EDC OCEANS of DATA INSTITUTE

Big Data Expert Panel: Biographies

Kirk Borne

Professor of Astrophysics and Computational Science, George Mason University

Kirk Borne is a multidisciplinary data scientist and an astrophysicist, with specific expertise in scientific data mining. He is currently working on research, design, and development for the proposed Large Synoptic Survey Telescope (LSST). He has advised several federal agencies on data mining and Big Data applications, including the Executive Office of the President, the Library of Congress, National Weather Service, FDA Office of Drug Safety, and the NITRD Big Data Senior Steering Group. He is a founding member of the new American Astronomical Society's Astroinformatics and Astrostatistics Working Group and the new International Astrostatistics and Astroinformatics professional society. Prior to moving to his position at George Mason University, he spent nearly 20 years in positions supporting NASA projects, including NASA's Data Archive Project Scientist for the Hubble Space Telescope, and as Project Manager in NASA's Space Science Data Operations Office. He received his BS degree in Physics from LSU and his PhD in Astronomy from Caltech.

Randy Bucciarelli

Programmer/Analyst, Coastal Data Information Program at Scripps Institution of Oceanography

Randy's experience involves analyzing and distributing massive datasets detailing spatial and temporal changes in the dynamic coastal environment. As part of the Southern California Coastal Observing System, which is part of the U.S. Integrated Ocean Observing System, he works regularly with local, state, and federal agencies, educators, scientists, and the general public to disseminate and make sense of 'Big Data'. The program he works for collects and processes large time-series datasets spanning decades such as high resolution wave measurements, high frequency radar, and Light Detection and Ranging (LIDAR). He specializes in automation, open source formats, Geographic Information Systems, and metadata generation. He holds a BS in Earth Systems Science and is currently pursuing a MS in Computational Science at San Diego State University.

Tim Chadwick

Principal Engineer, Dynamic Network Services, Inc.

Tim Chadwick is the Principal Engineer of Infrastructure at Dynamic Network Services, Inc. Tim drives Database as a Service at Dyn, which aligns the best practices for different data problems, standardizes provisioning and management and makes capacity planning easier and more predictable. Technologies used at Dyn include MySQL, PostgreSQL, Cassandra and Hadoop. Tim has been a primary contributor to data-oriented businesses and applications at scale for over 15 years. Industries of his focus include manufacturing, higher-ed, media, medical and Internet performance. He has implemented the gamut of database technologies, mostly focused in the open source space. Tim is a member of the Association for Computing Machinery (ACM), the Computing Technology Advisory Board at University of New Hampshire, Manchester, and teaches special topics classes such as Intro to Big Data there as well.

Ben Davison
Quantitative User Experience Researcher, Google Search
Ben is a data scientist in Google Search. His goal is to extract meaning from user activities in Google Search. The largest-n work involves A/B studies with billions of users on everything other than the main list of links or ads. This informs product teams and leadership on what users really want to see on the page. Ben started research with psychometric studies for accessibility in academia, where getting enough participants was a key problem. He has since discovered that big data has the reverse problem: spurious findings and results that are difficult to put into action. He found a distance in meaning not present in small-n studies. His current work in logs analysis, surveys and user needs is triangulating on meaningful results and improving methods for product-driven research. Ben holds a PhD in Human-Centered Computing from Georgia Tech.

Lucy Drotning
Associate Provost of Planning and Institutional Research, Columbia University
Lucy Drotning is associate provost for the Office of Planning and Institutional Research (OPIR) at Columbia University, which provides the provost and other senior administrators with planning information that is used in managing the university and in making decisions about its policies and goals. Planning information includes historical and current information about Columbia, comparisons across peer institutions, and data from surveys of faculty, staff, students and alumni. In addition to gathering and organizing relevant facts and figures, OPIR uses these data to carry out research and analyses regarding issues of importance to the university. Prior to working at Columbia, Drotning worked in the Office of Budget and Planning at the University of Michigan. She has a PhD in political science from the University of Rochester, and begins a certification program in data sciences at Columbia's Institute for Data Sciences and Engineering in Fall 2014.

Ryan Kapaun
Crime Analyst, Eden Prairie Police Department
Ryan Kapaun has been the crime analyst for the Eden Prairie Police Department (MN) since 2007. Prior to that, he worked for five years as an intelligence analyst with the State of Minnesota, assigned to the Division of Homeland Security and Emergency Management, Bureau of Criminal Apprehension and the Minnesota Joint Analysis Center. Kapaun is a past president of the Minnesota Association of Criminal Intelligence Analysts (MACIA) and in 2010 he was named the Minnesota Crime Analyst of the Year for his role in collecting and analyzing intelligence data for the Eden Prairie Police Department.

Juan Miguel Lavista Ferres
Principal Data Scientist at Bing/ Microsoft
Juan Miguel Lavista Ferres is currently the principal data scientist for Microsoft Data Science team (DnA), where he works with a team of data scientists searching for insights in petabytes of data. Juan joined Microsoft in 2009 to work for the Microsoft Experimentation Platform (EXP) where he designed and ran randomized control experiments across different Microsoft properties. In Microsoft, Juan also worked as part of the Bing Data Mining team. Before joining Microsoft, Juan was the CTO and co-founder of

alerts.com. Juan has 2 computer science degrees from the Catholic University in Uruguay, and a graduate degree in Data Mining from Johns Hopkins University.

Shannon McWeeney
Head of the Division of Bioinformatics and Computatational Biology, Oregon Health & Science University
Shannon McWeeney is the head of the Division of Bioinformatics and Computatational Biology in the Department of Medical Informatics and Clinical Epidemiology, and a faculty member in the Division of Biostatistics, Department of Public Health and Preventive Medicine. She is the director of the Oregon Health & Science University (OHSU) Knight Cancer Institute Bioinformatics Shared Resource and the OCTRI Translational Bioinformatics Program. The focus of her research is on the development and application of statistical and computational methodologies for functional genomics data such as gene expression (microarray and RNA-seq), chIP-Seq and proteomics. Applications of this work have included such complex traits as diabetes, cancer and alcoholism. McWeeney has a PhD in Statistical Genetics and a MSE in Computational Biology.

Jay Parker
Earth Scientist, Jet Propulsion Laboratory, California Institute of Technology
Jay Parker has developed methods for analysis and modeling of Earth remote sensing records for thirty years. Part of this work has been to craft and apply fast and accurate numerical models to geophysical observations, in a number of subdisciplines. These include simulation of vortex formation in the ionosphere, algorithms for radar wave scattering, measuring concentrations of greenhouse gasses using high-fidelity simulation of emission spectra, and models of ocean wind behavior seen through reflection of satellite GPS transmissions. Parker's current activities include detection of tiny fault motions using an airborne radar, modeling the permanent changes to landscapes from earthquakes, and testing the limits of the global reference frame when stressed by the very largest earthquakes. He is currently Co-Investigator for the NASA GeoGateway project, which aims to provide access to many kinds of geodetic imaging observations through a web-enabled map-based search and analysis gateway, and Principal Investigator for the NASA Application of Uninhabited Aerial Vehicle Synthetic Aperture Radar Imaging program, which seeks to use radar images to update the estimates of earthquake hazard in California. Parker has a PhD and MS in Electrical Engineering from the University of Illinois, and is also a graduate of Caltech (BSEE).

Steve Ross
Corporate Editor, Broadband Communities Magazine
Steven Ross is a long-time technology writer and former Columbia University professor and is also the Corporate Editor of Broadband Communities Magazine (www. bbcmag.com), the leading source of information on digital and broadband technologies for ultra-connected communities. Ross is known for his expertise in distance learning and computer-assisted reporting, also known as "analytic journalism". He is a Fellow of the American Institute of Chemists, has written a statistics book, and has taught statistics in academic and industrial settings. While at Columbia, he was on the full-time faculty of the Graduate School of Journalism, which he helped computerize in the 1980s and where he established

some of the world's first courses in analytic journalism. He was teaching spreadsheet modeling and online database access as early as 1979 at Columbia, in a pioneering national reporting course. Ross is an expert on broadband technologies and has domestic and international expertise in land-use planning, product safety, business, and education. Since 1993, Ross has been conducting annual surveys on media and cyberspace, with special emphasis on how emerging technologies shape Internet use. He has consulted on these issues for the Pulitzer Prize board at Columbia and for some of the world's largest media companies and foundations. Ross holds a BS in Physics from Rensselaer Polytechnic Institute and an MS in Journalism from Columbia.

Kartik Shah

Principal Consultant, Strategix Solutions

Kartik has over 15 years of experience focused on Enterprise Software, Telecom and Wireless Sectors. As a Principal Consultant at Strategix Solutions, Kartik has helped companies with Big Data and Enterprise Data Integration challenges, with a particular focus in Telecommunications and Utilities. At a key Telecom client, Kartik has been involved with handling and harvesting insight from their customer usage data. Before Strategix, Kartik has played several technical and architect roles for MobileQ and Embarcadero. Kartik has received a BA Sc in Computer Engineering from the University of Waterloo and an MBA in Strategy & Marketing from the University of Toronto, Rotman School of Management.