Introduction to Proceedings of Workshop on Big Data and Urban Informatics sponsored by National Science Foundation

Piyushimita (Vonu) Thakuriah, University of Glasgow, Workshop Chair
Nebiyou Yonas Tilahun, University of Illinois at Chicago, Workshop Co-Chair
Moira Zellner, University of Illinois at Chicago, Workshop Co-Chair

The Workshop on Big Data and Urban Informatics held at the University of Illinois at Chicago on August 11 and 12, 2014 and sponsored by the National Science Foundation provided a unique opportunity to bring together urban social scientists and data scientists interested in the use of Big Data to address urban challenges. A background paper titled "*Big Data and Urban Informatics: Innovations and Challenges to Urban Planning and Knowledge Discovery*" documenting the major motivations for the workshop is given in Chapter 2 of these proceedings and is also available online at *https://urbanbigdata.uic.edu/files/2015/11/Background.pdf*.

There has been a lot of buzz about Big Data, and the term has been used to mean different types or attributes of data. In the urban sector, Big Data is generated in different ways, allowing an exploration and analysis of patterns and dynamics underpinned by complex urban challenges that cannot be studied only by using data from traditional surveys, census or other conventional data systems. We use Big Data to mean data that are generated through transactional, operational, planning and social activities, which are not specifically designed for research, or the linkage of such data to traditional data. The resultant complexities associated with the generation, access and use of such data require special considerations for technology and information management, methods of analysis, and the underlying scientific paradigm and political economy supporting inquiry. Big Data provides significant opportunities for the advancement of Urban Informatics, which is the exploration and understanding of urban patterns and dynamics undertaken from both a theory-driven perspective, as well as an empirical data-driven perspective for conceptual insights and knowledge discovery regarding urban systems, and for improved urban operations, planning and policy, citizen engagement, governance, and innovations.

Urban Big Data can be generated through sensor systems in urban infrastructure and moving objects such as cars. It can be user-generated through social media, human computation tasks and citizen science projects. It can refer to administrative data accessed through open data portals or through restricted access to highly confidential governmental records, as well as to private sector data on consumer transactions on goods and services and service agreements. It can also be highly unstructured data from arts and digital humanities col- lections. Additionally, it encompasses hybrid data sources such as linked survey and sensor data. There are now significant user communities with interest in these types of urban data sources spanning across a range of disciplines, as well as private industry and governments.

We organized an initial workshop, which took place on March 28th and 29th, 2013 in the University of Illinois at Chicago, to help generate ideas for a larger event. Twenty-one academics and practitioners and three PhD students (listed in Appendix A) participated in this workshop, representing 9 institutions in the US, UK and Australia. The disciplines represented were Urban Planning, Computer Science, Civil Engineering, Economics, Statistics, and Geography. The participants of the first workshop were identified during the proposal-writing phase and served on the program committee of the main workshop, and helped to delineate the purpose driving this event.

Our main motivation for the main workshop, which took place in August 2014, was to convene researchers and professionals working on this interdisciplinary topic, and to organize a community with interests on theoretical developments and applications demonstrating the use of urban Big Data, and the next-generation of Big Data services, tools and technologies for Urban Informatics. We were interested in research results as well as idea pieces and works in progress that highlighted research needs and data limitations. We sought papers that clearly create or use novel, emerging sources of Big Data for urban and regional analysis in transportation, environment, public health, land-use, housing, economic development, labor markets, criminal justice, population demographics, urban ecology, energy, community development, and public participation.

The scope of the workshop ranged across five key strands of the Big Data debate in the urban context:

Theoretical developments and knowledge discovery in urban systems: Theoretical developments, knowledge discovery and hypothesis generation about urban dynamics and processes, and understanding of urban systems and their social, behavioral, political, mobility and economic aspects including models of transactions, incentives, collaboration, and cooperation, and behavioral or organizational change, as well as epistemological positions and scientific paradigms regarding modes of inquiry, and investigations into the validity of approaches used and limits of knowledge about urban systems derived from datafication;

Information management for Urban Informatics: Information management for Urban Informatics including solutions to managing urban information such as data infrastructure management solutions, search and querying, data security and privacy, resource discovery, and language and execution environments, and computational approaches relating to information gathering, management and distribution of urban Big Data;

Measurement, analysis and methodological research: Approaches for urban Big Data information retrieval, extraction, linkage, analytics and visualization; advancements in exploratory and predictive analytics; agent-based modeling and complex systems; statistical inference, data quality and related issues such as missing data, endogeniety and selection biases; assessment of the extent to which urban Big Data may be able to add to traditional survey-based urban social science research, as well as examples of Big Data linkage with census, survey and administrative data; evolving goodness of fit metrics for models and data; methods to handle verification and validation, and sensitivity analysis; and experimental design;

Planning, policy analysis and operational uses of urban Big Data: Empirical research exemplifying Big Data for improved planning, management and governance in the urban sectors including uses for decision-making, and development of indicators to monitor economic and social activity, and for urban sustainability, transparency, livability, social inclusion, place-making, accessibility and resilience;

Institutional issues, organizations, networks and infomediaries in urban Big Data: Studies of locational privacy, trust management and information security relating to urban Big Data; analysis of social networks and sensing systems involved in urban data; studies of civic hacking networks, new types of digital entrepreneurship, organizational assessments of open data portals and city dashboards, questions of political economy relating to information networks and power structures, data access, data confidentiality and security, data governance framework, provenance and ethics.

The main workshop on Big Data and Urban Informatics took place in Aug 11-12, 2014 at the University of Illinois at Chicago and was attended by approximately 150 researchers, educators, practitioners and students from urban social sciences and data sciences interested in the use of Big Data to address urban challenges. Participants were from 11 countries – USA, UK, China, Canada, Japan, Portugal, Greece, Ireland, Israel, Italy, and Australia –representing 91 different institutions. We received 91 initial submissions of which a total of 68 papers were presented over a course of two days, ranging across the five themes.

Papers presented can be classified into 12 broad categories as follows:

1. Analytics of User-Generated Content (9 papers)
2. Data behind Urban Big Data (5 papers)
3. Urban Plan-Making (4 papers)
4. Changing Organizational and Educational Perspectives with Big and Open Urban Data (4 papers)
5. Urban Knowledge Discovery- General (12 papers)
6. Urban Knowledge Discovery – Transportation (8 papers)
7. Urban Knowledge Discovery – Cities, Land Use, Energy (4 papers)
8. Health and Well-Being (5 papers)
9. Urban Data Management (5 papers)
10. Livability & Sustainability (4 papers)
11. Insights into Social Equity (4 papers)
12. Emergencies and Crisis Informatics (4 papers)

Selected papers from the workshop will be published, after additional peer-review, in an edited volume titled *Seeing Cities Through Big Data - Research, Methods and Applications in Urban Informatics* anticipated to be released towards the end of 2015. Papers presented at the workshop have also been released through the workshop's website at http://urbanbigdata.uic.edu/proceedings and are publicly available for researchers. The research team also used social media and academic and non-academic mailing lists to announce both the conference as well as the release of the conference proceedings.