INFO-I590 (34881)
Topics: Management, Access, and Use of Big and Complex Data
Dr. Beth Plale (plale@indiana.edu)

Data is abundant and its abundance offers potential for new discovery, and economic and social gain. But data can be difficult to use. It can be noisy and inadequately contextualized. There can be too big a gap from data to knowledge, or due to limits in technology or policy not easily combined with other data. This course will examine the underlying principles and technologies needed to capture data, clean it, contextualize it, store it, access it, and trust it for a repurposed use. Specifically the course will cover the 1) distributed systems and database concepts underlying noSQL and graph databases, 2) best practices in data pipelines, 3) foundational concepts in metadata and provenance plus examples, and 4) developing theory in data trust and its role in reuse.

- Fall 2014 Course Syllabus and Website
  http://www.soic.indiana.edu/graduate/degrees/data-science/online-certificate/index.html