



Ethics in AI: A Challenging Task

Ricardo Baeza-Yates, Institute for Experiential AI @ Northeastern University

May 21, 2021, 9-10am PST

In the first part we cover four current specific challenges: (1) discrimination (e.g., facial recognition, justice, sharing economy, language models); (2) phrenology (e.g., biometric based predictions); (3) unfair digital commerce (e.g., exposure and popularity bias); and (4) stupid models (e.g., Signal, minimal adversarial AI). These examples do have a personal bias but set the context for the second part where we address four generic challenges: (1) too many principles (e.g., principles vs. techniques), (2) cultural differences (e.g., Christian vs. Muslim); (3) regulation (e.g., privacy, antitrust) and (4) our cognitive biases. We finish discussing what we can do to address these challenges in the near future.

Join via this zoom link: <https://washington.zoom.us/j/95485497113>

Ricardo Baeza-Yates is Research Professor at the Institute for Experiential AI of Northeastern University. He is also part-time professor at Universitat Pompeu Fabra in Barcelona and Universidad de Chile in Santiago. Before, he was VP of Research at Yahoo Labs, based in Barcelona, Spain, and later in Sunnyvale, California, from 2006 to 2016. He is co-author of the best-seller Modern Information Retrieval textbook published by Addison-Wesley in 1999 and 2011 (2nd ed), that won the ASIST 2012 Book of the Year award. From 2002 to 2004 he was elected to the Board of Governors of the IEEE Computer Society and between 2012 and 2016 was elected for the ACM Council. Since 2010 is a founding member of the Chilean Academy of Engineering. In 2009 he was named ACM Fellow and in 2011 IEEE Fellow, among other awards and distinctions. He obtained a Ph.D. in CS from the University of Waterloo, Canada, in 1989, and his areas of expertise are web search and data mining, information retrieval, bias on AI, data science and algorithms in general. <http://www.baeza.cl/>



**Responsible AI
Systems &
Experiences**

**UNIVERSITY of
WASHINGTON**