Brian Ondov, Ph.D.

brian.ondov@gmail.com • ondov.net

Postdoctoral Fellow

National Library of Medicine

Seeking

Research position in Deep Learning for Natural Language Processing

Research

Citations: 4k+

H-index: 13

Project	Published	Year
Medical Text Consumerization	*(in preparation)*	
Adversarial Perceptual Proxies	IEEE VIS (TVCG)	2020
Perceptual Proxies of visual comparison	IEEE VIS (TVCG)	2019
Mash Screen - Genome containment estimation	Genome Biology	2019
Face to Face - Evaluating visual comparison	IEEE VIS (TVCG)	2018
Mash - Fast genome distance using MinHash	Genome Biology	2016
Gingr - Visualization of genome alignments	Genome Biology	2014
Krona - Interactive metagenomic visualization	BMC Bioinformatics	2011

Education

Degree	Major	School	Year
PhD	Computer Science	University of Maryland College Park	2021
MS	Bioinformatics	Georgia Institute of Technology	2009
BS	Computer Science	Rensselaer Polytechnic Institute	2005

Work experience

Employer	Title	Years
National Library of Medicine	Post-doctoral Fellow	2021–*2022*
National Human Genome Research Institute	Pre-doctoral Fellow	2017–2021
National Human Genome Research Institute	Bioinformatics Scientist	2016–2017
National Biodefense Analysis Center	Bioinformatics Engineer	2010–2016
J. Craig Venter Institute	Intern	2008–2008
Vicarious Visions (Activision)	Associate Programmer	2005–2007

Skills

Language	Applications
Python	TensorFlow, PyTorch, NumPy, Scikit-Learn
C++	STL, Qt, OpenGL
Perl	OOP, XML, CGI
Linux/Unix	Git/GitHub, Automake, SGE
JavaScript	D3, HTML5, MTurk

Education

Certificates

- Deep Learning Specialization, deeplearning.ai (Coursera)

University of Maryland College Park 2017-2020

Doctor of Philosophy, Computer Science

- **Dissertation:** Developed interactive perceptual experiments to study the efficacy of different comparative data visualization strategies, and used human-driven optimization approach to study underlying perceptual patterns.
- **Coursework:** Machine Learning, Neural Machine Translation, Computational Linguistics I & II, Geog. Info. Sys. & Spatial DBs, Interactive Data Analytics, Applied Mechanism Design for Social Good, Social Media Analysis

Georgia Institute of Technology 2007-2009

Master of Science, Bioinformatics

- Developed tools for analyzing *Bacillus anthracis* RNA-Seq data under Professor Nicholas Bergman
    - **SOCS** - An alignment tool for ABI Solid data capable of handling bisulfite-modified bases in color space (sole engineer)
- Received Bioinformatics department award for outstanding student research

Rensselaer Polytechnic Institute 2001-2005

Bachelor of Science, Computer Science

- Developed an educational game aimed at primary school students (independent study)
- Co-oped as programmer for Vicarious Visions
- Upsilon Phi Epsilon computing honor society

Outreach

- **Mentor**, NIH Graduate Data Science Summer Program, 2019
- **Guest Instructor**, UMD CBCB Undergrad Summer Internship Program, 2019
- **Panelist**, *Becoming a Data Scientist*, NIH Graduate and Professional School Fair, 2019
- **Peer reviewer**, Genome Biology/Nature/Bioinformatics/EMNLP, ongoing
- Extracurricular:
    - **Director,** NIH *Nerds In Harmony* (choir), 2020-2021
    - **Instructor,** Yoga, UMD Recreation & Wellness Group Fitness, 2019-2020
    - **Donor Ambassador**, Red Cross, 2020-*present*

Experience

National Library of Medicine 2020-*present*

Post-doctoral IRTA

- Currently developing system for “translating” medical lingo to consumer-friendly language
  - Used deep semantic embeddings to mine professional-level and consumer-level corpora for parallel pairs of sentences or passages
  - Developed machine-assisted Cloze procedure, which generates multiple-choice distractors using transformers, to use for user studies of comprehension

---

National Human Genome Research Institute 2016-2020

- **Pre-doctoral IRTA (2017-*present*)**
  - Performed retroactive screen of SRA metagenomes for recent microbial genomes
  - Performed perceptual user studies for comparative visualizations (dissertation)
- **Bioinformatics Scientist (2016-2017, contracted by Columbus)**
  - Continued development of Mash, including screening algorithm
  - Designed and developed lab website [genomeinformatics.github.io]

National Biodefense Analysis & Countermeasures Center 2010-2016

Bioinformatics Engineer

- Developed tools to expand the forensics capabilities of the Genomics group
  - **Mash** - a fast sequence distance estimator using the MinHash algorithm
  - **Gingr** - a visualization companion for *Parsnp* that can display alignments and trees with thousands of genomes
  - **Krona** - a browser-based metagenomic visualization tool
- Received three Outstanding Performance Awards

J. Craig Venter Institute 2008

Intern, Bioinformatics

- Developed a web-based phylogeny generator and viewer for the *Pathema* online resource, including form submission with Perl CGI backend, server side phylogeny computation using Muscle/T-Coffee, and rendering of phylogenetic tree with SVG

Vicarious Visions 2005-2007

Associate Programmer

- Developed GameBoy Advance and Nintendo DS games, including game logic, input processing, and graphics
  - *Guitar Hero* DS (programming and hardware prototyping), *Spiderman 3 DS, Over the Hedge* GBA, *Madagascar Penguins GBA (as Lead Engineer), Batman Begins GBA*
- Received two Manager’s Awards for exceptional performance

Honors and Awards

- NIH Intramural Research Training Award, 2017-present
- Honorable Mention for Best Paper, IEEE VIS, 2019
- Invited Speaker, AT&T Labs Graduate Student Symposium
- Student Research Award, Georgia Tech bioinformatics masters program, 2009
- Upsilon Phi Epsilon computing honor society, 2004

Skills

- Python
    - Machine learning with TensorFlow/PyTorch/Scikit-Learn/NumPy
- C++
    - Data processing and serialization
    - Standard Template Library data structures and algorithms
    - GUI development with the Qt framework
    - Packaging and deployment
- Perl
    - Parsing and manipulation of various bioinformatics data formats
    - XML manipulation
    - Server side CGI scripting
- Linux/Unix
    - Version control with Git and GitHub
    - Building and installation of software packages
    - Sun Grid Engine job submission
    - File manipulation with standard command line tools
- HTML/JavaScript
    - Interactive graphics with D3, HTML5 *canvas*
    - Mechanical Turk SDK
    - Web design with CSS
    - Static CMS with Jekyll
- Swift
    - iOS graphical app development
- Graphic design
    - Designed logo for BullEx, a manufacturer of firefighting training equipment
        - Used from 2004 until 2012, when the company was purchased by Lion
    - Designed logos for Mash, Parsnp, Gingr, and Krona
- Fitness instruction
    - Yoga Alliance RYT® 200 Registered Yoga Teacher
    - UMD RecWell Group Fitness Instructor

Publications

- **Ondov BD**, Yang F, Kay M, Elmqvist N, Franconeri S. Revealing Perceptual Proxies with Adversarial Examples. IEEE Trans. Vis. Comput. Graph. 2020 Oct 23.
- Jardine N, **Ondov BD**, Elmqvist N, Franconeri S. The Perceptual Proxies of Visual Comparison. IEEE Trans. Vis. Comput., Graph. 2019 Aug 20. **[*Hon. Mention for Best Paper*]**
- **Ondov BD**, Starrett GJ, Sappington A, Kostic A, Koren S, Buck CB, Phillippy AM. Mash Screen: high-throughput sequence containment estimation for genome discovery. Genome biology. 2019 Dec 1.
- **Ondov BD**, Jardine N, Elmqvist N, Franconeri S. Face to Face: Evaluating visual comparison. IEEE Trans. Vis. Comput. Graph. 2018 Aug 20.
- Bazinet AL, **Ondov BD**, Sommer DD, Ratnayake S. BLAST-based validation of metagenomic sequence assignments. PeerJ. 2018 May 28.
- **Ondov BD**, Treangen TJ, Melsted P, Mallonee AB, Bergman NH, Koren S, Phillippy AM. Mash: fast genome and metagenome distance estimation using MinHash. Genome Biol. 2016.
- Treangen TJ, Maybank RA, Enke S, Friss MB, Diviak LF, Karaolis DK, Koren S, **Ondov B**, Phillippy AM, Bergman NH, Rosovitz MJ. Complete Genome Sequence of the Quality Control Strain Staphylococcus aureus subsp. aureus ATCC 25923. Genome Announc. 2014 Nov 6.
- Treangen TJ, **Ondov BD**, Koren S, Phillippy AM. The Harvest suite for rapid core-genome alignment and visualization of thousands of intraspecific microbial genomes. Genome Biol. 2014.
- Treangen TJ, Koren S, Sommer DD, Liu B, Astrovskaya I, **Ondov B**, Darling AE, Phillippy AM, Pop M. MetAMOS: a modular and open source metagenomic assembly and analysis pipeline. Genome Biol. 2013 Jan 15.
- Passalacqua KD, Varadarajan A, Weist C, **Ondov BD**, Byrd B, Read TD, Bergman NH. Strand-specific RNA-seq reveals ordered patterns of sense and antisense transcription in Bacillus anthracis. PLoS One. 2012.
- **Ondov BD**, Bergman NH, Phillippy AM. Interactive metagenomic visualization in a Web browser. BMC Bioinformatics. 2011 Sep 30.
- **Ondov BD**, Cochran C, Landers M, Meredith GD, Dudas M, Bergman NH. An alignment algorithm for bisulfite sequencing using the Applied Biosystems SOLiD System. Bioinformatics. 2010 Aug 1.
- Passalacqua KD, Varadarajan A, **Ondov BD**, Okou DT, Zwick ME, Bergman NH. Structure and complexity of a bacterial transcriptome. J Bacteriol. 2009 May.
- **Ondov BD**, Varadarajan A, Passalacqua KD, Bergman NH. Efficient mapping of Applied Biosystems SOLiD sequence data to a reference genome for functional genomic applications. Bioinformatics. 2008 Dec 1.