DATA LAB @ CCIS

Everybody is talking about Big Data. Our lab is designing the algorithms and technology that power the management and analysis of big data, complex data, diverse data, and fast data (streams). We focus on scalable analysis and exploration techniques for data-driven science, uncertain data, graph data management, and parallel data processing.

The DATA Lab @ Northeastern is headed by Professors Wolfgang Gatterbauer and Mirek Riedewald, and will welcome another faculty member in Fall 2018. We collaborate with colleagues in computer science (e.g., machine learning, theory, formal methods) and various domain sciences (e.g., neurobiology, ornithology, physics).

Areas of Investigation
- Exploratory data analysis
- Scalable parallel data processing
- Data provenance
- Probabilistic inference
- Predictions over network data
- Multidimensional data
- Management of uncertain data
- Data cleaning and integration

Example Research Topics
- How can we parallelize data-intensive computation in the cloud and prove properties such as near-optimal load distribution?
- How can we extend relational databases to manage new types of data?
- How can we approximate hard problems in probabilistic inference?
- How can we predict labels of nodes in large networks in a few seconds?
- How can we help scientists make new discoveries about the structure of the brain, the migration of wild bird species, and the impact of environmental change on flora and fauna?

Related PhD Programs
- PhD in Computer Science
- PhD in Network Science

Join Our Growing Lab
We have 5 open and fully funded PhD positions. Create the algorithms, systems, and theoretical foundations of tomorrow’s big-data analysis.

Learn More About Data Lab @ CCIS
db.ccis.northeastern.edu

Northeastern University
College of Computer and Information Science

Wolfgang Gatterbauer
Associate Professor
w.gatterbauer@northeastern.edu

Mirek Riedewald
Associate Professor
m.riedewald@northeastern.edu