

Scraped Content

Title: Reaffirming IBM's commitment to the Rome Call for AI ethics - IBM Research

IBM, other leaders in industry, and representatives from many of the world's major religions, came together in Hiroshima, Japan to discuss developing AI ethically. There have been moments throughout history where the impacts of a new technology have been world-altering. Perhaps this is why the Vatican, along with leaders from most major religions across the world, chose to host a gathering to discuss the implications for future development of AI in Hiroshima, Japan. Last year, representatives from the Abrahamic religions came together at the Vatican to sign the Rome Call for AI Ethics, which IBM first signed with other industry and government leaders when it was launched by the Vatican in 2020. It's a document where the signatories committed to pursue an ethical approach to AI development and promote the human-centric and inclusive development of AI, rather than replacing humanity. At Hiroshima this year, the Rome Call was signed by representatives of many of the great Eastern religions, and past signees like IBM reaffirmed their commitment. It was organized by the Vatican's Renaissance Foundation, in collaboration with the Pontifical Academy of Life, Religions for Peace Japan, the UAE's Abu Dhabi Forum for Peace, and the Chief Rabbinate of Israel's Commission for Interfaith Relations. Representing IBM at the event was Senior Vice President and Director of IBM Research Darío Gil, who also attended last year's event in Italy. "Artificial intelligence is a horizontal technology with implications for every sector, country, and value system. We believe it must be developed through the collaboration of many, diverse institutions," Gil said in a speech during the event. "Because more than any other technology we've invented, artificial intelligence is just a reflection of its makers. And it needs to be governed like us, too." IBM was there when the very field of AI started, Gil reminded those in attendance. IBM co-hosted the seminal conference with Dartmouth in 1956 where the topic was first discussed by researchers in earnest, and it has been contributing to the field ever since. For Gil, like many others, the major inflection moment in AI came with the advent of foundation models. With AI models that could be tuned to carry out different tasks relatively easily, we saw the explosion of generative AI, and more

Scraped Content

recently, AI agents. But without ethical and open development of these tools ? with as wide a community to weigh in as possible ? large swathes of people will be left behind by the AI revolution. It?s why in 2018 IBM introduced its Principles for Trust and Transparency. It was also one of the first major companies to create an AI Ethics Board to govern the internal processes, tools, guidelines, education, and risk assessments for the company?s AI development and usage.

This approach has been infused into the bedrock of AI development across IBM. It?s at the heart of watsonx, IBM?s data and AI platform. Its governance and data tools were built with the same ideas and minds that work on the AI Ethics Board, and can be found in the recently open-sourced Granite family of AI models that IBM has built. The Granite LLM model was recently recognized by Stamford to be one of the most transparent major models available today. This commitment to an open approach to AI development was also a driving factor in why IBM co-founded the AI Alliance in December 2023. The group now consists of more than 100 companies, academic institutions, government agencies and research labs around the world, including Meta, Sony, NASA, Harvard, and the Cleveland Clinic. The Alliance aims to accelerate open-source innovation to improve trust in AI to ensure it benefits the entirety of our society. ?We?re able to work together because of our collective belief in the power of what a diverse set of institutions can achieve,? Gil said. Acknowledging that there was still much work that needed to be done to spur innovation forward in a way that leaves no one sidelined, Gil was hopeful about what a meeting like this meant for the future. ?So much of human history has been shaped by differences and disagreements over faith,? Gil said. ?But on this issue, the world?s major religions are uniting.? [Subscribe to our Future Forward newsletter and stay up to date on the latest research news](#) [Subscribe to our newsletter](#)
Date 15 Jul 2024 Authors Mike Murphy Topics AI Fairness, Accountability, Transparency Share