Bill Howe   *Curriculum Vitae – February 2016*

(404) 514-3003 · billhowe@cs.washington.edu · http://homes.cs.washington.edu/~billhowe

Employment
- **University of Washington**: Associate Director and Senior Data Science Fellow, eScience Institute, 2013–present.
- **University of Washington**: Affiliate Associate Professor, Principal Research Scientist, member of the Graduate Faculty, Computer Science & Engineering, 2013–present.
- **University of Washington**: Program Director and Faculty Chair, Data Science Masters Degree, 2015–present.
- **University of Washington**: Affiliate Assistant Professor, Senior Research Scientist, Computer Science & Engineering, 2009–2014.
- **Oregon Health and Science University**: Research Scientist, 2007–2009.
- **Portland State University**: Graduate Research Assistant, 2001–2006.
- **Deloitte, Microsoft, Schlumberger, Siebel**: Consultant, 1999-2001.

Education
- **Portland State University**, Ph.D. in Computer Science, 2007, *GridFields: Model-driven Data Transformation in the Physical Sciences*, under David Maier.
- **Georgia Institute of Technology**, B.S. in Industrial and Systems Engineering, 1999.

Grants
- Lead PI, **Department of Defense and PNNL** Polystore Data System, 2015-2016, $500,000
- Lead PI, **Macarthur Foundation** Advancing the Seattle MetroLab, 2016-2017, $250,000
- Co-PI **NSF Big Data Hub, Western Region**, 2015-2018, $1,200,000, with Ed Lazowska (UW), Ariel Rokem (UW), Mike Franklin (UC Berkeley), and Mike Norman (SDSC)
- Lead PI, **Department of Defense and PNNL** Unified Services for Federated Big Data Environment, 2015-2016, $500,000
- Lead PI, **Macarthur Foundation** Towards a Seattle Metro Consortium for Urban Data Science, 2014-2016, $300,000
- Co-PI **UW Provost's Office** Urban@UW, 2015-2017, $750,000, with Thaisa Way (UW), Vikram Jandalaya (UW), Thom Dunning (PNNL)
- Co-PI, **Gordon and Betty Moore Foundation and the Alfred P. Sloan Foundation**, Data Science Environments, 2014-2019, $32.8M, with NYU and UC Berkeley
- Co-PI, **Intel Science and Technology Center for Big Data**, 2014-2016, $125,000, one part of large distributed group across MIT, Brown, UW, and Portland State
- Lead PI, **PNNL Subcontract** Hybrid Architectures for Graph Query, 2014-2015, $200,000
- Co-PI, **NSF Grant** IIS-1247469, BIGDATA:A Formal Foundation for Big Data Management, 2013-2016, $2,966,667, with Dan Suciu (UW) and Magda Balazinska (UW)
- Lead PI, **PNNL Subcontract** Hybrid Languages for Graph Query, 2013-2014, $100,000
- Co-PI, **NSF Grant** ACI-1216879, Collaborative Research: Conceptualizing An Institute for Empowering Long Tail Research, 2012-2014, $500,000, with Ian Foster (UChicago), Bryan Heidorn (UArizona), Christine Borgman (UCLA), Carl

Kesselman (USC)

· Lead PI, **NSF Grant** IIS-1064505, Collaborative Research: Database-As-A-Service for Long Tail Science 2011-2013, $800,000, with Mike Cafarella (UMich), David Maier (Portland State), and Dan Suciu (UW).

· Co-PI, **EMC Grant** Large-Scale Visual Analytics, 2012-2014, $200,000, with Magda Balazinska (UW)

· Co-PI, **NSF Grant** CCF-1047815, Relational Data Markets in the Cloud, 2011-2013, $370,000, with Magda Balazinska (UW) and Dan Suciu (UW)

· Lead PI, **NSF Grant** CCF-1060213, EAGER: Scalable Algebraic Visualization in the Cloud, 2010-2011 $117,567

· Co-Lead, **Gordon and Betty Moore Foundation**, Data-Intensive Science in the Long Tail, 2010-2013, $750,000, with Ed Lazowska

· Lead PI, **NSF Grant** IIS-0844572, Collaborative Research: Where the Ocean Meets the Cloud: Ad Hoc Longitudinal Analysis and Collaboration Over Massive Mesh Data, 2009-2011, $500,000, with Claudio Silva (NYU) and Juliana Freire (NYU).

Honors

· **Best of VLDB 2010**, selected for publication in VLDB Journal special issue, for "Haloop: Efficient iterative data processing on large clusters." (currently the most cited paper from VLDB 2010)

· **Microsoft Jim Gray Seed Grant** 2010, $25,000

· **Microsoft Jim Gray Seed Grant** 2007, $40,000

· **Department Dissertation Award**, Maseeh College of Engineering and Computer Science, Portland State University, May 2007

· **Best of VLDB 2004**, selected for publication in VLDB Journal special issue, for "Algebraic manipulation of scientific datasets"

Publications

See also https://scholar.google.com/citations?user=dQ-x9NQAAAAJ

Peer-Reviewed Conference and Journal Publications

· Shrainik Jain, Dominik Moritz, Bill Howe, Ed Lazowska. **SQLShare: Results from a Multi-Year SQL-as-a-Service Experiment** SIGMOD (2016).

· Kanit Wongsuphasawat, Dominik Moritz, Anushka Anand, Jock Mackinlay, Bill Howe, Jeffrey Heer. **Voyager: Exploratory analysis via faceted browsing of visualization recommendations** Visualization and Computer Graphics, IEEE Transactions on 22(1) (2016).

· Jeremy Hyrkas, Bill Howe, **MusicDB: Relational Approach for Numeric Longitudinal Music Analytics** International Society on Music Information Retrieval (under review) (2016).

· Dominik Moritz,Daniel Halperin,Bill Howe and Jeffrey Heer. 2015. **Perfopticon: Visual Query Analysis for Distributed Databases**. *EuroVis* (2015).

· Hyrkas, Jeremy, Daniel Halperin, and Bill Howe. 2014. **Time-Varying Clusters in Large-Scale Flow Cytometry**, *Innovative Applications of Artificial Intelligence* (2014)

· Alawini, Abdussalam, David Maier, Kristin Tufte, and Bill Howe. 2014. **Helping scientists reconnect their datasets**. *Proc. of the 22nd Scientific and Statistical Database Management Conference SSDBM* (2014).

· Upadhyaya, Prasang, Nick R. Anderson, Magdalena Balazinska, Bill Howe, Raghav

Kaushik, Ravishankar Ramamurthy, and Dan Suciu. 2013. **Stop that query! The need for managing data use**. *Sixth Biennial Conference on Innovative Data Systems Research CIDR* (2013).

· Halperin, Daniel, Konstantin Weitz, Bill Howe, Francois Ribalet, Mark A. Saito, and E. Virginia Armbrust. 2013. **Real-time Collaborative Analysis with (almost) Pure SQL: A Case Study in Biogeochemical Oceanography**. *Proc. of the 21st Scientific and Statistical Database Management Conference (SSDBM)* (2013).

· Perry, Daniel, Bill Howe, Alicia M. F. Key, Cecilia Aragon. 2013. **VizDeck: Streamlining Exploratory Visual Analytics of Scientific Data**. *iConference 2013 Proceedings* (2013) pages 338–350. doi:10.9776/13206

· Koutris, Paraschos, Prasang Upadhyaya, Magdalena Balazinska, Bill Howe, and Dan Suciu. 2013. **Toward Practical Query Pricing with QueryMarket**. *SIGMOD conference* (2013), 613–624.

· Koutris, Paraschos, Prasang Upadhyaya, Magdalena Balazinska, Bill Howe, and Dan Suciu. 2012. **Query-based data pricing**. *Proceedings of the 31st ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems PODS* (2012), 167–178.

· Kwon, YongChul, Magdalena Balazinska, Bill Howe, and Jerome A. Rolia. 2012. **SkewTune: Mitigating Skew in Mapreduce Applications**. *SIGMOD conference* (2012), pages 25–36 (currently the most-cited paper from SIGMOD 2012).

· Shaw, Marianne, Bill Howe, Paris Koutris, and Dan Suciu. 2012. **Optimizing Large-Scale Semi-Naive Datalog Evaluation in Hadoop**. *Datalog 2.0* (2012).

· Howe, Bill, Alicia Key, Daniel Perry, and Cecilia R. Aragon. 2012. **VizDeck: A card game metaphor for fast visual data exploration**. *Proceedings of the 30th International Conference on Human Factors in Computing Systems* (2012), pages 1667–1672.

· AlSayyad, Yusra, Simon K. Krughoff, Bill Howe, Andrew J Connolly, Magdalena Balazinska, and Jones, Lynne. 2011. **Towards efficient and precise queries over ten million asteroid trajectory models**, *Proc. of the 23rd Scientific and Statistical Database Management Conference (SSDBM)* (2011), pages 568–570.

· Balazinska, Magdalena, Bill Howe, and Dan Suciu, **Data Markets in the Cloud: An Opportunity for the Database Community**, *Proceedings of the VLDB Endowment* PVLDB, 4, 12, 2011.

· Vo, Huy, Brian Summa, Joao Comba, Juliana Freire, Bill Howe, Claudio Silva, and Valerio Pascucci. 2011. **Parallel visualization on large clusters using mapReduce**. *IEEE Symposium on Large-Scale Data Analysis and Visualization, LDAV* (2011)

· Howe, Bill, Garret Cole, Emad Souroush, Paraschos Koutris, Alicia Key, Nodira Khoussainova, and Leilani Battle. 2011. **Database-as-a-service for Long Tail Science**. *Proc. of the 23rd Scientific and Statistical Database Management Conference (SSDBM)* (2011).

· Grochow, Keith, Mark Stoermer, James Fogarty, Charlotte Lee, Bill Howe, and Ed Lazowska. 2010. **COVE: A Visual Environment for Multidisciplinary Ocean Science Collaboration**, *IEEE Sixth International Conference on eScience (eScience)*, pages 269–276.

· Kwon, YongChul, Dylan Nunley, Jeffrey P Gardner, Magdalena Balazinska, Bill

Howe, and Sarah Loebman. 2010. **Scalable clustering algorithm for N-body simulations in a shared-nothing cluster**, *Proc. of the 22nd Scientific and Statistical Database Management Conference (SSDBM)*, (2010), pages 132–150.

· Kwon, YongChul, Magdalena Balazinska, Bill Howe, and Jerome Rolia. 2010. **Skew-resistant parallel processing of feature-extracting scientific user-defined functions**. *Proc. of the ACM Symposium on Cloud Computing (SOCC)* (2010).

· Bu, YingYi, Bill Howe, Magdalena Balazinska, and Michael Ernst. 2010. **HaLoop: Efficient iterative data processing on large clusters**. *Proc. of International Conf. on Very Large Databases (VLDB)* (Selected for Best of VLDB Issue of VLDB Journal; currently most-cited paper from VLDB 2010, a premier database conference)

· Grochow, Keith, Bill Howe, Mark Stoermer, and Ed Lazowska. 2010. **Client + cloud: Seamless architectures for visual data analytics in the ocean sciences**. *Proceedings of the 22nd international conference on scientific and statistical database management (SSDBM)* (2010).

· Howe, Bill, Harrison Green-Fishback, and David Maier. 2009. **Scientific Mashups: Runtie-Configurable Data Product Ensembles**. *Proceedings of the 21st international Conference on Scientific and Statistical Database Management (SSDBM)* (2009).

· Howe, Bill, Peter Lawson, Renee Bellinger, Erik Anderson, Emanuele Santos, Juliana Freire, Carlos Scheidegger, Antonio Baptista, and and Claudio T. Silva. 2008. **End-to-End eScience: Integrating Workflow, Query, Visualization, and Provenance at an Ocean Observatory**. *4th iEEE international Conference on eScience (eScience)* (2008).

· Howe, Bill, David Maier, and Laura Bright. 2007. **Smoothing the ROI curve for scientific data management applications**. *Proceedings of the Third Biennial Conference on Innovative Data Systems Research (CIDR)* (2007).

· Howe, Bill and David Maier. 2005. **Retrofitting a Data Model to Existing Environmental Data**. *Proceedings of the 17th international conference on Scientific and Statistical Database Management (SSDBM)* (Santa Barbara, CA, June 2005).

· Howe, Bill, and David Maier. 2004. **Algebraic Manipulation of Scientific Datasets**. *Proceedings of the 30th International Conference on Very Large Data bases* (2004) (Selected for Best of VLDB Issue of VLDB Journal).

· Marianne Koch, Lois Delcambre, Patricia Toccalino, Eric Landis, Fred Phillips, Tim Tolle, Len Shapiro, Nicole Steckler, David Maier, Mathew Weaver, Shawn Bowers, Balbinder Banga, Jason Brewster, Afrem Gutema, Sudarshan Murthy, Bill Howe, Rupa Tummala, Julia Norman, Kirsten Zillman, David Drake, Craig Palmer, Ashley Burt. 2003. **The Forest Portal: a Multidisciplinary Project**, *Proceedings of the 2003 annual national conference on Digital government research*, (2003), pages 1–4.

· Weaver, Mathew, Bill Howe, Lois Delcambre, Tim Tolle, and David Maier. 2002. **Representing, exploiting, and extracting metadata using metadata++**, *Proceedings of the 2002 annual national conference on Digital government research*, (2002), pages 1–7.

Journal Articles

· Seung-Hee Bae, Daniel Halperin, Jevin West, Martin Rosvall, Bill Howe. **Scal-**

able and Efficient Flow-Based Community Detection for Large-Scale Graph Analysis Transactions on Knowledge Discovery from Data (under review) 2016

· Poshen Lee, Jevin West, Bill Howe. **Viziometrics: Analyzing Visual Patterns in the Scientific Literature** Journal of the Association for Information Science and Technology (JASIST) (under review) 2016

· Jeremy Hyrkas, Sophie Clayton, Francois Ribalet, Daniel Halperin, E. Virginia Armbrust, Bill Howe. **Scalable clustering algorithms for continuous environmental flow cytometry** Bioinformatics 32(3) 2016

· Balazinska, Magdalena, Bill Howe, Dan Suciu, **The Database Group at the University of Washington**, *ACM SIGMOD Record*, 43, 1 (2014) pages 39–44

· Kwon, Kai Ren YongChul, Magdalena Balazinska, and Bill Howe. 2013. **Hadoop's adolescence: An analysis of hadoop usage in scientific workloads**. *PVLDB* (2013).

· Howe, Bill, Francois Ribalet, Daniel Halperin, Sagar Chitnis, and E. Virginia Armbrust. 2013. **Collaborative Science Workflows in SQL**. *Computing in Science & Engineering, Special Issue on Science Data Management.* 15, 3 (May/June 2013) pages 22–31, doi:10.1109/MCSE.2013.42

· Howe, Bill and Daniel Halperin. 2012. **Advancing Declarative Query in the Long Tail of Science**. *IEEE Data Engineerin Bulletin*, 35, 3. (2012). pages 16–26.

· Bu, YingYi, Bill Howe, Magdalena Balazinska, and Michael Ernst. 2012. **The HaLoop approach to large-scale iterative data analysis**. *Special Issue: Best of VLDB 2010* VLDB Journal, 21, 2 (2012).

· Barga, Roger, Bill Howe, David Beck, Stuart Bowers, William Dobyns, Winston Haynes, Roger Higdon, Chris Howard, Christian Roth, Elizabeth Stewart and others. 2011. **Bioinformatics and data-intensive scientific discovery in the beginning of the 21st century**, *Omics: A Journal of Integrative Biology*, 15, 4 (2011), pages 199–201

· Balazinska, Magdalena, Bill Howe, and Dan Suciu, **Data Markets in the Cloud: An Opportunity for the Database Community**, *Proceedings of the VLDB Endowment* PVLDB, 4, 12, 2011. (Also in conference)

· Wiley, Keith, Andrew J. Connolly, Jeffrey P. Gardner, K. Simon Krughoff, Magdalena Balazinska, Bill Howe, YongChul Kwon, Yingyi Bu. 2010. **Astronomy in the Cloud: Using MapReduce for Image Coaddition**, *CoRR*, abs/1010.1015, 2010

· Baptista, Antonio, Bill Howe, Juliana Freire, David Maier, and Claudio T. Silva. 2008. **Scientific exploration in the era of ocean observatories**. *Computing in Science and Engineering.* 10, 3 (2008), 53–58.

· Howe, Bill, and David Maier. 2004. **Logical and Physical Data Independence for File-based Scientific Applications**. *IEEE Data Eng. Bull.* 27, 4 (2004), 30–37.

· Howe, Bill David Maier, and Antonio Baptista. 2003. **A language for spatial data manipulation**. *Journal of Environmental Informatics.* 2, 2 (December 2003).

· Howe, Bill and David Maier. 2004. **Algebraic manipulation of scientific datasets**. *VLDB Journal: The International Journal On Very Large Data Bases (Special Issue: Best of VLDB 2004).* 14, 4 (2004).

Tech Reports

Myers, Brandon, Daniel Halperin, Jacob Nelson, Mark Oskin, Luis Ceze, and Bill Howe. **Radish: Compiling efficient query plans for distributed shared memory**.

Book Chapters

· Howe, Bill, 2014. **Reproducibility, Virtual Appliances, and Cloud Computing**, Chapter 10 in Victoria Stodden, Friedrich Leisch, Roger D. Peng (Ed.) *Implementing Reproducible Research*, Chapman and Hall/CRC.

· Balazinska, Magdalena, Bill Howe, Parachos Koutris, Dan Suciu and Prasang Upadhyaya. 2013. **A Discussion on Pricing Relational Data**, In Search of Elegance in the Theory and Practice of Computation, pages 167–173, Springer Berlin Heidelberg.

· Howe, Bill, 2013. **Transforming Data into the Appropriate Science Context**, Chapter 11 in Terence Critchlow, Kerstin Kleese van Dam (Ed.) *Data-Intensive Science*, Chapman and Hall/CRC.

· Howe, Bill and Magda Balazinska, 2012. **Beyond MapReduce: New Requirements for Scalable Data Processing**, Chapter 8 in Ian Gorton and Deborah K. Gracio (Ed.), *Data-Intensive Computing*, Cambridge University Press.

Peer-Reviewed Demonstration Papers

· Halperin, Daniel, Victor Teixeira de Almeida, Lee Lee Choo, Shumo Chu, Parachos Koutris, Dominik Moritz, Jennifer Ortiz, Vaspol Ruamviboonsuk, Jingjing Wang, Jingjing and Andrew Whitaker, Shengliang Xu, Magdalena Balazinska, Bill Howe, Dan Suciu, **Demonstration of the Myria big data management service**, *Proceedings of the 2014 ACM SIGMOD International Conference on Management of Data*, (2014), pages 881–884.

· Upadhyaya, Prasang, Nick R. Anderson, Magdalena Balazinska, Bill Howe, Raghav Kaushik, Ravishankar Ramamurthy, and Dan Suciu. 2013. **The power of data use management in action**. *SIGMOD* (2013), pages 1117–1120.

· Key, Alicia, Bill Howe, Daniel Perry, and Cecilia R. Aragon. 2012. **VizDeck: Self-organizing dashboards for visual analytics**. *SIGMOD conference* (2012), 681–684.

· Howe, Bill, Garret Cole, Nodira Khoussainova, Leilani Battle. 2011. **Automatic Example Queries for Ad Hoc Databases**, *Proceedings of the 2011 ACM SIGMOD International Conference on Management of data.* (2011), pages 1319–1322

· Howe, Bill, and David Maier. 2005. **Querying and Visualizing Gridded Datasets for eScience**. *21st international conference on data engineering (ICDE '05) (demo)* (Tokyo, Japan, 2005).

Peer-Reviewed Workshop Papers

· Shrainik Jain, Dominik Moritz, Bill Howe. **High Variety Cloud Databases**, Workshop on Cloud Data Management (CloudDM) (co-located with ICDE) [invited keynote] (2016)

· Poshen Lee, Jevin West, Bill Howe. **VizioMetrix: A Platform for Analyzing the Visual Information in Big Scholarly Data** BigScholar Workshop (co-located at WWW) 2016

· Joslyn, Cliff, Sutanay Choudhury, David Haglin, Bill Howe, Bill Nickless, Bryan

Olsen. 2013. Massive-Scale Cyber Traffic Analysis: A Driver for Graph Database Research, *First International Workshop on Graph Data Management Experiences and Systems* GRADES (2013).

· Myers, Brandon, Jeremy Hyrkas, Daniel Halperin, and Bill Howe. 2013. **Compiled plans for in-memory path-counting queries**. *Proceedings of the 1st international workshop on In Memory Data Management and Analytics, IMDM* (2013), 25–37.

· Howe, Bill and Garret Cole. 2010. **SQL is dead; long live SQL: Lightweight query services for ad hoc research data**. *4th Microsoft eScience Workshop* (2010)

· Loebman, Sarah, Dylan Nunley, YongChul Kwon, Bill Howe, Magdalena Balazinska, Jeffrey P Gardner, **Analyzing massive astrophysical datasets: Can Pig/Hadoop or a relational DBMS help?**, *IEEE International Conference on Cluster Computing* CLUSTER (2009), pages 1–10.

· Suciu, Dan, Andrew J Connolly, and Bill Howe. 2009. **Embracing Uncertainty in Large-Scale Computational Astrophysics**. *Proceedings of the Workshop on Management of Uncertain Data*, (2009), pages 63–77.

· Howe, Bill, Huy Vo, Claudio Silva, and Juliana Freire. 2009. **Query-driven visualization in the cloud with mapReduce**. *Fourth workshop on Ultrascale Visualization* (2009).

· Howe, Bill, David Maier, Nicolas Rayner, and James Rucker. 2008. **Quarrying dataspaces: Schemaless profiling of unfamiliar information sources**. *Data engineering workshop, 2008. iCDEW 2008. iEEE 24th international conference on* (2008), 270–277.

· Howe, Bill, and Roger Barga. 2008. **The eScience Appliance: Provisioning an Inexpensive Bottom-Up Cyberinfrastructure**. *2nd microsoft eScience workshop, in conjunction with eScience 2008* (2008).

· Bright, Laura, David Maier, and Bill Howe. 2006. **Managing the forecast factory**. *Proceedings of the ICDE Workshop on Workflow and Data Flow for Scientific Applications (SciFlow)* (2006).

· Howe, Bill, Kuldeep Tanna, Paul Turner, and David Maier. 2004. **Emergent semantics: Towards self-organizing scienitifc metadata**. *Proceedings of Semantics for a Networked World: Semantics for Grid Databases* (Paris, France, June 2004).

· Howe, Bill and David Maier. 2002. **Modeling data product generation**, Workshop on Data Derivation and Provenance (Chicago, IL, 2002).

Peer-Reviewed Abstracts, Talks, Posters

· Ren, Kai, Garth Gibson, YongChul Kwon, Magda Balazinska and Bill Howe. 2012. Hadoop's Adolescence; A Comparative Workloads Analysis from Three Research Clusters,, *High Performance Computing, Networking, Storage and Analysis (SCC)*, pages 1453–1453, 2012

· AlSayyad, Yusra, Simon Krughoff, Andrew J. Connolly, L. Jones, Tomas Budavari, and Bill Howe. 2011. Efficient Prediction of Asteroid Positions from Solar System Models, *Bulletin of the American Astronomical Society*, 43 (2011).

· Fishback, Harrison-Green, and Bill Howe. 2008. Scientific mashups: Runtime-configurable data product ensembles. *4th iEEE international conference on e-science (eScience) (poster)* (2008).

· Bill Howe and Nicholas Hagerty and Ethan Van Matre and David Maier and

Antonio Baptista and Charles Seaton, Paul Turner. 2007. The Ocean Appliance: Complete Platform Provisioning for Low-Cost Data Sharing. *Proceedings of the OCEANS 2007 mTS/IEEE conference* OCEANS (2007).

Teaching

University of Washington
· Program Director and Faculty Chair, Masters of Science in Data Science, 2015
· *Introduction to Data Science* (Coursera), Summer 2014, 100,000+ students.
· *Introduction to Data Science* (Coursera), Spring 2013, 100,000+ students.
· Founder, University Education Outreach Certificate in Data Science, 2013
· Founder, University Education Outreach Certificate in Cloud Computing, 2013
· *Introduction to Data Programming*, Winter 2013.
· *Introduction to Data Science* (PCE Certificate), Fall 2012.
· *CSE 190p Introduction to Data Programming with Applications*, Summer 2012.
· *Data-Intensive Computing in the Cloud* (PCE Certificate), Fall 2012.
· *CS599c Scientific Data Management*, Winter 2011.
· *CSE344 Introduction to Data Management*, Fall 2010

Portland State University
· *Scientific Data Management*, Spring 2006.

Students

· **Brandon Myers**, CSE (with Mark Oskin), Ph.D. expected June 2016 (query compilation, high-peformance computing)
· **Poshen Lee**, EE, Ph.D. expected August 2016 (machine vision
· **Shrainik Jain**, CSE, Ph.D. expected June 2018 (polystores, high-variety databases)
· **Dylan Hutchison**, CSE, Ph.D. expected June 2020.
· **Dominik Moritz** (co-advised with Jeff Heer), Ph.D. expected June 2019.
· **YongChul Kwon** (co-advised with Magda Balazinska), Ph.D. received June 2011. Skew-Resistant Parallel Computation. now with Microsoft.
· **Jeremy Hyrkas**, CSE (Big data applications, analytics for science) now at Microsoft
· **Yongchul Kwon** CSE Big data systems, fault-tolerance, skew handling. Now at Microsoft
· **Yingyi Bu** CSE Big data systems and languages. Moved to UC Irvine
· **Tanmay Modak** iSchool Masters Visual analytics, machine learning
· **Elton Dias** iSchool Masters Visual analytics, machine learning
· **Abhigyan Kaustubh** iSchool Masters Visual analytics, machine learning

Postdocs

· **Seung-Hee Bae**, CSE postdoc (Graph clustering algorithms, HPC + big data)
· **Sophie Clayton**, eScience postdoc (Environmental flow cytometry, analytics)
· **Dan Halperin**, CSE Postdoc and Senior Research Scientist. Now at Google
· **Andrew Whitaker**, CSE Research Scientist. Now at Amazon
· **Marianne Shaw**, CSE Postdoc. Now at Tableau

Staff

· **Sarah Stone**, PhD Oceanography, Program Manager
· **Micaela Parker**, PhD Oceanography, Program Manager

- **Rob Fatland**, PhD Physics, Director of Cloud and Data Solutions
- **Jacob Vanderplas**, PhD Astronomy, eScience Data Scientist
- **Ariel Rokem**, PhD Neuroscience, eScience Data Scientist
- **Bernease Herman**, eScience Data Scientist
- **Valentina Staneva**, PhD Applied Math, eScience Data Scientist

Panels/Reports
- Co-PI, NSF Western Big Data Hub
- Co-founder, Urban@UW
- Director, Seattle MetroLab
- Chair, *Should we all be teaching intro to data science instead of intro to databases?*, Panel, Proceedings of the 2014 ACM SIGMOD international conference on Management of data.
- Co-Organizer, *Designing good algorithms for MapReduce and beyond*, Proceedings of the Third ACM Symposium on Cloud Computing. 2014.
- Co-Organizer, *Education and career paths for data scientists*, Panel, Magdalena Balazinska, Susan B Davidson, Bill Howe, Alexandros Labrinidis, Proceedings of the 25th International Conference on Scientific and Statistical Database Management. 2013.
- Chair, High-Performance Computing Meets Databases, 2011
- Chair, High-Performance Computing Meets Databases, 2010

Recent Talks
- Over the last three years I gave invited talks at Los Alamos National Labs, Illinois, UC Berkeley, Indiana, Penn State, Intel, the Department of Defense, the American Geophysical Union, the Joint Statistical Meetings, Pacific Northwest National Labs, UCBerkeley, the Allen Institute for Brain Science, the Macarthur Foundation, Microsoft Research
- Keynote, Cloud Data Management Workshop (with ICDE), 2016, Helsinki
- Invited Talk, Urban Data Science @ UW, NIAC Workshop on Smart Cities, 2016
- Invited Talk, High Productivity Computing, Los Alamos National Lab, November 2015
- Invited Talk, High Productivity Computing, Chesapeake Large-Scale analytics Conference, October 2015
- Invited Talk, High Productivity Computing, Neurofutures Symposium, Portland, August 2015
- Invited Talk, ITIC and Macarthur workshop "Partnership for Sustainable Smart Cities of the Future"
- Invited Talk, National Research Council Meeting on "Training the Big Data Workforce"
- Keynote at eXtremely Large Databases, South America, 2014
- Invited Talk at JSM panel on Big Data Education: Big Data Curricula at the UW eScience Institute.
- Invited Talk on Myria and scalable graph clustering at the MMDS workshop at UC Berkeley, 2014.
- Keynote at eResearch New Zealand, eScience and Data science at the University of Washington's eScience Institute, 2013
- For a partial list of presentations, see http://www.slideshare.net/billhoweuw

Program Committees and Editing

- Program Chair, eScience 2016
- Program Committee, HotCloud 2014
- Program Committee, eScience 2014
- Program Committee, SSDBM 2013
- Program Committee, LDAV 2013
- Chair, Workshop on HPC meets Databases, co-located with Supercomputing 2012
- Program Committee, ICDE 2013
- Program Committee, PVLDB 2012-2013
- Demo Co-chair, SSDBM 2012
- Program Committee, ScienceCloud 2012
- Chair, Workshop on HPC meets Databases, co-located with Supercomputing 2011
- Editorial Board, Journal of Data Semantics
- Organizing Committee, XLDB 2011
- Program Committee, LDAV 2011
- Program Committee, ScienceCloud 2011
- Co-Chair, Workshop on Array Databases
- Registration Chair, SSDBM 2011
- Program Committee, SSDBM 2011
- Demonstrations Program Committee, SIGMOD 2011
- Program Committee, EDBT 2010
- Program Committee, SSDBM 2010
- Program Committee, IIMAS Workshop, 2008
- Reviewer, VLDB Journal, 2007
- Program Committee, dg.o 2006
- Program Committee, dg.o 2005
- Demonstrations Program Committee, SIGMOD 2005
- Student Session Program Committee, dg.o 2004

Activities

- Associate Director of the eScience Institute, responsible for management of the staff, operations, and program development.
- Program Director and Faculty Chair, Data Science Masters Degree. Led curriculum design, program design, hiring
- Co-founder, Urban@UW
- Director, Seattle MetroLab
- Launched a free campus-wide service based on our research data management system SQLShare in partnership with UW-IT
- Lead, Twitter Data Grant Program, University of Washington.
- Designed and implemented the UW eScience Incubation Program for Data-Intensive Science Projects
- Attended the National Research Council meeting on "Training the Big Data Workforce," 2014
- Founder, University Education Outreach Certificate in Data Science, University

of Washington, 2013

· Founder, University Education Outreach Certificate in Cloud Computing, University of Washington, 2013

· Designed and operated *Introduction to Data Science* online course on Coursera, attracting 200,000+ students across two offerings

Personal   5222 36th Ave
Seattle, WA 98105

Phone: (404) 514-3003
Email: billhowe@cs.washington.edu
Web: http://homes.cs.washington.edu/~billhowe
Citizenship: USA