Proceedings of 7th Transport Research Arena TRA 2018, April 16-19, 2018, Vienna, Austria

iKaaS Project: "Intelligent Knowledge-As-A-Service platform"

Sergio Fernández Balaguer[1*], Andrés Recio Martín[2] , Guadalupe Rodríguez Díaz[3], Patricia Cervigón[4], Consuelo de Garrastazu Díaz[5]

1. *Communication and Consultancy Department, EMT Madrid, C/Cerro de la Plata 4, 28007 Madrid, Spain*
2. *Tenchnology Department, EMT Madrid, c/Cerro de la Plata 4, 28007 Madrid, Spain*
3. *Atos Research & Innovation, C/ Albarracín 25, 28037 Madrid, Spain*
4. *Coordinadora Red Palinocam. D.G. de Salud Pública. Comunidad de Madrid, C/Ronda de Segovia 52, 28005 Madrid, Spain*
5. *Instituto de Salud Pública. Madrid Salud. Ayuntamiento de Madrid, Avda. Mediterráneo, 62 , 28007 Madrid, Spain*

Abstract

The project iKaaS, "intelligent Knowledge as a Service" is part of the European Program Horizon 2020 – EUJ2014 of the European Union. The aim is the development of innovative Cloud technologies to fulfil the new challenges surged with the large amount of data steaming from the Internet of Things, facing the requirements from business, industrial and social applications.

In this sense iKaaS aims to integrate these three technological fields to develop a new platform that leverage the advantages of IoT such as object virtualization and real time processing, Big Data analysis to build knowledge while offering their capabilities through a multi-cloud environment.

Platform features will be demonstrated through Smart City applications to promote the self-management of health and security of citizens and improving information systems and data analysis to achieve a more intelligent city, focused in areas like epidemiological surveillance. The results of the project are tested in three use cases, including the so-called "Environmental health service in Madrid City (Spain)", which includes transport and mobility.

Keywords: Cloud ; Internet of Things (IoT) ; Big Data

* *Tel.: +34 914068831 fax: +34 914068801*
E-mail address: sergio.fernandez@emtmadrid.es

1. Introduction

One of the biggest challenges in today's society is to drive technological progress to a place where everyone feels secure while taking advantage of opportunities. The large data volumes (Big Data) and the connection to the "Internet of Things" (IoT) are trends that will influence and affect the future development of cloud computing systems. Collecting information, processing and computing large volumes of data generated and delivered by and to highly distributed devices (sensors and actuators) creates new challenges for services and data that may also be hosted in different countries and must be executed beyond borders taking into account the different regulations. These requirements will have a strong impact on the underlying cloud infrastructure, which will require efficient management of these large volumes of globally distributed data.

At present, for example, municipalities are very concerned about the health and safety of their citizens. In most cities, there is a great amount of information coming from different sensors, but often it is not possible to make sense of all that information and use it properly. It would be very useful to be able to efficiently use this information and the knowledge that can be derived from it to help citizens, especially those with health problems, to avoid over-contaminated or dangerous areas well in advance the situation becomes risky for your health. It would also be desirable to assure them of any security or privacy concerns they may have about this process.

But not only may the administrations need this kind of services. For example they can also be demanded and used by large companies or even SMEs that provide services to clients that move around the world efficiently using federated Clouds services and this way, follow their clients in any part of the world.

2. iKaaS project

iKaaS, "Intelligent Knowledge-as-a-Service Platform" is a research and innovation project funded by the European Union within a specific call to promote Japanese-European cooperation. The iKaaS project involves 15 partners including European and Japanese hi-tech companies, universities and public entities. The project started in October 2015 and will end in September 2017.

The focus of the joint research is the integration of three technological fields (IoT, Big Data and Cloud Computing) to develop a platform (see Figure 1), which will allow to take advantage of the IoT power, such as virtualization and Real-time processing using large data analysis techniques, and at the same time will work on the generation of knowledge that will be provided to users through applications that will run in a Multi-Cloud environment.

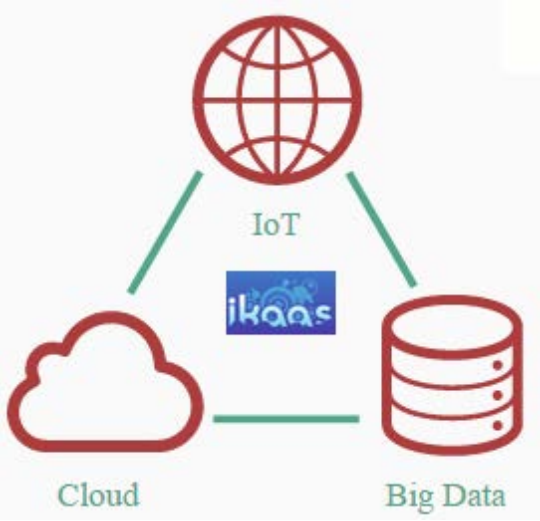

Fig. 1 iKaaS concept. Source: iKaaS Project

Thus, the iKaaS project is aimed at developing an "intelligent being" that preserves privacy and protects resources such as large volumes of data or an analysis engine built on top of a Multi-Cloud infrastructure, which will be fed with ubiquitous large-scale data collected from heterogeneous sensor networks and other data sources, including cyber-physical systems, portable sensors or social networks. An additional goal will be to provide this data and analysis engine (knowledge base) as building blocks for various heterogeneous environment-related applications, lifestyle recommendations, future urban planning, research and academic analysis, location services, etc.

The joint consideration of all the technologies poses great challenges with respect to the architecture that must support them in an effective way. Given the dispersed nature of the world of IoT devices, a system must be created that allows Cloud-based approaches for the management of this topological dispersion and distributed support for data storage and processing. It is also necessary to take into account the heterogeneous nature of the devices, so that, in order to become an ecosystem, a layer of "homogenization" must be introduced to clearly record and explain the operations that can be carried out, the operations that support each device and what control action are able to perform.

In the world of IoT, security and privacy are of the utmost importance because if appropriate measures are not taken, personal information or information intended for specific groups can be exposed to unauthorized entities. In addition, if we want to facilitate the transfer of data between different borders, it is necessary to coordinate the real privacy and data protection frameworks between the participating member countries. The iKaaS platform should include points of demarcation of the real responsibility of the holders and the transfer or receipt of resources to avoid such problems.

The quality of communication networks is a factor that also needs to be monitored and guaranteed. This can be done through a reputed service.

As mentioned above, IoT devices can generate large amounts of data (Big Data), so scalable and efficient forms for storage, retrieval and processing are needed. In addition to that, in order to reduce the overhead in data transmission, it is convenient to perform these operations as close to the data source as possible. In order to maximize the "coverage" and the distribution of resources, it has been thought to develop a Multi-Cloud environment, consisting of a Global Cloud and Local Clouds, for which specific security and privacy issues must be addressed.

The platform must also allow the easy introduction of new services, as well as the reuse of the components, thus allowing the generation of new knowledge. It is also important that the management of it and of the services is easy and efficient, without requiring frequent and manual reconfigurations. This poses a system of automatic management and orchestration of services.

Taking into account all these requirements, the architecture of the iKasS platform has been designed and the functional components that must conform it have been determined and can be seen in Figure 2:

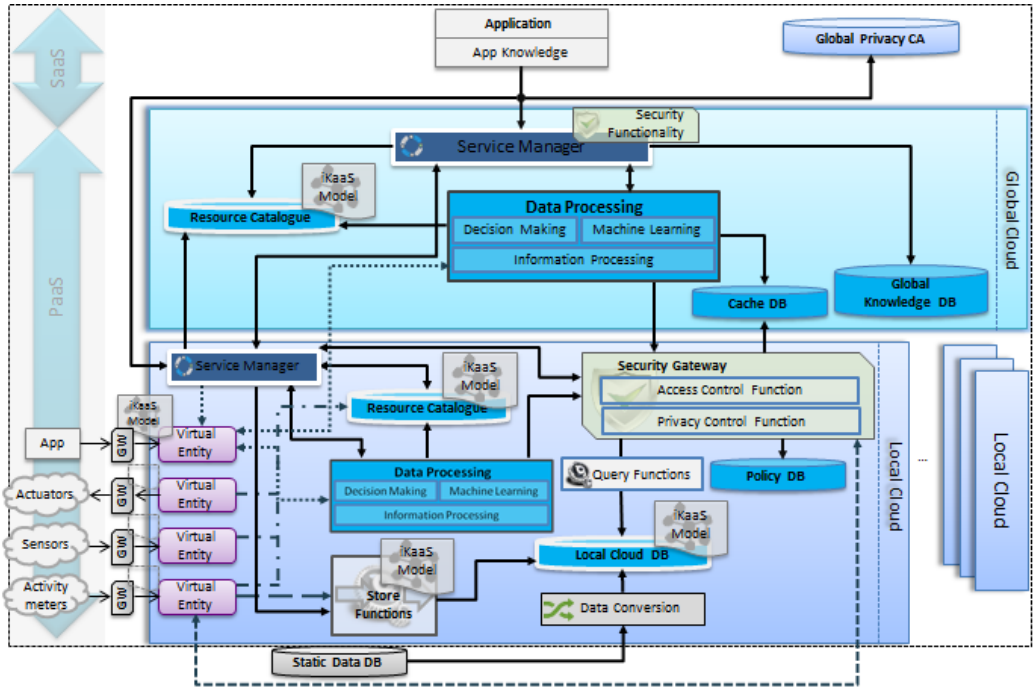

Fig. 2 iKaaS architecture. Source: iKaaS Project

Ultimately, the iKaaS platform will enable many players to become part of an ecosystem of shared services and resources in the cloud, targeting both providers and consumers. The features of the platform will be demonstrated through Smart City applications that promote self-management of the health and safety of citizens, as well as the improvement of a data analysis information system for a smarter life in the city, focusing on areas such as epidemiological surveillance, as it is the following Madrid use case which will be explain more in detail:

3. Madrid Environmental Health Service

3.1. Context

Madrid is the capital and largest city of Spain, with a population of 3.2 million inhabitants, with an average age of 43.5 years. The city has two distinct areas: a central urban area with a high concentration of people, services and traffic, and a suburban area with larger green spaces, separated from the central area by the M-30 ring road.

Although Madrid is not an industrial city, its importance as a large logistics, administrative and service center raises problems related to the quality of air that must be treated accordingly. As a big city, it is especially sensitive to environmental problems and therefore one of the biggest challenges and priorities for the next years is to ensure proper management of the environmental impact on the health of the citizens of Madrid, preserving the environment.

Pollution, particulate matter and pollen are one of the main factors affecting the health of citizens today. These particles pose a major threat to respiratory health by increasing exposure to respiratory diseases or directly aggravating respiratory diseases. Older people and children are the most affected strata in our industrialized society, but they are not the only ones. Healthy people are also affected in their daily life, which increases the risk of developing respiratory diseases. In the city of Madrid, for example, winter and spring are the seasons in which the need for emergency assistance for asthma increases. The seasonal nature of this type of respiratory disease is the result of high pollen levels of existing species but can also be combined with a high level of air pollutants.

3.2. iKaaS in the Madrid use case

The Environmental Health Service of the city of Madrid, through the iKaaS project, will focus on pollutants and pollen and on how to improve current mechanisms to measure them and improve the accuracy of information to be provided to citizens, with the aim to reduce health-related effects, especially in the allergic and hypersensitive population. This will provide real-time information on pollution levels and pollen and its distribution in the city of

Madrid. It is also intended to offer other services such as complementary aerobiological data, pollen maps in the air and health advice for citizens.

By deploying a new network of environmental sensors using the infrastructure - electronic panels and buses - of the Municipal Transport Company (EMT), it will be possible to increase the number of geographical points where measurements are taken with a limited number of sensors.

These new sensors, together with the environmental network already deployed by the Community of Madrid (Regional Government) and the City Council, will feed the iKaaS system; This will make it possible to provide real-time information to citizens and through the ability to analyze large volumes of iKaaS data and build a new source of knowledge about pollution, pollen "and" citizens "for" citizens.

"And", since it will also be possible to control how people move around the city through the use of, for example, smartphones, and to know if people take into account the new information provided.

"For", as it will provide more information that will be more accurate than the one which is currently provided.

To achieve this goal, milestones have been set:

- Increase the current number of city environmental sensors including new mobile sensors in the system located on board public buses.
- Develop a local platform that will integrate all environmental devices and sensors.
- Collect and store all historical data and values in real time in a local cloud.
- Define a protocol for the exchange of data in the city of Madrid for mobility, traffic and environmental data.
- Define the semantic data layers and public data models in order to produce public and open knowledge.
- Create an open API connector for the exchange of such data.

Figure 3 shows an overview of the approach within the iKaaS project of the Environmental Health Service for the city of Madrid.

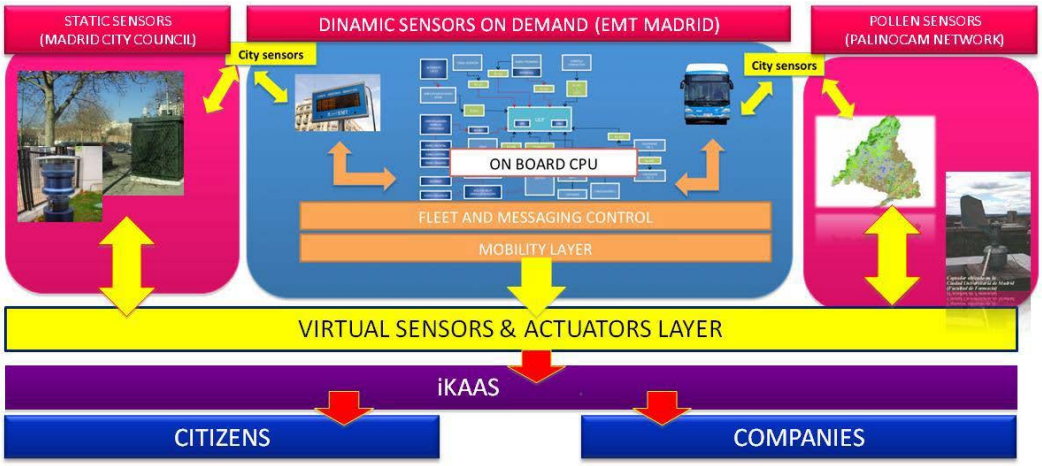

Fig. 3 General overview of the iKaaS Environmental Health Service for the City of Madrid. Source: iKaaS Project. Madrid Use Case

3.3. Specific services for the Madrid Environmental Health Service

Within the iKaaS project, the following specific services have been defined:

- **User warnings about pollution and pollen**

This service will provide warnings via mobile devices to subscribed users to the service regarding contamination levels and / or pollen. These warnings will be based on each specific user profile.

The system will also provide warnings based on the geographic position of the user, warning them when approaching an area with high pollution concentration and / or pollen particles.

- **Warnings on EMT electronic panels**

This service will provide real-time warnings to citizens about contamination levels in the geographic position in which the panel is placed. These warnings will be made taking into account the existing regulation in the health care system.

- **Pollution and pollen maps**

This service will provide maps of the city with distribution of pollution and pollen.

This service will provide real-time maps and provide predictions for contamination and distribution of pollen particles.

- **Recommendations**

This service will provide recommendations regarding the most appropriate areas and dates / times to move around the city to avoid exposure to environmental risk factors.

This service will also provide information about health care centers, pharmacies, etc. based on the geographical position of the user.

3.4. Data Processing and Knowledge Generation

As part of the project's knowledge generation effort, a forecast and prediction system is being developed with the aim of using the space-time characteristics of nitrogen dioxide concentrations in the Madrid fixed monitoring stations to predict future values throughout the city.

This system uses geostatistical techniques for the data modelling recorded in the fixed monitoring stations, which together with the State Meteorological Agency's (AEMET) numerical predictions of pollution and predictions provided by the ECMWF (European Center for Medium- Range Weather Forecasting) will allow the production of pollution forecasts for the entire city. These forecasts can be used to define the healthiest route in terms of air quality between two points. Data from the different available sources will be recovered, cleaned and standardized.

In the task of recovering and cleaning the data from the sensors, the Complex Event Processing (IoE Lab, (2015)) developed by Atos is especially important, allowing the real-time processing and analysis of the data or events collected by the sensors and discriminate them according to complex patterns or events that are constructed manually by context experts. For example, one of the problems when working with sensors is that measurements are taken asynchronously and not all sensors provide values at the same time and with the same frequency. This makes an agreement necessary to decide when an entry can be considered to be available. This fact is directly related to the nature of the data; the ambient temperature changes more slowly in time than the wind speed, therefore, a wind sensor measurement speed will produce new values more frequently than a temperature measurement, and this will have a lower frequency than, for instance, a sensor to control the heart rate of a person.

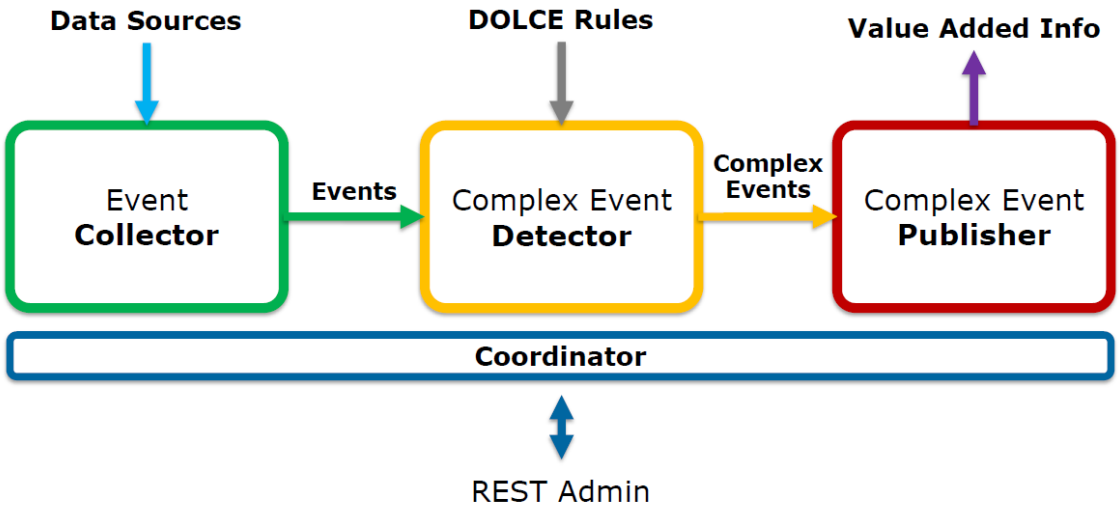

Fig. 4 CEP functional architecture. Source: IoE Lab., Atos Research and Innovation

To make the prediction in places where there are no stations and therefore no real observations, a space-time approach is being used. The interpolation will be carried out through a combination of regression, automatic models and kriging.

This combination can be done in several ways: real data and numerical predictions can be combined in a regression or machine learning model and then apply kriging to interpolate data or model residues, or certain variants of kriging that allow the inclusion of exogenous variables, such as regression-kriging or universal kriging. The best way to choose which one to use in the prediction service is being studied.

Once the model is set up and trained using historical records, it will be possible to obtain future values for nitrogen dioxide in different forms. On the one hand, the forecasts for the locations of the fixed stations will be produced, which allows to interact with the air quality protocol of the City of Madrid. On the other hand, a grid will be obtained that covers the whole city with the prognostic values that will allow integrating the predictions with the geographic information system used in iKaaS.

Finally, a third product is the possibility of checking the predictions against the measurements obtained by the mobile pollution sensors installed in urban buses. Being able to calculate the forecasts for the time and place where these mobile sensors measure real values will give an idea of how reliable the calculated values are, therefore, a valuable tool for evaluating the system.

4. Conclusions

iKaaS has designed an open, adaptable and secure as a “whole” service framework for incorporating the deployment of optimal services including migration and parallelization, as well as distributed object management, associated storage, processing and communication of Data, aimed at allowing the reuse of applications across different domains and platforms, as well as knowledge as a service.

iKaaS has also developed mechanisms that allow the notion of knowledge-as-service (“Kaas”, Knowledge-as-a-service).

iKaaS takes security, privacy and trust into account in a holistic and integrated way as part of the “whole”, i.e. as a service architecture. This provides more security to owners of ICT objects. The platform allows sensitive data to be processed within the domain of the data provider (Local Cloud), so only indirect information (i.e. knowledge) is distributed.

Finally, iKaaS strives to interact with related standardization bodies in order to contribute to the harmonization of international standardization efforts as well as to bring the developed technologies to market.

Thanks to iKaaS, the city of Madrid improves the way of providing environmental health information including the best suitable routes for those citizens and users with health limitations, optimizing their mobility within the city.

5. References

Cervigón Morales, P.; Bardón Cristóbal, M.A.; Fernández Recio, A.; Garrastazu, C.; Mateo M. (2015) "Horizonte 2020. iKaaS, EU_JP. Plataforma de servicios y autogestión de la salud de los ciudadanos". Revista de Salud Ambiental, volumen 15, pp. 67.

Fernández, A (2016). "Smart Solutions for cities session (MaaS): Open platform for transport data in Madrid" Proceedings of ECOMM 2016.

Seira Hidano, Shinsaku Kiyomoto, Yosuke Murakami, Panagiotis Vlacheas, Klaus Moessner (2015). "Design of a Security Gateway for iKaaS Platform". Cloud Computing. Springer International Publishing, pp. 323-333.

Kazuo Hashimoto, Keiji Yamada and other (2014): "iKaaS Data Modelling: A Data Model for Community Services and Environmental Monitoring in Smart City". Autonomic Computing (ICAC), 2015 IEEE International Conference on. pp 301-316.

Masahiro Hiji, Yuichi Hashi (2014): "CESP- Community Environment Platform of community-care 3D representation of Urban Space for Community Services". IEICE Technical Report, Vol 114, nº 286, pp. 11-15 Chiangmai, ThailandC.

Panagiotis Vlacheas, Vera Stavroulaki, Andreas Georgakopoulos, Dimitrios Kelaidonis, Abdur Rahim, Klaus Moessner, Yutaka Miyake, Shinsaku Kiyomoto, Keiji Yamada, Kazuo Hashimoto (2015). "An overview and main benefits of an Intelligent Knowledge-as-a-Service Platform". Proceedings of WWRF 34th Meeting.

Satoru Izumi, Kazuyoshi Matsumoto, Tetsuro Sato, Naoki Nakamura, Takuo Suganuma, and

Norio Shiratori, (2014). "Design of Green-oriented Management Information Base (G-MIB)", The Special Interest Group Technical Reports of Information Processing Society, Vol.2012, CDS 4-11 No.9, pp.1-6.

V. Foteinos, D.Kelaidonis, G.Poulios, P.Vlacheas, V.Stavroulaki, P.Demestichas, (2013) "Cognitive Management for the Internet of Things: A Framework for Enabling Autonomous Applications," in Vehicular Technology Magazine, IEEE, vol.8, no.4, pp.90,99.

Vlacheas, P., R. Giaffreda, V. Stavroulaki, D. Kelaidonis, A. Somov, V. Foteinos, G. Poulios, A. R.Biswas, and P. Demestichas, (2013) "Enabling smart cities through a cognitive management framework for the internet of things", IEEE Communications Magazine, vol. 51, no. 6, pp. 102-111.

Vlacheas Panagiotis, Stavroulki V. and other (2015): "An overview and main benefits of an intelligent Knowledge-as-a-Service Platform". Proceedings of WWRF 34th Meeting.

Wahlber, I.; Jämsä, T. (2015) "D5.1 Definition of the Application Scenarios". Specific deliverable of iKaaS project.

Wei Wang, Fang Yao, Suparna De, Klaus Moessner, Zhili Sun (2015) "A ranking method for sensor services based on estimation of service access cost". Information Sciences, Elsevier, Volume 319, Pages 1–17.

Web site iKaaS project: http://ikaas.com/

Twitter iKaaS project: @iKaasH2020