Published at the 7th IEEE International Conference on
Network of the Future, November 16-18, 2016

# InterSCity: Addressing Future Internet Research Challenges for Smart Cities

Daniel Macêdo Batista, Alfredo Goldman, Roberto Hirata Jr., Fabio Kon
Department of Computer Science
University of São Paulo
Email: {batista,gold,hirata,kon}@ime.usp.br

Fabio M. Costa
Instituto de Informática
Universidade Federal de Goiás
Email: fmc@inf.ufg.br

Markus Endler
Department of Informatics
PUC-Rio
Email: endler@inf.puc-rio.br

***Abstract*—The Future Internet will integrate large-scale systems constructed from the composition of thousands of distributed services, while interacting directly with the physical world via sensors and actuators, which compose the Internet of Things. This Future Internet will enable the realization of the Smart Cities vision, in which the urban infrastructure will be used to its fullest extent to offer a better quality of life for its citizens. Key to the efficient and effective realization of Smart Cities is the scientific and technological research covering the multiple layers that make up the Internet. This paper discusses the research challenges and initiatives related to Future Internet and Smart Cities in the scope of the InterSCity project. The challenges and initiatives are organized in three fronts: (1) Networking and High-Performance Distributed Computing; (2) Software Engineering for the Future Internet; and (3) Analysis and Mathematical Modeling for the Future Internet and Smart Cities. InterSCity aims at developing an integrated open-source platform containing all the major building blocks for the development of robust, integrated, sophisticated applications for the smart cities of the future.**

***Index Terms*—Smart Cities, Internet of Things, Future Internet, Big Data, Machine Learning, Mathematical Modeling.**

## I. Introduction

The term "Smart City" has gained traction in recent years; it has been used by researchers from multiple fields to refer to the use of modern approaches to use city resources in a more intelligent way. Some extrapolate the software context, focusing on social or business aspects. Regarding software systems, many authors define a Smart City as the integration of social, physical, and IT infrastructure to improve the quality of city services [1], [2]. Other authors focus on a set of Information and Communication Technology (ICT) tools used to create an integrated environment [2], [3], [4].

When evaluating the various definitions related to IT, integration is a frequently cited aspect and we consider the technologies associated with the Future Internet an essential aspect to guarantee such a integration. We are aligned with the vision that the city infrastructure must provide an integrated environment, facilitating the development, deployment, and operation of interoperable smart city applications from multiple domains. Based on that, in our view [5]:

> *"a Smart City is a city in which its social, business, and technological aspects are supported by Information and Communication Technologies to improve the experience of the citizen within the city. To achieve that, the city must provide public and private services that operate in an integrated, affordable, and sustainable way."*

Despite the existence of multiple smart city initiatives in different countries around the world [1], [6], [7], [8], [9], these deployments are often based on custom systems that are neither interoperable, nor portable across cities, extensible, or cost-effective [10]. To resolve these limitations, the research community must address key challenges in the areas of Networking and High-Performance Distributed Computing, Software Engineering, Machine Learning, Data Analysis, and Mathematical Modeling, among others. These challenges involve highly-specialized knowledge, requiring a deep understanding of different areas, but they also require close cooperation among specialists from multiple fields working in a collaborative way.

InterSCity[1] is a multidisciplinary project aiming at addressing the key research challenges around smart cities from a software infrastructure point of view. Its goal is to develop an integrated open-source platform containing all major building blocks for the development of robust, integrated, sophisticated applications for the smart cities of the future. This paper aims at motivating researchers and graduate students to engage in innovative research on Smart Cities, helping to advance the state-of-the-art.

The paper is organized as follows. Section II discusses in more detail the objectives of the InterScity project. Section III presents the research challenges that will need to be addressed to meet those objectives. Section IV presents our reference architecture for Smart City platforms, and Section V concludes the paper and considers the results that are expected from the InterSCity project.

## II. Requirements and Objectives

To better define InterSCity's objectives, we conducted an extensive literature review on Smart City Platforms [5]. After analyzing 47 projects, we derived the requirements for Smart City Software Platforms, which are divided into functional and non-functional. The most relevant functional requirements are



[1]http://interscity.org

[big] data management, application run-time support, wireless sensor network management, data processing and HPC, external/open data access, service management, advanced software engineering methods and tools, and city models. The most relevant non-functional requirements are interoperability, scalability, security, privacy, context awareness, runtime adaptation, extensibility, configurability, and energy awareness. More details about each requirement can be found in [5].

To meet these requirements and work towards our Smart City Software platform vision, the main objectives of the InterSCity project have been defined as follows.

1) The development of analytic mathematical models for the representation, study and analysis of large structured collections of objects in future Smart City environments. These models will be instrumental for the efficient management and planning of resource usage and deployment in smart cities, with the goal of optimizing performance, maximizing sustainability of systems and promoting savings of energy and other resources.
2) Design, implementation and evaluation of new software architectures for distributed, large-scale, self-configurable and self-adapting systems based on intelligent algorithms and machine learning techniques.
3) Design, implementation and evaluation of new communication protocols and mechanisms to process large amounts of multimedia data and streams (text, audio, image, video, etc.) with guaranteed quality of service (QoS).
4) Scheduling algorithms for large scale environments involving large numbers of tasks and resources in multiple clouds, mobile devices and networks of sensors and actuators.
5) Development and experimentation with new methods, techniques and software engineering tools and IDEs to support the development, testing and deployment of high-quality, complex, large-scale distributed systems for Smart Cities.
6) Application of the technologies developed within the project in real-world scenarios involving real smart city services, aiming at improving quality of life for citizens and providing solutions for more efficient management of large and mega cities. This will be achieved both via partnerships with city governments and through technology transfer to the public and private sectors, especially considering startups.

## III. Research Challenges

To face the scientific and technological challenges and provide innovative solutions to the problems around the Future Internet and Smart Cities, the InterSCity project will develop scientific and technological research covering the different layers that compose networked services and applications. The project will contribute to the development of the Future Internet and to the development of reusable technologies and methods for Smart Cities, with special focus on problems such as urban mobility in cities of the developing world. More specifically, as illustrated in Figure 1, research is being developed at three levels that overlap and interact with each other:

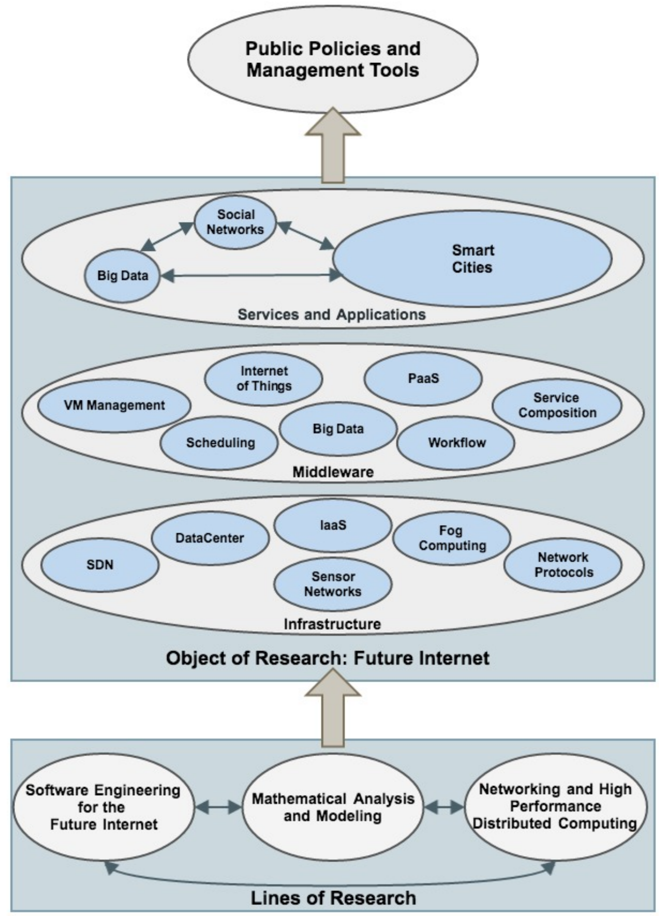


Figure 1. Technical-Scientific Structure of the InterSCity Project

- **Infrastructure** includes basic support for computer networks, protocols, cloud datacenters, software defined networks, and sensor networks.
- **Middleware** includes software systems for the support and management of virtual machines, service composition, communication, Big Data, Internet of Things, processes scheduling, workflows, data stream management, and energy management.
- **Services and applications** include social networks, collaborative networks, advanced Big Data mechanisms to process large volumes of multimedia streams, and specific support for Smart Cities.

Notably, infrastructure and middleware services will provide a basis for the development of high level integrated services and applications in a number of interrelated domains. This will ultimately enable strategic decision making and the planning of public policies, and will subsidize the development and operation of tools to manage the different everyday aspects of smart cities.

The objects of research shown in Figure 1 will be tackled from three different perspectives, as shown at the bottom of the figure. The respective research challenges and initiatives are discussed next.

### *A. Networking and High-Performance Distributed Computing*

With fast growing deployment of sensors and smart devices in all sectors of economy and our daily life, the Internet of Things (IoT) is becoming a reality [11]. This can be defined as a global network of heterogeneous devices and objects, addressable in a uniform way and interacting through standard communication protocols, thereby extending the ubiquitous connectivity already present in the current wireless Internet (3G/4G and WiFi) to all sorts of devices, including sensors and appliances with processing capacity and wireless interfaces (the so-called *Smart Objects*). IoT also encompasses the direct object-to-object or machine-to-machine (M2M) communication, without the intervention of people, and at global scale. According to some reliable projections, until 2020, the number of nodes/objects in the IoT will exceed 50 billion, encompassing applications in practically all sectors of the economy, such as manufacturing, commerce, healthcare, financial systems, transport, security, homes, and smart buildings [12].

To support large-scale M2M communication and the processing of large volumes of sensor data from *Smart Objects*, IoT systems count on cloud-based infrastructure services for mobility (known as *Mobile Cloud*). Thus, the majority of the proposed architectures for IoT consists of the following four general network layers/types: opportunistic/ad hoc networks that operate at low power and short ranges (*Edge Networks*); long-range networks between hubs and cloud servers (*Device-to-Cloud*); networks between servers of a data-center (*Intra-Cloud*); and networks between centers (*Inter-Cloud*), as shown in Figure 2.

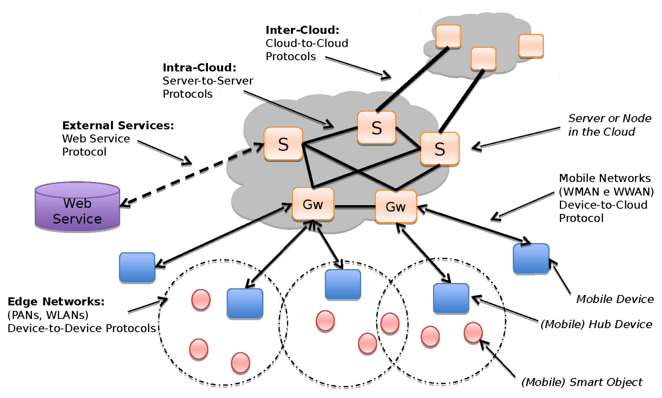


Figure 2. Network layers for IoT

Due to the increasing amount of Smart Objects and their geographical dispersion and heterogeneity, as well as the diversity of the networks involved, the IoT presents many challenges, mainly related to interoperability, discovery, selection, and access to sensors and actuators in Smart Objects, efficient communication at all levels, and the development and maintenance of applications and systems.

In particular, IoT poses major challenges related to the capacity to process continuous flows of sensor data generated by a large number of Smart Objects. This characterizes the *Data Stream Processing* paradigm, where data is volatile and is only relevant immediately after its generation, while the queries on this data are permanent, as they may need to be processed repeatedly over long periods of time. The efficient processing and the management of these flows of data are transversal concerns of the IoT that need to be investigated urgently, as they will affect the data models of future IoT systems.

Solutions for the majority of these challenges require the development of Middleware infrastructures for IoT [13], in addition to dealing with heterogeneity, interoperability, compensation for different capacities and volatility of nodes and networks, and the efficient processing of data flows. These middleware systems should also provide a set of abstractions (concepts and APIs) that assist in the project, execution, and maintenance of IoT applications.

There are also different models being successfully applied in distinct areas, such as weather forecast, traffic control, smart buildings, and energy and waste management. Many of those models are treated separately. However, in order to optimize the design, evolution, and operation of cities, it is essential to quantify, understand and formalize their interaction. Nowadays, examples of such sophisticated (composite) models already use powerful supercomputers, pointing at the need for further research on exascale computing.

Some techniques, tools and mechanisms being used to face these challenges are the SDDL (Scalable Data Distribution Layer) middleware [14], nsRFIDSim [15], the MapReduce programming model, workflow management systems, Complex Event Processing (CEP), emerging hardware patterns for Fog Computing, Software Defined Networking (SDN), Network Function Virtualization (NFV), and the FITS testbed (Future Internet Testbed with Security) [16].

### *B. Software Engineering for the Future Internet*

Building complex large-scale systems such as the ones needed to comply with the requirements presented in Section II will inevitably demand a disciplined approach based on Software Engineering. However, the current technologies and methodologies provided by existing practices in Software Engineering do not completely address the challenges involved in the development of software systems for the Future Internet and Smart Cities. This new scenario involves thousands of people developing systems that must be integrated in a complex manner, resulting in even more complex systems of systems [17]. This trend will lead to an explosive growth in the degree of distribution and heterogeneity of these systems, as well as in their scale, both in terms of number of users and number of machines. In this context, it will become increasingly common to create systems composed of humans, services and "things", which interact among themselves to perform distributed activities. The complexity of large-scale systems, combined with issues such as interoperability, heterogeneity, mobility, adaptability, security, and privacy constitute some of the challenges imposed by future Smart Cities.

Determining the techniques and methods that are suitable in this context is still an open issue. The emergence of a new development scenario such as described above causes new workflows to be created in an ad hoc way, many of them susceptible to machine and human errors, as well as to rework and integration failures. To avoid these undesirable effects, different areas of Software Engineering must be revisited and extended to adapt to the new challenges. This will require, in many cases, a complete reconception of the methods and the development of new techniques, approaches and tools [18]. To that end, the following elements are needed: (1) architectural patterns, software architectures and approaches to deal with the complexity of Future Internet systems, while considering the need for openness, flexibility, and extensibility; (2) new tools, methods and metrics to evaluate the internal and external quality of software for the Future Internet, making the reuse of code increasingly likely, reducing maintenance costs, enhancing usability, and enabling the evolution of software to incorporate new functionality and to adapt to changing contexts and needs; (3) new protocols, techniques and processes enabling the construction of robust, resilient systems with strong support for fault tolerance; and (4) new tools, IDEs, processes and guidelines to support collaborative development of large-scale systems and applications for future Smart Cities, with emphasis on creating and fostering distributed open source software communities and on the widespread use of this kind of software.

Some techniques, tools, and mechanisms currently being used to face these challenges are agile software development, experimental software engineering, mining of software repositories, discovery and management of intra- and inter-systems dependencies between artifacts and modules, analysis of sociotechnical networks, automated classification of software quality and threats, service-oriented architectures, and Web services composition.

### *C. Analysis and Mathematical Modeling for the Future Internet and Smart Cities*

The era of scarcity of real data, when researchers and practitioners would, most of the time, depend on data generated by simulation, seems to be coming to an end, except for a few specific areas. Currently, the big challenge is having large amounts of data to deal with. Particularly for smart cities, Big Data plays a fundamental role [19]. The following factors have caused this data explosion: fast development and widespread adoption of sensor technology (such as position and motion sensors, cameras, and a variety of electromagnetic and chemical sensors); production and availability of large amounts of content on the Internet; and the large number of computers and other kinds of smart devices connected to the Internet.

This abundance of data will help to understand several phenomena that affect the daily life, such as those involving knowledge of physical phenomena (in the environment, telecommunications, transport, etc.), knowledge of biological phenomena (such as new medical techniques), and knowledge of social phenomena (politics, economy, education, city management, etc.), with the potential to promote significant advances in the quality of life.

However, due to the nature of sensors, noise and missing data are a fact of life, demanding new mathematical methods and ideas to avoid making *big* mistakes and *big* bad decisions with Big Data. This raises the need for the continued evolution of storage and information retrieval techniques, as well as for advances in data sampling, missing data analysis and inputation, anomaly detection, statistical analysis, and high performance computing.

The analysis of medium to large datasets may still be feasible nowadays. However, it may not be practical in big data scenarios without statistical sampling techniques on very large datasets [20], [21] (due to time and space complexity). Recent advances on sampling techniques, combined with database technology and methods, mainly for storage and retrieval, will probably become promising research areas in databases.

On the other hand, the classical hypothesis test, typically used for small and non-representative samples, may cause bad generalizations on big datasets. Therefore, sampling representative data from big datasets will also be an important problem to be solved in the coming years to achieve good accuracy and robustness of the estimates.

The type of data analysis required tends to generate even more data, as it is based on mathematical transformations, whose intermediate results are generally stored during the analysis process. This will allow for further meta-analyses, namely the combination of results from several previous analyses.

In the context of Smart Cities, an example of a particular kind of data that will need intensive processing is video. It is expected that large amounts of videos (essentially, temporally ordered sets of images) of traffic and urban mobility will be made available in real time for analysis. Efficient techniques need to be developed in this area for the management and manipulation of raw and processed data.

The Future Internet requires performing pattern recognition on all sorts of media and data, in environments that are full of noise, imprecision and uncertainties, which in turn can be modeled by Fuzzy Sets (FS) and imprecise probability models. A recent technique that has proven to be very effective to lower the classification error employs FS to generate tailored kernels for SVM (Support Vector Machine) classifiers as a way to model this kind of uncertain data [22]. In addition, methods were created for discovering outliers and anomalous subsets of labeled or unlabeled data [23]. An application of these techniques is the study of bus ticketing data (e.g., data about users and bus lines, date, and time of day) to discover behavior patterns for different kinds of users, such as students, workers and occasional users. For example, the analysis of the data by gender and age (modeled as intervals) can reveal specific behavioral patterns about senior users.

There are other techniques, tools and mechanisms that

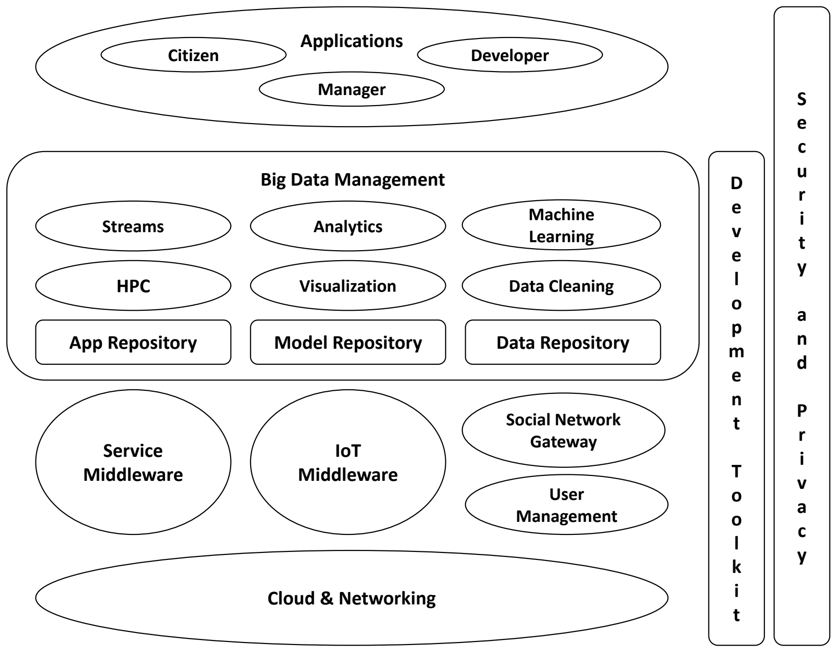


Figure 3. Reference Architecture for Smart City Platforms

have been used to deal with big datasets and that will open new areas of research, or improve existing ones. A few of these areas include: visualization techniques and algorithms, clustering algorithms, automatic design of image processing operators, computer-aided image segmentation, computational learning techniques, graph databases, and scalable algorithms for tractable fragments of the problem of probabilistic satisfiability (PSAT).

## IV. Reference Architecture

To guide our work, as part of a comprehensive literature review on Smart City Platforms that analyzed 23 different platforms [5], we derived a reference architecture for Future Smart City Software Platforms. Figure 3 presents the elements required for the development of a highly effective software platform to facilitate the construction of highly scalable, integrated Smart City applications.

The lowest level component of the reference architecture is *Cloud and Networking*, which is responsible for the management and communication of city network nodes in a scalable and extensible way. This component is in charge of identifying all the devices connected to the platform, including servers, sensors, actuators, and user devices.

On top of the Cloud and Networking infrastructure, the reference architecture includes the *IoT Middleware* and the *Service Middleware* components. The former is in charge of managing the city IoT network and enabling effective communication of the platform with user devices, city sensors, and actuators. The latter is in charge of managing the services that the platform provides to applications, performing operations such as publishing, enacting, monitoring, composing, and choreographing those services.

To provide better services to citizens, it is important for the platform to have access to some user data and preferences, which is the role of the *User Management* component. However, to ensure user privacy, this data must be properly protected, and permission to store it must be acquired from the user. Moreover, as the city platform will have many applications, it can be helpful to offer a single sign-on mechanism.

Social networks will play a major role in Smart Cities. They can be used to discover data about city conditions, and can be an efficient communication channel between the platform and its users. Therefore, it is important to allow the integration of the Smart City platform with existing social networks. This is the role of the *Social Network Gateway*.

The *Big Data Management* module manages all the data in the platform, including data collected from the city and data generated by the platform. To this extent, it has three repositories: (1) an *App Repository* to store applications, including their source/binary code, images, and associated documents; (2) a *Model Repository* to store city models, such as traffic models, sensor network models, data models, city maps, and energy distribution models; and (3) a *Data Repository* to store data collected from sensors, citizens, and applications.

Besides data storage, the Big Data Management module is also responsible for the processing of city data. There are two types of data processing that are suitable for different situations: *Stream processing* and *Analytics*, to perform real-time analytics and data-flow processing; and *Batch processing (HPC)*, to analyze large datasets. Moreover, this module must be capable of performing useful pre-processing tasks, such as data filtering, normalization, and transformation.

The Big Data module also has a *Machine Learning* component, which facilitates understanding of the city by automatically building behavior models of city processes and making predictions of city phenomena. Since a Smart City will produce enormous amounts of data, information *Visualization* is an important component, along with a *Data Cleaning* component for deleting data that is no longer needed and archiving old data on slower, high capacity stores.

The platform must also provide an application development toolkit, including tools such as an Integrated Development Environment (IDE), libraries, and frameworks, as well as a Smart City Simulator for debugging and experimenting with applications before actual deployment. Finally, a number of non-functional properties, notably security and privacy, must be supported across all layers.

We are currently working on the incremental implementation of this architecture, following an Agile Software Development method, and reusing open source tools, libraries, and frameworks as much as possible. Our initial prototype, which is based on Ruby on Rails, is available as open source software at *https://gitlab.com/groups/smart-city-software-platform* and more information about the project can be found at *http://interscity.org*.

## V. Conclusions and Expected Results

The realization of effective Smart City applications depends on research findings in the areas of Networking and High-Performance Distributed Computing, Software Engineering, and Analysis and Mathematical Modeling. The solutions being developed in these three areas in the InterSCity project shall generate reusable open source tools, libraries, frameworks, systems, and applications. This software will be available to the society and will serve as a basis for the elaboration of further projects focusing on the Future Internet, targeting the academic community, governments, established companies, and startups. In addition, we will seek to collect and create open datasets from different areas of a smart city to serve as a basis for experimentation within the area.

We expect that the technological and scientific research under development within the context of the InterSCity project will generate concrete results in the broader area of Smart Cities, as well as in specific applications related to areas such as urban mobility and intelligent transportation systems, public safety, pollution control and air quality, entertainment, and water, energy and waste management. Specifically, we shall develop a generic open source platform for data management, communication and processing of activities in Smart Cities, which can be used by corporations and governments for the enhancement of several aspects of urban life.

### *Acknowledgements*

This paper received contributions from the professors, graduate students, and industrial and governmental partners that compose the InterSCity research team, to whom we are sincerely grateful. The project will be funded by CNPq, proc. 465446/2014-0, and FAPESP.

## References

[1] A. Caragliu, C. Del Bo, and P. Nijkamp, "Smart cities in Europe," *Journal of Urban Technology*, vol. 18, no. 2, pp. 65–82, 2011.

[2] R. G. Hollands, "Will the real smart city please stand up? intelligent, progressive or entrepreneurial?" *City*, vol. 12, no. 3, pp. 303–320, 2008.

[3] D. Washburn, U. Sindhu, S. Balaouras, R. A. Dines, N. M. Hayes, and L. E. Nelson, "Helping CIOs understand 'smart city' initiatives," *Growth*, vol. 17, 2009.

[4] R. E. Hall, B. Bowerman, J. Braverman, J. Taylor, H. Todosow, and U. Von Wimmersperg, "The vision of a smart city," in *2nd International Life Extension Technology Workshop, Paris*, 2000.

[5] E. F. Z. Santana, A. P. Chaves, M. A. Gerosa, F. Kon, and D. S. Milojicic, "Software platforms for smart cities: Concepts, requirements, challenges, and a unified reference architecture," *Submitted for publication (available at http://arxiv.org/abs/1609.08089)*, 2016.

[6] C. Manville, G. Cochrane, J. Cave, J. Millard, J. K. Pederson, R. K. Thaarup, A. Liebe, M. Wissner, R. Massink, and B. Kotterink, "Mapping smart cities in the EU," RAND Corporation, Tech. Rep. EP-50486, 2014.

[7] P. Liu and Z. Peng, "China's smart city pilots: A progress report," *Computer*, vol. 47, no. 10, 2014.

[8] M. Fortes, V. Ferreira, G. Sotelo, A. Cabral, W. Correia, and O. Pacheco, "Deployment of smart metering in the Búzios City," in *Transmission & Distribution Conference and Exposition-Latin America (PES T&D-LA), 2014 IEEE PES*. IEEE, 2014, pp. 1–6.

[9] I. Janajreh, L. Su, and F. Alan, "Wind energy assessment: Masdar City case study," *Renewable Energy*, vol. 52, pp. 8–15, 2013.

[10] "International Technical Working Group on IoT-Enabled Smart City Framework," NIST, 2016, https://pages.nist.gov/smartcitiesarchitecture/. Accessed Sept. 27, 2016.

[11] H. Schaffers, N. Komninos, M. Pallot, B. Trousse, M. Nilsson, and A. Oliveira, "Smart cities and the future internet: Towards cooperation frameworks for open innovation," in *The Future Internet*, ser. LNCS. Springer Berlin Heidelberg, 2011, vol. 6656, pp. 431–446.

[12] L. Atzori, A. Iera, and G. Morabito, "The internet of things: A survey," *Computer networks*, vol. 54, no. 15, pp. 2787–2805, 2010.

[13] C. Perera, A. Zaslavsky, P. Christen, and D. Georgakopoulos, "Context aware computing for the internet of things: A survey," *Communications Surveys & Tutorials, IEEE*, vol. 16, no. 1, pp. 414–454, 2014.

[14] L. David, R. Vasconcelos, L. Alves, R. André, and M. Endler, "A DDS-based middleware for scalable tracking, communication and collaboration of mobile nodes," *Journal of Internet Services and Applications*, vol. 4, no. 1, pp. 1–15, 2013.

[15] R. P. B. Mota and D. M. Batista, "An NS-2 module for simulating passive RFID systems," in *Proceedings of the 11th IEEE/IFIP International Conference on Embedded and Ubiquitous Computing (EUC)*. IEEE, 2013, pp. 2263–2270.

[16] I. M. Moraes, D. M. Mattos, L. H. G. Ferraz, M. E. M. Campista, M. G. Rubinstein, L. H. M. Costa, M. D. de Amorim, P. B. Velloso, O. C. M. Duarte, and G. Pujolle, "FITS: A flexible virtual network testbed architecture," *Computer Networks*, vol. 63, no. 0, pp. 221 – 237, 2014, Special issue on Future Internet Testbeds – Part II.

[17] P. H. Feiler, K. Sullivan, K. C. Wallnau, R. P. Gabriel, J. B. Goodenough, R. C. Linger, T. A. Longstaff, R. Kazman, M. H. Klein, L. M. Northrop, and D. Schmidt, *Ultra-Large-Scale Systems: The Software Challenge of the Future*, 1st ed. Software Engineering Institute, 2006.

[18] I. Gorton, A. B. Bener, and A. Mockus, "Software engineering for big data systems," *IEEE Software*, vol. 33, no. 2, pp. 32–35, 2016.

[19] E. A. Nuaimi, H. A. Neyadi, N. Mohamed, and J. Al-Jaroodi, "Applications of big data to smart cities," *Journal of Internet Services and Applications*, vol. 6, no. 25, 2015.

[20] S. Kumar, M. Mohri, and A. Talwalkar, "Sampling methods for the nyström method," *Journal of Machine Learning Research*, vol. 13, no. Apr, pp. 981–1006, 2012.

[21] A. Kleiner, A. Talwalkar, P. Sarkar, and M. I. Jordan, "A scalable bootstrap for massive data," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 76, no. 4, pp. 795–816, 2014. [Online]. Available: http://dx.doi.org/10.1111/rssb.12050

[22] J. G. Díaz, R. H. Jr., and S. Canu, "Positive definite kernel functions on fuzzy sets," in *IEEE International Conference on Fuzzy Systems*, 2014.

[23] J. G. Díaz, S. Canu, and R. H. Jr., "Support measure data description for group anomaly detection," in *ODDx3 Workshop on Outlier Definition, Detection, and Description, at 21st ACM SIGKDD Intl. Conf. on Knowledge Discovery and Data Mining, Sydney, Australia*, 2015.