ANNA SHCHERBINA
annashch@stanford.edu
(310)689-6688

EDUCATION

Stanford University, Stanford, CA — 2019 (expected)
Doctor of Philosophy in Biomedical Informatics

Massachusetts Institute of Technology, Cambridge,MA — 2012
Master of Engineering in Computer Science (Concentration in Computer Systems) (GPA: 4.6/5.0)

Massachusetts Institute of Technology, Cambridge, MA
Bachelor of Science in Computer Science (GPA: 4.2/5.0) — 2011
Bachelor of Science in Biological Engineering (GPA: 4.2/5.0) — 2011

EXPERIENCE

Research

Stanford University, Stanford, CA — Jun 2015-present
PhD student in Professor Euan Ashley's lab

- Data mining and machine learning analysis of accelerometry, activity state, heart rate, and phenotypic data from the My Heart Counts mobile health study (50K subjects). Developed algorithms for automatic feature generation and selection for unsupervised learning from the data.
- Clinical validation of leading mobile health monitors.
- Neural network approaches to detect arrhythmias in real time from ECG and PPG signal data.

Jan 2016-present

PhD student in Professor Anshul Kundaje's lab

- Developed supervised convolutional neural network classifiers to identify transcription factor motif networks involved in pluripotency induction in the heterokaryon system as well as stem cell differentiation in response to DMSO treatment.
- Developed deep learning algorithms to study differences in chromatin accessibility across multiple cell types in single-cell ATAC-seq data.

MIT Lincoln Laboratory, Lexington, MA
Bioinformatician — 2012-2015

- Developed software and algorithms for metagenomic sample analysis, pathogen identification, SNP/STR calling, biogeographic ancestry prediction, and kinship identification from NextGen sequencing data.
- Identified gene networks and pathways significant in healing of muscoskeletal injury by analyzing ChIP-Seq and RNA-Seq data.
- Organized two bioinformatics challenge days to foster community collaboration and problem solving in the field of Metagenomics.

Co-op through the MIT VI-A MEng. Program — 2010-2012

- Differentiated authentic forensic DNA profiles from artificial profiles synthesized in a laboratory through machine learning techniques.
- Developed a software toolkit in C# for genetic kinship identification using STR profile analysis.

Memento Security, Inc., Burlington, MA — Jan 2010
Software Development Intern

- Developed software for identity theft detection and social security number validation. Used the C# programming language to code modules for database development and records management.

MIT Energy Initiative, Cambridge, MA — Jun 2009-Jan 2010
Undergraduate Researcher

- Created time-lapse movies, processed images, and analyzed data pertaining to the use of biofilm as a means to enhance carbon dioxide sequestration in underground aquifers.

MIT Media Lab, Affective Computing Research Group, Cambridge, MA — Jun-Sep 2008
Undergraduate Researcher

- Created a graphical user interface in Python and used the software package Protel to create PCBs for biometric sensors, thus furthering the research group's work on articles of clothing that measure and store biometric data for the purpose of autism therapies.

Teaching

- Teaching assistant for Stanford University's course “Deep learning for genomics” (CS273B) — Fall 2016
- Teaching assistant for Stanford University's course “Modeling Biological Systems: Ontology, Terminology, Problem Solving” (BIOMEDIN210) — Winter 2017

PUBLICATIONS

A. Shcherbina, et al. Accuracy in wrist-worn sensor-based measurements of heart rate and energy expenditure in a diverse cohort. *JACC (under review).* Preprint: https://doi.org/10.1101/094862

E. Ashley, **A. Shcherbina**, et al. Feasibility of obtaining measures from a Smartphone App: The MyHeart Counts Cardiovascular Heatlh Study. JAMA Cardiol. 2017;2(1):67-76. doi:10.1001/jamacardio.2016.4395

Aguilar C, Pop R, **Shcherbina A**, Meissner A, et al. Transcriptional and chromatin dynamics of muscle regeneration after severe trauma. *Stem Cell Reports.* 2016 Nov 8. 7(5):983-997. doi: 10.1016/j.stemcr.2016.09.009.

Gafney-Stomberg E, Lutz J, **Shcherbina A.**, Ricke D, Petrovick M, Cropper T, Cable S, McClung J. Associations between single gene polymorphisms and bone biomarkers and response to calcium and vitamin D supplementation in young adults undergoing military training. J Bone Miner Res. 2016 Sep 28. doi: 10.1002/jbmr.3008.

Shrikumar A, Greenside P, **Shcherbina A**, Kundaje A. Not just a black box: learning important features through propagating activation differences. ArXiv e-prints [Internet]. 2016 May 1; 1605 (Preprint)

D. Ricke, J. Harper, **A. Shcherbina**, N. Chiu. Integrated Biomedical System. BioRxiv e-prints [Internet]. 2016 April 25; doi: http://dx.doi.org/10.1101/050138.

D. Ricke, **A. Shcherbina**, N. Chiu. Evaluating performance of metagenomic characterization algorithms using *in silico* datasets generated with FASTQSim. BioRxiv e-prints [Internet]. 2016 March 31; doi: http://dx.doi.org/10.1101/046532.

A. Shcherbina. FASTQSim: platform-independent data characterization and in silico read generation for NGS datasets. BMC Research Notes. 2014. 7(1):533.

C. Aguilar, **A. Shcherbina,** D. Ricke, R. Pop, C. Carrigan, C. Gifford, M. Urso, M. Kottke, A. Meissner. In vivo Monitoring of Transcriptional Dynamics After Lower-Limb Muscle Injury Enables Quantitative Classification of Healing. Scientific Reports. 2015 Sep 18;5:13885. doi: 10.1038/srep13885.

D. Ricke, **A. Shcherbina.** DAWN: Rapid Larse-Scale Protein Multiple Sequence Alignment and Conservation Analysis. IEEE HPEC, Sep. 2015.

D. Ricke, **A. Shcherbina,** N. Chiu, J. Harper, M. Petrovick, T. Boettcher, C. Zook, J. Bobrow, E. Wack. Sherlock’s Toolkit: A Forensic DNA Analysis System. IEEE International Conference. Mar. 2015.

A. Shcherbina, D. Ricke, N. Chiu, T. Boettcher, C. Zook, J. Bobrow, M. Petrovick, E. Schwoebel. KinLinks: Software tool for kinship analysis and pedigree generation from NGS datasets. BioRxiv e-prints [Internet]. 2016 April 06; doi: http://dx.doi.org/10.1101/046938.

J Isaacson, E. Schwoebel, **A. Shcherbina**, D. Ricke, J. Harper, M. Petrovick, J. Bobrow, T. Boettcher, B. Helfer, C. Zook, E. Wack. Robust detection of individual forensic profiles in DNA mixtures. Forensic Sci. Int. Genetics. 2014. 14C:31-37.

S. Dodson, D. Ricke, J. Kepner, N. Chiu, **A. Shcherbina**. Rapid sequence identification of potential pathogens using techniques from sparse linear algebra. IEEE International Conference. 2015.

POSTERS

A. Shcherbina, E. Schwoebel, M. Petrovick, D. Ricke, J. Harper, T. Boettcher, C. Zook, J. Bobrow, N. Chiu, B. Helfer, E. Wack. Integrated Forensics DNA Data Analysis System. Homeland Protection Workshop Series, Lexington, MA, Dec. 9-12, 2014.

A. Shcherbina, C. Aguilar, D. Ricke. Translating Molecular Ontology Networks into Quantitative Diagnostics. Women in Machine Learning Symposium, Montreal, Quebec, Canada, Dec. 8, 2014.

M. Petrovick, E. Schwoebel, D. Ricke, T. Boettcher, C. Zook, J. Bobrow, **A. Shcherbina,** J. Harper, E. Wack. Promega 25th International Symposium on Human Identification, Phoenix, AZ, Sep. 29-Oct. 2, 2014.

M. Petrovick, E. Schwoebel, D. Ricke, J. Bobrow, T. Boettcher, J. Harper, B. Helfer, J. Isaacson, **A. Shcherbina**, C. Zook, E. Wack. Next-generation sequencing of STR and SNP loci for human identification and mixture analysis. Promega 24th International Symposium on Human Identification, Atlanta, GA, Oct 7-10, 2013.

A. Lapadula, **A. Shcherbina**, V. Pillalamarri, J. Palmer. STRvillage: A Large-Scale Virtual Pedigree Generator. Promega 21St International Symposium on Human Identification, San Antonio, TX, Oct 11-14, 2010.

CONTRIBUTED TALKS

J. Israeli, CS Foo, **A. Shcherbina**, A. Kundaje. How to train your DragoNN (Deep Regulatory Genomic Neural Network). Jun 9 2016. ENCODE Users Meeting. Stanford University.

D. Ricke, **A. Shcherbina**, C. Aguilar. Bioinformatics analysis of muscle healing. Thomson Reuters Systems Biology Symposium, Cambridge, MA, Jun 25, 2014.

A. Shcherbina, D. Ricke, E Wack, C Rozenzweig, P Carcel, E Fochler. $1M Metagenomic Algorithm: MetaScope update. Algorithms for Threat Reduction Program Review. Boulder, CO, Apr 10, 2014.

W. M. Durham, A. Leombruni, M McKinely**, A. Shcherbina**, R. Stocker. A microfluidic study of biofilms on topographically complex surfaces. American Physical Society, Division of Fluid Dynamics Meeting. Minneapolis, MN, Nov 2009.

VOLUNTEERING/LEADERSHIP

Instructor for Stanford's SAILORS (Stanford Artificial Intelligence Laboratory Outreach Summer Program)	Summ. 2016
MIT Biological Engineering-Biomedical Engineering Society member	2007-2011
MIT Undergraduate Biochemistry Association	2007-2011

SKILLS

Computer: Python, Java, Scala, C#, C++, MATLAB, R, SQL, Unix, LaTeX, Git, SVN, HTML, Microsoft Office.
Language: Russian (native speaker), Spanish (fluent)