

# MARK HAMILTON

Website: [www.mhamilton.net](http://www.mhamilton.net)

[mhamilton723@gmail.com](mailto:mhamilton723@gmail.com)

## EDUCATION

---

**Massachusetts Institute of Technology** 2019-Present  
Ph.D Student in Computer Science, NSF GRFP Fellow GPA: 5.0/5.0  
Thesis: *Unsupervised Structure Discovery with Foundation Models* Advisor: William T Freeman

**Yale University** 2012 - 2016  
Bachelors of Science in Mathematics and Physics with Distinction Magna Cum Laude, GPA: 3.9/4.0  
Thesis: *Language Independent Automated Theorem Proving* Advisor: Gregg Zuckerman

## WORK EXPERIENCE

---

**Microsoft New England Research and Development** August 2016-Current  
*Senior Software Engineer* Cambridge, MA

- Created and lead “SynapseML”, An open-source distributed ML product in Microsoft Fabric, Azure Synapse Analytics, and Azure Databricks.  
**4.9k stars, 118 contributors, 3M downloads.** Website: <https://aka.ms/spark>
- Lead and manage two engineers, serve as tech lead for a group of 6 engineers
- My work to create 5k free and open audiobooks was selected as a TIME top 200 Invention of 2023: <https://aka.ms/audiobook>
- My work on Snow Leopard recognition project featured as international Microsoft advertising campaign and commercial: <https://aka.ms/leopards>
- Designed and taught a 14 lecture deep learning course to hundreds of Microsoft employees.
- Presented to tens of thousands of people in Keynotes across the globe
- Developed and deployed machine learning systems for : The Metropolitan Museum of Art, Coca Cola, NASCAR, The Museum of Fine Art, The Snow Leopard Trust, Air Shepherd, Shell, Aclara Energy, Jabil Manufacturing, Microsoft Dynamics CRM, Seeing AI, and others

## SELECTED RESEARCH EXPERIENCE

**SEE 5 OTHERS** 

---

**MIT Computer Vision Group** 2019-Present  
Advisor: William T. Freeman Massachusetts Institute of technology

Studied and worked on a variety of research projects aiming to discover and extract structure from the world without human supervision. Examples include discovering language from natural videos, unsupervised semantic segmentation, unsupervised upsampling of visual foundation models, and formalizing model explanations and credit assignment for search engines and contrastive learners.

**Yale Coifman Applied Math Group** 2016-2017  
Main Collaborators: Uri Shaham, Kelly Stanton Yale University

Developed methods for semi-supervised translation using maximum mean discrepancy distance, Generative Adversarial Networks, and word embeddings. Created and implemented algorithms for adversarial training of word2vec by leveraging a connection to PMI matrix factorization. Unified MMD networks and GANs by using adversarially trained neural functions in an MMD network.

- [Hamilton MT](#), Hershey J., Zisserman A., Freeman W. *Separating the “Chirp” from the “Chat”*: Self-supervised Visual Grounding of Sound and Language. In Review, Computer Vision and Pattern Recognition (CVPR) 2024
- [Hamilton MT\\*](#), Fu S.\*, Brandt L, Feldman A, Zhang Z. Freeman W. *FeatUp: A Model-Agnostic Framework for Features at Any Resolution*. In Review, International Conference on Learning Representations (ICLR) 2024
- [Hamilton MT\\*](#), Walsh B.\*, Newby G., Wang X., Ruan S., Zhao S. , He L. , Zhang S. , Dettinger E. , Freeman W, Weimer M. *Large-Scale Automatic Audiobook Creation* . INTERSPEECH Show and Tell 2023. **TIME Top 200 Invention of 2023**
- [Hamilton MT](#), Zhang Z, Hariharan B, Snavely N, Freeman W. *Unsupervised Semantic Segmentation by Distilling Feature Correspondences*. International Conference on Learning Representations (ICLR), 2022
- [Hamilton MT](#), Lundberg S, Zhang L, Fu S, Freeman W. *Axiomatic Explanations for Visual Search, Retrieval, and Similarity Learning*. International Conference on Learning Representations (ICLR), 2022
- [Hamilton MT](#), Fu S, Lu M, Bui J, Bopp D, Chen Z, Tran F, Wang M, Rogers M, Zhang L, Hoder C, Freeman W. *MosAIC: Finding Artistic Connections across Culture with Conditional Image Retrieval*. NeurIPs, 2020 Competitions and Demonstrations Track
- [Hamilton MT](#), Gonsalves N, Lee C, Raman A, Walsh B, Prasad S, Banda D, Zhang L, Zhang L, Freeman WT. *Large Scale Intelligent Microservices* IEEE Big Data, 2020, <https://arxiv.org/pdf/2009.08044.pdf>
- Hamilton KA, [Hamilton MT](#), Johnson D, Jjemba P, Bukhari Z, LeChevallier M, Haas CN, Gurian PL *Risk-based critical concentrations of Legionella pneumophila for indoor residential water uses*. **Editors Choice and Cover Article**, Environmental Science & Technology, 2019
- Erickson E, [Hamilton MT](#). *Companies and the Rise of Economic Thought: The Institutional Foundations of Early Economics in England, 1550 to 1720*. American Journal of Sociology, **Granovetter Best Paper Honorable Mention**, 2017 [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3075219](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3075219)

---

## RECENT AWARDS

2023	Winner	TIME Top 200 Inventions of 2023
2021	Winner	NSF Graduate Research Fellowship
2021	Winner (\$60k)	Systems That Learn Fellowship
2020	Winner (\$100k)	Systems That Learn Fellowship
2019	Distinguished Speaker	IEEE High Performance Extreme Computing
2019	Editors Choice Article	American Chemical Society
2019	Cover Article	Environmental Science and Technology
2019	Honorable Mention	Granovetter Award for Best Article in Economic Sociology
2018	Featured Presentation	Microsoft AI for Earth Summit
2016	Winner	Howard L. Schultz Prize for Experimental Physics
2015	Winner	Yale College Science and QR Center Intl. Fellowship
2015	Winner	Pierson College Richter Fellowship
2014	Two Time Winner	Yale University Tetelman Fellowship
2014	Winner	Yale University Science Scholars Fellowship

## LANGUAGES AND SOFTWARE

---

### Languages

- **Scientific:** Python, R, Matlab, Mathematica
- **Functional and Object Oriented:** Scala, Java, C#
- **Scripting:** Bash, CMD, Powershell
- **Web:** Javascript, Typescript, HTML, CSS
- **Native:** C++, C

### Software

- **Deep Learning:** PyTorch, Pytorch Lightning, Tensorflow, Horovod, Keras, Theano, CNTK
- **Distributed Computing:** Spark, Kubernetes, Azure Databricks, CosmosDB, HDInsight, TORQUE, Azure Blob, Azure Data Lake, Virtual Kubelet, Azure Container Instances
- **Probabilistic Programming:** pyMC2/3, MC2D, Edward, Infer.NET
- **Machine Learning and Scientific Computing:** Scikit-Learn, SparkML, NLTK, Azure Cognitive Services, Numpy, Pandas, Lephare
- **Engineering:** Azure Pipelines, Github Actions, Docker, Kubernetes, IntelliJ, PyCharm, VSCode, Git, SSH, RDP, Certificates, VPN Gateways, Web Serving, Web Clients, Helm, SWIG, JNI, SBT, Maven, Sockets, HTTP, Container Registries, Vagrant, Visual Studio
- **Publishing:** L<sup>A</sup>T<sub>E</sub>X, React, Overleaf, EndNote, WebFlow, GIMP, InkScape

## RELEVANT COURSEWORK

---

Machine Learning	Statistics	Linear Algebra
Computer Vision	Information Theory	Distributed Systems
Shape Analysis + Geometry	Neural Networks	Probabilistic Graphical Models
Vector Analysis	Abstract Algebra	Real Analysis
Complex Analysis	Multivariable Calculus	Complex Systems
Classical Physics	Quantum Physics	Particle Physics and Field Theory

## TEACHING

---

### Modern Deep Learning

2016-2017

~ 200 *Students, Microsoft New England Research and Development*

Cambridge, MA

Created and taught a 14 week course on deep learning theory and applications. Topics covered: network architectures (FF, RNN, Conv, ResNet, LSTM), network inversion, Deep Dream, distribution metrics (MMD, EMD, etc), GANs, information theory/geometry, language models and embeddings, stochastic matrix factorization, deep reinforcement learning, deep-q learning, Alpha Go, Neural Turing Machines, optimization methods, optical flow.

For a detailed syllabus see: <https://mhamilton.net/files/mdls.pdf>

**LEADERSHIP**

<b>Role</b>	<b>Project</b>	<b>Date</b>
Lead 2 Engineers	<i>SynapseML Core Team</i>	2021-Current
Tech Lead for 6 Engineers	<i>SynapseML vTeam</i>	2021-Current
Advised 1 Masters	<i>Unsupervised Hierarchical Object Discovery</i>	2023-Current
Advised 1 Undergrad	<i>Generalized Shapley Value Theory</i>	2023-Current
Workshop Organizer	<i>CVPR Workshop on Multimodal Learning for Earth and Environment</i>	2023
Advised 1 Masters	<i>Unsupervised Learning for Remote Sensing</i>	2022-2023
Workshop Organizer	<i>CVPR Workshop on Multimodal Learning for Earth and Environment</i>	2022
Advised 1 Undergrad	<i>Upsampling Deep Vision Backbones</i>	2022
Advised 1 Undergrad	<i>Exploring Gender and Race Biases in the NFT Market</i>	2022
Advised 1 Undergrad	<i>Deep Independent Component Analysis</i>	2021
Led 6 Engineers	<i>Cross-Cloud benchmarking of distributed algorithms</i>	2021
Led 8 Interns	<i>Conditional KNNs for Cross Cultural Art Discovery</i>	2020
Advised 1 Intern	<i>Distributed Speech To Text and Heterogeneous Machine Learning Orchestration</i>	2020
Led 7 Interns	<i>Deep Learning for Cultural Institutions</i>	2019
Led 5 Interns	<i>Gen Studio: A Deep Art experience for the Metropolitan Museum of Art</i>	Winter 2019
Led 6 Engineers	<i>Distributing Custom Search Engine Creation</i>	2018-2019
Led 6 Engineers	<i>Unsupervised Object Detection for Individual Snow Leopard Identification</i>	2018
Advised 1 Intern	<i>Distributing Tensorflow on Spark</i>	2017
Advised 1 Intern	<i>Automating Energy Meter Quality Assurance with SSD Networks</i>	2017
Advised 2 Interns	<i>Deploying Spark on Azure Kubernetes Service</i>	2017
Advised 1 Intern	<i>Implementing Grid LSTMs in the Cognitive Toolkit</i>	2016-2017