

OBJECTIVE

Generalist researcher with a diverse background in applied natural language processing and cloud infrastructure interested in evaluating, analyzing and uncovering model failure in generative language models. Opportunity to work within a multi-disciplinary research team with real-world data to influence product development considered a significant plus.

EDUCATION

- | | |
|---|----------------------------|
| University of Toronto
Ph.D. in Electrical and Computer Engineering
– Thesis: “Predictor Virtualization: Teaching Old Caches New Tricks” | Toronto, Canada
2012 |
| University of Toronto
M.A.Sc. in Electrical and Computer Engineering
– Thesis: “XToPSS-efficient XML filtering against large numbers of XPath expressions” | Toronto, Canada
2004 |
| University “Politehnica” of Bucharest
B.S. in Computer Science and Engineering
– Valedictorian | Bucharest, Romania
2002 |

WORK EXPERIENCE

- | | |
|---|--------------------------------------|
| IBM Research
Senior Research Scientist | Yorktown Heights, NY
2012-Current |
| Bias measurement and analysis of NLP systems
Technical Lead | 2021-Current |
- **Generative language models auditing and red-teaming** - Coordinating research across several lab locations responsible for auditing open-source and in-house generative language models. Design new datasets that uncover unwanted social bias related to stigmatized conditions; develop red-teaming infrastructure and conduct red-teaming experiments to discover different model failures; evaluate bias in generative language models [7, 10].
 - **The interplay between bias and robustness in NLP systems** - Using language models to suggest lexical variations of existing datasets along with adversarial filtering, create challenging samples that uncover bias in downstream tasks (e.g., NLI) [7, 6]. Propose new bias measures that take into account the type of model error.
 - **Fairness analysis of language models for toxic text prediction** - Analyzed the performance and fairness characteristics for more than a dozen language models for toxic text prediction. Applied post-processing methods for bias mitigation and showed promising results [1].
 - **Doctoring the note: Using AI to improve clinical care** - MIT-IBM Collaboration, co-PI with Dr. Marzyeh Ghassemi: Demonstrate that clinical notes contain a non-negligible level of Black patient bias and can be used to infer patient’s race with high accuracy, while clinicians cannot determine patient’s race using the same information. This hidden bias can have consequences in downstream tasks using clinical notes [11].

- **Fairness Auditor** - RPI-IBM Collaboration, co-PI with Dr. Kristin Bennett, work in progress: Generate synthetic healthcare data and analyze its fairness-utility characteristics. Generate datasets with various degrees of bias and shift to stress-test existing methods for bias mitigation [2, 4, 3, 8].
- **Dataset Debugging** - RPI-IBM Collaboration, work in progress: Use weak supervision methods to detect problematic data samples and audit toxic text datasets [12].
- **Biasly** - MILA-IBM Multidisciplinary collaboration: Collection, annotation and model training with a large dataset for detecting misogyny in movie scripts.

Science for social good projects in partnership with non-profit organizations

Technical Lead

2017-2021

- **Drug repurposing for treating cancer** - Partner: RebootRx - Built an end-to-end pipeline for filtering biomedical articles and extracting information related to the efficacy of generic drugs for treating cancer. Lead all components of the pipeline, from developing crowd-sourcing forms and collecting data to building language models based predictors for filtering and information extraction, along with a web-based application to browse the prediction results [13, 15, 5, 14, 16]. The pipeline was used by our partner clinical practitioners to timely select documents from PubMed and create a COVID dataset relevant to cancer patients ([Reboot: Covid Cancer Project](#)).
- **Document filtering for biomimicry** - Partner: Biomimicry Institute - Built an end-to-end pipeline for filtering articles relevant to biomimicry. Developed an internal crowd-sourcing framework using serverless computing to collect data annotations for biomimicry articles and build models for categorization of such documents [20].
- **Semantic understanding of data science programs** - Partner: Accelerated Cure Project for Multiple Sclerosis - Envisioned ways in which data scientists could better collaborate to work on open-ended problems. Developed a technique for expressing the semantic meaning of a data science program [18, 19, 17, 24, 23].

Cloud infrastructure and serverless computing

Individual contributor

2014-2017

- **Serverless platform: OpenWhisk and IBM Cloud Functions** - Part of a research team that built a prototype for a serverless platform. The prototype was open-sourced as [OpenWhisk](#). Worked in a large team of researchers and software engineers and extended the prototype to a full fledged product featured as [IBM Cloud Functions](#). Responsible for different components of the system related to service name resolution, collecting, storing and monitoring component logs, and chaining the execution of serverless functions [22, 21, 25, 27].

High-level programming language compilation for heterogeneous architectures

2012-2014

Individual Contributor

- Part of a research team that designed and implemented a high-level programming language and a compiler for heterogeneous architectures comprising CPUs, GPUs and FPGAs. Responsible for the dispatch of code in heterogeneous runtime systems [29, 30]. Used machine learning to predict potential speedups from GPU execution [28].

University of Toronto

Toronto, Canada

Graduate Student Researcher

2002-2011

- **Predictor Virtualization:** Designed, implemented and evaluated a method for increasing the perceived capacity of on-chip hardware predictors such as memory caches and branch predictors [31, 33, 34]. A similar method was independently proposed and implemented in real hardware in the IBM Z series mainframe computers.
- **Publish-subscribe systems:** Worked in the area of distributed systems and developed different algorithms for document matching and routing in publish-subscribe systems [35, 37, 36, 39].

IBM Research

Yorktown Heights, NY

Research Intern

June-Aug 2010

- Scheduling policies for processor instruction queues in the mainframe Z series. Work presented to the mainframe processor design team and patented.

Intel Labs

Research Intern

Santa Clara, CA

Jan-July 2009

- Analysis of on-chip cache replacement policies for Intel processors.

Microsoft

Student Consultant

Bucharest, Romania

2000-2002

- Liaison between the local Microsoft subsidiary and the Computer Science and Engineering Department at “Politehnica” University of Bucharest. Participated in training and European summits with the European Microsoft Student Consultant cohort. Part of the support staff for the Microsoft laboratory in the department. Taught C and C++ courses to first and second year students.

PUBLICATIONS

- [1] Ioana Baldini, Dennis Wei, Karthikeyan Natesan Ramamurthy, Mikhail Yurochkin, and Moninder Singh. “Your Fairness May Vary: Pretrained Language Model Fairness in Toxic Text Classification”. In: *Findings of ACL 2022*.
- [2] Karan Bhanot, Ioana Baldini, Dennis Wei, Jiaming Zeng, and Kristin P. Bennett. “Downstream Fairness Caveats with Synthetic Healthcare Data (poster)”. In: *CHIL’22: Conference on Health, Inference, and Learning 2022*.
- [3] Karan Bhanot, Ioana Baldini, Dennis Wei, Jiaming Zeng, and Kristin P. Bennett. “Evaluating Fairness of Synthetic Healthcare Data Models, Poster presentation (poster)”. In: *AMIA 2022 Annual Symposium*.
- [4] Karan Bhanot, Ioana Baldini, Dennis Wei, Jiaming Zeng, and Kristin P. Bennett. “Stress-testing Fairness Mitigation Techniques under Distribution Shift using Synthetic Data”. In: *ACM SIGKDD Workshop on Ethical Artificial Intelligence: Methods and Applications, KDD 2022*.
- [5] Sejal Dua, Ioana Baldini, Dmitriy A. Katz-Rogozhnikov, Emily van der Veen, Allison Britt, Pradeep Mangalath, Laura B. Kleiman, and Catherine Del Vecchio Fitz. “Biomedical Corpus Filtering: A Weak Supervision Paradigm With Infused Domain Expertise”. In: *Workshop on Scientific Document Understanding co-located with AAAI 2021*.
- [6] Chhavi Yadav, Ioana Baldini, Payel Das, and Kush R. Varshney. “Belief Injection for Adversarial Bias Attacks in Natural Language Inference”. In: *under review*.
- [7] Ioana Baldini, Chhavi Yadav, Payel Das, and Kush R. Varshney. “Keeping Up with the Language Models: Robustness-Bias Interplay in NLI Data and Models”. In: *TrustNLP 2023, colocated with ACL 2023*. 2023.
- [8] Karan Bhanot, Ioana Baldini, Dennis Wei, Jiaming Zeng, and Kristin P. Bennett. “Stress-testing Bias Mitigation Algorithms to Understand Fairness Vulnerabilities”. In: *AIES ’23: AAAI/ACM Conference on AI, Ethics, and Society*. 2023.
- [9] Tim Draws, Karthikeyan Natesan Ramamurthy, Ioana Baldini, Amit Dhurandhar, Inkit Padhi, Benjamin Timmermans, and Nava Tintarev. “Explainable Cross-Topic Stance Detection for Search Results”. In: *Proceedings of the 2023 Conference on Human Information Interaction and Retrieval*. New York, NY, USA: Association for Computing Machinery, 2023. ISBN: 9798400700354. DOI: 10.1145/3576840.3578296. URL: <https://doi.org/10.1145/3576840.3578296>.
- [10] Stefania Druga, Rishi Bommasani, Mihaela Vorvoreanu, and Ioana Baldini. “Language Models and Society: Bridging Research and Policy”. In: *ACM Conference on Fairness, Accountability, and Transparency (ACM FAccT)*. 2023.

- [11] Hammaad Adam, Ming-Ying Yang, Kenrick Cato, Ioana Baldini, Charles Senteio, Leo Anthony Celi, Jiaming Zeng, Moninder Singh, and Marzyeh Ghassemi. “Write It Like You See It: Detectable Differences in Clinical Notes by Race Lead to Differential Model Recommendations”. In: *AIES '22: AAAI/ACM Conference on AI, Ethics, and Society*. 2022.
- [12] Kofi Arhin, Ioana Baldini, Dennis Wei, Karthikeyan Natesan Ramamurthy, and Moninder Singh. “Ground-truth, whose truth? Examining the difficulties with annotating toxic text datasets”. In: *Data-Centric AI Workshop colocated with NeurIPS 2021*, to appear. 2021.
- [13] Ioana Baldini, Mariana Bernagozzi, Sulbha Aggarwal, Mihaela A. Bornea, Saksham Chawla, Joppe Geluykens, Dmitriy A. Katz-Rogozhnikov, Pratik Mukherjee, Smruthi Ramesh, Sara Rosenthal, Jagrati Sharma, Kush R. Varshney, Laura B. Kleiman, Pradeep Mangalath, and Catherine Del Vecchio Fitz. “Exploring the Efficacy of Generic Drugs in Treating Cancer”. In: *AAAI*. 2021.
- [14] Lu Cheng, Dmitriy A. Katz-Rogozhnikov, Kush R. Varshney, and Ioana Baldini. “Automated Meta-Analysis: A Causal Learning Perspective”. In: *CoRR* abs/2104.04633 (2021).
- [15] Diego García-Olano, Yasumasa Onoe, Ioana Baldini, Joydeep Ghosh, Byron C. Wallace, and Kush R. Varshney. “Biomedical Interpretable Entity Representations”. In: *Findings of the Association for Computational Linguistics: ACL/IJCNLP 2021*. 2021.
- [16] Shivashankar Subramanian, Ioana Baldini, Sushma Ravichandran, Dmitriy A. Katz-Rogozhnikov, Karthikeyan Natesan Ramamurthy, Prasanna Sattigeri, Kush R. Varshney, Annmarie Wang, Pradeep Mangalath, and Laura B. Kleiman. “A Natural Language Processing System for Extracting Evidence of Drug Repurposing from Scientific Publications”. In: *The Thirty-Second Innovative Applications of Artificial Intelligence Conference, IAAI 2020, co-located with AAAI 2020*. 2020.
- [17] Yaoli Mao, Dakuo Wang, Michael J. Muller, Kush R. Varshney, Ioana Baldini, Casey Dugan, and Aleksandra Mojsilovic. “How Data Scientists Work Together With Domain Experts in Scientific Collaborations: To Find The Right Answer Or To Ask The Right Question?”. In: *ACM Human Computer Interaction* (2019).
- [18] Evan Patterson, Ioana Baldini, Aleksandra Mojsilovic, and Kush R. Varshney. “Semantic Representation of Data Science Programs”. In: *International Joint Conference on Artificial Intelligence, IJCAI*. 2018.
- [19] Evan Patterson, Ioana Baldini, Aleksandra Mojsilovic, and Kush R. Varshney. “Teaching machines to understand data science code by semantic enrichment of dataflow graphs”. In: *CoRR* abs/1807.05691 (2018). URL: <http://arxiv.org/abs/1807.05691>.
- [20] Yuanshuo Zhao, Ioana Baldini, Prasanna Sattigeri, Inkit Padhi, Yoong Keok Lee, and Ethan Smith. “Data Driven Techniques for Organizing Scientific Articles Relevant to Biomimicry”. In: *Conference on AI, Ethics, and Society, AIES 8*. Ed. by Jason Furman, Gary E. Marchant, Huw Price, and Francesca Rossi. 2018.
- [21] Ioana Baldini, Paul C. Castro, Kerry Shih-Ping Chang, Perry Cheng, Stephen Fink, Vatche Ishakian, Nick Mitchell, Vinod Muthusamy, Rodric Rabbah, Aleksander Slominski, and Philippe Suter. “Serverless Computing: Current Trends and Open Problems”. In: *Research Advances in Cloud Computing*. 2017.
- [22] Ioana Baldini, Perry Cheng, Stephen J. Fink, Nick Mitchell, Vinod Muthusamy, Rodric Rabbah, Philippe Suter, and Olivier Tardieu. “The serverless trilemma: function composition for serverless computing”. In: *ACM SIGPLAN International Symposium on New Ideas, New Paradigms, and Reflections on Programming and Software, Onward!* 2017.
- [23] Evan Patterson, Ioana Baldini, Aleksandra Mojsilovic, and Kush R. Varshney. “Machine Representation of Data Analyses: Towards a Platform for Collaborative Data Science”. In: *2017 AAAI Spring Symposia*, 2017.

- [24] Evan Patterson, Robert N. McBurney, H. Schmidt, Ioana Baldini, Aleksandra Mojsilovic, and Kush R. Varshney. “Dataflow representation of data analyses: Toward a platform for collaborative data science”. In: *IBM Journal of Research and Development* (2017).
- [25] Ioana Baldini, Paul C. Castro, Perry Cheng, Stephen J. Fink, Vatche Ishakian, Nick Mitchell, Vinod Muthusamy, Rodric Rabbah, and Philippe Suter. “Cloud-native, event-based programming for mobile applications”. In: *Proceedings of the International Conference on Mobile Software Engineering and Systems, MOBILESoft*. 2016.
- [26] Islam Atta, Xin Tong, Vijayalakshmi Srinivasan, Ioana Baldini, and Andreas Moshovos. “Self-contained, accurate precomputation prefetching”. In: *International Symposium on Microarchitecture, MICRO*. 2015.
- [27] Alexander Tarvo, Peter F. Sweeney, Nick Mitchell, V. T. Rajan, Matthew Arnold, and Ioana Baldini. “CanaryAdvisor: a statistical-based tool for canary testing (demo)”. In: *Proceedings of the 2015 International Symposium on Software Testing and Analysis, ISSTA*. Ed. by Michal Young and Tao Xie. 2015.
- [28] Ioana Baldini, Stephen J. Fink, and Erik R. Altman. “Predicting GPU Performance from CPU Runs Using Machine Learning”. In: *International Symposium on Computer Architecture and High Performance Computing, SBAC-PAD*. 2014.
- [29] Erik R. Altman, Joshua S. Auerbach, David F. Bacon, Ioana Baldini, Perry Cheng, Stephen J. Fink, and Rodric M. Rabbah. “The Liquid Metal Blokus Duo Design”. In: *International Conference on Field-Programmable Technology, FPT*. 2013.
- [30] Joshua S. Auerbach, David F. Bacon, Ioana Burcea, Perry Cheng, Stephen J. Fink, Rodric M. Rabbah, and Sunil Shukla. “A compiler and runtime for heterogeneous computing”. In: *The 49th Annual Design Automation Conference 2012, DAC*. Ed. by Patrick Groeneveld, Donatella Sciuto, and Soha Hassoun. 2012.
- [31] Ioana Burcea, Livio Soares, and Andreas Moshovos. “Pointy: a hybrid pointer prefetcher for managed runtime systems”. In: *International Conference on Parallel Architectures and Compilation Techniques, PACT*. 2012.
- [32] Mohammad Sadoghi, Ioana Burcea, and Hans-Arno Jacobsen. “GPX-matcher: a generic boolean predicate-based XPath expression matcher”. In: *14th International Conference on Extending Database Technology*. 2011.
- [33] Ioana Burcea and Andreas Moshovos. “Phantom-BTB: a virtualized branch target buffer design”. In: *Proceedings of the 14th International Conference on Architectural Support for Programming Languages and Operating Systems ASPLOS*. 2009.
- [34] Ioana Burcea, Stephen Somogyi, Andreas Moshovos, and Babak Falsafi. “Predictor virtualization”. In: *Proceedings of the 13th International Conference on Architectural Support for Programming Languages and Operating Systems ASPLOS*. 2008.
- [35] Ioana Burcea, Hans-Arno Jacobsen, Eyal de Lara, Vinod Muthusamy, and Milenko Petrovic. “Disconnected Operation in Publish/Subscribe Middleware”. In: *5th IEEE International Conference on Mobile Data Management MDM*. 2004.
- [36] Ioana Burcea and Hans-Arno Jacobsen. “L-ToPSS - Push-Oriented Location-Based Services”. In: *Technologies for E-Services Workshop, TES*. 2003.
- [37] Ioana Burcea, Milenko Petrovic, and Hans-Arno Jacobsen. “I know what you mean: semantic issues in Internet-scale publish/subscribe systems”. In: *The first International Workshop on Semantic Web and Databases, Co-located with VLDB*. 2003.
- [38] Hubert Ka Yau Leung, Ioana Burcea, and Hans-Arno Jacobsen. “Modeling location-based services with subject spaces”. In: *Conference of the Centre for Advanced Studies on Collaborative Research*. 2003.

- [39] Milenko Petrovic, Ioana Burcea, and Hans-Arno Jacobsen. “S-ToPSS: Semantic Toronto Publish/Subscribe System”. In: *Proceedings of 29th International Conference on Very Large Data Bases, VLDB*. 2003.

SELECTED SCHOLARSHIPS AND AWARDS

- IBM Research Division Award (contributions to OpenWhisk) 2017
- IBM PhD Fellowship 2010
- Google Canada Anita Borg Scholarship 2010
- NSERC (NSF equivalent) Canada Graduate Scholarship 2005–2007
- University “Politehnica” of Bucharest, Computer Science and Engineering, Valedictorian 2002
- Prizes at Romanian National Math Olympiad 1995-1997

SELECTED PROFESSIONAL SERVICE

- **Machine Learning and Artificial Intelligence (2017-Current):** Reviewer for conferences such as AAAI, ICLR, NeurIPS, EMNLP, ACL, ARR. Selected top reviewer several times for NeurIPS and ICLR.
- **IBM Research Computer Science Council Member (2019-2021):** Reviewed and decided funding for long term projects at IBM Research.
- **AAAI Spring Symposium 2019 organizer (with Evan Patterson):** Organized a symposium on AI for collaborative, open science; reviewed papers, invited speakers and coordinated the one-day event.
- **Computer Architecture and Systems (2004-2017):** Reviewer for several conferences such as ASPLOS (Architectural Support for Programming Languages and Operating Systems), ISCA (International Conference on Computer Architecture), PACT (Parallel Architectures and Compilation Techniques), Micro (International Symposium on Microarchitecture), IIWC (International Symposium on Workload Characterization), Supercomputing.

TEACHING

Teaching assistant for more than a dozen courses during graduate and undergraduate studies, spanning a variety of subjects, such as numerical programming, data structures and algorithms, compilers, computer architecture, introduction to programming.

SELECTED DIVERSITY, INCLUSION AND OUTREACH

- TechTogether Boston, Challenge lead 2020
Proposed a challenge, judged participating teams, chose winning team who interned with IBM Research. Mentored 4 undergraduate students during summer internships, one of them is now a full time IBM employee.
- Mentor for Girls Who Code 2017
Presented my research and mentored high school students participating in the summer program.
- Founder of Wonder Years - New Mom Support Group 2015–Current
Initiated a support group for new mothers returning to work. Advocated for mom’s rights in the workplace.
- New Employee Technical Fair, Organizer 2012-2013
Organized research show-case fair for new employees to get exposed to a range of research projects and find mentorships.
- ACM Distinguished Lecture Series, Organizer 2010
Organized and hosted a distinguished lecture series at University of Toronto with the goal of supporting women in tech