

Big Data: New Paradigm or “Sound and Fury, Signifying Nothing”?

Panel Organizer: Andrei Broder*
Google

Michael Franklin
University of California, Berkeley

Eric Xing
Carnegie Mellon University

Lada Adamic
Facebook

Maarten de Rijke
University of Amsterdam

Kai Yu
Baidu

PANEL OVERVIEW

The Gartner’s 2014 Hype Cycle released last August moves Big Data technology from the Peak of Inflated Expectations to the beginning of the Trough of Disillusionment when interest starts to wane as reality does not live up to previous promises. As the hype is starting to dissipate it is worth asking what Big Data (however defined) means from a scientific perspective: Did the emergence of gigantic corpora exposed the limits of classical information retrieval and data mining and led to new concepts and challenges, the way say, the study of electromagnetism showed the limits of Newtonian mechanics and led to Relativity Theory, or is it all just “sound and fury, signifying nothing”, simply a matter of scaling up well understood technologies? To answer this question, we have assembled a distinguished panel of eminent scientists, from both Industry and Academia: Lada Adamic (Facebook), Michael Franklin (University of California at Berkeley), Maarten de Rijke (University of Amsterdam), Eric Xing (Carnegie Mellon University), and Kai Yu (Baidu) will share their point of view and take questions from the moderator and the audience.

Categories and Subject Descriptors

A.0 General Literature, GENERAL: Conference proceedings; E.0 Data, GENERAL

Keyword: big data

*Moderator’s email: broder@google.com

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage, and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s). Copyright is held by the author/owner(s).

WSDM’15, February 2–6, 2015, Shanghai, China.

ACM 978-1-4503-3317-7/15/02.

<http://dx.doi.org/10.1145/2684822.2697027>



Andrei Broder is a Google Distinguished Scientist. From 2005 to 2012 he was a Fellow and VP for Computational Advertising at Yahoo!. Previous positions include Distinguished Engineer at IBM and VP for Research and Chief Scientist at AltaVista. Broder has authored more than a hundred papers and was awarded more than forty US patents. His current research interests are centered on personalization, computational advertising, web search, context-driven information supply, and randomized algorithms. He is a member of the US National Academy of Engineering and a Fellow of ACM and of IEEE. Other honors include the ACM Paris Kanellakis Theory and Practice Award and a PhD Honoris Causa from the Technion.



Lada Adamic leads the Product Science group within Facebook’s Data Science Team. She is also an adjunct associate professor at the University of Michigan’s School of Information and Center for the Study of Complex Systems. Her research interests center on information dynamics in networks: how information diffuses, how it can be found, and how it influences the evolution of a network’s structure. Her projects have included identifying expertise in online question and answer forums, studying the dynamics of viral marketing, and characterizing the structural and communication patterns in online social media. She has received an NSF CAREER award, a University of Michigan Henry Russell award, the 2012 Lagrange Prize in Complex Systems.



Michael Franklin is the Thomas M. Siebel Professor of Computer Science and Chair of the Computer Science Division at the University of California, Berkeley. He has over 30 years of experience in the database, data analytics, and data management fields as a researcher, lab director, faculty member, entrepreneur, and software developer. Prof. Franklin is also the Director of the Algorithms, Machines, and People

Laboratory (AMPLab) at UC Berkeley. The AMPLab currently works with 23 industrial sponsors including founding sponsors Amazon Web Services, Google, and SAP, and received a National Science Foundation CISE “Expeditions in Computing” award, announced as part of the White House Big Data research initiative in March 2012. AMPLab is well-known for creating a number of popular systems in the Open Source Big Data ecosystem including Spark, Mesos, Shark, GraphX and MLlib, all parts of the Berkeley Data Analytics Stack (BDAS). Prof. Franklin is also a co-PI and Executive Committee member for the Berkeley Institute for Data Science, part of a multi-campus initiative to advance Data Science Environments. He is an ACM Fellow, a two-time winner of the ACM SIGMOD “Test of Time” award, and recipient of the outstanding Advisor Award from the Computer Science Graduate Student Association at Berkeley.



Maarten de Rijke is full professor of Information Processing and Internet in the Informatics Institute at the University of Amsterdam. He holds MSc degrees in Philosophy and Mathematics (both cum laude), and a PhD in Theoretical Computer Science. He worked as a postdoc at CWI, before becoming a Warwick Research Fellow

at the University of Warwick, UK. He joined the University of Amsterdam in 1998, and was appointed full professor in 2004. De Rijke leads the Information and Language Processing Systems group, one of the world’s leading academic research groups in information retrieval. During the most recent computer science research assessment exercise, the group achieved maximal scores on all dimensions. His research focus is on intelligent information access, with projects on self-learning search engines, semantic search, and social media analytics.



Dr. Eric Xing is a professor in the School of Computer Science at Carnegie Mellon University. His principal research interests lie in the development of machine learning and statistical methodology, and large-scale computational system and architecture, for solving problems involving

automated learning, reasoning, and decision-making in high-dimensional, multimodal, and dynamic possible worlds in complex systems. Professor Xing received a Ph.D. in Molecular Biology from Rutgers University, and another Ph.D. in Computer Science from UC Berkeley. His current

work involves, 1) foundations of statistical learning, including theory and algorithms for estimating time/space varying-coefficient models, sparse structured input/output models, and nonparametric Bayesian models; 2) framework for parallel machine learning on big data with big model in distributed systems or in the cloud; 3) computational and statistical analysis of gene regulation, genetic variation, and disease associations; and 4) application of statistical learning in social networks, data mining, and vision. Professor Xing has published over 200 peer-reviewed papers, and is an associate editor of the Journal of the American Statistical Association, Annals of Applied Statistics, the IEEE Transactions of Pattern Analysis and Machine Intelligence, the PLoS Journal of Computational Biology, and an Action Editor of the Machine Learning journal, and the Journal of Machine Learning Research. He is a member of the DARPA Information Science and Technology (ISAT) Advisory Group, a recipient of the NSF Career Award, the Alfred P. Sloan Research Fellowship, the United States Air Force Young Investigator Award, and the IBM Open Collaborative Research Faculty Award.



Dr. Kai Yu is Head of Institute of Deep Learning (IDL) at Baidu , and also Senior Engineering Director , in charge of Baidu Image Search. Dr. Yu has published a number of papers with over 7000 citations. He received

many awards including the Best Paper Runner-up Award of ICML-2013, the first prize of PASCAL VOC 2009, and the first prize of ImageNet competition 2010. In 2013 and 2014, his team has won three Baidu Highest Achievement Awards, for having made significant contributions to speech recognition, computer vision, online ads, and web search ranking. Since 2014, he has led a series of high-impact projects, including Baidu Brain, BaiduEye, Autonomous Driving, DuBike etc. Dr. Yu is elected to China “Top 1000 Talents”, a prestigious group of technology leaders of the country. He is an Adjunct Professor of Nanjing University, Beijing University of Posts and Telecommunications, and Institute of Computing Technology, Chinese Academy of Sciences. Before joining Baidu, he worked at Microsoft, Siemens, NEC and was an Adjunct Faculty at Department of Computer Sciences at Stanford University. He received his B.Sc and M.Sc degree from Nanjing University and his Ph.D. degree from University of Munich.