

# Daniel Cohen

Email : dan.cohenCS@gmail.com

Website: dscohen.github.io

## RESEARCH INTERESTS

---

Information Retrieval, Natural Language Processing, Machine Learning focused on: text matching, ranking and relevance, deep learning, question answering, domain transfer, reinforcement learning, and reasoning under uncertainty. I have consistent publications with respect to these areas (SIGIR, EMNLP, ICML among others).

## EDUCATION

---

- **Ph.D. Computer Science.** University of Massachusetts Amherst. Amherst, MA. 2015 - 2020
- **M.S. Computer Science.** University of Massachusetts Amherst. Amherst, MA. 2015 - 2017
- **B.A. Computer Science and Mathematics.** New York University. NY, NY. 2013 - 2015

## EXPERIENCE

---

- **Research Scientist** - Dataminr Aug. 2022 - present
  - I work on developing methods to improve the detection of critical events in a high volume, online setting.
- **Postdoctoral Researcher** - Brown University (Prof. Carsten Eickhoff) Dec. 2020 - July 2022
  - I executed an independent research plan in addition to guiding current PhD and undergraduate students with their own research agenda and ensuring grant deliverables. My work focused on developing methods to improve decision making in uncertain ranking situations.
- **Research Assistant** - UMass Amherst (Prof. W. Bruce Croft) Sep. 2015 - Dec. 2020
  - I developed novel approaches for training and ensuring robustness of neural models in the realm of information retrieval and question answering.
- **Site Project Manager** - UMass Amherst (IARPA MATERIAL) Sep. 2017 - Dec. 2020
  - I managed the UMass Amherst responsibilities of the greater multi-university team with biweekly updates, reports, and yearly evaluations. We concluded the grant with the best information retrieval scores among all teams.
- **Research Intern** - Microsoft Research (Dr. Fernando Diaz, Dr. Bhaskar Mitra) Feb 2019 - Jun.2019
  - I incorporated well studied search approaches to identify promising policies for task transfer in reinforcement learning. While previous approaches rely on simultaneous learning or expensive comparisons, our approach builds an index of policies to allow for efficient lookup via dense retrieval.
- **Research Intern** - Microsoft Research (Dr. Katja Hofmann, Dr. Bhaskar Mitra) May 2017 - Aug.2017
  - I worked on the task of adapting models to new domains and markets with minimal labeled data. Our proposed solution leveraged adversarial learning to allow for a robust domain agnostic model that can function gracefully across new out of distribution samples.
- **Research Assistant** - New York University (Prof. Mohamed Zahran) Jun. 2014 - Jan. 2015
  - With the growing trend of using GPUs as general compute devices, I introduced a toolkit to identify both processor and memory bottlenecks in user's CUDA code when executed. We implemented a simulated GPU to allow for any GPU architecture to be used in the toolkit.

## PUBLICATIONS

---

- [1] George Zerveas, Navid Rekabsaz, **Daniel Cohen**, and Carsten Eickhoff. Mitigating bias in search results through set-based document reranking and neutrality regularization. In *SIGIR '22*, Madrid, Spain, 2022. ACM.
- [2] **Daniel Cohen**, Kevin Du, Bhaskar Mitra, Laura Mercurio, Navid Rekabsaz, and Carsten Eickhoff. Inconsistent ranking assumptions in medical search and their downstream consequences. In *SIGIR '22*, Madrid, Spain, 2022. ACM.

- [3] Geoorge Zerveas, Navid Rekabsaz, **Daniel Cohen**, and Carsten Eickhoff. Coder: An efficient framework for improving retrieval through contextualized document embedding reranking. *EMNLP '22*, 2022.
- [4] Oleg Lesota, Navid Rekabsaz, **Daniel Cohen**, Klaus Antonius Grasserbauer, Carsten Eickhoff, and Markus Schedl. A modern perspective on query likelihood with deep generative retrieval models. In *ICTIR '21*, Montreal, Canada, 2021. ACM.
- [5] **Daniel Cohen**, Bhaskar Mitra, Oleg Lesota, Navid ReKab-Saz, and Carsten Eickhoff. Not all relevance scores are equal: Efficient uncertainty and calibration modeling for deep retrieval models. In *SIGIR '21*, Montreal, Canada, 2021. ACM.
- [6] **Daniel Cohen**. Allowing for the grounded use of temporal difference learning in large ranking models via sub state updates. In *SIGIR '21*, Montreal, Canada, 2021. ACM.
- [7] Scott Jordan, Yash Chandak, **Daniel Cohen**, Mengxue Zhang, and Philip S. Thomas. Evaluating the performance of reinforcement learning algorithms. In *ICML*, volume 97 of *PMLR*, Vienna, Austria, 12–18 Jul 2020. PMLR.
- [8] Yen-Chieh Lien, **Daniel Cohen**, and W. Bruce Croft. An assumption-free approach to the dynamic truncation of ranked lists. In *ICTIR 2019*, pages 79–82, Santa Clara, CA, USA, October 2-5 2019. ACM.
- [9] **Daniel Cohen**, Scott M. Jordan, and W. Bruce Croft. Learning a better negative sampling policy with deep neural networks for search. In *ICTIR 2019*, page 19–26, New York, NY, USA, 2019. ACM.
- [10] Scott M. Jordan, **Daniel Cohen**, and Philip S. Thomas. Evaluating reinforcement learning algorithms using cumulative distributions of performance. In *NeurIPS - Workshop on Critiquing and Correcting Trends in Machine Learning*. Montreal, Canada, December 3-8 2018.
- [11] Constantine Lignos, **Daniel Cohen**, Yen-Chieh Lien, Pratik Mehta, W. Bruce Croft, and Scott Miller. The challenges of optimizing machine translation for low resource cross-language information retrieval. In *EMNLP-IJCNLP*, pages 3497–3502, Hong Kong, China, November 2019. ACL.
- [12] **Daniel Cohen**, Brendan O'Connor, and W. Bruce Croft. Understanding the representational power of neural retrieval models using nlp tasks. In *ICTIR '18*, page 67–74, New York, NY, USA, 2018. ACM.
- [13] **Daniel Cohen**, Scott M. Jordan, and W. Bruce Croft. Distributed evaluations: Ending neural point metrics. In *SIGIR - LND4IR Workshop*, SIGIR '18, New York, NY, USA, 2018. ACM.
- [14] **Daniel Cohen**, Liu Yang, and W. Bruce Croft. Wikipassageqa: A benchmark collection for research on non-factoid answer passage retrieval. In *SIGIR '18*, page 1165–1168, New York, NY, USA, 2018. ACM.
- [15] **Daniel Cohen**, John Foley, Hamed Zamani, James Allan, and W. Bruce Croft. Universal approximation functions for fast learning to rank: Replacing expensive regression forests with simple feed-forward networks. In *SIGIR '18*, page 1017–1020, New York, NY, USA, 2018. ACM.
- [16] **Daniel Cohen**, Bhaskar Mitra, Katja Hofmann, and W. Bruce Croft. Cross domain regularization for neural ranking models using adversarial learning. In *SIGIR '18*, page 1025–1028, New York, NY, USA, 2018. ACM.
- [17] **Daniel Cohen** and W. Bruce Croft. A hybrid embedding approach to noisy answer passage retrieval. volume 10772 of *ECIR 2018*, pages 127–140, Grenoble, France, March 26-29 2018. Springer.
- [18] **Daniel Cohen** and W. Bruce Croft. End to end long short term memory networks for non-factoid question answering. *ICTIR 2016*, pages 143–146. ACM, September 12-16 2016.
- [19] **Daniel Cohen**, Qingyao Ai, and W. Bruce Croft. Adaptability of neural networks on varying granularity ir tasks. In *SIGIR - Neu-IR Workshop*, SIGIR '16, New York, NY, USA, 2018. ACM.

---

## TECHNICAL SKILLS

- **Programming:** Java, Python, C, Julia
- **Theory:** Math, Statistics, Probability Theory, Reinforcement Learning, Machine Learning, Deep Learning, NLP, IR
- **Common Tools:** Numpy, Scipy, Scikit-learn, Pandas, Theano, Tensorflow, PyTorch, NLTK, Galago, PyLucene, COSMOS, Linux/Unix, Bash, SQL, git

---

## AWARDS AND GRANTS

- Best full paper, ICTIR

- SIGIR Student Travel Grant 2019
- Bloomberg Data Science Research Grant, co-writer. 2018
- Best short paper, SIGIR 2018
- SIGIR Student Travel Grant 2018
- IARPA MATERIAL Research Grant, co-writer. 2017
- SIGIR Student Travel Grant 2016
- Best poster - NYU Undergraduate Research Conference 2016

#### PROFESSIONAL ACTIVITIES AND SERVICE

---

- **Program Committee Member** SIGIR, EMNLP, EACL, CIKM, WWW, WSDM, ACL, KDD, AAI, CIKM
- **Guest Lecturer** - NYU ACM 2016
- **Treasurer** - NYU ACM 2013-2014