**Richard (Tommy) Guy**

Seattle, WA, USA guyrt.github.io

[richardtguy84@gmail.com](mailto:richardtguy84@gmail.com) 425-435-7409

I enjoy challenging positions that require deep technical expertise typically involving both Computer Science and Data Science. I want to work on real-world problems where ownership of the solution is expected of everyone involved. The ideal position combines opportunities for technical, product, and business leadership. The absolute best position also allows for open-source collaboration.

**MAJOR TECHNICAL SKILLS: RESEARCH TO SHIPPING TO MEASURING**

* Data Engineering in massive production environments, A/B Testing, Statistical Machine Learning
* Production Machine Learning deployments including Deep Learning inference and RAG+LLM systems.
* Python: Pandas/numpy/scipy/pytorch, C#, Spark, [In the past C++, Matlab, R]
* Web development: Django, ASP.net, TS, and shipping on AWS and Azure
* Technical leadership experience.
* Major experience in technical writing and communication

**RESEARCH AND WORK EXPERIENCE**

**Principal Architect – AI, Data Science**

Microsoft (Business Applications Platform, Dataverse)

*Worked as the “bridge” between Science teams developing traditional and LLM based ML products and the platform team building their in-production environments.*

* Architected and built out platforms to train and run inference on large scale customer data in a compliant environment. Worked with product teams and Machine Learning specialists to define the “job to be done”, iterate on model and features, and ultimately ship at scale. Our system evolved to support Spark-based batch compute, Azure OpenAI, and Azure Machine Learning based jobs.
* Lead Architect on our LLM/GPT based efforts to “meet the moment” in generative AI. Led team to design, build, ship, and iterate on RAG-based features that leverage the enormously valuable data in Dynamics databases. My role included teaching non-ML specialists and significant product ideation as well as technical guidance. We shipped our first RAG features to market in 6 weeks and are actively iterating.
* Responsible AI Champion for my 500-person organization. Focused on identifying harms, assisting on mitigations, and defining best practices to ensure continuous testing. We focus on shipping highly valuable *but trustworthy* ML often in complex regulatory environments.

**Principal Data Scientist**

Microsoft (Web Data Team) *July 2019 to Oct 2020*

*Collaborated across teams in Bing to improve the web index including improving safety and trustworthiness and improving signals for personalization.*

* Developed and implemented novel metrics strategy that became the main ship metric for Bing Safety team. Previously, the spam team was using a relevance metric for ship decisions, but some spam shipments hurt relevance as previously defined (e.g. protect users from relevant but unsafe content.) With a new metric that was better aligned to overall goals, we were able to (a) ship several techniques that had previously been blocked, (b) improve team morale by giving a clearer goal, and (c) better explain to leadership how the spam team improved Bing’s product.
* Developed pipeline tools to improve model iteration efficiency by 10x. Worked with members to the spam team to incorporate these tools.
* Led efforts to introduce Attention models for detecting machine generated content.
* Led collaboration with relevance team to promote trusted sites for “high risk” queries. As part of this work, I was active on all aspects of the project from project conception to implementation to experimentation.

**Principal Data Scientist (Previously Data Scientist 2, Sr. Data Scientist)**

Microsoft (Analysis & Experimentation) *October 2013 to June, 2019*

*Leader in trustworthy A/B platform development and implementation at Microsoft. Preferred projects that required both technical and educational components. Focused on changing cultures in addition to changing code.*

* Technical Lead on projects that resulted in the first successful A/B tests in Exchange, Outlook, and OneDrive Consumer. This involved direct input on all aspects of the project including metrics development, data pipeline, user randomization, and statistics.
* Scaled A/B testing in Exchange/Outlook from first experiment to 30/month. They now run close to 500/month.
* Collaborated with multiple groups including the Office AI/ML teams to define and measure KPIs for determining A/B test success.
* Mentored data scientists across Microsoft in A/B testing, ML, and data mining roles.
* Lead engineer on multiple efforts to improve infrastructure and product offering with focus on analysis and metrics compute.
* Speaker at multiple internal and external conferences on multiple topics in A/B testing.
* Consistent strong reviews as a technical and thought leader on the A&E team. HiPO participant for FY16. Winner of Individual Contributor award for January, 2016.

**Lead Data Scientist**

Wave Apps *March 2012 to August 2013*

*Manage Wave’s data set and data insights team. Develop analytics platform and provided actionable insight to executives. Architect and develop data-focused consumer applications.*

* Designed, architected, and developed machine learning applications using our data, including *categorization, accounting automation*, and *fraud detection.*
* Architected Wave’s data mining warehouse, which is accessed org-wide.
* Educated organization on statistical best practices and lean analytics methods.
* Made open source contributions to Pandas (CSV input, pivot table edge cases) and Statsmodels (multiple). **I would strongly value a role that includes more open-source development.**

**Software Engineering Course Development** *September 2010 to 2015*

Software Carpentry

*Developed course material and taught software engineering best practices primarily to PhD level scientists.*

* Wrote lesson material on numerical programming with NumPy and Matlab.
* Led many workshops in 4 countries.

**Statistical Genetics Research Assistant** *Aug. 2008 to Aug. 2010 + consulting to 2013*

Wake Forest University Baptist Medical Center

*Primary developer in a Biostatistics lab working on high throughput statistical genetics.*

* Managed all phases of designing, coding, testing, and deploying a fully parallel, statistical genetics software suite in C++ with MPI.
* Led teams of between 2 and 7 programmers.
* Mastered C++, Perl, MPI multithreading, SQL, R, and software engineering in Unix.
* Developed novel Machine Learning Algorithms for genetic associations analysis.

**TEACHING EXPERIENCE**

*[2016,2017,2019] Community Data Science Course (COM597):* I teach a course on Data Science at the University of Washington in the Communication Leadership MA program for aspiring Product Management professionals. Students do not have a technical background, and I focus on the skills necessary to interpret statistical results and to work effectively with Data Scientists in professional settings.

*[2014-2016] Community Data Science Workshops*: Help develop and lead series of workshops introducing Data Science to several hundred researchers in multiple departments. The series is the subject of an upcoming book chapter.

*[2010-2012] University of Toronto*: Numerical Methods and Algorithm Design and Analysis TA

*[2007-2009] Wake Forest University*: Calculus II TA, Multivariate Calculus Tutor

**OTHER PROJECTS**

**Verbal Victor for iPhone (2010-2012)**

*Verbal Victor is an Augmented Communication Device for children with speech and fine motor disorders with a professor at Wake Forest University.*

* Supervised a team of 5 developers for product development and deployment.
* Received coverage on CNN and in several major North American newspapers.
* Sold over 50K gross revenue.

**EDUCATION**

**Graduate Work** in Computer Science, University of Toronto

*Research Area*: Urban navigation for visually impaired pedestrians.

Note: I left UofT in 2013 to pursue private-sector opportunities.

**M.S. Computer Science**, Wake Forest University, 2010

Thesis: *Machine Learning for Biostatisticians: A Hypothesis Driven Approach*

**M.A. Mathematics**, Wake Forest University, 2009

Thesis: *Some new results on Composition-Delay Equations with Asymptotically Periodic Solutions*

**B.A. Mathematics, Philosophy and Religion**, Appalachian State University, 2007.

*Summa Cum Laude, Honors Thesis*  Minor: Computer Science

**MAJOR HONORS AND AWARDS**

* 2009 Gordon A. Melson Outstanding Master's Student Award. Given to one student annually.
* Student commencement speaker for the College of Arts and Sciences, 2007, Appalachian State University. This award is given to the top overall graduate.

**LEADERSHIP ROLES**

* Technical lead on multiple projects at Microsoft on projects involving up to dozens of participants.
* Led team of 3 scientists at Wave Apps in close contact with executives
* Executive on Graduate Student Committee for the Department of Computer Science, University of Toronto, 2011.

**FOR FUN**

Ultramarathoner with multiple 50 to 100km races completed (and two longer ones attempted…). Recently, interested in backcountry skiing, ski mountaineering, and bike touring. See [this writeup](https://github.com/guyrt/guyrt.github.com/blob/master/Cougar.md) for of a project to find the most efficient way to run every trail at Cougar Mountain in one push.

Frequent traveler. Favorites were Mexico City and Argentine Patagonia.