

Wisdom of Crowds or Wisdom of a Few?

Ricarddo Baeza-Yates
Yahoo! Labs
Sunnyvale, USA

ABSTRACT

In this keynote we give an introduction to wisdom of crowds in the Web, the long tail of web content, and the bias involved in the generation of user generated content (UGC). This bias creates the wisdom of ad hoc crowds or the wisdom of a few. Although it is well known that user activity in most settings follows a power law, that is, few people do a lot, while most do nothing, there are few studies that characterize well this activity. In a recent analysis of social network data we corroborated that a small percentage of the active users (passive users are the majority) represent at least the 50% of the UGC. As a sub-product, we also found a lower bound for the digital desert, the content in the Web that nobody reads. These results implies that most of the wisdom comes from a few users, which is not that surprising, as the Web is a reflection of our own society, where economical or political power also is in the hands of minorities.

Keywords

Social Personalization and Search, Wisdom of the Crowds

1. BIO

Ricardo Baeza-Yates is VP of Research for Yahoo Labs leading teams in United States, Europe and Latin America since 2006 and based in Sunnyvale, California, since August 2014. During this time he has lead the labs in Barcelona and Santiago de Chile. Between 2008 and 2012 he also oversaw the Haifa lab. He is also part time Professor at the Dept. of Information and Communication Technologies of the Universitat Pompeu Fabra, in Barcelona, Spain. During 2005 he was an ICREA research professor at the same university. Until 2004 he was Professor and before founder and Director of the Center for Web Research at the Dept. of Computing Science of the University of Chile (in leave of absence until today). He obtained a Ph.D. in CS from the University of Waterloo, Canada, in 1989. Before he obtained two masters (M.Sc. CS & M.Eng. EE) and the electronics engineer degree from the University of Chile in Santiago. He is co-author of the best-seller Modern Information Retrieval textbook, published in 1999 by Addison-Wesley with a second enlarged edition in 2011, that won the ASIST 2012 Book of the Year award. He is also co-author of the 2nd edition of the Handbook of Algorithms and Data Structures, Addison-Wesley, 1991; and co-editor of Information Retrieval: Algorithms and Data Structures, Prentice-Hall, 1992, among more than 500 other publications.

From 2002 to 2004 he was elected to the board of governors of the IEEE Computer Society and in 2012 he was elected for the ACM Council. He has received the Organization of American States award for young researchers in exact sciences (1993), the Graham Medal for innovation in computing given by the University of Waterloo to distinguished ex-alumni (2007), the CLEI Latin American distinction for contributions to CS in the region (2009), and the National Award of the Chilean Association of Engineers (2010), among other distinctions. In 2003 he was the first computer scientist to be elected to the Chilean Academy of Sciences and since 2010 is a founding member of the Chilean Academy of Engineering. In 2009 he was named ACM Fellow and in 2011 IEEE Fellow.