

March 4th, 2020

Bias on the Web

Ricardo Baeza-Yates

NTENT & Northeastern Univ. at SV

Abstract

The Web is the most powerful communication medium and the largest public data repository that humankind has created. Its content ranges from great reference sources such as Wikipedia to ugly fake news. Indeed, social (digital) media is just an amplifying mirror of ourselves. Hence, the main challenge of search engines and other websites that rely on web data is to assess the quality of such data. However, as all people has their own biases, web content as well as our web interactions are tainted with many biases. Data bias includes redundancy and spam, while interaction bias includes activity and presentation bias. In addition, sometimes algorithms add bias, particularly in the context of search and recommendation systems. As bias generates bias, we stress the importance of debiasing data as well as using the context and other techniques such as explore & exploit, to break the filter bubble. The main goal of this talk is to make people aware of the different biases that affect all of us on the Web. Awareness is the first step to be able to fight and reduce the vicious cycle of web bias. For more details see the article of same title in Communications of ACM, June 2018



Short bio



Ricardo Baeza-Yates is, since June 2016, CTO of NTENT, a semantic search technology company based in California, USA. He is also the Director of Data Science Programs at Northeastern University, Silicon Valley campus, since August 2017. 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 January 2006 to February 2016. Until 2004 he was the founding director of the Center for Web Research at the University of Chile. 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, data science and algorithms in general.