

Maciej Pacula

CONTACT INFORMATION 12 Valley St 617-308-8738
Apt. 504 maciej.pacula@gmail.com
Everett, MA 02149 http://mpacula.com

INTRODUCTION A hands-on leader with a passion for Machine Learning & Data Science, committed to engineering robust software systems that support growth and innovation. I lead a team of software engineers, bioinformaticians and data scientists building systems for next-generation cancer sequencing and personalized cancer care at the Massachusetts General Hospital.

EDUCATION **Massachusetts Institute of Technology**, Cambridge, MA
Master of Engineering, Computer Science, June 2011

- Thesis: “Evolutionary Algorithms for Compiler-Enabled Program Autotuning”
- GPA: 5.0/5.0. Relevant coursework: Machine Learning, Advanced Natural Language Processing, Probability and Computation (MCMC, statistical physics, phylogenetics), Large-Scale Symbolic Systems

Bachelor of Science, Computer Science, June 2010

- GPA: 4.6/5.0. Relevant coursework: Artificial Intelligence, Probabilistic Systems Analysis, Autonomous Robotics, Discrete Mathematics, Biology, Linear Algebra, Compiler Design and Implementation, Software Engineering, Software Performance Engineering, Signals and Systems, Semantics & Pragmatics, Psycholinguistics

PROFESSIONAL EXPERIENCE **MGH Center for Integrated Diagnostics**, Boston, MA **June 2015 - present**
Team Lead, Bioinformatics & Machine Learning

- Leading two teams: CID Informatics and Shared Pathology Services, comprising seven software engineers, bioinformaticians and data scientists building tools for next-generation cancer sequencing and personalized cancer care, as well as supporting computational efforts within the larger Pathology department
- Kickstarting and leading data science within the department; building databases, tools and systems that combine clinical and molecular data in order to enable discovery and analysis, with the goal of developing clinical decision support systems that improve patient outcomes
- Participating in collaborative research projects with investigators at MGH, the Broad Institute and other partner institutions
- Co-managing (with CID Director of Technology Development) collaborations with outside vendors - InterSystems and EMC and others

SynapDx, Lexington, MA **March 2014 - May 2015**
Principal Data Scientist

- Led machine learning to develop a blood test for autism using next-generation sequencing technology and other molecular data
- Main developer of SynapDx’s cloud machine learning pipeline

Raytheon BBN Technologies, Cambridge, MA **March 2012 - March 2014**
Staff Scientist

- Technology lead on the DARPA DCAPS program. Coordinated research, engineering, data annotation and technology transition from research to production
- Developed novel algorithms to automatically detect and label psychological distress signals in multimodal communications using Natural Language Processing and Machine Learning (see publications below)

DataXu, Boston, MA **June 2011 - March 2012**
Software Development Engineer

- Part of the "Learning System" team: developed the computational infrastructure to support big data Machine Learning for advertising optimization on a Hadoop stack
- Selected projects: high-performance reverse geolocation, static analysis for refactoring of legacy code

Microsoft Corp., Redmond, WA **June 2010 - August 2010**
Software Development Intern

- Member of the Natural Language Group within Microsoft Office
- Designed and implemented a fully functional Windows Phone 7 version of the group's project management software Huddle.

DataXu, Cambridge, MA **Summer 2008 & 2009**
Software Development Intern

- Worked on real-time bidding algorithms for online advertising, including a geolocation library that could sustain 2 million queries per second with near-linear parallel scalability
- Prototyped the company's highly parallel behavioral targeting engine.

WindowGain, Watertown, MA **March 2008 - May 2008**
Software Consultant

- Part of a team that developed Linux-based rendering software for multi-projector displays.
- Worked on automatic calibration and the remote management interface, including overhauling the original architecture for scalability and maintainability.

SKILLS

- Excellent communication skills with proven ability to communicate complex technical topics to C-level management and non-technical staff
- Track record of innovation with publications in scientific journals spanning compilers, machine learning, natural language processing, and biotechnology; co-inventor on a USPTO patent
- Capable software engineer proficient in Python and Java, with experience in C/C++, Haskell, Matlab and Lisp.
- Experience in architecting and overseeing the development of complex, production-grade software systems
- Start-up and cloud (AWS) experience; I thrive in high-stakes, high-pressure environments

PATENTS

Ute Geigenmuller, Doris Damian, **Maciej Pacula**, Mark A. DePristo. *Methods and Systems for Determining Autism Spectrum Disorder Risk*. United States Patent 9,176,113. Granted November 3, 2015.

PUBLICATIONS

Jochen K. Lennerz, Heather M. McLaughlin, Jason M. Baron, David Rasmussen, Meini Sumbada Shin, Nancy Berners-Lee, Julie Miller Batten, Kathryn J. Swoboda, Manish K. Gala, Harland S. Winter, Jeremy D. Schmahmann, David A. Sweetser, Marianne Boswell, **Maciej Pacula**, Albrecht Stenzinger, Long P. Le, William Hynes, Heidi L. Rehm, Anne Klibanski, Stephen W. Black-Schaffer, Jeffrey A. Golden, David N. Louis, Scott T. Weiss, A. John Iafrate. *Health Care Infrastructure for Financially Sustainable Clinical Genomics*. The Journal of Molecular Diagnostics. September 2016. Volume 18, Issue 5, Pages 697706.

M. Pacula, T. Meltzer, M. R. Crystal, A. Srivastava, B. Marx. *Automatic Detection of Psychological Distress Indicators and Severity Assessment in Crisis Hotline Conversations*. 39th International Conference on Acoustics, Speech and Signal Processing (ICASSP 2014). Florence, Italy. May 2014.

S. Saleem, **M. Pacula**, R. Chasin, R. Kumar, R. Prasad, M. Crystal. *Automatic Detection of Psychological Distress Indicators in Online Forum Posts*. Asia Pacific Signal and Information Processing Association (APSIPA 2012). Hollywood, CA. December 2012.

S. Saleem, R. Prasad, S. Vitaladevuni, **M. Pacula**, M. Crystal. *Automatic Detection of Psychological Distress Indicators and Severity Assessment from Online Forum Posts*. 24th International Conference on Computational Linguistics (COLING 2012). Mumbai, India. December 2012.

J. Ansel, **M. Pacula**, Y. L. Wong, C. Chan, M. Olszewski, U.M. O'Reilly, S. Amarasinghe. *SiblingRivalry: Online Autotuning Through Local Competitions*. International Conference on Compilers Architecture and Synthesis for Embedded Systems (CASES 2012). Tampere, Finland. October 2012.

M. Pacula, J. Ansel, S. Amarasinghe, U.M. O'Reilly. *Hyperparameter Tuning in Bandit-Based Adaptive Operator Selection*. European Conference on the Applications of Evolutionary Computation (EvoApplications 2012). Malaga, Spain. April 2012.

J. Ansel, **M. Pacula**, S. Amarasinghe, U.M. O'Reilly. *An Efficient Evolutionary Algorithm for Solving Incrementally Structured Problems*. Annual Conference on Genetic and Evolutionary Computation (GECCO'11). Dublin, Ireland. July 2011.

M. Pacula. *Evolutionary algorithms for compiler-enabled program autotuning*. Master's thesis. Massachusetts Institute of Technology. Cambridge, MA. May 2011.

REFERENCES

Available upon request