

# SOCIAL DATA ANALYTICS (SODA)

---

SODA 501: Big Social Data: Approaches and Issues

3 Credits

Interdisciplinary integration of computational, informational, statistical, visual analytic, and social scientific approaches to the creation of big social data. This course addresses computational, informational, statistical, visual analytic, and social scientific approaches to the creation of data that are both "social" (about, or arising from, human interactions) and big (of sufficient scale, variety, or complexity to strain the informational, computational, or cognitive limits of conventional social scientific approaches to data collection or analysis). Examples include text, image, audio, video, intensive spatial and/or longitudinal data, data with complex network, hierarchical and/or other relational information, data from distributed sensors and mobile devices, digitized archival data, and data exhaust from sources like social media. Possible topics include sources of social data, data structures and formats for social data, data collection and manipulation technologies, data linkage and alignment, ethics and scientific responsibility in human subjects research, experimental and observational data collection design for causal inference, measurement of latent social concepts, reliability and validity, search and information retrieval, nonrelational and distributed databases, and standards for data preservation and sharing.

SODA 502: Social Data Analytics: Approaches and Issues

3 Credits

Interdisciplinary integration of computational, informational, statistical, visual analytic, and social scientific approaches to learning from big social data. This course addresses the interdisciplinary integration of computational, informational, statistical, visual analytic, and social scientific approaches to learning from data that are both "social" (about, or arising from, human interactions) and "big" (of sufficient scale, variety, or complexity to strain the informational, computational, or cognitive limits of conventional social scientific approaches to data collection or analysis). Topics include alternative scientific models for learning from data (Bayesian inference, causal inference, statistical / machine learning, visual analytics, measurement modeling), analytics issues with big data (variable selection, parallel computing, algorithmic scaling, ensemble modeling, validation), analytics issues with particular structures and channels of social data (network data, geospatial data, intensive longitudinal data, text data), and issues of scientific responsibility and ethics in analysis of big social data.