

Maryann Foley
19330904
CSU33012
Biography of a Software Engineer

IBM Watson is best known for its Jeopardy run, the culmination of over 5 years of work, in which it defeated the show's reigning champions. However, since its inception as a single-use system similar to Deep Blue, a chess computer that was able to defeat grandmaster Garry Kasparov in 1997, the uses of Watson have expanded exponentially, with its natural language processing capabilities being put to use in many industries, from education to retail. The current CEO of IBM and former VP of Cloud and Cognitive Software, Arvind Krishna, has been a leader in pushing for the use of Watson technology for good, in particular in Open Source libraries focused on transparency in Artificial Intelligence.

Born in 1962, Krishna grew up in Andhra Pradesh in South-Eastern India. He went on to attend the Indian Institute of Technology Kanpur, where he received a degree in Electrical Engineering in 1985. He immediately went on to the University of Illinois, Urbana-Champaign where he was awarded a Master's degree and PhD in Computer and Electrical Engineering. It was while working towards his doctorate that he joined IBM as a researcher, a role that would eventually lead him to become the director of IBM's Research department.

In his technical roles, Krishna started in microprocessor creation but transitioned into software engineering and was integral in the creation of IBM's security, database management, and hybrid cloud software. In his time as an executive, he also spearheaded IBM's \$34 billion acquisition of RedHat, which was largely purchased for its hybrid cloud products.

One of his most innovative advancements at the company was overseeing the creation of and advocacy for Hyperledger, which is a platform that uses a blockchain structure that is applied to more than just cryptocurrencies. His persistence in his belief in the product's future was apparent, and earned him a spot on Wired's "25 Geniuses Who Are Creating the Future of Business."

When discussing his plans for the future of IBM under his leadership, he emphasized the significance of Hybrid Clouds and Artificial Intelligence, and the power that IBM has to revolutionize these industries. Watson as an Artificial Intelligence research project was seen as a failure to many- big promises of AI disrupting every industry have not come to fruition. However, the shift of Watson's engineering research power towards tools used to build AI applications for third-parties is not to be underestimated. In particular, efforts to make AI more ethical through recently released open source libraries has the potential to transform AI into a sustainable and dependable aspect of daily life.

The importance of transparency in Artificial Intelligence struck me after learning about the ousting of Timnit Gebru from Google Research Ethical AI team last year, after presenting a research paper that presented critical arguments about the public perception of language

models. Her previous studies on the racial biases of facial recognition software shocked me, and since learning about the measurable inequality of machine learning models, I have been interested in ways that companies are trying to mitigate these issues. Unfortunately, it seems that many large industries have not done enough, in particular the healthcare industry. Diagnosis of conditions is far more accurate for white men than any other group; algorithms to detect skin cancer work significantly worse for darker skin tones and chest X-rays are inaccurate when it comes to biological females, just to name a few examples. Engineering is still a very uniform field, with less than a quarter of engineers at top companies being female and less than ten percent of engineers being Latinx or Black. Because of this, bias can be overlooked when assembling training data.

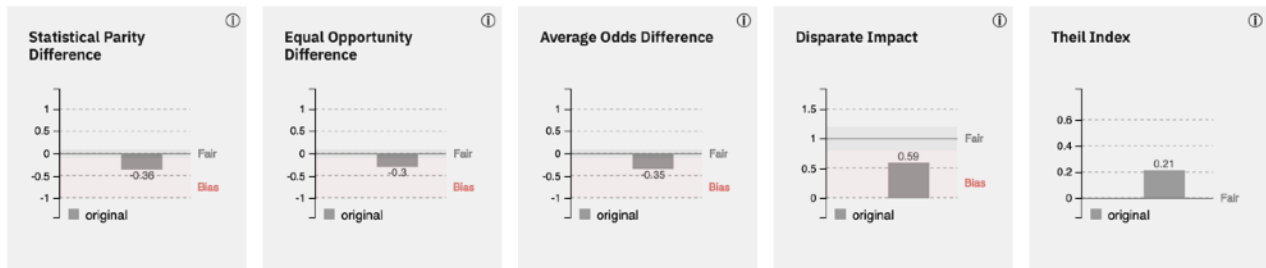
The potential for biases within machine learning models that may have detrimental or fatal consequences for individuals is a terrifying concept. IBM Research under Arvind Krishna's leadership has addressed this, creating an open source toolkit for ML models that checks the model for biases and gives potential solutions. By creating this tool, named AI Fairness 360, that is easily usable for developers, it makes bias detection much more accessible. An example of its assessments on demo data provided by the team is shown below. AI Fairness 360 is only one of a multitude of open source tools that IBM has created to make the process of AI creation more transparent and therefore more trustworthy.

Protected Attribute: Sex

Privileged Group: **Female**, Unprivileged Group: **Male**

Accuracy with no mitigation applied is 66%

With default thresholds, bias against unprivileged group detected in 4 out of 5 metrics



In addition to advancements in ML bias detection, Krishna announced last summer that IBM would stop producing and providing facial recognition software in response to public concerns about government surveillance and the biases that have been proven to exist within this software. In 2019, a federal study showed that “Asian and African American people were up to 100 times more likely to be misidentified than white men.” Growing up in the United States, I have been keenly aware of the inequalities that already exist in the American justice system, and systems that work to reinforce these injustices must be avoided. Krishna's devotion to creating a better environment through Machine Learning models is a breath of fresh air in a world where it seems like many companies are solely driven by profits. His history as a researcher has no doubt led to this philosophy.

Works Cited

- Harwell, Drew. "Federal Study Confirms Racial Bias of Many Facial-Recognition Systems, Casts Doubt on Their Expanding Use." *The Washington Post*, WP Company, 21 Dec. 2019, <https://www.washingtonpost.com/technology/2019/12/19/federal-study-confirms-racial-bias-many-facial-recognition-systems-casts-doubt-their-expanding-use/>.
- Kaushal, Amit. "Health Care AI Systems Are Biased." *Scientific American*, Scientific American, 17 Nov. 2020, <https://www.scientificamerican.com/article/health-care-ai-systems-are-biased/>.
- Panch, Trishan, et al. "Artificial Intelligence and Algorithmic Bias: Implications for Health Systems." *Journal of Global Health*, International Society of Global Health, Dec. 2019, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6875681/>.
- "Ai Fairness 360." *IBM Developer*, IBM, <https://developer.ibm.com/open/projects/ai-fairness-360/>.
- "Arvind Krishna." *IBM Newsroom*, IBM, <https://newsroom.ibm.com/Arvind-Krishna>.
- "A Precision Regulation Approach to Controlling Facial Recognition Technology Exports." *THINKPolicy Blog*, IBM, 11 Sept. 2020, <https://www.ibm.com/blogs/policy/facial-recognition-export-controls/>.