Workshop on

Big Data and Urban Informatics:
e-Infrastructure for Urban Social Science Research

Dates: March 28 and 29, 2013
Location: University of Illinois at Chicago

BACKGROUND MATERIAL

Vonu Thakuriah
University of Glasgow

March 21, 2013

1 Short Description

The workshop has two major goals: (1) to establish a multi-disciplinary research group to consider decadal-scale research questions on how Big Data, together with other urban data, can be used for urban social science research (analytics, applications, models and theory-building) and (2) to plan for a second (larger) workshop on the subject. Other topics of discussion will include the design of an Urban Informatics E-Infrastructure (UIE) for urban social science research, a journal special issue, a whitepaper and related topics.

Participants of the workshop are from the following disciplines: urban planning, geography, public administration and management, economics, civil engineering, statistics, survey methodology and computer science.

Informal interactions and presentations highlighting multiple perspectives will, together with group discussions and roundtables, help tease out how to best go about achieving the workshop goals.

2 Background

Urban and regional analysis involves a complex range of methods and tools and spans a broad range of areas including transportation, environment, health, land-use, housing, economy, labor markets, population demographics, ecology, energy and water usage. The goals of such research are wide-ranging; a far from complete list includes discovering novel patterns in urban systems, evaluating outcomes of policy interventions, understanding social dynamics of citizens and socio-

political institutions, determining strategies to manage risk, and for overall public management, planning and decision support towards long-term sustainable and socially-just outcomes.

Currently, such research activities mostly use data from sources such as the census; federal, state and local administrative datasets; and surveys. Specialized survey efforts that yield insights into new and emerging urban processes and risk factors are few and far between. Administrative data that are used to study problems are not always enough to adequately discover gaps, needs and demands. Response rates of traditional surveys are declining and although surveys will continue to play a central role in urban research, alternative approaches should be examined to supplement the data infrastructure for urban social science research.

The convergence of pervasive sensing, wireless networking and location-aware technologies from vast numbers of mobile and static sensors offer new opportunities regarding ways in which urban systems of the future can be designed and used. Information that is available anywhere and anytime is being greatly fostered by rapidly growing sensor systems enabled by personal devices such as cell phones and smartphones and by microblogging and other forms of social media. These, along with sensors in the transportation, energy, utility, building, weather and environmental management infrastructure, can lead to "zettabyte" scale sensor-based information on cities. Such "Urban Big Data" has primarily stimulated extensive research on urban sensing and knowledge discovery in urban spaces primarily from a computational perspective.

Big Data presents significant opportunities for existing urban and regional analysis models and urban policy analysis tools. Research is needed on how the data needs to be structured in order to be used as input into such analysis methods. A synergistic question is on ways in which the methods themselves may need to be retrofitted to accommodate Big Data. New opportunities that cut across a spectrum of urban disciplines and lead to transformative urban social science research may be stimulated by Big Data particularly with respect to human dynamics, social processes, institutional transformation and individual and group decision-making.

The workshop will focus on how data and tools may be presented in an integrated e-infrastructure (UIE), so that future theory-building and experimental work can occur rapidly. The UIE will potentially consist of three components: a Big Data resource, Urban Informatics Analysis Tools and Emerging Technologies to facilitate the use of the e-infrastructure by researchers.

Approaches by which Big Data can be utilized, *when integrated with traditional sources of data such as surveys and administrative data*, have not been systematically examined in urban social science research. Most research on the urban space use traditional data or more recently Big Data but not a combination of the two. The opportunity provided by the workshop for networking and inter-active research agenda-building among members of multiple disciplines is likely to stimulate new collaborations and ideas regarding urban research and applications. A discussion between members of the urban research community and the computational and information science communities will be beneficial for this purpose.

3   Workshop Objectives in Summary

The major objective of the workshop are to:

1. Establish a multi-disciplinary research community on Urban Informatics and Big Data;

2. Generate concrete ideas for an e-infrastructure (Urban Informatics e-Infrastructure, UIE) which will house data and tools to facilitate urban social science research;

3. Stimulate decadal-scale data-intensive research questions on how Big Data can be leveraged for urban social science research;

4. Develop educational and training plans and practices;

5. Disseminate our work to a broad audience of researchers and practitioners to ensure sustainability of the research community over time.

4   Topics for Potential Consideration in the Workshop

The list of potential topics to be considered in Workshop 1 is given below. While the list is long and we will not be able to cover all topics, we may consider including some of the topics in the Call for Workshop 2 presentations. the journal special issue and the whitepaper.

1. Big Data and other resources for urban social science research that could potentially be included in the UIE:

   (a) Identify data-intensive questions of importance to urban social scientists and practitioners;

   (b) Data from mobile devices and Web 2.0, sensor systems in the built environment and urban infrastructure, vehicular sensing systems;

   (c) Outputs from models, administrative data, census and existing survey data;

   (d) Multi-sensor fusion;

   (e) Data quality metrics and metadata.

2. Tools and methods:

   (a) Data mining and knowledge discovery for urban dynamics;

   (b) Agent-Based Modeling and large-scale simulations of urban systems using Big Data;

   (c) Planning support systems, visualization and participatory planning and decision-support tools;

   (d) Sustainable choice impact models;

   (e) Urban sensing, community-based volunteered information system formation and sustainability.

3. Strategies to develop a UIE prototype:

   (a) Data exchange, integration, storage and dissemination;

   (b) Web services and e-Research tools provision;

(c) Information extraction and retrieval;

(d) Approaches to form on-line virtual community: publish research data, share knowledge;

(e) Tools to rapidly deploy and access software applications;

(f) Privacy, information security and management;

(g) Data quality issues;

(h) E-infrastructure sustainability models and approaches.