

# David Wadden

## Curriculum Vitae

Paul G. Allen Center for Computer Science & Engineering

University of Washington

✉ [dwadden@cs.washington.edu](mailto:dwadden@cs.washington.edu)

📄 [dwadden.github.io](https://dwadden.github.io)

## Research Interests

I study natural language processing, with a focus on information extraction and fact-checking. I like to work on applications in science and health.

## Education

since 2016 **PhD Computer Science**, *University of Washington*, Seattle.

2016–2019 **MS Computer Science**, *University of Washington*, Seattle.

2006–2010 **BA Physics**, *Amherst College*, Amherst.

With Distinction

Phi Beta Kappa

## Publications

### Conference publications

- 2021 **David Wadden**, Tal August, Qisheng Li, Tim Althoff. The Effect of Moderation on Online Mental Health Conversations. *ICWSM 2021*.
- 2020 **David Wadden**, Shanchuan Lin, Kyle Lo, Lucy Lu Wang, Madeleine van Zuylen, Arman Cohan, Hannaneh Hajishirzi. Fact or Fiction: Verifying Scientific Claims. *EMNLP 2020*.
- 2019 **David Wadden**, Ulme Wennberg, Yi Luan, Hannaneh Hajishirzi. Entity, Relation, and Event Extraction with Contextualized Span Representations. *EMNLP 2019*.
- 2019 Yi Luan, **David Wadden**, Luheng He, Hannaneh Hajishirzi, Mari Ostendorf. A General Framework for Information Extraction using Dynamic Span Graphs. *NAACL 2019*.

### Journal papers

- 2019 The ALS Stratification Consortium. Stratification of amyotrophic lateral sclerosis patients: a crowdsourcing approach. *Scientific Reports* **9**, Article number: 690 (2019).
- 2018 Oana M Enache, David L Lahr, Ted E Natoli, Lev Litichevskiy, **David Wadden**, Corey Flynn, Joshua Z Gould, Jacob K Asiedu, Rajiv Narayan, Aravind Subramanian. The GCTx format and `cmap{Py, R, M, J}` packages: resources for the optimized storage and integrated traversal of dense matrices of annotated dense matrices. *Bioinformatics*, bty784.
- 2017 Aravind Subramanian et al. (**David Wadden** author #17 of 52). A Next Generation Connectivity Map: L1000 Platform And The First 1,000,000 Profiles. *Cell* 171(6):1437-1452.e17, 2017.  
Ian Smith, Peyton Greenside, Ted Natoli, David L. Lahr, **David Wadden**, Itay Tirosh, Rajiv Narayan, David E. Root, Todd R. Golub, Aravind Subramanian, John G. Doench. Evaluation of RNAi and CRISPR technologies by large-scale gene expression profiling in the Connectivity Map. *PLoS Biol* 15(11):e2003213.

---

## Research Experience

- Since Fall '18 **Graduate Research Assistant**, UNIVERSITY OF WASHINGTON, Seattle, WA.
- **Supervisor:** Professor Hannaneh Hajishirzi
  - **Research focus:** NLP methods for scientific document understanding.
- Fall-Spr '19-20 **Research Intern**, ALLEN INSTITUTE FOR AI, Seattle, WA.
- **Supervisors:** Professor, Hannaneh Hajishirzi, Kyle Lo
  - **Research focus:** Dataset creation and modeling for scientific fact-checking.
- Win-Spr '18 **Graduate Research Assistant**, UNIVERSITY OF WASHINGTON, Seattle, WA.
- **Supervisor:** Professor William Noble
  - **Research focus:** Recurrent neural network models for peptide retention time prediction.
- Spr-Sum'17 **Modeling Intern**, ALLEN INSTITUTE FOR CELL SCIENCE, Seattle, WA.
- **Supervisor:** Molly Maleckar, PhD
  - **Research focus:** Time series modeling of fluorescence microscopy imaging.
- Spr'14-Spr'15 **Research Assistant**, MIT PROBABILISTIC COMPUTING PROJECT, Cambridge, MA.
- **PI:** Vikash Mansinghka, PhD
  - **Research focus:** The Venture system for probabilistic programming.
- Spr'11-Spr'14 **Associate Computational Biologist**, THE BROAD INSTITUTE OF MIT AND HARVARD, Cambridge, MA.
- **PI:** Professor Todd Golub
  - **Supervisor:** Aravind Subramanian, PhD
  - **Research focus:** Computational tools for analyzing gene expression data.

---

## Industry Experience

- Sum'18 **Software Engineering Intern**, GOOGLE, Kirkland, WA.
- **Supervisor:** Evgeny Skvortsov, PhD
  - **Internship focus:** Designed and implemented a machine learning approach to forecast and report ad reach into affinity (i.e. interest) categories. Trained machine learning models using TensorFlow.

---

## Awards

- 2010 William Warren Stifler Prize in Physics, Amherst College
- 2007 Bassett Physics Prize, Amherst College

---

## Relevant Coursework

Machine learning, natural language processing, probabilistic graphical models, statistical inference, information theory, probability theory, convex optimization.

---

## Programming Languages and Software Tools

Python, JavaScript, R, C, Matlab, Scheme.

PyTorch, Tensorflow, AllenNLP, scikit-learn, Pandas, Git.

---

## Hobbies and interests

- 2014-2015 **Emerging Artist**, BOSTON LYRIC OPERA, Boston, MA.
- Spent a year working as an opera singer.