

# Democratizing Data-Intensive Science

**Abstract:** Advances from data science (and data-intensive science) appear to be derived primarily from the composition, integration, and broad application of existing techniques and technologies rather than (solely) the development of new techniques. But this problem of technology "delivery" receives relatively little research attention.

At the UW eScience Institute and in the UW Database Group, we are building platforms to democratize advanced data management, curation, and analytics across all fields of science and across all levels of expertise.

In this talk, I'll describe our findings from a multi-year deployment of a database-as-a-service system called SQLShare, and recent results in the context of the Myria project, a federated data management and analytics system that supports multiple backend engines, iteration as a first-class citizen, new algorithms, built-in visualization and performance profiling, and a language interface that balances imperative and declarative features.

I'll wrap up with a tour of our efforts to develop organizational infrastructure to complement the software infrastructure, including an incubator program for interdisciplinary projects, new educational initiatives, and cross-campus collaborations in data-intensive science.

**Bio:** Bill Howe is the Associate Director of the UW eScience Institute and holds an Affiliate Faculty appointment in Computer Science & Engineering. His research interests are in data management, curation, analytics, and visualization in the sciences. Howe has received two Jim Gray Seed Grant awards from Microsoft Research for work on managing environmental data, has had two papers selected for VLDB Journal's "Best of Conference" issues (2004 and 2010), and co-authored what are currently the most-cited papers from both VLDB 2010 and SIGMOD 2012. Howe serves on the program and organizing committees for a number of conferences in the area of databases and scientific data management, and developed a first MOOC on data science that attracted over 200,000 students across two offerings. He has a Ph.D. in Computer Science from Portland State University and a Bachelor's degree in Industrial & Systems Engineering from Georgia Tech.