Yuxiao Wen

New York, NY (open to relocation) | yuxiaowen@nyu.edu | (347) 833-8643 | yw99.github.io | GitHub | LinkedIn

SUMMARY

Ph.D. candidate in Computer Science at NYU working on **reinforcement learning**, **bandits**, **causal inference**, and **experimentation**. Applied research at Spotify (calibrated recommendations shipped to production) and Amazon Web Services (RL-based ad floor pricing). Published 7 NeurIPS/ICML/ICLR/AISTATS papers (5 first-author, 1 solo), including deep learning theory. Focused on data-efficient learning and evaluation across ads, recommendation, and LLMs.

EDUCATION

New York University *New York, NY*

Ph.D., Computer Science · Advisors: Yanjun Han and Zhengyuan Zhou · GPA: 3.95/4.0 *Expected May 2027*

B.A., Honors Mathematics; B.A., Computer Science · GPA: 3.88/4.0 · Magna cum laude *May 2021*

INDUSTRY EXPERIENCE

Isara Lab (AI startup) — Research Scientist Intern (Part-time) *Nov 2026 – Present | New York, NY*

- Post-training a 27B Qwen LLM (vLLM, LoRA, GRPO) for commodity price forecasting; built a statistical forecaster that beats the random-walk benchmark on **74%** of commodities.

Amazon Web Services — Applied Scientist Intern *May 2025 – Aug 2025 | Seattle, WA*

- Designed an RL floor-pricing policy, evaluated in a real-time bidding simulator on replay logs (**>1M** requests/min).
- Estimated **>3%** publisher revenue uplift in simulation at unchanged fill rate.
- Analyzed large-scale bidder logs for cohort behavior patterns, informing per-cohort floor prices and policy features.

Spotify — Machine Learning Engineer Intern *Jun 2024 – Aug 2024 | New York, NY | [Blog]*

- Built and shipped the homepage calibration system to production, achieving **+3.93%** overall and **+36.6%** podcast impression-to-stream rates in online A/B tests; co-authored a RecSys 2025 workshop paper.
- Designed a calibrated exploration strategy for contextual-bandit recommendations on Spotify Home (**>5M** daily users), reducing the user engagement cost of exploration while preserving data quality for training.

Amazon Web Services — Software Development Engineer Intern *Summers 2021 & 2020 | Seattle, WA*

- Built a Java query parser and 5 search interfaces; migrated ~10K data files with AWS Athena (Presto) and Lambda.

SELECTED RESEARCH

Sample-efficient evaluation of adaptive policies [solo] *2026 – Present*

- Designed an exact A/B/n test that shares feedback across adaptive policies (ranking, bidding, LLMs) while each policy's trajectory remains distributed exactly as if it ran alone. On 6 search-ranking bandit policies (MSLR-WEB10K), used up to ~55% fewer user interactions than independent A/B/n with lower error; evaluating LLMs on RewardBench (12 models) and MMLU-Pro, cut interactions by up to ~50% and variance by up to ~67% (*under review, ICLR 2027*).

Exploration for multi-item decisions *2022 – Present*

- Developed provably efficient exploration for systems that recommend several items at once (carousels, slates, ad slots) or also observe feedback on unchosen items (*NeurIPS 2024; ICML 2025 [solo]; AISTATS 2026*).

Estimating & bidding on incremental value *2024 – Present*

- Developed bidding algorithms that learn the incremental value of ad exposure while adapting decisions from sequential causal feedback, with provable regret guarantees and limited holdouts. Second-price auctions in search (*ICML 2026*); first-price auctions in ads (*under review, Management Science*); related: contextual pricing (*ICLR 2026*).

Open-source LLM agents — PaperTalk & LLM-Emerald *2026* | [*PaperTalk*] [*LLM-Emerald*]

- Built PaperTalk, a Codex skill for evidence-grounded paper analysis via context engineering, and LLM-Emerald, a long-horizon multimodal agent for Pokémon Emerald with hierarchical planning and verifier-gated skills.

SELECTED PUBLICATIONS (* equal contribution; αβ alphabetical order)

Bandits, reinforcement learning & evaluation

[12] **Y. Wen**. Fast A/B/n Testing: Exact Multi-Policy Comparison via Tree-Coupled Feedback Sharing. *Under review, ICLR 2027*.

[11] Q. Bai*, Y. Han*, **Y. Wen***, W. Xu, Z. Zhou. Functional Critics for Off-Policy Actor-Critic: From Stability to Exploration. *Under review, ICLR 2027.*

[10] **Y. Wen**, Y. Han, Z. Zhou. Optimal Arm Elimination Algorithms for Combinatorial Bandits. *AISTATS 2026.*

[9] **Y. Wen**. Adversarial Combinatorial Semi-bandits with Graph Feedback. *ICML 2025.*

[8] D. Feijer, H. Abdollahpouri, S. Gupta, A. Clare, **Y. Wen**, et al. Calibrated Recommendations with Contextual Bandits. *RecSys 2025 CONSEQUENCES Workshop.*

[7] **Y. Wen**, Y. Han, Z. Zhou. Stochastic Contextual Bandits with Graph Feedback: From Independence Number to MAS Number. *NeurIPS 2024.*

Advertising, pricing & causal learning

[6] **Y. Wen**, Z. Hu, Y. Han, Y. Yao, Z. Zhou. The (Marginal) Value of a Search Ad: An Online Causal Framework for Repeated Second-price Auctions. *ICML 2026.*

[5] Y. Han*, X. Xu*, **Y. Wen**, et al. Semi-Parametric Contextual Pricing with General Smoothness. *ICLR 2026.*

[4] **Y. Wen**, Y. Han, Z. Zhou. Joint Value Estimation and Bidding in Repeated First-price Auctions. *Under review, Management Science*.

[3] Z. Hu, **Y. Wen**, Y. Yao, J. Zhang, Z. Zhou.(αβ) Learning to Bid with Unknown Private Values in Budget-Constrained First-Price Auctions. *Preprint, 2026.*

Deep learning

[2] A. Jacot, S. H. Choi, **Y. Wen**. How DNNs Break the Curse of Dimensionality: Compositionality and Symmetry Learning. *ICLR 2025.*

[1] **Y. Wen**, A. Jacot. Which Frequencies do CNNs Need? Emergent Bottleneck Structure in Feature Learning. *ICML 2024.*

Also: scientific ML (Physica D 2024, first author; SIAM J. Sci. Comput. 2023). Full list: Google Scholar.

HONORS & AWARDS

Best Student Paper: Finalist, POMS-HK 2026; Third Prize, POMS-China 2026

NYU AI Safety Fellowship, 2026

NYU CS Outstanding Performance Award (2 recipients/year), 2021

SKILLS & SERVICE

Programming & Tools: Python, C++, Java, SQL, Spark, Hive, Presto, BigQuery, pandas, PyTorch, TensorFlow, JAX, scikit-learn, XGBoost, LightGBM, CUDA, vLLM, Claude Code, Codex

Machine Learning: Deep Learning (CNNs, Transformers), Reinforcement Learning (PPO, SAC, TD3), Learning to Rank, Contextual Bandits, Recommender Systems, Data-Efficient Learning, Calibration

LLMs & Agents: Supervised Fine-tuning, GRPO, LoRA, Reward Modeling, Agent Orchestration, LLM Evaluation

Causal Inference & Experimentation: A/B Testing, Off-policy Evaluation, Treatment Effect Estimation

Ads & Operations Research: Bid Optimization, Floor Pricing, Budget Pacing, Dynamic Pricing, Revenue Management

Reviewer: ICML (Silver Reviewer Award), NeurIPS, ICLR, AISTATS, JMLR, MOR, Management Science