ODF best practices towards a GDL Smart City

Ramon O. Guardado-Medina, Victor M. Larios, Andere Salma H. Patarroyo. and Javier Larios, IEEE Members

IEEE Guadalajara Open Data Framework for Smart Cities Working Group

Abstract— The aim of this paper is to present the status of the Open Data Framework (ODF) around the world, describe various trends and demonstrate why ODF is important to the infrastructure of smart cities and, in particular, how a GDL Smart City should define a proper model. In the first part, this work discusses smart cities and the definition of Open Data, showing the most common standards used to build an infrastructure for data analysis. In the following part, we present a general overview of the World Bank and the different open data initiatives worldwide (including Europe, Africa, Asia, Oceania, America, and Latin America), concluding with a comment about Mexico and, more specifically, the Guadalajara Smart City.

Index Terms— Open data, Smart Cities, Open Data Framework.

1 Introduction

THIS document has the goal of introducing the concept of Open Data Framework (ODF) worldwide, in relation to the context of Smart Cities, Standards, Tools and Methodology. In order to understand the meaning of ODF, its basic concepts and its principles, we provide some examples that may help in understanding its usage. We include a review of different continent context levels from a government perspective to then enable the analysis of ODF in America, starting in North America with the US and then on through Latin America to focus on Mexico, and then present the Guadalajara ODF initiative for Smart Cities.

2 Context

2.1 Smart Cities

Cities everywhere are flourishing because new technologies make them even more valuable and effective as face-to-face gathering places.
Sensors, software, digital networks, and remote controls will automate the things we now operate manually. Where there is now a waste of resources in the city, there will be opportunities to achieve efficiency. Where there is volatility and risk, there will be predictions and early warnings for better quality of life [1].
Due to the novelty of the paradigm, it is hard to articulate a unique definition for the "Smart City" concept because there is not a clear understanding of it, although there is a quite general understanding of what issues are important in such an initiative. A Smart City can be seen as a system with livable conditions where development, usage and management of available resources are performed in an intelligent way in the pursuit of a sustainable environment; various definitions can be found in [2]. In order to achieve a Smart City solution, it is important to identify the primordial trends to invest in, which require a deeper look to develop sustainable and "optimal" solutions. Nevertheless, there is general agreement that the integration of technology (IT infrastructure) is considered a key factor for a smart solution [3].

"Smarter cities use the possibilities of new technologies, such as data collection and modeling and create new insights and help decision making for the city" [3].

2.2 Open Data

According to [4], The Open Definition gives full details on the requirements for 'open' data and content. Open data make up the building blocks of open knowledge. Open knowledge is what open data become when they are useful, usable and used. Open data have the potential to empower citizens, change how government works, and improve the delivery of public services while generating significant economic value [5].

- *R.O. Guardado is with the Department of Information Technologies, Guadalajara University, Periférico Norte N° 799, Núcleo Universitario Los Belenes, C.P. 45100, Zapopan, Jalisco, México. E-mail: ramon.o.guardado-medina@ieee.org .*
- *Víctor M. Larios Rosillo is a full Professor and the Director of the Smart Cities Innovation Center at CUCEA UDG, E-mail: victor.m.lariosrosillo@ieee.org.*
- *Salma H. Patarroyo is an IEEE Volunteer of the ODF Working Group, E-mail: salmap16@hotmail.com*
- *Javier Larios is a Professor and Researcher at the Smart Cities Innovation Center at CUCEA UDG, E-mail: Javier.larios@ieee.org.*

xxxx-xxxx/0x/$xx.00 © 20015 IEEE Published by the IEEE-CCD Smart Cities

The key features of openness are the following:
Availability and access: Data must be available as a whole and at no more than a reasonable reproduction cost, preferably by downloading over the Internet. The data must also be available in a convenient and modifiable form.
Reuse and redistribution: The data must be provided under terms that permit re-usage and redistribution, including intermixing with other datasets. The data must be machine-readable.
Universal participation: Everyone must be able to use, reuse and redistribute the data; there should be no discrimination against fields of endeavor or against persons or groups. For example, 'non-commercial' restrictions that would prevent 'commercial' usage, or restrictions of usage for certain purposes (e.g., only for education), are not allowed.
Primacy: Datasets should be primary source data, allowing users to verify that the information was collected properly and recorded accurately.
Machine readability: Data should be available in machine-friendly formats.
Timeliness: Information should be released as quickly as it is gathered. Real-time info updates would maximize the utility the public can obtain from this information [6].
Licensing: Data must be "explicitly licensed in a way that permits commercial and non-commercial use and re-use without restrictions" [7].

3 Standards

The standards have three branches: 1) file structures that include JSON, XML and CSV formats; 2) protocols over TCP/IP, such as CAP (Common Alert Protocol) simple version, RESTFUL (Representational state transfer), MQTT and XMPP as machine-to-machine (M2M) "Internet of Things" connectivity protocols; and 3) type of storage criteria as the flow of continuous data and periodic subset of data [8]. It is important to note that IEEE is working on a standard for IoT architectures (IEEE P2413) and has already developed standards for communication protocols among IoT sensor/actuator networks (IEEE 21451).

3.1 Tools and methodology aim to enable the physical infrastructure for storage repositories (Big Data or compatible), ensure sufficient network bandwidth for the access, provide a scalable architecture, and establish mirrors or backups of information in private or public data clouds. A web portal to provide access to the repository is considered in order to manage and support users, security, privacy, and the implementation of policies. The methodology must include the management and the scale of data repositories [8].

The following are several examples:

- Hive (AL, Apache): Data Warehouse System.
- Pig (AL, Apache): A platform for analyzing large data sets that consists of a high-level language for expressing data analysis programs coupled with infrastructure for evaluating these programs.
- Drill (AL, Apache): A distributed system for interactive analysis of large-scale datasets (inspired by Google's Dremel).
- Ambari (Apache): Aimed at making Hadoop management simpler by developing software for provisioning, managing, and monitoring Apache Hadoop clusters.
- ZooKeeper (DC): An effort to develop and maintain an open-source server that enables highly reliable distributed coordination
- StarCluster (CM, MIT)
- Mesos (CM, Apache): A cluster manager that provides efficient resource isolation and sharing across distributed applications or frameworks. It can run Hadoop, MPI, Hypertable, Spark, and other applications on a dynamically shared pool of nodes.
- Whirr (CM, Apache): A set of libraries for running cloud services.
- Helix (CM, LinkedIn): A generic cluster management framework used for the automatic management of partitioned, replicated and distributed resources hosted on a cluster of nodes.

In addition to such platform-envisioning flows of unstructured information, other proposals will be related to NoSQL databases explored by the IEEE Smart Cities Metrics Working Group at Guadalajara City. One of the most important questions is What is the best infrastructure to build in Guadalajara?, Guadalajara has an existing IEEE Working Group in Open Data Framework, which has been working with the Smart Cities Innovation Center at CUCEA UDG to investigate this topic. The goal is to propose the best model for the GDL Smart City to be tested in its Living Lab for Smart Cities in cooperation with the Innovation, Science and Technology Ministry of the Jalisco State government.

4 The State of Open Data

4.1 The World Bank Vision for Open Data

The World Bank primarily promotes the Open Government among governments all over the world, and for that to happen, there must be Open Data. In order to be open, data must be *technically open*. This means that data should be retrieved and meaningfully processed by a computer application. Data must be *legally open*. This is understood as being explicitly licensed in a way that permits commercial and non-commercial use and reuse without restrictions [7]. There are many benefits from opening data, such as economic benefits, improved public services and more transparent and accountable government [10].

The World Bank provides sources that make these statements countable. According to McKinsey&Company, $3 trillion to $5 trillion a year could be generated in additional value in seven sectors as a result of Open Data

[5].

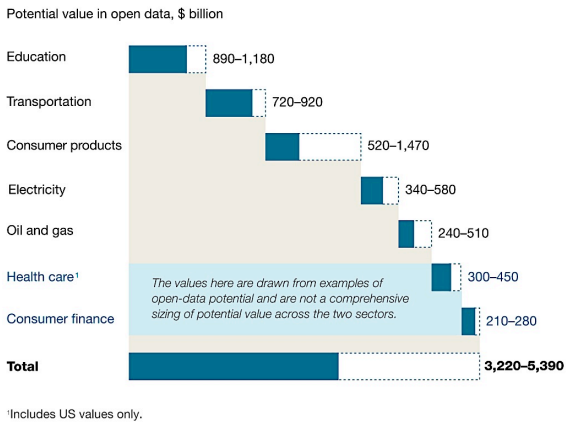


Fig 1. Potential Value in Open Data (McKinsey&Company)

Nevertheless, investments in technology and expertise are required to use the data effectively. Work has to be done by governments, companies, and consumers to craft policies that protect privacy and intellectual property [5].

The usage of open data has a large potential economic value, as it increases efficiency and brings new products, services and cost savings. As for the consumers, it brings convenience and better products as well as new channels for feedback in order to improve the quality of goods and services. By creating transparency and enabling experimentation, it supports or replaces human decision making while promoting innovative business models, products, and services [5].

4.2 European policies for Open Data to add value

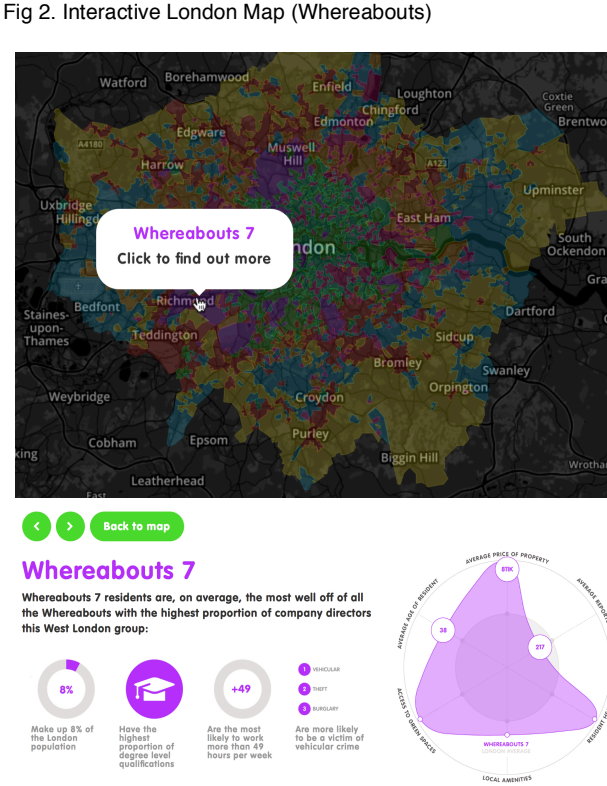
According to [9], The Commission's work in the area of Open Data is focused on generating value through re-use of a specific type of data: public sector information, sometimes also referred to as government data. That is all the information that public bodies produce, collect or pay for. Examples include geographical information, statistics, weather data, data from publicly funded research projects, and digitized books from libraries.

They support Open Data for 4 reasons:

- Public data have significant potential for re-use in new products and services.
- Addressing societal challenges: having more data openly available will help us discover new and innovative solutions.
- Achieving efficiency gains through sharing data inside and between public administrations.
- Fostering participation of citizens in political and social life and increasing the transparency of government.

For example, in the United Kingdom, there is an ongoing experiment by the Future Cities Catapult named *Whereabouts London*. Its main objective is to explore how Open Data can be used to improve future cities by grouping neighborhoods based on their ways of living rather than where they live. It operates under an open source license, allowing people to access to the information and process it for their own particular needs [10].
Figure 2 shows an interactive map of London, divided by its boroughs, and statistics about the people leaving there, including degree level qualifications, average price of properties, average age of residents, average reported crime rates, and access to green spaces. This is a clear example of some of the ideas that come up using Open Data and their utility in real life [10].

Fig 2. Interactive London Map (Whereabouts)



4.3 Africa applies Open Data to reduce poverty

The African Development Bank Group (AfDB) is committed to supporting statistical development in Africa as a sound basis for designing and managing effective development policies for reducing poverty on the continent. Reliable and timely data are critical to setting goals and targets as well as evaluating project impact. Reliable data constitute the single most convincing way of getting people involved in what their leaders and institutions are doing. Reliable data also help them to get involved in the development process, thus giving them a sense of ownership of the entire process. The AfDB has a large team of researchers who focus on the production of statistical data on economic and social situations. The data produced by the institution's statistics department constitute the background information in the Bank's flagship development publications. In addition to its own publication,

the AfDB also finances studies in collaboration with its partners [11] [12].

4.4 Asia is using Open Data as a Knowledge Economy Development Strategy

Currently, few governments in the region include Open Data in their national Information Communication Technology (ICT) or economic development strategies. In 2008, during the OECD Ministerial Meeting in Seoul, the opening of public sector information was discussed, as was the Future of the Internet Economy.

The Seoul Declaration states that the Internet economy has become a new source of growth, with the potential to boost the entire economy, to foster innovation, competitiveness and user participation, and to contribute effectively to the prosperity of society as a whole. Governments around the world were looking for ways to revive the economy after the 2008 financial crisis, and opening public sector information was considered a way to create value from data that can be freely used, reused and distributed by anyone. Since 2010, there have been thematic sessions and workshops focusing on public sector information and Open Data policy development in Asia at international and regional meetings of the Internet Governance Forum, with representatives from government, academia, business and civil society. At these meetings, questions about a more comprehensive overview on Open Data developments in Asia came up. This study aims to provide an overview of the current situation of Open Data policies and practices in North and South-East Asia. The countries included are, in alphabetical order, Cambodia, Hong Kong, Indonesia, Japan, South Korea, Laos, Malaysia, Myanmar, Philippines, Singapore, Taiwan, Thailand and Vietnam. The selected countries are part of the group categorized as "East Asia and the Pacific" in the widely used Knowledge Economy Index of the World Bank [13].

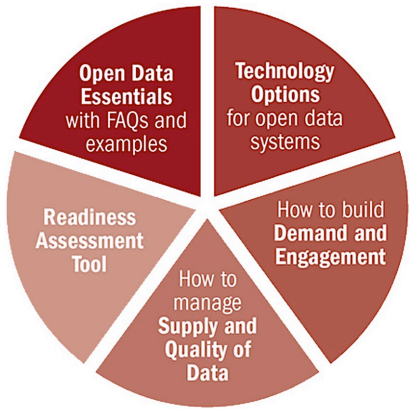


Fig 3. Open Data Readiness Dimensions (World Bank)

4.5 Oceania improving e-Government with Open Data

In this part of the world, Australia has important status because the Government Initiatives are interesting; for example, they have some programs for building infrastructures to analyze topics and to show important metrics. The website data.gov.au was created and is maintained by the Department of Finance, but the data sets have been created and provided by different government agencies, so the quality or timeliness of the data sets is not guaranteed. The site receives feedback to correct any problems identified, related to either functionality or data availability [14].

4.6 North America and U.S. Cities using Open Data as a resource for innovation

Chicago: Fostering civic hackers

This city generates interest in its Open Data by focusing time and attention on its civic hackers, the developers, analysts, and business people who mine public data for solutions to persistent problems and new services they can offer. Chicago has the most committed civic hacker communities in the world. Constant support of and investment in developers have led to partnerships and the development of businesses that contribute to the economy using Open Data [15].

In 2010, the city's portal came to life, and in 2012, Mayor Emanuel signed an Executive Order stating that every city agency had to contribute data to it. This portal offers nearly 600 datasets that are constantly updated. For example, the traffic dataset is refreshed every 10 minutes using the GPS data from its city buses.

The most interesting part of this portal is that between November 2012 and November 2013, data downloads from the Chicago open data portal grew nearly 200%. Usage means participation and innovation [15].

According to Tom Schenk, the City of Chicago Director of Analytics and Performance Management, "by giving developers longer histories of data and more frequent updates, the ability to see patterns and offer useful services is increased" [15].

New Orleans: Understanding Social Dynamics

For the five consecutive years between 2008 and 2012, the city had the highest murder rate of all cities in the country. In 2011, the murder rate was nearly 20% higher than that of the second highest city.

Intense analysis of data helped the Bloomberg Philanthropies Innovation Delivery Team to understand the situation and put in place some creative solutions that led to a 19% drop in murders by 2013. This was the lowest annual total since 1985 and the lowest rate since 1999 [16].

New York City: focused on mobility and workforce innovation

The last decade has seen an increasing realization among economists and policymakers that innovation has become the central economic growth driver and a key to improved standards of living [16]. In 2011, with Mayor Michael R. Bloomberg, the first Digital Roadmap was introduced with the goal of making this city the number one digital city in the U.S., in both the private and public sec-

tors [17].
According to NYC Digital, potential projects include developing new databases, web and mobile applications, and emergency-related informational maps using city data, analysis of impacted populations, and data sharing with other governments or utilities [18]. The foundation of a digital city is the connectivity of its people, and that is the reason why the first goal was to connect New Yorkers. By 2013, 99% of New Yorkers had residential access to high-speed broadband [17].
In addition to promoting Hackathons focusing on developing apps that improved job and economic mobility, lifelong learning, healthy living and sustainability for city residents, a linking of volunteer technologists with lifesaving digital projects came to life. Code Corps is the teamwork between the public and private sectors during emergency situations such as hurricanes and blizzards. One example application is the availability of accurate, actionable Open Data, such as evacuation zone maps and shelters. Entities such as Google, The New York Times and WNYC.org have collaborated [17].
As for the technology sector, NYC has increased considerably their investments and workforce, as shown in Fig. 4

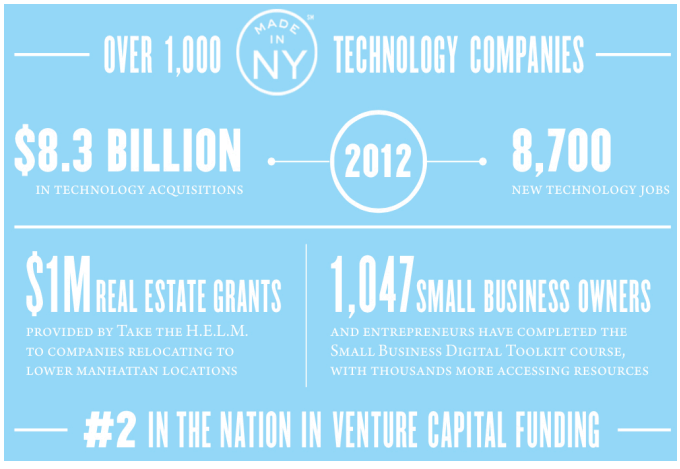


Fig 4. 2013 Roadmap (New York City's Digital Leadership)

4.7 Latin America using ODF as a tool for researchers

Based on this premise, the Open Data for Development in Latin America and the Caribbean project has been implemented in partnership with W3C Brazil and the European Commission for Latin America and the Caribbean (ECLAC), within the scope of the Observatory for the Information Society in Latin America and the Caribbean (OSILAC) and the International Development and Research Center of Canada (IDRC).

Open Data for the Development of Public Policies in Latin America and the Caribbean (OD4D) has 6 specific objectives:

To map out the main initiatives in Latin America and the Caribbean for structured economic, social and environmental data sharing and to design a methodological framework to examine the relationship between opening data and the quality of public policies.
To study and discuss alternative strategies to foster technical training in governmental agencies and observatories in the region, thus implementing open data repositories for the design, monitoring and assessment of public policies.

To support research networks in Latin America and the Caribbean in producing new information and creating innovative applications and services based on Open Data.
To examine the relationship between more inclusive economic development and the opening of data in key economic segments.
To raise awareness among the community of public policy makers, public servants and researchers of the potential of Open Data and appropriate strategies for its successful implementation.
To assess the potential of Open Data strategies in the design and implementation of public policies aimed at promoting economic development and social inclusion in Latin American countries and in the Caribbean. [19]

4.8 Mexico is starting to define an ODF model

The Open Data initiative has been built in an open and collaborative manner, in the same way that datos.gob.mx is a place for everyone. On this open source site, a forum has been created to receive proposals for how to continuously improve. Some areas of work in the country include the following:

- Maternal mortality.
- Registration IMSS (Social Security) insured.
- Poverty and social backwardness.
- Floods and registration during and after hurricanes.
- Emergency response in natural disasters.
- Encouraging Parents' Engagement in Public Education Reform.

4.8.1 Guadalajara is looking to test an ODF approach in their Living Labs network

According to [8], the goals of the Open Data Framework are to provide an open data platform, enabling people to build on Open Data from Guadalajara living labs. Such data sources must have national legislation regarding freedom of information, with standard reporting protocols. The creation of incentives for the Open Data Framework is needed in order to promote the data sharing for goods between business and citizens and to promote collaboration among universities, government, and industry. The framework of the Guadalajara Smart City Metrics resides over components that proceed from different layers, based on standards, tools, methodologies and stakeholders. These will be integrated into the context of local collaboration in working groups

5 Concluding Remarks and Perspectives

New technologies are emerging and playing a fundamental role in the building of Smart Cities. While we have learned about the usage of the Open Data Framework among continents and initiatives for keeping the freedom of information in Mexico, we found ODF to be the fundamental basis for Guadalajara's Smart City initiative.

The majority of this paper focuses on what is happening around the world with ODF and how communities are building their infrastructures to display the existing metrics, keeping in mind that such infrastructures may vary according to the specific needs of their users as well as the technical analysis of the same data.

Therefore, the Smart City initiative in the city of Gudalajara provides an opportunity to study different areas (Mobility, Health, Government, etc.) that provide information, and these are fed back by specialists to allow for better final decisions that bring improvements in different sectors.

Acknowledgments

The authors wish to thank the IEEE Guadalajara Section volunteers for their support in sharing its vision and taking care of the working group initiative. Special thanks are extended to the Universidad de Guadalajara and its PhD program in Information Technologies, which provided PhD Students, the advice of professors, facilities and access, as well as to its Smart Cities Innvoation Center at the CUCEA campus.

References

[1] M. Batty, K. W. Axhausen, F. Giannotti, A. Pozdnoukhov, A. Bazzani, M. Wachowicz, G. Ouzounis, and Y. Portugali, "Smart cities of the future," The European Physical Journal Special Topics, vol. 214, no. 1, pp. 481–518, Nov. 2012.

[2] Chourabi, H. et al, Understanding Smart Cities: An Integrative Framework, System Science (HICSS), IEEE, (2012) .

[3] C. Hafedh, N. Taewoo, W. Shawn, G. Ramon, M. Sehl, N. Karine, P. Theresa, and J. Hans, (2012). Understanding Smart Cities: An Integrative Framework. Proceedings of the 45th Hawaii International Conference on System Sciences.

[4] Open Knowledge https://okfn.org/opendata/

[5] McKinsey&Company, http://www.mckinsey.com/insights/business_technology/open_data_unlocking_innovation_and_performance_with_liquid_information?cid=other-eml-alt-mgi-mck-oth-2910/

[6] Sunlight Foundation, http://sunlightfoundation.com/policy/documents/ten-open-data-principles/

[7] The World Bank http://www.worldbank.org/en/about/what-we-do

[8] Vázquez-Castañeda, C, Estrada-Guzman, E, Towards the preparation of the Guadalajara's Smart City Metrics Structure, IEEE Guadalajara Metrics for Smart Cities Working, 2014

[9] European Commission http://ec.europa.eu/digital-agenda/en/open-data-0

[10] Whereabouts, http://whereaboutslondon.org/#/

[11] Open Data for Africa http://opendataforafrica.org/about/us

[12] Waltraut Ritter, An overview of Open Data Policies and Practices in 13 Countries, August 2014.

[13] data.gov.au/about/contact/

[14] Socrata, http://www.socrata.com/case-study/chicago-growing-open-data-economy/

[15] Bloomberg Philanthropies, http://www.bloomberg.org/program/government-innovation/innovation-delivery-teams/#problem

[16] Open Knowledge Foundation (Open data portal for Latin America), http://blog.okfn.org/2012/11/06/open-data-portal-for-latin-america/#sthash.QXCr8qSI.dpuf

[17] New York City's Digital Leadership: 2013 Roadmap, http://www.nyc.gov,

[18] Anthony M. Townsend, Smart Cities: Big Data, Civic Hackers, and the Quest for a New Utopia, W.W. Norton & Company 2013 First Edition.

[19] The World Bank, http://data.worldbank.org/about/open-government-data-toolkit/knowledge-repository

Guardado-Medina, Ramon O. is a PhD Student in Information Technologies at the University of Guadalajara campus CUCEA. He has a post-graduate degree and undergraduate degree in electronics systems from the Instituto Tecnológico de Ciudad Guzmán. He now works as a Volunteer in the IEEE Smart Cities Initiative Guadalajara Pilot and the Volunteer Leader of the Open Data Framework.

V.M. Larios has received his PhD and a DEA (French version of an MS program) in Computer Science at the Technological University of Compiegne, France and a BA in Electronics Engineering at the ITESO University in Guadalajara, Mexico. He is the Volunteer Leader of the IEEE Smart Cities Initiative Guadalajara Pilot http://smartcities.ieee.org, a Full Professor and the Director of the Smart Cities Innovation Center Information Systems Department CUCEA Universidad de Guadalajara, and a Volunteer and council member for the Guadalajara Ciudad Creativa Digital A.C. http://ccdguadalajara.com.

Patarroyo Andrade, Salma H. has a B.S. in International Business from the Universidad Anáhuac México Norte with an academic merit scholarship. She was an intern at PIDES: Social Innovation at the Open Government Department 2014. She is currently a volunteer at the IEEE Smart Cities Initiative Guadalajara and Open Data Framework Working Group.

J.Larios Has a bachelor's degree in Mass Media and Corporate Communication Studies. He started a career by developing skills in Public Affairs (Public Relations) by analyzing audiences' behaviors and their needs to achieve a better understanding between a focused audience and the business unit required. He is currently working as an Instructor at the Smart Cities Innovation Center CUCEA UDG.