A Middleware Framework for Urban Data Management

Larissa Romualdo Suzuki
University College London
Gower Street
London, WC1E 6BT, UK
larissa.romualdo.11@ucl.ac.uk

Anthony Finkelstein
University College London
Gower Street
London, WC1E 6BT, UK
a.finkelstein@ucl.ac.uk

David Gann
Imperial College London
Faculty Building
South Kensington Campus
London SW7 2AZ, UK
d.gann@imperial.ac.uk

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.

UbiComp'13 Adjunct, September 8–12, 2013, Zurich, Switzerland.
Copyright © 2013 978-1-4503-2215-7/13/09…$15.00.
http://dx.doi.org/10.1145/2494091.2499223

Abstract
The domain of inquiry of this research is the collection, organization, integration, distribution and consumption of knowledge derived from urban open data, and how it can be best offered to application cities' stakeholders through a software middleware. We argue that the extensive investigation proposed in this research will contribute to a growing body of knowledge about data integration and application in smart cities, and offer opportunities to re-think an integrated urban infrastructure.

Author Keywords
Smart cities; software architecture; big data; value chain.

ACM Classification Keywords
D.2.11 [Software Engineering] Software Architectures - Domain-specific architectures - smart cities

Introduction
The infrastructure of cities has evolved through many vintages of technology that developed along their own path, often separately. The lack in connecting its component systems, which depend upon the other, often makes city utilities and services operate sub-optimally, limiting the creation of new value-added services, increasing transport costs, damaging existing logistics chains and economic models. Digital technologies offer a new wave of opportunities to mitigate some of these impacts and create a balance between social, environmental and economic

Smart city big data

Today, many areas of science are facing hundred- to thousand-fold increases in data volumes when comparing with just one decade ago [7]. This huge volume of data is mainly due to the world explosion of Internet connected devices which are expected to reach 25 billion by 2015 and 50 billion by 2050. In a similar fashion, during the past decades decisions about city infrastructures and services have been made based mostly on traditional structured data stored in relational databases. With the decreasing in the cost of computing (power and storage) and the proliferation of wireless and sensors networks, pervasive technologies begun to be embedded into city environment interconnecting the digital and physical realms transforming cities into pervasive urban environments [8].

opportunities that will be delivered through smart city planning, design, and construction [1].

Smart city represents a very important new concept, and it is no surprise that it has received so much attention from researchers, private organizations, government, and investors. Lead by this new "buzzword", many people have assumed that the use of technology will change everything and taking a top down approach, i.e. infusing modern technology without considering the main consumers of the services, is the best tactic to follow. That may be a natural reaction, especially coming from business companies which sell software solutions; however, it can be a dangerous one. By failing to follow appropriate business strategies, many cities have reduced the likelihood that they could succeed. Take the case of one of the first digital cities initiatives, Blacksburg Electronic Village, which had its activity decreased after only two years as a result of the disagreement between stakeholders (technology providers and users) in regarding the goals and expectations of community networks [2]. The experience of the South American city of Lima [3] demonstrated that technology diffusion not necessarily means acceptance nor equal take up. Some cities have used technology to shift the basis of smart cities away from its essence. While there are evidences on the impact of ICTs on urban environment, there is a belief that information technology will on its own deliver a smart city [4]. Actually, smart cities are much more than cables, broadband, sensors, actuators and data.
To date, many attempts have been made towards designing smart cities data management solutions. However, the missing insights on the impacts that technologies, stakeholders/users requirements, and big data will have in the overall city environment operations has led the creation of isolated "smart cities", each with its own silos of data, standard agreements, and technological approaches (e.g. [5-6]). This research suggests that what hinders cities from creating real smart cities goes beyond technical hurdles and social effects, yet it also encompasses a comprehensive overview of the complete data "supply chain", the users and system design requirements.

This research entails investigating the technical, institutional and logistical obstacles that complicate the development of such software middleware, the smart city stakeholders and their expectations with regards the middleware, the semantic challenges introduced by the data heterogeneity, and the nonexistence of a common model. In order to identify the relevant urban data activities and processes that enables the creation of value, we investigates a smart city data value chain which serves as building blocks to the architectural design of the proposed middleware framework.

Research Approach

As an alternative to the prevailing top-down approaches adopted in previous studies, this research explores a "middle-out" approach putting both the data and the stakeholders at the center of our interest. This work investigated the user's requirements and expectations with regards the middleware using KAOS goal-oriented modeling [13]. This investigation suggests the smartness of a city relies on interconnecting its various component systems, managing and coordinating a networked movement of voluminous data transmitted from data suppliers to data consumers. We argue that smart cities can be created through the formation of a global-scale knowledge network between cities' stakeholders (e.g. government agencies, non-government agencies and citizens). Although a rich pool of data can significantly improve an urban services, the high **volume**; the different **velocity** in which the data arrives, is stored, and retrieved; the disparate heterogeneous sources which originates a **variety** of

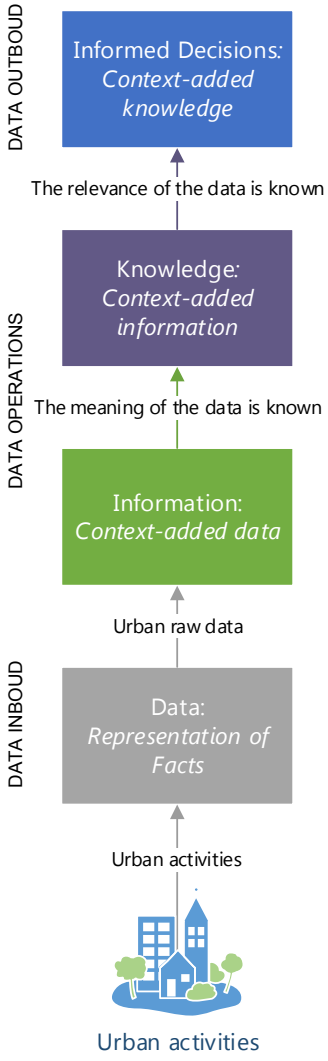

Figure 1. Urban Data Value Chain

different data formats; the **variability** in the meaning and context of the data; **vulnerability** in terms of security, trust and privacy of data; the lack in data **verification** resulting to poor data quality and redundant information; the issues surrounding **visibility** of fixes and mobile data sources (e.g. Sensors); and the difficult task of integrating the data to unlock **value** has posed a significant challenge in the integration of events that will trigger city services [8-12]. This problem will certainly become worse as the amount of data exchange increases in the system-of-system [1]. These data characteristics identified in this work are refereed as "the **8 v's** of the urban big data", and are the fundamental requirements of the smart city software architecture. In order to identify the data transformations necessary to bring urban data together using a suitable data model and successfully collect, organize, integrate, distribute and use of knowledge derived from urban big data, this paper introduces the urban data value chain [14].

The value chain framework provides means to understand the influence of data on cities with an approach for breaking down the sequence (chain) of business functions into the strategically relevant activities through which value is added to products and services. The proposed value chain enables understanding the activities and processes involved in unlocking value from the data that create competitive advantage, and manage those activities to create smart cities. The simplified primary structure in a smart city can be described in terms of three sectors: inbound logistics, operations, and outbound logistics. The simplified value chain is illustrated in Figure 1, and the three first sectors are defined as follows:

- **Inbound logistics**: involves all the activities involved in the creation and gathering of data. Urban activities triggers data creation in various systems, sensors and actuators in an urban environment. In this point data can be gathered using standard protocols and interfaces. It will require continuously data live stream acquisition from different data sources, providing accurate locating sensing without compromising the local resources (e.g. sensor battery).
- **Operations**: involves all the activities involved in enriching the raw data gathered in the inbound logistics. It includes data quality verification and provenance, privacy protection, semantic annotation and modelling in a standard and established format to provide uniform access to data.
- **Outbound Logistics**: concerns to both city- and global-wide delivery of data with added value. It involves integrating and analyzing data from both internal and external data repositories. The information that will pass through this point of the value chain can be transformed into knowledge which can be consumed by humans and machines to produce informed decisions.

Conclusions and Future Research

The overall objective of this research is to investigate new processes, strategies and mechanisms to create a rich urban data environment through the development of a middleware framework. The value chain analysis explicitly recognizes the interdependencies and profits cost efficiencies from the exploitation of linkages between value activities of a city, representing a powerful tool for strategic thinking to derive an effective software architectural design for smart cities. For instance, changes in the standard formats (value activity) may significantly influences the activities involved in operations and outbound logistics (value activity). Establishing new data standards without the platform having means to support it will affect the reliability of the platform. Another issue to consider is a service not be real-time available-to-promise/capable-to-promise and fulfill users' needs. Time-sensitive processes in a city, such as

fraudulent operation, traffic flows, infrastructure monitoring, emergencies or tragedies, needs high data/service availability in order to detect any problem before it actually happens. The veracity of data can either empower the city or lead to poor understanding, decisions and provision of services. Ensuring and maintaining data quality is a challenge that software integration solution will be required to address in order to provide quality assurance. Currently we are evaluating the proposed middleware framework considering several architectural designs and datasets to replicate the conditions of a city-wide distribution and consumption of data. We expect the proposed middleware framework will provide ways to users to integrate data from a variety of sectors and distribute it across different domains, value chains and stakeholders, thereby supporting intelligent decisions in the urban environment and the creation of valuable new businesses.

Acknowledgements

The authors would like to thank the Engineering and Physical Science Research Council (EPSRC) for their financial and technical support for this research work.

References

[1] Gann, D. M., Dodgson, M. & Bhardwaj, D. 2011. Physical-digital integration in city infrastructure. IBM J. Res. Dev., 55, 90-99

[2] Ishida, T., Aurigi, A. & Yasuoka, M. 2005. World digital cities: beyond heterogeneity. Proceedings of the Third international conference on Information Technologies for Social Capital: cross-Cultural Perspectives. Amsterdam, The Netherlands: Springer-Verlag.

[3] Graham, S. & Marvin, S. 2001a. Splintering Urbanism, Networked infrastructures, technological mobilities and the urban condition. Taylor & Francis.

[4] Eger, J. 1997. Cyberspace and Cyberplace: building the smart communities of tomorrow. San Diego Union-Tribune, Insight, 2.

[5] Zhu, D., Li, Y., Shi, J., Xu, Y. & Shen, W. 2009. A service-oriented city portal framework and collaborative development platform. Information Sciences, 179, 2606-2617.

[6] Yong Woo, L. & Seungwoo, R. 2010. U-city portal for smart ubiquitous middleware. In: Advanced Communication Technology (ICACT), 7-10.

[7] Hey, T. & Trefethen, A. 2003. The Data Deluge: An e-Science Perspective. Grid Computing. John Wiley & Sons.

[8] Zikopoulos, P. C., Eaton, C., Deroos, D., Deutsch, T. & Lapis, G. 2012. Understanding big data. Analytics for enterprise class hadoop and streaming data. New York: McGraw Hill.

[9] Buneman, P., Khanna, S. & Tan, W. C. 2001. Why and Where: A Characterization of Data Provenance. Proceedings of the 8th International Conference on Database Theory. Springer-Verlag.

[10] Reades, J., Calabrese, F. & Ratti, C. 2009. Eigenplaces: Analysing Cities using the Space-Time Structure of the Mobile Phone Network. Environment and Planning B: Planning and Design, 36, 824-836.

[11] Haslhofer, B. & Klas, W. 2010. A survey of techniques for achieving metadata interoperability. ACM Comput. Surv., 42, 1-37.

[12] Mcdaniel, P. & Mclaughlin, S. 2009. Security and Privacy Challenges in the Smart Grid. Security & Privacy, IEEE, 7, 75-77.

[13] Dardenne, A., van Lamsweerde, A. and Fickas, S. 1993. Goal Directed Requirements Acquisition, Science of Computer Programming, 20, 3–50.

[14] Porter, M. E. 1998. Competitive advantage: Creating and sustaining superior performance, Free press.