

Kirk D. Borne

<http://www.linkedin.com/in/kirkdborne> , <http://classweb.gmu.edu/kborne>

Contact: kborne@gmu.edu , (mobile) 443-474-4204

PhD, Data Scientist, Top Big Data Influencer, Data Mining Specialist, Professor of Astrophysics & Computational Science, George Mason University, Fairfax, VA

Background

Data Scientist, Data Mining Specialist, Big Data Researcher, Consultant, Public Speaker (given nearly 200 invited talks in the US and worldwide on Data Science, Big Data, and Science topics).

Transdisciplinary Data Scientist - data science education, research, applications, and consulting across a variety of fields - transcending traditional discipline boundaries.

Specialties: Data Science Education & Research, Data Mining, Data Analytics, Big Data, X-Informatics (Discovery Informatics, Science Informatics, and more), Scientific databases, Scientific data mining, Computational-X (Computational Science, and more), Astroinformatics¹ (Data Science for Astronomy), Observational Astronomy (ground-based and space-based), Computational Astrophysics, Higher Education, Curriculum Development, Science Education and Public Outreach, Project Management, Proposal-writing, Business development.

(11/14/2013): #1 Big Data influencer on Twitter (out of 150,000+):

<http://www.onalytica.com/blog/posts/onalytica-big-data-influencers-q4-13>

(7/31/2013): #1 Big Data influencer on Twitter (out of 125,000+):

<http://www.onalytica.com/blog/post/2013/07/31/Big-Data-and-Influencers.aspx>

(5/21/2013): #2 on List of Big Data Top Influencers on Twitter:

http://www.bigdatarepublic.com/author.asp?section_id=2642&doc_id=263560&

(4/16/2013) #11 in Big Data Top 100:

http://www.bigdatarepublic.com/author.asp?section_id=2642&doc_id=262105&

(2013) Top 10 - Data Science A-List: <https://alist.traackr.com/datascience>

Winner of the 2013 GMU College of Science Dean's Annual Impact Award for Faculty Excellence.

¹ co-creator of the field of Astroinformatics: see <http://www.springerlink.com/content/007642412771162j/> and <http://arxiv.org/abs/0909.3892>

Professional Work Experience

Professor of Astrophysics and Computational Science

George Mason University

2003 – Present (10 years)

Summary: Conducts Research, Teaches Courses, and Advises Students (undergraduate and graduate) in these areas: Data Science, Data Mining, Analytics, Big Data, Informatics, Statistics, Scientific Databases, Astroinformatics, Computational Astrophysics, Astronomy, and Computational Science (Modeling & Simulation).

Kirk Borne was the co-creator of the GMU Data Science BS degree program (2006)². He developed the curriculum and several courses for the Data Science undergraduate program and has served as academic advisor for the program (2010-present)³. He has also created and taught online distance education courses since 2003: Data Mining (graduate), Scientific Databases (graduate), and Data Ethics (undergraduate). He is dissertation advisor, dissertation committee chairman, and/or dissertation committee member for over 40 Ph.D. students in the GMU Computational Science and Informatics program.

He has carried out scientific research for over 30 years⁴. His research was initially in the fields of astronomy and computational astrophysics, but since 2000 he has developed & applied data mining & data science algorithms and methodologies on large data collections in the sciences and in other disciplines (which is now called Big Data Analytics), including large astronomy databases, earth science remote sensing data collections, medical data, financial data, and more⁵. He was perhaps the first person to develop the concept of “unknown unknowns” in data mining research and applications (in science, national security, and elsewhere)⁶. He is investigating both new and existing algorithms for the discovery of patterns, correlations, associations, and outliers (surprises) within highly dimensional (complex) datasets. These techniques are being used in several research areas, including: scaling relations and parameterizations in data-driven applications, characterization and classification of time-varying events, and novelty detection within massive datasets.

He is investigating Big Data Analytics connections with other data-intensive disciplines, to explore opportunities for advancing research and education in a broad context, through the investigation and application of data science methods (informatics and statistics) in a manner that transcends discipline boundaries. With his extensive experience in Big Data and Data Science, he has consulted for numerous businesses (large and small) and has advised several federal agencies on data mining and Big Data applications, including the Executive Office of the

² Borne, K. et al., "The New Undergraduate Program in Computational and Data Sciences at GMU," Lecture Notes in Computer Science, 5545, 74-83 (2009).

³ <http://classweb.gmu.edu/kborne/courses.html>

⁴ http://scholar.google.com/citations?hl=en&user=SclrpfYAAAAJ&view_op=list_works&pagesize=100

⁵ https://www.researchgate.net/profile/Kirk_Borne/

⁶ Borne, K. "Science Requirements for Data Mining," <http://arxiv.org/abs/astro-ph/0008307> (2000)

President, the Library of Congress, National Agricultural Library, NOAA National Weather Service, NASA, NSF, NRC, FDA Office of Drug Safety, and the NITRD Big Data Senior Steering Group. He has given invited presentations to those agencies, as well as to the following: Homeland Security Transition Planning Office, NIH, NIST, CMS (Centers for Medicare and Medicaid Services), Institute for Defense Analyses (IDA), and the Logistics Management Institute (LMI).

He has given nearly 200 invited talks on the subjects of Data Science, Data Mining, Big Data, Analytics, and Astronomy, including the Conference Keynote Talk several times (including at the 2011 Medicare and Medicaid Statistics and Data Analysis Conference and the 2013 Army Applied Statistics Conference⁷).

He is one of the originators of the field of Astroinformatics, which is data science for astronomy research and education. As part of this research, he serves on several international committees (see below). In the education area, he is on the LSST Outreach Advisory Board (below) and is working with the Galaxy Zoo and Galaxy Merger Zoo citizen science projects.

He is working with education researchers in two areas: (1) to develop new pedagogies for learning that engage students' instinctive interests in information-rich environments (including social networks) through active experiences using Big Data in the classroom⁸; and (2) application of Big Data to track student performance and to design interventions that aid student learning, increase retention, and predict student performance challenges⁹.

- *Phi Kappa Phi National Honor Society*

Chair of Phi Kappa Phi Scholarship Committee, GMU chapter (2010-present)
Charter Faculty of GMU chapter of Phi Kappa Phi (2010)

Executive Committee Member

Astroinformatics and Astrostatistics International Organizations
2009 – Present (4 years)

Kirk Borne is contributing to the international and national efforts to promote astrostatistics and astroinformatics as robust research disciplines: organizing international research conferences on astrostatistics and astroinformatics, and contributed to the creation of three new international and national organizations, plus serves on other national committees:

- (1) Vice President, International Astrostatistics and Astroinformatics (IAA) professional society.
- (2) Senior Steering Committee member, AAS (American Astronomical Society) Working Group on Astroinformatics and Astrostatistics.

⁷ <http://classweb.gmu.edu/kborne/travel-schedule.htm>

⁸ Borne, K. et al., "The Revolution in Astronomy Education: Data Science for the Masses," <http://arxiv.org/abs/0909.3895> (2009).

⁹ K.Borne was one of 29 participants selected nationwide to participate in NSF Ideas Lab on "Data-Intensive Research to Improve Teaching and Learning" (Oct.2013) <http://www.nsf.gov/pubs/2013/nsf13565/nsf13565.htm>

(3) Organizing Committee member, IAU (International Astronomical Union) Working Group in Astrostatistics and Astroinformatics.

(4) Secretary, SIAM SIAG/DMA (SIAM Activity Group on Data Mining & Analytics).

Chairman, LSST Astroinformatics and Astrostatistics Research Collaboration Team

2009 – Present (4 years)

Kirk Borne is chairman of the informatics and statistics research team for the LSST project. The Large Synoptic Survey Telescope (LSST) is planned to generate a 20-petabyte science database and a 100-petabyte astronomical image archive -- the analysis of such an enormous data collection requires novel research methods and algorithms. Our collaboration team consists of over 50 scientists, including astronomers, computer scientists, data mining experts, and statisticians --- we are working together on solutions to the LSST petascale data challenges.

Astronomer and Data Scientist, Zooniverse Citizen Science Projects

2009 – Present (4 years)

Kirk Borne is developing and applying machine learning, data mining, and exploratory data analysis methods on the research data products that are being generated by the enormously popular <http://Zooniverse.org> citizen science volunteer community.

Board Member, Outreach Advisory Board

[LSST Large Synoptic Survey Telescope](#)

2005 – Present (8 years)

Kirk Borne is advising the LSST project team regarding the education and public outreach programs, activities, goals, data products, data system design, interfaces, and learning objectives that will support the mission of the LSST and that will be based on the massive petascale scientific database and astronomical imaging data collection that LSST will produce.

Board Member, UMUC Data Analytics Board of Advisors

[LSST Large Synoptic Survey Telescope](#)

2013 – Present

Kirk Borne serves on the Advisory Board for the new Master of Science in Data Analytics program at the University of Maryland University College¹⁰.

¹⁰ <http://www.umuc.edu/analytics/partners/index.cfm>

Vice President / Astronomer / Data Scientist

[ARIES Scientific](#)

2005 – Present (8 years) <http://www.aries-scientific.org/>

ARIES Scientific Inc. is a non-profit 501(c)(3) corporation, dedicated to the advancement of Science Research and STEM Education.

Kirk Borne has advised and consulted with numerous federal agencies and private companies in data science areas, including data mining and large-scale data management.

Data Analytics Scientist

[XLDB \(eXtremely Large Databases\) founding working group](#)

2007 – 2008 (1 year)

Kirk Borne contributed science use cases and requirements, data mining expertise, scientific databases experience, and LSST science database research inputs to the initial formation of the XLDB.org and SciDB.org projects.

Program Manager - Project Manager for the NASA Space Science Data Operations Office Contract

[QSS Group, Inc. and Perot Systems Government Services](#)

December 2005 – January 2008 (2 years 2 months) at NASA/GSFC

QSS Group, Inc. was acquired by PSGS (Perot Systems Government Services) in early 2007. SSDOO was NASA's Space Science Data Operations Office at the Goddard Space Flight Center.

As Program Manager, Kirk Borne managed a budget of \$30M on a 5-year scientific data services contract, with 70 staff members. Activities included science data management, database programming and management, data archiving, metadata development, web interface design and development, search and visualization tools development, documentation, research support and publication, helpdesk support, photo lab management, NASA mission support & data processing, systems management and development, and more. His management team received exceptionally high marks for project management, subcontract management, EEO compliance, staff development & training, Performance-Based Contracting metrics performance, and more.

Corporate activities included business development, corporate proposals, training, and organizing the Science Data Centers Symposium (every 2 years).

Adjunct Associate Professor - Data Mining

[University of Maryland University College](#)

2002 – 2007 (5 years)

Kirk Borne taught graduate course in Data Mining each semester, as part of the Database Technologies Master's Degree Program. He taught several semesters online via distance education. Completed training via a graduate education course in Distance Education.

Senior Scientist

[US National Virtual Observatory](#)

2001 – 2007 (6 years) Virtually Everywhere

Kirk Borne provided contributions in these areas: science use cases, scientific research, data science, distributed data mining, astronomical data mining, scientific databases, astronomical catalogs, semantic linked data, event characterization and classification, large sky surveys, science education and public outreach.

Co-Director, Institute for Science and Technology at Raytheon (I-ST@R)

Raytheon Technical Services Company

2002 – 2003 (1 year) Lanham, MD

Kirk Borne founded and then became Co-Director For Space Sciences of the new Institute for Science and Technology at Raytheon (I-ST@R). He was funded to carry out research into scientific data mining: (1) data mining of large astronomical databases; (2) data mining of satellite imagery to detect wildfires, as inputs to climate change models; (3) development of semi-supervised Machine Learning algorithms; (4) distributed data mining across large databases -- algorithms for distributed PCA, distributed outlier detection, and distributed text annotation and text mining.

Astrophysics Department Manager for Hughes STX, Astronomy Data Center and Astrophysics Data Facility (Founder & co-Director of the Institute for Science & Technology at Raytheon IST@R)

[NASA Goddard Space Flight Center](#)

1995 – 2003 (8 years) at NASA

Hughes STX later became Raytheon STX, which then became Raytheon ITSS (Information Technology & Scientific Services Company).

As technical line manager, Kirk Borne directly supervised and managed a staff of ~35 scientists, programmers, database staff, web designers & developers, user interface designers & developers, archive specialists, data curators, documentation specialists, and helpdesk support staff.

Corporate activities included business development, corporate proposals, training, and organizing the Science Data Centers Symposium (every 2 years).

Details: Managed an annual budget of up to \$3 million while providing outstanding customer service to the NASA task monitors in many areas: the ASCA science data processing center, the

U.S. ROSAT Science Data Center, the XTE science data processing center, the Astronomical Data Center, the multi-mission archive, archive interface development, object-oriented databases, constellation spacecraft operations, interoperable data systems, astrophysics data acquisition and user support (for optical, UV, infrared, and X-ray astronomy missions), and more. His staff also produced the extremely popular and widely distributed Multi-Wavelength Milky Way poster, which appears in nearly 100,000 classrooms and offices around the country. In 2002, NASA closed the ADC and ADF, at which time (2002-2003) He founded and became Co-Director of the new Institute for Science and Technology at Raytheon (I-ST@R), while continuing line management duties over my data center staff.

Training: Six Sigma. CMM (Capability Maturity Model). 7 Habits of Highly Successful People. PBC (Performance-Based Contracting). Business Proposal Development.

Scientist - Astronomer - Hubble Space Telescope Data Archive Project Scientist

Space Telescope Science Institute

1985 – 1995 (10 years) Baltimore, MD

From 1985 through 1995, Kirk Borne worked at the Hubble Space Telescope Science Institute (STScI) in Baltimore, Maryland. In 1991, he was co-recipient for the NASA Goddard Space Flight Center Certificate of Recognition for the successful deployment and initiation of science operations of the Hubble Space Telescope.

From 1992 through 1995, he was the Hubble Space Telescope DADS (Data Archive and Distribution System) Project Scientist and the StarView (HST User Interface) Project Scientist. He managed a group of 3-6 archive operators during the early deployment years of the HST data archive system. He provided guidance and advice to NASA management in the design and development of the system. In 1994, he and his group received the STScI Group Achievement Award for design, development, and deployment of the ST-DADS (Archive) Project.

From 1985 through 1992, he supported the proposal selection system at STScI. He supported external and internal scientists, supported the proposal TAC (telescope allocation committee) in their proposal reviews, and he provided detailed reports and statistics to NASA HQ related to the science program of the HST. He also developed single-handedly the entire Phase I HST proposal submission system, including databases, user interfaces, submission software (which every HST astronomer in the world used from 1990 through 1992), TAC reports, and more. He supervised a group of 6 user support staff in the planning of the HST proposal review meetings, and supervised as many as 6 additional personnel each year during the few weeks around the time of the proposal review meetings.

Carnegie Fellow - Postdoctoral Fellow - Astronomy and Astrophysics Research

[Carnegie Institution of Washington Department of Terrestrial Magnetism](#)

1983 – 1985 (2 years)

Kirk Borne carried out research in observational astronomy and theoretical astrophysics, involving imaging and spectroscopic data analysis, modeling, numerical simulation, evolutionary algorithms, and scientific programming.

Postdoctoral Teaching Fellow - Astronomy

[University of Michigan](#)

1981 – 1983 (2 years)

Kirk Borne carried out research in observational astronomy and theoretical astrophysics, involving imaging and spectroscopic data analysis, modeling, numerical simulation, evolutionary algorithms, and scientific programming. He taught astronomy courses to undergraduates.

Education

California Institute of Technology

1975 – 1981

M.S., Astronomy (1980)

Ph.D., Astronomy (1983)

National Science Foundation (NSF) Graduate Fellow

Louisiana State University

1972 – 1975

B.S., Physics (1975)

Summa Cum Laude; 4.0 GPA

University Medal for Highest Academic Honors

Honors & Awards

Astrophysics Department Manager

Raytheon Information Technology & Scientific Services

1999 - Sabbatical Research Award

Adjunct Associate Professor - Data Mining

University of Maryland University College

2007 - Phi Kappa Phi Faculty Mentor Award (UMUC)

Additional Honors & Awards

1972 American Mathematical Society National Achievement Award
1972 Allied Chemical Corporation National Achievement Award
1975 Keen-Morris Prize: Outstanding Student of Physics (LSU)
1975 Sigma Xi Outstanding Student of Science (LSU)
1975 LSU University Medal for Highest Academic Honors
1975-1978 National Science Foundation (NSF) Graduate Fellowship
1990 Hubble Space Telescope Individual Achievement Award
1993 Hubble Space Telescope Data Archive Group Achievement Award
2013 (July) #1 Big Data influencer on Twitter:
<http://www.onalytica.com/blog/post/2013/07/31/Big-Data-and-Influencers.aspx>
2013 (November) #1 Big Data influencer on Twitter:
<http://www.onalytica.com/blog/posts/onalytica-big-data-influencers-q4-13>
2013 GMU College of Science Dean's Annual Impact Award for Faculty Excellence

Business References (academic references upon request)

- (1) Richard (Dick) Bishop: dbishop98@gmail.com
Former President and COO, QSS Group Inc.
Executive Advisor, Board of Advisors, Earth Resources Technology Inc.
(mobile) 703-216-4511
- (2) Paul Thompson, Cornell Technical Services, COO: pthompson@cts-llc.com
Former Vice President, Hughes STX (later, Raytheon STX)
(mobile) 703-216-7356, (office) 703-648-3834
- (3) Ashok Kaveeshwar, Science & Technology Corp., Exec. VP: ashokkaveeshwar@aol.com
Former President, Hughes STX (later, Raytheon STX)
(mobile) 301-980-0661, (office) 410-309-0815
- (4) Ken Klenk, Science Systems Consulting, Principal Consultant: klenk@consultssc.com
Former VP, Raytheon Information Technology & Scientific Services
(mobile) 605-376-0411, (office) 605-336-9158
- (5) Lou Mayo, Adnet Systems, Program Manager: astronomer2go@verizon.net
(mobile) 301-641-8206, (office) 301-286-0165

Research and Professional Areas of Interest:

- _ Data Science, Big Data Analytics – Education and Research
- _ Large Scientific Database: Design, Data Systems, Science, Education
- _ Data Science, Discovery Informatics, & Astroinformatics
- _ Scientific Data Mining
- _ Machine Learning Algorithms for Knowledge Discovery in Large Databases

Pre-proposal submitted by Kirk Borne for participation in the 2013 NSF Ideas Lab

Brief Summary of Professional Background

I am a Transdisciplinary Data Scientist and Astrophysicist. Officially, I am Professor of Astrophysics and Computational Science in GMU's School of Physics, Astronomy, & Computational Sciences, but my passions are Big Data and using Data-in-Education, transcending any and all discipline boundaries. I am the undergraduate advisor for the GMU Computational & Data Sciences program. I have a physics BS degree from LSU and an astronomy PhD from Caltech. I have been at GMU since 2003. Previously, I spent nearly 20 years in data-intensive 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. I have extensive experience in Big Data and Data Science, including nearly 14 years of expertise in data mining. I have advised several federal agencies on data mining and Big Data applications, including the Executive Office of the President, the Library of Congress, the National Agricultural Library, National Weather Service, FDA Office of Drug Safety, and the NITRD Big Data Senior Steering Group. I am currently working on Big Data and education research, design, and development for the proposed NSF/MREFC Large Synoptic Survey Telescope (LSST). I am a founding member of the American Astronomical Society's Astroinformatics and Astrostatistics Working Group, a founding member of the new International Astrostatistics and Astroinformatics professional society, and on the editorial boards of several research journals. I have published nearly 200 research papers and book chapters, and I have given almost 200 invited talks at conferences and universities worldwide. In these roles, I focus on achieving big discoveries from big data, specifically in the context of "using data in the classroom." I promote the use of information and data-centric experiences with Big Data in the STEM education pipeline at all levels – I believe in data literacy for all!

Relevant Expertise

Throughout my career, I have engaged in activities that are particularly relevant to the Ideas Lab: I have worked with data as a research scientist and I have contributed to enhancements in the STEM education pipeline. In addition to winning numerous astronomical and data mining research grants, I was also PI on one NASA education grant, co-PI on another NASA education grant, and co-PI on two NSF education grants. During 2001-2007, I was a senior member of NSF's astronomy big data project NVO (National Virtual Observatory), contributing primarily in the areas of data mining and education/public engagement (EPE). Since 2005, I have been an active member of NSF's next big data astronomy project: the LSST. For LSST, I have contributed for many years (as an unfunded team member) to the big data research objectives and also (as a funded collaborator) to the development of the project's EPE program. For the latter, I made substantial contributions to the education user interface design, requirements analysis, use cases, objectives, goals, strategic vision, database design, data products specification, user community engagement, education logic models, "Understanding by Design" program definition, and more. After being a funded member of the education team for about 6 years, I joined the LSST OAB (Outreach Advisory Board), which is a multi-disciplinary external team of astronomers, technologists, formal educators, informal educators, and education professors. In addition to these activities, I am a member of the NSF-funded Zooniverse.org Citizen Science project, hosting a suite of online informal education experiences – the general public contributes to the discovery of knowledge and advancement of science through a variety of activities that crowd-source many of the time-consuming labor-intensive Big Data mining and characterization tasks that are far beyond the scope of various projects' science teams.

Experience with Teams

I have 25+ years of experience with teams. I was Senior Scientist on the NVO project, consisting of 70+ scientists and technologists, with whom I served on several project teams. I am Chair of the 50-member LSST Informatics and Statistics Research Collaboration Team. I also serve on the LSST Outreach Advisory Board – as a team, we develop formal and informal education programs for LSST Big Data. I was elected to two committees of the American Astronomical Society: the Dynamical Astronomy Division and Employment Committees – I was the chair of the latter one year. I was department manager for 8 years and then Project Manager for 2 years at NASA’s Space Science Data Operations Office, where I formed, facilitated, and participated in countless technical, scientific, and administrative teams. I attended a SERC-sponsored workshop in 2006, participating in a successful team effort to develop a DataSheet for “Using Data in the Classroom.”

Communication Skills with Non-Experts

This is one of my strongest skills. I have given presentations in grade school classes and public outreach venues for decades, with very positive feedback. I often get outstanding reviews from general education (non-expert) students who take my introductory data science courses. I gave a TedX talk on Big Data, with YouTube views of my talk surpassing all previous TedX talks by speakers from my institution. I have been able to communicate with non-experts and explain difficult concepts for as long as I can remember. For example: in high school, I tutored students in Math and Chemistry; and in college, as an undergraduate, I tutored teachers who were returning to college for Mathematics Education Masters degrees. In 2003, I successfully facilitated a professional development program for Oklahoma’s elementary school geography teachers – developing novel ideas for the use of Hubble Space Telescope data in the classroom.

Contributions to Creativity and Innovation with Strangers

I am often labeled as a creative, innovative person, and for being able to make people at ease in discussions of technical material. I was invited to the PD workshop mentioned above based on exactly such a recommendation from a NASA Education Program official to the workshop host. In this case, not only were the teachers unknown to me, but also: their field (geography) was not my field, the geography education standards were foreign to me, we were from different parts of the country, their expertise (elementary education) was totally unrelated to mine (Astrophysics Data Mining), and the teachers were from Native American Indian schools in Oklahoma’s tribal nations. Despite the outward appearance of wide gulfs between us, we all bonded wonderfully, developed several lesson ideas and classroom exercises, and had fun creating and innovating for a whole week – it was one of the richest experiences of my professional life.

Motivation and Personal/Professional Outcomes from the Ideas Lab

I have been driven to teach Data Mining in schools for many years – at least since 1998, when I received a NASA Education IDEAS grant for my Project AstroData concept – using cool science (Astronomy) and even cooler data (Hubble Telescope images), as a means to teach STEM concepts and thinking skills to K-12 students. I have become a social media world leader in Big Data (@KirkDBorne ranked #1 influencer on Twitter), where I engage others with these ideas daily. I recognize the unparalleled value of data-driven thinking, processes, and decision-making everywhere, especially in education. I am passionately motivated to explore and contribute to the fields of Learning Analytics, Personalized Learning, and Education Data Mining. The Ideas Lab offers me the opportunity to fulfill a dream, professional and personal, to share and grow my expertise within an education-focused setting, with the goal of making data-driven transformative changes in teaching and learning.