. R. Jiang, D. Mahajan, E. Sumbul, I. Wang, & C.-J. Wu
   NeurIPS, **2025**.

3. *Aligned Multi-Objective Optimization.*
   Yonathan Efroni, Ben Kretzu, Daniel R. Jiang, Jalaj Bhandari, Zheqing Zhu, & Karen Ullrich
   ICML, **2025**.
   Preliminary version: Optimization for ML Workshop at NeurIPS 2024

4. *Exploiting Structure in Offline Multi-Agent RL: The Benefits of Low Interaction Rank.*
   Wenhao Zhan[§], Scott Fujimoto, Zheqing Zhu, Jason D. Lee, Daniel R. Jiang, & Yonathan Efroni
   ICLR, **2025**.

5. *On the Linear Speedup of Personalized Federated Reinforcement Learning with Shared Representations.*
   Guojun Xiong, Shufan Wang, Daniel R. Jiang, & Jian Li
   ICLR, **2025**.
   Preliminary version: Deployable RL Workshop at RLC 2024

6. *Pearl: A Production-Ready Reinforcement Learning Agent.*
   Z. Zhu, R. de Salvo Braz, J. Bhandari, D R. Jiang, Y. Wan, Y. Efroni,
   L. Wang, R. Xu, H. Guo, A. Nikulkov, D. Korenkevych, F. Cheng, U. Dogan, Z. Wu, & W. Xu
   Journal of Machine Learning Research, **2024**.

7. *Weakly Coupled Deep Q-Networks.*
   Ibrahim El-Shar[§], Daniel R. Jiang
   NeurIPS, **2023**.

8. *Dynamic Subgoal-based Exploration via Bayesian Optimization.*
   Yijia Wang[§], Matthias Poloczek, & Daniel R. Jiang
   Transactions on Machine Learning Research, **2023**.
   Preliminary version: Task-Agnostic Reinforcement Learning Workshop at ICLR 2019

9. *On Noisy Evaluation in Federated Hyperparameter Tuning.*
   K. Kuo, P. Thaker, M. Khodak, J. Nguyen, D. R. Jiang, A. Talwalkar, & V. Smith
   MLSys, **2023**.

10. *Dynamic Inventory Repositioning in On-Demand Rental Networks.*
    (α-β) Saif Benjafaar, Daniel R. Jiang, Xiang Li, & Xiaobo Li
    Management Science, **2022**.

11. *Interpretable Personalized Experimentation.*
    H. Wu[*], S. Tan[*], W. Li, M. Garrard, A. Obeng, D. Dimmery, S. Singh, H. Wang, D. R. Jiang, & E. Bakshy
    KDD, **2022**.

12. *Multi-Step Budgeted Bayesian Optimization with Unknown Costs.*
    Raul Astudillo[§], Daniel R. Jiang, Maximilian Balandat, Eytan Bakshy, & Peter I. Frazier
    NeurIPS, **2021**.

13. *Efficient Nonmyopic Bayesian Optimization via One-Shot Multi-Step Trees.*
    Shali Jiang[*], Daniel R. Jiang[*], Maximilian Balandat[*], Brian Karrer, Jake Gardner, & Roman Garnett
    NeurIPS, **2020**.

14. *BoTorch: A Framework for Efficient Monte-Carlo Bayesian Optimization.*
    M. Balandat, B. Karrer, D. R. Jiang, S. Daulton, B. Letham, A. G. Wilson, & E. Bakshy
    NeurIPS, **2020**.

15. *Lookahead-Bounded Q-Learning.*
    Ibrahim El-Shar[§] & Daniel R. Jiang
    ICML, **2020**.

16. *Optimistic Monte Carlo Tree Search with Sampled Information Relaxation Dual Bounds.*
    Daniel R. Jiang, Lina Al-Kanj, & Warren B. Powell
    Operations Research, **2020**.

17. *Feedback-Based Tree Search for Reinforcement Learning.*
    Daniel R. Jiang, Emmanuel Ekwedike, & Han Liu
    ICML, **2018**.
    Selected for long talk; 8.6% acceptance

18. *Risk-Averse Approximate Dynamic Programming with Quantile-Based Risk Measures.*
    Daniel R. Jiang & Warren B. Powell
    Mathematics of Operations Research, **2018**.
    Editor's Pick for INFORMS ICYMI in Dec 2017

19. *Shape Constraints in Economics and Operations Research.*
    Andrew L. Johnson & Daniel R. Jiang
    Statistical Science, **2018**.

20. *An Approximate Dynamic Programming Algorithm for Monotone Value Functions.*
    Daniel R. Jiang & Warren B. Powell
    Operations Research, **2015**.

21. *Optimal Hour-Ahead Bidding in the Real-Time Electricity Market using Approximate Dynamic Programming.*
    Daniel R. Jiang & Warren B. Powell
    INFORMS Journal on Computing, **2015**.

Refereed Extended Abstracts and Proposals

1. Workshop proposal: *Deployable RL: From Research to Practice*
   (α-β) J. Bhandari, Y. Efroni, M. Ghavamzadeh, D. R. Jiang, A. Pacchiano, Y. Wan, K. Zhang, & A. Zhou
   RLC, **2024**.

2. Workshop proposal: *2nd Workshop on Multi-Armed Bandits and Reinforcement Learning: Advancing Decision Making in E-Commerce and Beyond*
   Chu Wang, Yingfei Wang, Haipeng Luo, Daniel R. Jiang, Jinghai He, & Zeyu Zheng
   KDD, **2023**.

3. *Towards Green, Accurate, and Efficient AI Models Through Multi-Objective Optimization*
   Udit Gupta, Daniel R. Jiang, Maximilian Balandat, & Carole-Jean Wu
   Tackling Climate Change with Machine Learning: Global Perspectives and Local Challenge Workshop (Proposal Track) at **ICLR**, **2023**.

4. *Adaptive Policies for Staggered Rollout of Large-Scale Online Experiments*
   Daniel R. Jiang, Jelena Markovic, Shubhankar Ray, Adam Obeng, Max Balandat, & Eytan Bakshy
   CODE @ MIT, **2021**.

5. *Learning Efficient Interpretable Policies on Experimental Data*
   H. Wu, S. Tan, W. Li, M. Garrard, H. Wang, D. R. Jiang, A. Obeng, & E. Bakshy
   CODE @ MIT, **2021**.

6. Workshop proposal: *Multi-Armed Bandits and Reinforcement Learning: Advancing Decision Making in E-Commerce and Beyond*
   Daniel R. Jiang[†], Haipeng Luo[†], Chu Wang[†], & Yingfei Wang[†]
   KDD, **2021**.

7. *Hierarchical Reinforcement Learning for Naloxone Procurement and Distribution*
   Yijia Wang & Daniel R. Jiang
   Women in Machine Learning Workshop at **NeurIPS**, **2017**.

8. *An ADP Algorithm for Optimal Hour-Ahead Bidding in the Real-Time Electricity Market with Battery Storage*
   Daniel R. Jiang & Warren B. Powell
   RLDM, **2013**.

9. *A Comparison of Approximate Dynamic Programming Techniques on Benchmark Energy Storage Problems*
   Daniel R. Jiang, Thuy V. Pham, Warren B. Powell, Daniel F. Salas, & Warren R. Scott
   IEEE Symposium on ADP and RL, **2014**.

Technical Reports

1. *Structured Actor-Critic for Managing Public Health Points of Dispensing*
   Yijia Wang & Daniel R. Jiang
   Technical Report, 2022.

2. *Practicality of Nested Risk Measures for Dynamic Electric Vehicle Charging*
   Daniel R. Jiang & Warren B. Powell
   Technical Report, 2017.

Patents

1. *Method for Bidding Battery Storage into Hour-Ahead Energy Markets*
   Warren B. Powell & Daniel R. Jiang
   U.S. Patent No. 9,965,802, 2018.

Awards

2024   Winner of the Kaggle & NFL Big Data Bowl: *Uncovering Missed Tackle Opportunities*
     Matthew Chang, Katherine Dai, Daniel Jiang, Harvey Cheng
     Winner out of around 300 submissions; received $25,000 and invited to present at the 2024 NFL Scouting Combine. The project was productionized by NFL Next Gen Stats for TV broadcasts.

2018 – 2022   National Science Foundation ECCS Award 1807536, $350,892.
     Dynamic Risk-Averse Optimization of Distributed Energy Resource Aggregators
     PI: Daniel R. Jiang, co-PI Jeff Kharoufeh

2018   Long talk presentation at ICML 2018 for *Feedback-Based Tree Search for Reinforcement Learning*.
     8.7% of 2,473 submissions.

2017 – 2018   Tencent AI Lab Faculty Award, $50,000.
     Advanced Reinforcement Learning Methods for Gameplay AI
     PI: Daniel R. Jiang

2015   Wu Prize for Excellence (Sir Gordon Y.S. Wu '58), Princeton University.
     The Wu Prize for Excellence is awarded to upper year graduate students who have performed at the highest level as scholars and researchers.

2006   USA Math Olympiad (USAMO) Qualifer.
     Top 0.2% (approximately 400/200,000) of participating US high school students in AIME & AMC.

Mentoring Experience

Intern & Post-doc Supervision at Meta

2025 –   Tal Lancewicki, Tel Aviv University (post-doc).

2024   Wenhao Zhan, Princeton University (research intern).
     Co-mentored with Yonathan Efroni.

2022 – 2023   Jiayue Wan, Cornell University (research intern).

2020 – 2021   Raul Astudillo, Cornell University (research intern).

Student Supervision at University of Pittsburgh

2017 – 2022   Ibrahim El-Shar, Ph.D. (now: Research Scientist at Apple).
     Thesis: *Exploiting Structure and Relaxations in RL and Stochastic Optimal Control*

2016 – 2022   Yijia Wang, Ph.D. (now: Applied Scientist at Amazon).
     Thesis: *Structured Strategies for Learning and Exploration in Sequential Decision Making*

Teaching Experience

University of Pittsburgh

2016 – 2021   *Decision Models*, IE 1086/2086 (5 times).

> *Advanced undergraduate/master's level; data-driven decision making under uncertainty with applications in energy, inventory, revenue management, healthcare, and finance.*
>
> *Nominated for IISE Annual Award for Excellence in the Teaching of Operations Research.*

2017, 2018   *Approximate Dynamic Programming & Reinforcement Learning*, IE 3186.

> *Ph.D. level; fundamental theory of ADP and RL following Bertsekas's textbooks.*

2018   *Reinforcement Learning*, IE 2186.

> *Master's level; introductory course on RL, following Sutton & Barto textbook.*

2017, 2018   *Innovate: Global Issues in Engineering and Business*, one-week study abroad.

> *Undergraduate level; taught by Prof. Jayant Rajgopal and covers global engineering and business. The week-long study abroad portion visits Beijing, Suzhou, and Shanghai, China.*

Princeton University

2015   *Probability and Stochastic Systems*, ORF 309.

> *Course on probability with weekly precepts. Nominated by students for Outstanding TA Award.*

2013 – 2016   *ORFE Senior Thesis Writing Group*.

> *Mentoring group for undergraduate seniors working on their senior theses.*

2013, 2014   *Operations and Information Engineering*, ORF 411.

> *Senior level course on decision-making in OR.*

2013   *Optimal Learning*, ORF 418.

> *Senior level course on sequential information collection (bandits, Bayesian optimization).*

Talks and Seminars

Research Seminars

2025   *Improving Generative AI Ad Text on Facebook using Reinforcement Learning*

> Meta (11/2025).
>
> Keynote Speaker, Symposium on AI in Business and Society, UIC Business School (05/2025).

2022   *Faster Reinforcement Learning by Freezing Slow States*

> University of Pittsburgh (12/2022).
>
> HKUST Business School (05/2022).
>
> Duke Fuqua (04/2022).

2021 – 2023   *Bayesian Optimization with a Budget and Unknown Costs*

> Netflix, Machine Learning and Inference Research (06/2023).
>
> Amazon Science Seminar (08/2022).
>
> University of Pittsburgh (04/2021).

2021   *Structured Actor-Critic RL for Managing Public Health Points-of-Dispensing*

> Singapore Management University (04/2021).

2019 – 2020   *Sampled Lookahead Approaches in Dynamic Programming and Reinforcement Learning*

> National University of Singapore (12/2020).
>
> University of Tennessee (03/2020).

UIC Business School (09/2019).

2018   *Feedback-based Tree Search for Reinforcement Learning*
ICML Long Talk (8.6% acceptance rate), Stockholm Sweden (07/2018).

2018   *Advances in Monte-Carlo Tree Search Methods*
Core Data Science, Facebook (06/2018).
Systems and Industrial Engineering Seminar, University of Arizona (04/2018).

2016   *Risk-Neutral and Risk-Averse ADP Methods for Bidding in the Energy Market*
Princeton University (05/2016).
University of Pittsburgh (01/2016).
CMU Tepper School of Business (01/2016).

General Talks

2023   *The Importance of Foundations, Balance, and Adaptivity as an OR & AI Practitioner*
Keynote Speaker, IISE Pittsburgh Chapter Awards Dinner, University of Pittsburgh (11/2023).

2020 – 2021   *Making Decisions using Artificial Intelligence*
SWE Code Day (for middle-schoolers), University of Pittsburgh (11/2021).
SWE Code Day (for middle-schoolers), University of Pittsburgh (11/2020).

Service

2025   Review committee for the *INFORMS Harvey J. Greenberg Research Award*.

2024   Co-organizer of RLC Workshop: *Deployable RL: From Research to Practice*.

2021, 2023   Co-organizer of KDD Workshop: *MAB and RL: Advancing Decision Making in E-Commerce and Beyond*.

2020   Co-organizer of *Facebook Adaptive Experimentation Workshop*.

2018 – 2021   Associate Editor, *Operations Research Letters*.

2017   Mentor, *Women in Machine Learning (WiML) Workshop* at NeurIPS 2017.

Selected Reviewing

2017 – 2026   *Operations Research*

2019 – 2022   *Management Science*

2021 – 2025   *NeurIPS*

2023 – 2026   *ICML*

2022 – 2025   *ICLR*

2023   *AISTATS*

2026   *AAAI*

2024   *Journal of Machine Learning Research*

2022 – 2025   *Transactions on Machine Learning Research*

2017 – 2020   *Manufacturing & Service Operations Management*

2020 – 2025   *INFORMS Journal on Computing*

2021 – 2023   *INFORMS Journal on Data Science*

2019 – 2020   *IEEE Transactions on Automatic Control*

Media

09/16/2025   *What Exactly Are A.I. Companies Trying to Build? Here's a Guide.* **The New York Times**.

08/11/2025   *Import AI 424: Facebook Improves Ads With RL.* **Jack Clark, Import AI Newsletter**.

09/10/2024   *NFL Tries to Tackle Tackling With a New Next Gen Statistic.* **AP News**.

09/10/2024   *Beyond the Box Score: AWS and NFL AI-Powered Tackle Analysis.* **AWS Blog**.

09/04/2024   *Next Gen Stats: New Advanced Metrics [...] for the 2024 NFL Season.* **National Football League**.

03/04/2024   *2024 Big Data Bowl Winners.* **National Football League**.

05/01/2019   *Facebook Open-Sources Ax and BoTorch to Simplify AI Model Optimization.* **Techcrunch**.

11/05/2018   *Pitt Researcher Uses Video Games to Unlock New Levels of AI.* **Swanson Newsroom**.

12/12/2018   *Swanson School Professor Using Video Games to Test AI.* **Pitt University Times**.

11/12/2013   *The Story Behind "What Would I Say?"* **The New Yorker**.

Other Awards & Honors

2022   Transportation Science Meritorious Service Award (for reviewing), 2022.

2009   Eta Kappa Nu Honors Society, Outstanding Junior in Electrical Engineering Award.
     Awarded to the student with the highest GPA in the department (1 of approx. 200 students).

05/2008   First Place, Purdue Department of Mathematics *Problem of the Week* Competition.
     Winner of a semester-long, campus-wide mathematical problem solving contest as a freshman.

2007 – 2011   MathCounts Scholarship (full tuition, room, and board), Purdue University.

2007 – 2011   Indiana Resident Top Scholar (full tuition, room, and board), Purdue University.

05/2007   Valedictorian, West Lafayette High School.

Past Research & Internship Experience

06 – 08/2010   Undergraduate Researcher, Computer Science, Stanford University.
     Advised by Prof. Monica Lam. Developed an algorithm to use email data to generate graph-based representations of social affiliations to improve sharing granularity on social networks.

05 – 08/2009   Engineering Intern, Boston Scientific, St. Paul, MN.
     Implemented search algorithms for error finding software for implantable cardiac devices.

06 – 08/2007   Undergraduate Researcher, Electrical Engineering, Purdue University.
     Advised by Prof. Hong Tan. Developed a virtual Chinese calligraphy brush using a haptic device.

Personal

2019 – 2021   Tutor, Boys and Girls Club, Menlo Park, CA.

12/2013   3 on 3 Intramural Basketball Champions, Princeton University Dillon Intramural Sports.

11/2013   Co-creator of social media webapp *What Would I Say?*, Hack Princeton.
     17 million pageviews, 9 million unique visitors, winner of "Best Facebook Integrated Hack," and the subject of articles in The New Yorker, CNN, The Telegraph, and Time.

05/2013   Princeton University Science Action Most Viral Video Award for *Charged*.
     Co-written, directed, shot, edited with Omer Malik, Harvey Cheng, Matt Chang, Yasmin Afsar.