

Aaron Ferber, PhD

Email: aaron.ferber@cornell.edu

Mobile: +1(609)304-6665

data \rightleftharpoons decisions \rightleftharpoons deployment

Website: <https://www.cs.cornell.edu/~aaronferber/>

EDUCATION

- **Cornell University** Ithaca, NY
Postdoctoral Associate August 2023 – Present
- **University of Southern California** Los Angeles, CA
Computer Science, Ph.D.; GPA: 4.0 December 2017 – August 2023
- **Georgia Institute of Technology** Atlanta, GA
Computer Science Engineering, Ph.D. (transferred); GPA: 4.0; Presidential Fellow September 2017 – December 2017
- **Cornell University** Ithaca, NY
Computer Science, Operations Research Double Major, BS; GPA: 3.4; Dean's list September 2013 – May 2017

EXPERIENCE

- **Cornell AI for Science Institute** Ithaca, NY
Postdoctoral Associate, Technical Instructor September 2023 - Present
Topics: AI for Science, Large Language Models, Computer Vision, Discrete Optimization, Conservation, Chemistry, Material Science, Metabolomics, Neuroscience
Mentors: Carla P. Gomes, Frank Schroeder
Mentees: Laura Greenstreet, Christian Belardi, Delia Qu, Brendan Rapazzo, Yuanqi Du, Utku Umur Acikalin, Yingheng Wang, Yimeng Min, Victoria Mao, Marc Grimson, Shufeng Kong PhD
 - **Academic Project Lead:** Guide lab, implement collaborations, and manage projects for 9 PhD students, 1 postdoc, and 1 masters student. Give lectures and hold technical office hours for 17 non-CS Schmidt postdocs.
 - **Multimodal LLM for Science:** Fine tune Multimodal LLM to generate domain-relevant explanations of classifications for scientific tasks. Categorizes 10k water bodies with interactive deployment to Portuguese-speaking stakeholders in Brazil. Detects and explains Eelgrass wasting disease from 1k images with deployment from Southern California to Alaska.
 - **Computational Chemistry:**
 - * Predicting mass spectra and metabolism from chemical structure
 - * Inferring metabolome structure from mass spectra
 - * Foundation models for metabolomics
 - * Predicting density of states (DOS) for condensed matter
 - **Conservation:**
 - * Invasive species prediction and management with citizen science data
 - * Satellite detection of aquaculture ponds
 - * Automatic visual detection of Eelgrass wasting disease
 - **General:**
 - * AI method for microscopic imaging (Ptychography) via diffusion models
 - * Decoding visual stimuli from mouse brain activity
 - * Identifying neural assemblies with contrastive learning
 - * Accelerating vehicle routing solvers with attention models
 - * Fast combinatorial optimization via unsupervised transformers
- **USC Center for Artificial Intelligence in Society** Los Angeles, CA
PhD Student September 2017 - August 2023
Topics: Finance, Wildlife Trafficking, Virtual Reality, Recommendation Systems
Mentors: Bistra Dilkina, Milind Tambe, Yisong Yue, Meredith Gore, Burcu Keskin, Bryan Wilder

- **Artificial Decision Intelligence:** Developed differentiable discrete optimization layers for training machine learning + optimization pipelines end-to-end.
- **Deployments:** wildlife conservation, finance, virtual reality device design, and recommendation systems
 - * 32% higher avg returns: Portfolio Optimization
 - * 182x more recommendations per second: Meta Advertisement Recommendation
 - * New nanoscale VR parts designed with AI + Photonics: Metaverse
 - * 13% higher wildlife trafficking detection rates
- **Faster Optimization:** Improved solver speed by identifying core problem structure with deep learning
- **Social Impact:**
 - * ML modeling of neighborhood crime impact on childhood stress
 - * Supply chain interdiction for wildlife trafficking focusing on Pangolin trade
 - * Embedded shortest path solver in deep neural network to predict wildlife trafficking routes
 - * Deployed tools to interdict illegal fishing / transshipment activities

• Meta AI FAIR

Menlo Park, CA

Research Intern

May 2022 - Jan 2023

Topics: Virtual Reality, Recommendation Engines

Mentors: Yuandong Tian, Benoit Steiner, Arman Zharmagambetov

- **SurCO:** Developed SurCo for nonlinear combinatorial optimization using differentiable optimization, improving Meta's existing recommendation system speed by 19% and inverse photonic design quality to theoretical optimality.
- **Constrained Generative Models:** Imposed constraints on generative AI end-to-end during training, obtaining 100% constraint satisfaction, and up to 261% improvement in generative performance.

• NEC Labs

Princeton, NJ

Research Intern

May 2021 - August 2021

Mentor: Alexandru Niculescu-Mizil

- **Video Object Tracking:** Designed a novel algorithm to train deep learning models with downstream reasoning modules end-to-end.

• Cornell University

Ithaca, NY

Data Scientist, Research Assistant

Jan 2014 - May 2017

Topics: Machine Learning, Optimization, Simulation

Mentor: David Shmoys, Shane Henderson, David Williamson, Peter Frazier

- **Bikeshare Optimization:** Optimized Citi Bike bikeshare system to increase daily trip capacity from 20,000 to > 60,000 trips
- **Demand Forecasting:** Designed demand prediction system to increase bike traffic prediction accuracy by 62%
- **Math Modeling:** Constructed mathematical models to expand system by 130 stations and demonstrate provable need for 2,000 more bikes
- **Code Improvement:** Reduced runtime of daily system-wide optimization from 24hr to 10min with linear algebra and parallelization
- **Deployed Web App:** Developed and improved web tools for Citi Bike staff with Google App Engine using Python, HTML, JavaScript, D3, and SVG

• Microsoft Azure

Redmond, WA

Machine Learning Intern

May - August, 2016 & 2017

Topics: Machine Learning, Cybersecurity

Mentors: Malcolm Davis, Ram Shankar Siva Kumar

- **Flagging Suspicious Accounts:** Designed and tested deep learning model to flag suspicious Azure service accounts, external logins, and privilege escalation
- **Thread Detection:** Demonstrated a 46% increase in real threat detection for external Azure endpoints
- **Red Team Bot:** Developed deep generative attack bot which automates penetration testing and external cybersecurity evaluation

- [1] **A. Ferber**, A. Zharmagambetov, T. Huang, B. Dilkina, and Y. Tian. “GenCO: Generating Diverse Designs with Combinatorial Constraints”. In: *Forty-first International Conference on Machine Learning*. 2024. URL: <https://openreview.net/forum?id=DiyE6OOGBa>.
- [2] E. C. Griffin, **A. Ferber**, L. Lafferty, B. B. Keskin, B. Dilkina, and M. Gore. “Interdiction of wildlife trafficking supply chains: An analytical approach”. In: *IIE Transactions* 56.3 (2024), pp. 355–373. URL: <https://www.tandfonline.com/doi/full/10.1080/24725854.2023.2255643>.
- [3] T. Huang, **A. Ferber**, A. Zharmagambetov, Y. Tian, and B. Dilkina. “Contrastive Predict-and-Search for Mixed Integer Linear Programs”. In: *Forty-first International Conference on Machine Learning*. 2024. URL: <https://openreview.net/forum?id=zatLnLvbs8>.
- [4] W. Huang, T. Huang, **A. Ferber**, and B. Dilkina. “Distributional MIPLIB: a Multi-Domain Library for Advancing ML-Guided MILP Methods”. In: *arXiv preprint arXiv:2406.06954* (2024).
- [5] L. Kong, Y. Du, W. Mu, K. Neklyudov, V. De Bortol, H. Wang, D. Wu, **A. Ferber**, Y.-A. Ma, C. P. Gomes, and C. Zhang. “Diffusion Models as Constrained Samplers for Optimization with Unknown Constraints”. In: *arXiv preprint arXiv:2402.18012* (2024).
- [6] A. Parjadis, Q. Cappart, B. Dilkina, **A. Ferber**, and L.-M. Rousseau. “Learning Lagrangian Multipliers for the Travelling Salesman Problem”. In: *Principles and Practice of Constraint Programming*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik. 2024.
- [7] A. Zharmagambetov, B. Amos, **A. Ferber**, T. Huang, B. Dilkina, and Y. Tian. “Landscape surrogate: Learning decision losses for mathematical optimization under partial information”. In: *Advances in Neural Information Processing Systems*. Vol. 36. 2024, pp. 27332–27350. URL: https://proceedings.neurips.cc/paper_files/paper/2023/hash/574f145eac328cc4aaf9358e27120eb5-Abstract-Conference.html.
- [8] **A. Ferber**, E. Griffin, B. Dilkina, B. Keskin, and M. Gore. “Predicting Wildlife Trafficking Routes with Differentiable Shortest Paths”. In: *International Conference on Integration of Constraint Programming, Artificial Intelligence, and Operations Research*. Springer Nature Switzerland Cham. 2023, pp. 460–476.
- [9] **A. Ferber**, T. Huang, D. Zha, M. Schubert, B. Steiner, B. Dilkina, and Y. Tian. “Surco: Learning linear surrogates for combinatorial nonlinear optimization problems”. In: *International Conference on Machine Learning*. PMLR. 2023, pp. 10034–10052.
- [10] M. L. Gore, E. Griffin, B. Dilkina, **A. Ferber**, S. E. Griffis, B. B. Keskin, and J. Macdonald. “Advancing interdisciplinary science for disrupting wildlife trafficking networks”. In: *Proceedings of the National Academy of Sciences* 120.10 (2023).
- [11] M. L. Gore, R. Hilend, J. O. Prell, E. Griffin, J. R. Macdonald, B. B. Keskin, **A. Ferber**, and B. Dilkina. “A data directory to facilitate investigations on worldwide wildlife trafficking”. In: *Big Earth Data* 7.2 (2023), pp. 338–348.
- [12] T. Huang, **A. Ferber**, Y. Tian, B. Dilkina, and B. Steiner. “Local branching relaxation heuristics for integer linear programs”. In: *International Conference on Integration of Constraint Programming, Artificial Intelligence, and Operations Research*. Springer Nature Switzerland Cham. 2023, pp. 96–113.
- [13] T. Huang, **A. Ferber**, Y. Tian, B. Dilkina, and B. Steiner. “Searching large neighborhoods for integer linear programs with contrastive learning”. In: *International Conference on Machine Learning*. PMLR. 2023, pp. 13869–13890.
- [14] B. B. Keskin, E. C. Griffin, J. O. Prell, B. Dilkina, **A. Ferber**, J. MacDonald, R. Hilend, S. Griffis, and M. L. Gore. “Quantitative investigation of wildlife trafficking supply chains: A review”. In: *Omega* 115 (2023), p. 102780.

- [15] A. Paul, D. Freund, **A. Ferber**, D. B. Shmoys, and D. P. Williamson. “Erratum to “budgeted prize-collecting traveling salesman and minimum spanning tree problems””. In: *Mathematics of Operations Research* 48.4 (2023), pp. 2304–2307.
- [16] **A. Ferber**, J. Song, B. Dilkina, and Y. Yue. “Learning pseudo-backdoors for mixed integer programs”. In: *International Conference on Integration of Constraint Programming, Artificial Intelligence, and Operations Research*. Springer International Publishing Cham. 2022, pp. 91–102.
- [17] M. Gore, B. Keskin, J. Macdonald, **A. Ferber**, E. Griffin, O. Prell, B. Dilkina, and R. Hilend. “Identifying datasets for global wildlife trafficking”. In: (2022).
- [18] U. Gupta, **A. Ferber**, B. Dilkina, and G. Ver Steeg. “Controllable Guarantees for Fair Outcomes via Contrastive Information Estimation”. In: *Proceedings of the AAAI Conference on Artificial Intelligence* 35.9 (May 2021), pp. 7610–7619. DOI: 10.1609/aaai.v35i9.16931. URL: <https://ojs.aaai.org/index.php/AAAI/article/view/16931>.
- [19] **A. Ferber**, B. Wilder, B. Dilkina, and M. Tambe. “Mipaal: Mixed integer program as a layer”. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 34. 02. 2020, pp. 1504–1511.
- [20] A. Paul, D. Freund, **A. Ferber**, D. B. Shmoys, and D. P. Williamson. “Budgeted prize-collecting traveling salesman and minimum spanning tree problems”. In: *Mathematics of Operations Research* 45.2 (2020), pp. 576–590.
- [21] A. Paul, D. Freund, **A. Ferber**, D. B. Shmoys, and D. P. Williamson. “Prize-collecting TSP with a budget constraint”. In: *European Symposium on Algorithms*. Schloss-Dagstuhl-Leibniz Zentrum für Informatik. 2017.

ORGANIZATION

- **Academic Project Lead:** Project lead for 11 CS researchers partnering with 17 Schmidt AI for Science postdocs.
- **LXAI Co-Organizer:** LatinX in AI NeurIPS 2023 Affinity Workshop Chair: ~400 in-person + interface with 11 sponsors
- **Computational Sustainability Co-Organizer:** NeurIPS 2024 + CompSust DC (2018, 2019): (46, 63) in-person
- **Quant Project Manager:** Lead for QuantSC 2022 FTX market making competition team of 4
- **INFORMS Session Chair:** INFORMS 2020 Fairness in Optimization and Machine Learning: 100 virtual
- **NeurIPS Workshop Co-Organizer:** NeurIPS 2020 Learning Meets Combinatorial Algorithms (LMCA): 120 virtual

TEACHING

- **Artificial Intelligence for Science @Cornell:** Lecturer: Fall 2024
- **Artificial Intelligence for Social Good @USC:** Lecturer, Teaching Assistant: Spring 2019, Fall 2019
- **Engineering Applications of Operations Research @Cornell:** Teaching Assistant: Spring 2015, Spring 2016
- **Introduction to Analysis of Algorithms @Cornell:** Teaching Assistant: Spring 2016
- **Introduction to Computing Using Python @Cornell:** Tutor: Fall 2013, Spring 2014, Fall 2014
- **Object-Oriented Programming and Data Structures @Cornell:** Tutor: Spring 2014, Fall 2014, Fall 2015

SERVICE / INDEPENDENT PROJECTS

- **LXAI:** Volunteer and Co-Organizer for LXAI (LatinX in AI)
- **Hygeia Medical:** Technical advisor to Hygeia Medical, a reproductive and public health NGO in Madagascar ([Link](#))
- **FinRL:** Mathematical optimization specialist & programmer for FinRL, a reinforcement learning for finance library
- **EHR Quality Improvement:** Quality improvement for Electronic Health Record for a Neurology practice in New Jersey
- **Omdena COVID Challenge:** Data Scientist for COVID measuring domestic violence & unemployment: ([Link](#))
- **COVID Testing Site Placement:** Algorithm designer / data scientist for COVID testing site placement tool: ([Link](#))

SKILLS

- **Expertise:** Deep Learning, Computer Vision, Natural Language Processing, Operations Research, Optimization
- **Programming and Markup Languages:** Python, C++, C, Java, Objective-C, Javascript (React), Matlab, Julia
- **Machine Learning & Data Analysis:** Tensorflow (Python + C), PyTorch, JAX, Scikit-Learn, Pandas, GeoPandas, Numpy (Python + C), Weights and Biases, AWS, Azure, GCP, Portfolio Optimization, Geocoding, PostGIS, BigQuery, Parallelization, Hyperparameter Optimization, FinRL, Alpaca, Qlib, Facebook LLaMA, BERT, Stable Diffusion, Pytorch Geometric, DGL
- **Operations Research Expertise:** Linear / Nonlinear / Discrete / Semidefinite Programming, Financial Optimization
- **Languages:** English (Native), Spanish (Fluent), French (C2 Level)