MARK HAMILTON

Website: `www.mhamilton.net`

EDUCATION

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

WORK EXPERIENCE

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

- Created and lead "Microsoft Machine Learning for Apache Spark", Microsoft's open source and distributed ML library for the Apache Spark ecosystem. **1.2k stars, 260 forks, 940k docker pulls**. Website: `aka.ms/spark`
- Created frameworks for elastic and distributed deep learning (CNTK on Spark), sub-millisecond latency distributed web serving (Spark Serving), and distributed intelligent micro-services (The Cognitive Services on Spark).
- Designed and taught a 14 lecture deep learning course to hundreds of Microsoft employees.
- Presented to tens of thousands of people during Microsoft Keynotes across the globe
- Developed and deployed machine learning systems for:
  - The Snow Leopard Trust: Deep Leopard identification
  - Air Shepherd: Deep reality simulation for UAV-based poacher detection
  - Shell: FPGA accelerated gas station hazard detection
  - Aclara Energy: Energy-meter predictive maintenance
  - Jabil Manufacturing: Computer-chip defect detection
  - Microsoft Dynamics CRM: "Social Selling" (Online Bayesian collaborative filtering)

TEACHING

Modern Deep Learning 2016-2017
$\sim$ 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`

RESEARCH EXPERIENCE

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.

KWARC Knowledge Representation Group 2015
PI: Dr. Michael Kohlhase, Adviser: Dr. Florian Rabe Jacobs University, Bremen Germany

Created a logic-independent, agent-based, automated theorem prover for the MMT module system. This prover can operate in almost all languages of mathematics and computer science. (PL, FOL, HOL, Category Theory, HOTT, Physical Systems, etc.) Applied this prover to a new type system that formalized physical systems with units.

Yale University Active Galactic Nuclei Group 2014-2015
PI: Dr. Meg Urry, Adviser: Dr. Stephanie LaMassa Yale University and Santiago, Chile

Created META, a machine learning model comparison and optimization software to estimate the spatial distribution of Luminous AGN. Used SQL to stitch together heterogeneous catalogs of galactic images to create a detailed spectral map of the "Stripe 82" region. Used Maximum Likelyhood methods to normalize and fit galactic spectral profiles.

The ATLAS Experiment at the Large Hadron Collider 2013
PI: Dr. Tobias Golling, Adviser: Dr. Mark Cooke Geneva, Switzerland

Designed and implemented an analysis that searched for the existence of the Vector-Like quark at the ATLAS particle detector at CERN. Optimized cut-based analyses for statistical power.

Columbia University Nanophotonics Laboratory 2011-2012
PI: Dr. Chee Wei Wong, Adviser: Richard Grote New York, NY

Designed subwavelength photonic crystals to improve thin film solar cells. Used Rigorous Coupled Wave Analysis (RCWA) and Finite Difference Time Domain (FDTD) simulations to estimate absorption and band gap structures of photonic crystals.

Queens College Nanophotonics Laboratory 2010
PI: Dr. Sajan Saini Queens, NY

Computed the reflectivity and band gap structures of 1D photonic crystals to analyze solar cell efficiency using the Transfer Matrix Method.

PUBLICATIONS

- Hamilton MT, Raghunathan S., Matiach I., Schonhoffer A., Raman A., Barzilay E., Thigpen M., Rajendran K., Mahajan J.S., Cochrane C., and Eswaran A. *MMLSpark: Unifying Machine Learning Ecosystems at Massive Scales*. Submitted to ICML, 2019. `https://arxiv.org/abs/1810.08744`
- Hamilton MT, Raghunathan S, Annavajhala A, Kirsanov D, de Leon E, Barzilay E, Matiach I, Busch M, Oprescu M, Sur R, Astala R, Wen T, Park CY. *Flexible and Scalable Deep Learning with MMLSpark*. Proceedings of Machine Learning Research, 2017 `http://proceedings.mlr.press/v82/hamilton18a.html`
- Bondi E, Fang F, Hamilton MT, Kar D, Dmello D, Choi J, Hannaford R, Iyer A, Joppa L, Tambe M, Nevatia R. *SPOT Poachers in Action: Augmenting Conservation Drones with Au-*

tomatic Detection in Near Real Time. Proceedings of the Thirtieth Annual Conference on Innovative Applications of Artificial Intelligence, 2018 http://teamcore.usc.edu/papers/2018/spot-camera-ready.pdf

- Hamilton MT, Sur RR. *Massively Scalable Neural Networks with CNTK on Spark.* Microsoft Journal of Applied Research, 2017.
- Hamilton MT. *Semi-supervised translation using MMD networks.* Submitted to ICML 2019 https://arxiv.org/abs/1810.11906
- Ananna TT, Salvato M, LaMassa S, Urry CM, Cappelluti N, Cardamone C, Civano F, Farrah D, Gilfanov M, Glikman E, Hamilton MT, Kirkpatrick A, Lanzuisi G, Marchesi S, Merloni S, Nandra K, Natarajan P, Richards G, Timlin J. *AGN Populations in Large Volume X-ray Surveys: Photometric Redshifts and Population Types found in the Stripe 82X Survey.* The Astrophysical Journal, 2017, http://iopscience.iop.org/article/10.3847/1538-4357/aa937d/meta
- Hamilton MT, Jia K, Cooke M, Golling T. *A study of the feasibility of a search for the single production of a vector-like top quark decaying to a Z boson and a top quark in* **pp** *collisions at* $\sqrt{s} = 8$ *TeV with the ATLAS detector.* ATL-COM-PHYS-2014-474, 2014
- 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, 2017 https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3075219
- Hamilton KA, Hamilton MT, Johnson W, Bukhari Z, Jjemba P, LeChevallier M, Haas CN.*Health risks from exposure to Legionella in reclaimed water aerosols: Toilet flushing, spray irrigation, and cooling towers.* Water Research, 2017 https://mhamilton.net/files/water-research.pdf
- Ahmed WA, Hamilton KA, Vieritz A, Powell D, Goonetilleke A, Hamilton MT, Gardner T. *Microbial risk from source-separated urine used as liquid fertilizer in sub-tropical Australia.* Microbial Risk Analysis, 2017, 5:53-64 http://dx.doi.org/10.1016/j.mran.2016.11.005
- Holland C., Hamilton MT, Samenez-Larkin G., *Heuristic Models Outperform Traditional Discounting Utility Models Across Multiple Discounting and Reward Types.* Society for Neuroeconomics, 2017 https://mhamilton.net/files/neuroecon.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.* Minor Revisions, Environmental Science & Technology, 2019

TECHNICAL REPORTS

- Hamilton MT, Rabe F, Kohlhase M, Zuckerman G. *Automated theorem proving in the module system for mathematical theories.* Accepted as a thesis for the B.S. in Mathematics and Physics, 2016, https://mhamilton.net/files/thesis.pdf
- Hamilton MT. *Category Theory and the Curry-Howard-Lambek Correspondence.* 2016, https://mhamilton.net/files/chl.pdf
- Hamilton MT. *Representations of the Heisenberg group and reproducing kernels.* 2016, https://mhamilton.net/files/hkernel.pdf
- Hamilton MT. *A General Tool for Learning-Algorithm Optimization and Comparison.* 2015, https://mhamilton.net/files/meta.pdf
- Hamilton MT. *pyTrading: A Framework for the Creation and Analysis of Automated Trading Strategies.* 2016, https://mhamilton.net/files/pytrading.pdf

KEYNOTE PRESENTATIONS

- Kumar R. *Developing for the Intelligent Cloud and the Intelligent Edge.* Spark and AI Summit Europe, London 2018. Built and presented keynote demo on unsupervised classification for Shell gas station hazard detection. `https://www.youtube.com/watch?v=N3ozCZXeOeU`
- Sirosh J. *The Microsoft AI platform: a State of the Union.* Microsoft Build Conference, Seattle WA 2018. Built and presented keynote demo on interactive deep learning for Jabil Manufacturing circuit board defect detection. `https://channel9.msdn.com/Events/Build/2018/BRK3224`
- Kumar R. *Developing for the Intelligent Cloud and the Intelligent Edge.* Spark and AI Summit, San Francisco CA 2018. Built keynote demo on realtime power grid maintenance detection with UAVs for eSmart Systems. `https://vimeo.com/274266764`
- Sirosh J, Hamilton MT, Ren S, Thomas A, Smith B. *Day 1 Keynote.* Microsoft Machine Learning and Data Science Conference (MLADS), Seattle WA Summer 2018. Built and presented keynote demo on the Cognitive Services on Spark for endangered animal recognition. Microsoft Internal: `https://resnet.microsoft.com/video/41365`
- Hamilton MT. *Flexible and Scalable Deep Learning with MMLSpark.* Keynote Speaker, International Conference on Predictive Applications and APIs, Boston MA 2017 `https://github.com/mhamilton723/Talks-and-Papers/blob/master/MMLSpark%20PAPIS.pdf`
- Hamilton MT, Sengupta R. *Deep Learning for Snow Leopard Conservation.* Keynote Presentation, Microsoft AI for Earth Summit, Seattle WA May 2018.
- Sirosh J. *How to Get Started with Microsoft AI.* Keynote Demo, Microsoft Connect(), New York, 2017. Built and presented keynote demo on high throughput distributed streaming systems for leopard classification `https://channel9.msdn.com/Events/Connect/2017/G102`
- Hamilton MT. *Saving Snow Leopards with Azure Machine Learning.* Keynote Demo, Microsoft Machine Learning and Data Science Conference (MLADS), Seattle WA, Summer 2017. Built and presented keynote demo on elastic deep learning and programming by example for leopard classification.
- Sirosh J. *Will AI help save the snow leopard?.* Built keynote demo of streaming leopard classifier. Strata Data Conference, New York 2017.
  `https://www.oreilly.com/ideas/will-ai-help-save-the-snow-leopard`

CONFERENCES AND TALKS

- Hamilton MT. *Spark Serving: Unifying Batch, Streaming, and RESTful Serving.* Accepted for Spark and AI Summit, San Francisco 2019.
- Hamilton MT. *The Azure Cognitive Services on Spark: Clusters with Embedded Intelligent Services.* Accepted for Spark and AI Summit, San Francisco 2019.
- Hamilton MT, Raman A. *Deep Reality Simulation For Automated Poacher Detection.* Spark and AI Summit Europe, London 2018. `www.aka.ms/deep-reality`
- Hamilton MT, Raman A. *Unsupervised Object Detection using the Azure Cognitive Services on Spark.* Spark and AI Summit Europe, London 2018.
  `www.aka.ms/spark-summit-cognitive-services`
- Hamilton MT. *Unifying Microsoft's ML Ecosystems at Massive Scales with MMLSpark.* International Conference on Predictive Applications and APIs, Boston, 2018. `https://papis2018.sched.com/event/637e8d7ce6eae199037ee022b14e39cc`

- Hamilton KA, Hamilton MT, Johnson W, Bukhari Z, Jjemba P, LeChevallier M, Haas CN. *Evaluating the health risks from exposure to Legionella in reclaimed water aerosols.* Presentation and Conference proceedings for International Water Association International Conference on Water Reclamation and Reuse, 2017. `https://mhamilton.net/files/legionella_IWA.pdf`
- Hamilton KA, Hamilton MT, Haas CN, Johnson W, Bukhari Z, Jjemba P, LeChavallier M. *Health Risks from Legionella in reclaimed water aerosols produced by cooling towers and spray irrigation.* Association of Environmental Engineering and Science Professors Research and Education Conference, Ann Arbor, MI 2017 `https://mhamilton.net/files/legionella_AEESP.pdf`
- Noelle LaCharite, George Earl, Shawn Roberts, Hamilton MT. *Azure AI - developing with Azure Cognitive Services and Azure Bot Service.* Microsoft Ignite, Seattle WA 2018.
- Hamilton MT. *Distributed AI for Earth: Using Microsofts Open-Source Spark Ecosystem for Wildlife Conservation.* Apache Spark Developer Meetup Boston MA August 2018. `https://www.meetup.com/Boston-Apache-Spark-User-Group/events/252940344/`
- Hamilton MT. *Deep Reality Simulation for Automated Poacher Detection.* AI Collaboration to End Wildlife Trafficking Online Workshop, Redmond WA 2018.
- Hamilton MT, et. al. *Deep Learning for Wildlife Conservation.* Conservation Asia. Bishkek Kyrgystan 2018.
- Hamilton MT, Raman A. *Distributed Deep Domain Adaptation for Automated Poacher Detection.* OReilly Artificial Intelligence Conference, San Francisco CA, 2018. `https://conferences.oreilly.com/artificial-intelligence/ai-ca-2018/public/schedule/detail/72155`
- Hamilton MT, Vankamandi S. *What's new in Azure Machine Learning.* Session, Microsoft Machine Learning and Data Science Conference (MLADS), Seattle WA, Summer 2017
- Pathak S, Siede F, Hamilton MT. *Tutorial on Massively Scalable Production Grade Deep Learning with the Microsoft Cognitive Toolkit.* Full Day Workshop, International Conference on Information and Knowledge Management, Singapore, 2017
- Hamilton MT, Wu H. *Using Spark and Computer Vision to Help Save Snow Leopards.* Taiwan Artificial Intelligence Annual Meeting, Taiwan 2017 `aka.ms/snow-leopard-taiwan`
- Hamilton MT. *Introduction to Deep Learning.* Harvard University Psychology Dept. Cambridge MA, 2017
- Hamilton MT. *Saving Snow Leopards with Deep Learning and Computer Vision on Spark.* Microsoft Machine Learning Community Talk Series, Online, 2017
- Hamilton MT, Raghunathan S. *Deep Learning with the Microsoft Cognitive Toolkit and MMLSpark.* Boston Data Science Meetup, Cambridge MA, 2017
- Hamilton MT. *Deep Learning for Unsupervised Translation.* SUMS Mathematics Society, Yale University 2016
- Hamilton MT. *Algebraic Data Types.* SUMS Mathematics Society, Yale University 2016
- Hamilton MT. *Categorical Biology.* SUMS Mathematics Society at Yale University 2016 `https://mhamilton723.github.io/files/catbio.pdf`

BLOGS AND NEWS

- Hamilton MT, et. al. *Deep Learning Without Labels.* Microsoft Machine Learning Blog 2018. `https://blogs.technet.microsoft.com/machinelearning/2018/10/03/deep-learning-without-labels/`

- Stiffler L. *Microsoft says AI is finally ready for broader use to help solve Earths environmental woes*. GeekWire 2018. https://www.geekwire.com/2018/microsoft-says-ai-finally-ready-broader-use-help-solve-earths-environmental-woes/
- Ho V. *How snow leopard selfies and AI can help save the species from extinction*. Microsoft Transform Features Article, 2018. https://news.microsoft.com/transform/snow-leopard-selfies-ai-save-species/
- Hamilton MT, Astala R. *Saving snow leopards with deep learning and computer vision on Spark*. Microsoft Customer Stories, 2018 https://customers.microsoft.com/en-us/story/snow-leopard-trust-nonprofit-azure
- Microsoft and Databricks *Image Data Support in Apache Spark* . Microsoft Machine Learning Blog, 2018. https://blogs.technet.microsoft.com/machinelearning/2018/03/05/image-data-support-in-apache-spark/
- Smith D. *Saving Snow Leopards with Artificial Intelligence* . Revolution Analytics Blog, 2017 https://blog.revolutionanalytics.com/2017/10/snow-leopards.html
- Hamilton M, Schonhoffer A. *Tensorflow Serving with a GPU Kubernetes cluster on Azure*. Github Gist 2017. https://gist.github.com/mhamilton723/b1d3402b1c3ee4aeffcaf45b467eba61

LANGUAGES, SOFTWARE

Languages

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

Software

- **Deep Learning**: Tensorflow, Horovod, Keras, Theano, CNTK, PyBrain, Matlab Deep Learning Toolbox
- **Distributed Computing**: Spark, Kubernetes, Azure Databricks, 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, Tensorflow Object Detection API, TF Model Zoo, Azure Cognitive Services, Numpy, Pandas, Lephare
- **DevOps**: Docker, Intellij, PyCharm, Git, SSH, MobaXTerm, RDP, pypi, Web Serving, Web Clients, Helm, SWIG, JNI, Putty, Azure Pipelines, SBT, maven, Sockets, HTTP, Container Registries, Helm Repos, Vagrant, Visual Studio
- **Publishing**: LaTeX, Overleaf, EndNote, WebFlow, GIMP, InkScape

RELEVANT COURSEWORK

Graduate Machine Learning	Statistics	Linear Algebra
Automated Decision Systems	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

RECENT AWARDS

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
2015	Four Time Winner	Sons of Norway Scholarship
2015	Four Time Winner	Ace Architecture and Engineering Scholarship
2014	Two Time Winner	Yale University Tetelman Fellowship
2014	Winner	Yale University Science Scholars Fellowship