

# Miguel Fuentes

Amherst, MA | 302-332-0176 | mmfuentes@umass.edu | miguel-fuentes.github.io | linkedin.com/in/miguel-fuentes-phd | github.com/Miguel-Fuentes

## Summary

---

Ph.D. Candidate in Computer Science (graduating May 2026) specializing in differentially private synthetic data, probabilistic modeling, and machine learning. Demonstrated impact through publications in top venues (NeurIPS, AISTATS) and open-source releases. Seeking Research Scientist roles leveraging expertise in privacy, generative models, and applied ML for real-world challenges.

## Education

---

<b>University of Massachusetts Amherst, PhD in Computer Science</b>	<i>Expected Graduation: May 2026</i>
• Advisor: Daniel Sheldon    Honors: Spaulding-Smith Fellowship	
<b>University of Delaware, BS in Mathematics &amp; BA in Computer Science</b>	<i>Aug 2016 – May 2020</i>
• Honors: University Honors Program General Honors Award	

## Skills

---

**Languages:** Python, MATLAB, R  
**Libraries & Frameworks:** JAX, PyTorch, NumPy, Matplotlib, Optax  
**Developer Tools:** Git, Docker, SLURM, LaTeX, Snakemake, uv  
**Open Source:** Contributed to [github.com/ryan112358/private-pgm](https://github.com/ryan112358/private-pgm) (10+ contributors, 100+ stars)

## Research Experience

---

<b>PhD Researcher, University of Massachusetts Amherst</b>	<i>Aug 2020 – Present</i>
• Developed a novel differentially private synthetic data algorithm that outperforms prior approaches (up to 40% error reduction); published at AISTATS 2024 and openly released. [ <a href="https://github.com/Miguel-Fuentes/JAM_AiStats">github.com/Miguel-Fuentes/JAM_AiStats</a> ]	
• Designed and implemented novel algorithms for efficient differentially private marginal query reconstruction, reducing workload error by 44× over baseline. See code: <a href="https://github.com/bcmullins/efficient-marginal-reconstruction">github.com/bcmullins/efficient-marginal-reconstruction</a> .	
• Led foundational work on BirdFlow, a probabilistic model for bird migration using >1B citizen science records; project success enabled group to secure NSF funding and broader adoption in ecological research.	
• Supported the development of a production-ready BirdFlow R package (as part of an NSF project), and maintain the core BirdFlow Python library to enable ongoing collaboration and reproducible science <a href="https://github.com/birdflow-science/BirdFlowPy">github.com/birdflow-science/BirdFlowPy</a> .	
• Built reproducible, large-scale experimental pipelines using SLURM, Snakemake, and Docker to streamline group research and accelerate development cycles.	

<b>Research Project Lead, Institute for Pure and Applied Mathematics</b>	<i>June 2019 – Aug 2019</i>
• Led a summer research project designing a privacy risk assessment framework for large-scale data analytics in Google Ads Data Hub, interfacing directly with a team of Google engineers.	
• Authored a comprehensive technical report (43 pages) and presented findings both to Google Ads Data Hub leadership and at the Joint Mathematics Meetings (JMM) 2020.	

<b>University of Delaware Summer Scholars Program</b>	<i>June 2018 – Aug 2018</i>
• Investigated activation regions for shallow feed-forward neural networks, demonstrating (constructively) the existence of novel non-convex regions; results supported by theoretical proof and empirical validation.	

## Leadership Experience

---

<b>Undergraduate Research Mentor</b>	<i>2023 – 2025</i>
• Supervised an undergraduate student for two years on research applying implicit differentiation to Wasserstein distance optimization, resulting in an undergraduate thesis and acceptance into a PhD program.	

## EMBER Research Mentor

2022 – 2023

- Mentored four undergraduate students in a project on optimizing BirdFlow model performance. Guided experimental design and coding; students delivered a technical report and public symposium presentation.

## Co-founder, UMass Differential Privacy Reading Group

2022 – 2024

- Initiated and organized a weekly reading group on foundational/contemporary topics in differential privacy, facilitating collaborations and ongoing research discussions.

## Selected Publications

---

### Efficient and Private Marginal Reconstruction with Local Non-Negativity

*NeurIPS 2024*

B. Mullins, M. Fuentes, Y. Xiao, D. Kifer, C. Musco, & D. Sheldon

<https://neurips.cc/virtual/2024/poster/93838>

*Introduced GReM-LNN, a principled post-processing method that incorporates non-negativity constraints, reducing marginal query error by up to 44× over baseline.*

### Joint Selection: Adaptively Incorporating Public Information for Private Synthetic Data AISTATS 2024 (oral)

M. Fuentes (first author), B. Mullins, R. McKenna, G. Miklau, & D. Sheldon

<https://proceedings.mlr.press/v238/fuentes24a/fuentes24a.pdf>

*Developed JAM-PGM, the first graphical-model-based mechanism to jointly use public/private sources; achieved up to 40% error reduction on benchmarks.*

### BirdFlow: Learning seasonal bird movements from eBird data Methods in Ecology and Evolution, 2023

M. Fuentes (first author), B. Van Doren, D. Fink, & D. Sheldon

<https://doi.org/10.1111/2041-210X.14052>

*Created a probabilistic model for large-scale migration forecasting; code and models used in Cornell Lab collaborations.*

*Full publication list:* scholar.google.com/citations?user=wy3LBgsAAAAJ&hl

## Presentations

---

- Oral: "Joint Selection: Adaptively Incorporating Public Information for Private Synthetic Data" at AISTATS 2024 (Valencia, Spain) and Theory and Practice of Differential Privacy (Boston, MA), 2024.