AGENCIA ESTATAL DE INVESTIGACIÓN - Convocatorias 2018
Proyectos de I+D de GENERACIÓN DE CONOCIMIENTO y Proyectos de I+D+i RETOS INVESTIGACIÓN

IP 1 (Nombre y apeliidos): Antonio Vallecillo Moreno

IP 2 (Nombre y apellidos): Carlos Canal Velasco

TÍTULO DEL PROYECTO (ACRÓNIMO): Digital Avatars: un framework para aplicaciones de computación social colaborativa (CoSCA)

TITLE OF THE PROJECT (ACRONYM): Digital Avatars: A Framework for Collaborative Social Computing Applications (CoSCA)

1. PROPUESTA CIENTÍFICA - *SCIENTIFIC PROPOSAL*

1.1. Context and research problem

Social computing (SC) is the area of Information Technology (IT) that deals with the interrelationships between social behavior and computer systems. At its beginning, it was mainly focused on the processing and analysis of social information, i.e., the fingerprint that people produce when performing their daily activities. Currently, the term has progressively evolved towards a wider meaning that includes the use of computer systems to support any type of social behavior, and where humans become the main protagonists—not only as beneficiaries but also as active players.

Although any aspect of SC is based on people, their activities, their habits, and their interactions with other people or with technology, a large part of SC activities are currently performed in contexts which are quite different and separated from the people who generate the data and contents, from the rest of the information sources, and from the applications' beneficiaries. For example, most current SC systems are based on server-centric applications whereby contents created by distributed users are transferred to cloud servers (Figure 1a). These servers employ this information to perform all the computations, including those of interest to the users, but also others that may permit obtaining benefits to the server owners different from those originally intended by the users (even misusing in some cases the users' information). This is also the model commonly used in the majority of applications for smart cities, where Social Computing is rapidly gaining relevance.

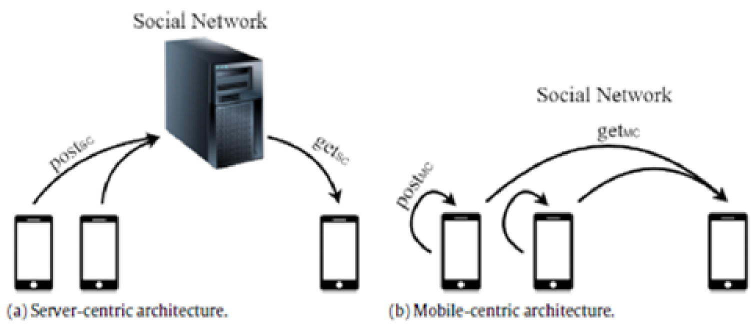


Figure 1. Client-server vs. P2P Models for mobile-based applications (from [BGV+17]).

In this project, we will explore additional models for SC, different from those client-server systems. In particular, we advocate the adoption of collaborative peer-to-peer models (Figure 1b) based on mobile devices, e.g., smartphones or tablets, as the main components of the system architecture. Specifically, we will focus on *mobile-based collaborative social computing* applications (CoSCA).

In our view, these collaborative computing models enable the empowerment of users, allowing them to take control of the information and contents they generate, and how all that information is accessed and exploited in a secure manner by third parties. In this way, this proposal defends a paradigm of SC focused on people and not on companies or institutions, and where users interact among themselves, collaborating in order to reach a common goal. This paradigm is known as People as a Service (PeaaS) [GMB+14]. PeaaS provides a conceptual model for application development focused on the mobile phone as a representative and interface to its owner.

In SC contexts, smartphones have already been proposed as an unrivaled candidate to host the *virtual persona* of their owners, able to reflect a social image of themselves that evolves through their interaction with the device [ROA09]. The progressive increase of the processing, storage and communication capacities of smartphones have enabled new architectural models in the development of mobile applications [AWM14]. Architectures based on P2P models gradually acquire a greater presence in fields such as social networks [WYJ+14], multimedia streaming [XJZ+15], or recommendation systems [YH13]. These architectures grant mobile devices a more important role in the system, instead of considering them as simple and passive clients of cloud-based services. Nevertheless, these mobile-based architectures pose very interesting challenges for developing software applications with them, due to the inherent limitations of the underlying technologies and devices, e.g., the limited memory and battery lifetime of smartphones.

Similarly, Edge Computing or Fog Computing also try to bring the cloud closer to mobile devices [AA16, SW14] as a means to reduce the necessary bandwidth when processing the collected data, and also to generate results closer to the data sources. These architectures that effectively combine both approaches provide a perfect alternative to centralized systems, eliminating most of the issues mentioned above, while reducing some of the problems inherent to plain smartphone-based computing mentioned above.

Measurement uncertainty is also an intrinsic property of any physical system in which this mobile-based computing takes place. The uncertainty can be due to different causes, such as unreliable data sources and communication networks, lack of precision in positioning systems and other measurement instruments, inaccurate data coming from faulty sensors or imprecise estimations, or the inability to determine whether a particular event has actually happened or not. The explicit representation and proper management of measurement uncertainty is a crucial issue that needs to be taken into account when developing the kinds of application we are concerned with in this project. This includes not only the (static) representation of the uncertainty, but also how it evolves and is propagated through the system operations, and considered when making decisions.

In our previous project DADIoT, we exploited that smartphones are able to collect information about the activities carried out by their owners. This information was then analyzed and organized in order to elaborate *virtual profiles* of their users with meaningful information about their habits and routines. Data was acquired using mobile phones, and data mining techniques permit extracting relevant information from them [LRU14, BGC+16]. Patterns and routines of behavior were inferred by analyzing this information, hence acquiring a first-hand knowledge of users' habits and preferences.

However, this information is decontextualized and misses some important aspects. For instance, the virtual profile of a user may tell us the places that she often visits, but it does not know what she does there, the transportation methods used for travelling to these places, or how many people are there at the same time. Hence, this kind of virtual profile is not enough to leverage effective collaborative social computing applications. We also need to make use of additional sources of information that will permit contextualizing these profiles. For example, virtual profiles can be significantly enriched by using the information provided by external sources. This includes information coming from the IoT devices the users interact with during their daily activities, streams of open data currently available in smart cities, or information obtained from other users' profiles. The advantages of combining all these sources of information are threefold. First, users will be able to make the most of their current contexts and environment. Second, virtual profiles will be enriched with a more elaborated knowledge about the users and their interactions with their environments, thus making them not only simple containers of people information but also behavioral engines of their environment/context relationships. Finally, IoT devices and smart city services could also learn from the users' interactions and adjust their operations to the owner's preferences. In our previous project Smart-FI [NCD+18] we have initially explored how the data coming from open data sets provided by smart cities, would enable relating the virtual profile of a user to the real-word context in which she moves. In this project we propose the use of complex event processing (CEP) techniques to deal with this kind of information

In this project proposal, we also introduce the concept of a ***Digital Avatar*** (DA). The digital avatar of a user extends the virtual profile with the ability to not only record information about the user, but also to interact with the environment and with the DAs of other users, and to modify the virtual profile accordingly. The DA will also be in charge of responding, on behalf of the user, to requests coming from other avatars or from other external applications. In these interactions, the DA will be responsible for ensuring the levels of privacy and security dictated by its owner. In short, a DA will serve as a proxy for its owner in the context of collaborative computing applications. It will also be able to take into account some of the uncertainties that happen during its operations in real environments.

Finally, we will show how this model of collaborative social media enables the development of further kinds of applications, in particular those able to understand the behavior of groups of users, beyond individuals. These applications can be very helpful for users to respect and adapt to the social behavior of their current environment.

1.2. A framework for developing mobile-based Collaborative Social Computing Applications using Digital Avatars

This project targets both the definition and development of users' DAs, and the specification, modeling, analysis and implementation of Collaborative Social Computing Applications (CoSCA) that can be built with them.

- **Digital Avatars.** DAs comprise both structure and behavior. The structure determines the information relevant both to the owner and to the applications she participates in, and will be persistently stored in the user's mobile phone. The behavior dictates how the DA reacts to external requests and stimuli (e.g., external events, environmental conditions, or context changes) and will be specified in terms of rules—similar to Event-Condition-Action (ECA) or Complex Event Processing (CEP) rules. Such rules can also be parameterized using values that can be changed, and that form part of the information stored and maintained in the avatar. The set of rules that define the behavior of a DA can be modified during its lifetime, for example when the DA is engaged in a new SC application, or when the user decides to incorporate a new behavior. Analogously, rule parameters can be automatically updated by the applications, for example to adapt that DA behavior to the user's observed practices and current preferences.
- **Mobile-based CoSCA Applications**. The use of Digital Avatars enables the development of different kinds of CoSCA applications, each one exploiting different features and exhibiting separate characteristics. We have identified at least three representative kinds of application scenarios that our project will support, depending on who directly benefits from it: individuals, collaborative groups, and third parties (e.g., companies or public administrations that develop applications for their customers or citizens). The first scenario consists in an application that exploits an individual DA and its interactions with its owner and its environment, to provide services to its user, with limited interaction with other DAs. The second scenario corresponds to a prototypical collaborative social computing application that is developed using DAs interacting among themselves to achieve a common goal. The last one involves a third-party system running on the cloud that uses the services and aggregates information provided by several DAs. Representative case studies of these kinds of CoSCA applications are later described in Section 1.5.

In this project we will devise and develop an architectural framework for the development of mobile-based collaborative social computing applications. This framework will provide a common architecture for these kinds of applications, as well as the required infrastructure and appropriate services for their implementation, deployment, analysis and execution.

The framework will make use of **mobile phone** infrastructures and development technologies as the basic supporting platforms. In addition, **Complex Event Processing** (CEP) will be employed to process the streams of data coming from heterogeneous sources; **mobile phone databases** to store the virtual profiles; **pattern languages** to represent the rules that determine the behavior of DAs, and **query languages** to access both the information stored in the DA and the rules governing its behavior.

Model-based Software Engineering (MBSE) [BCW17] will be used in the project for the definition of the high-level languages (through metamodels), the specification of the framework and its architecture, and the design and analysis of the collaborative social computing applications. The use of MBSE will also enable the semi-automated generation of the final code of the applications (both for the cloud/fog servers and mobile devices) using model transformations.

Another essential aspect of this project is the incorporation of **uncertainty** into the information managed by the applications, and into the patterns that define the behavioral rules of the DAs. The representation and management of uncertainty is required because of the inherent lack of precision in the measurement instruments that manage real data (e.g., GPS sensors or local clocks); lack of confidence in the data sources (e.g., defective sensors or unreliable networks), or incomplete or erroneous assumptions about the environment in which these applications operate.

As a research project, we also plan to study the **analysis and verification** of the applications being developed, as well as some of their properties, which are critical in the domain of SC. In particular, we will address the analysis of their performance, scalability, privacy, security, and reliability properties. We plan to use different techniques, including static and dynamic ones, depending on the property of interest. Simulation methods, and specifically statistical model checking, will play a key role in the project, given its appropriateness for the analysis of complex systems with uncertainty.

Relationship to the Spanish Scientific Research and Innovation Plan 2017-2020. Through the results of this project we aim at leveraging the introduction of novel and innovative techniques and technologies into the realm of social computing, which have been using rather traditional (software) engineering methods until recently. That introduction of novel IT techniques in these domains is fully aligned with the challenges of the Spanish Scientific Research and Innovation Plan for the period 2017-2020. In particular, **Challenge 7** (*Digital Economy, Society and Culture*) includes two specific strategic priorities to which this proposal would provide clear contributions:

> *7.i. "Internet of Things and their applications"*
> *7.iv. "Development, innovation and adoption of solutions and technologies bound to: (i) …cloud computing… ; (ii) …massive processing technologies of data…"*

1.3. Overall Goals and Objectives

This project aims at producing contributions at different levels.

- At the **conceptual level**, the project will provide methods and tools for the development of distributed collaborative social computing applications, as well as the definition of the supporting software architecture and the set of common services that the applications will use. Other scientific contributions of the project will be the definition of languages for the specification of virtual profiles and the information they store (both structural information, e.g., the user data; and behavioral information, i.e., the rules that define the user patterns of conduct); languages for querying and modifying such information; and script languages for interacting with the DAs, so that their services can be effectively used from other DAs or from server-based applications. The internal architecture of the DA's and skeleton implementations for them will be generated.
- At the **platform level**, the project will contribute in several ways. First, to the extension of CEP systems to deal with information structured not only as streams, but also as graphs of interconnected data through many different kinds of relationships. Second, to the representation, management and propagation of measurement uncertainty in the data handled by the application and in the users' behavior. Third, to the appropriate management of the information that mobile phones have to process, both transient (e.g.,

streaming events) and persistent (user personal data, information about her habits and behavioral patterns, etc.).

- At the **analysis level**, the project will provide methods and tools for the analysis of several aspects of the applications that are developed using this project's framework and libraries. In particular, we are interested in using quantitative techniques for the analysis of the performance, reliability and privacy of collaborative SC applications using statistical model checking.
- At the **engineering level**, the project will have to address the stringent requirements that the use of mobile phones impose on the applications and on the technologies they can use, in terms of memory, battery lifetime, network bandwidth, communications reliability, etc. These requirements will limit the choices of technologies currently available and will force us to look for adaptations and improvements that enable their use in the context of this project.

Given that these goals are too wide to be tackled in general, in this project we will focus on certain application scenarios, which are representative systems of the kinds application our framework will support, and for which our group has already some experience. These applications are the ones described below in Section 1.5, and will serve to illustrate the scope and implications of this project, to validate our proposal, and to prototype our ideas.

The ***specific objectives*** derived from the above overall goals are the following:

[O1] To develop an architectural framework and a set of common services to support the effective development of collaborative social computing applications, including the internal architecture of the DAs and skeleton implementations for them.

[O2] To define the appropriate modeling languages for the specification of virtual profiles and the information they store; for querying and modifying such information; as well as script languages for interacting with the DAs.

[O3] To provide extensions of CEP systems to deal with information structured not only as streams, but also as graphs of interconnected data, in contexts of reduced memory and low computing and network resources

[O4] To provide mechanisms and tools for the representation, management and propagation of uncertainty by the social computing applications.

[O5] To develop methods and tools for the analysis of the performance, privacy and reliability of SC applications.

[O6] To develop exemplar applications using the project that can serve to show-case the proposal and validate its results.

1.4. Supporting concepts and technologies

This section briefly describes the main concepts and supporting technologies that will be used in this project proposal.

1.4.1. Social Mining

Social mining is the process of representing, analyzing, and extracting patterns and tendencies from social media data. Its objective is to generate detailed profiles of individual users as well as of social groups [Rus11]. However, the sociological information currently available is scattered among the many existing social networks. Each of them focuses on a specific field of social activity, and thus they contain information that is partial and often contradictory, while none of them permits the representation of the full identity of their users [GA05]. For this reason, a number of researchers have explored alternative computational models and architectures for social mining, in particular those centered in mobile devices [RMM+12]. Currently, several research groups work on how to collect contextual information on smartphones' users in order to build more comprehensive virtual profiles (see [HSK09], [BCF+12], [MSS13] or [IJR16] for a review of the literature on this subject).

From these individual virtual profiles, collective sociological profiles can be developed, inferring habits, routines and tendencies of groups of people. These sociological profiles are built aggregating individual data.

1.4.2. Smart cities: the role of people

Local institutions from Smart Cities are currently offering open data collected from different sources, including real-time information about air quality and pollution, availability and schedules of city buses and other public transportation methods, etc. Their goal is to enable a more efficient management of the resources and services offered by the city, as well as improving the quality of life of its citizens [NP11]. This is the case, for instance, of the Málaga Town Hall, which is actively working on implementing these applications. However, providing this information is not enough to make a city smart [Coh12]. A global and more integrated point of view should be adopted, including a greater implication from people [MCC+14]. In this way, citizens become one of the main pillars in the development of smart cities.

To achieve this goal, we need to obtain information on how citizens use the services offered by the city, and about their user experience. This will enable a better planning of city infrastructures, adapting them to the context of the people that use them. In previous works (e.g., [MMG15]) we have proposed the already mentioned PeaaS model for IoT environments. It allows smartphones to collect and analyze information about their owners, in order to obtain information on their habits and routines, and to use this information for interacting with nearby IoT devices. In this project we plan to use this P2P collaborative model for the creation of a collaborative network of information, instead of the server-based approach we previously used for smart cities in [NCD+18]. This information will be offered as a service to third parties, including city authorities. The goal is to implement a shift on who is the owner of the information, which in our model belongs to the individuals generating it, instead of the applications making use of them.

1.4.3. Complex Event Processing (CEP)

CEP [EN10] is a form of Information Processing [CM12] whose goal is the definition and detection of situations of interest, from the analysis of low-level event notifications. CEP systems analyze multiple streams of simple events to detect occurrences of complex events that represent the high-level situations of interest, using declarative rules that define the complex events in terms of patterns of (simple or complex) events, their contents, and temporal relations. Event patterns are defined using Event Processing Languages (EPL), such as Esper, which provide mechanisms to query the streams of events.

The use of CEP to develop the DA is novel to this project, and it may greatly simplify the method used in our previous approach. Moreover, this project addresses two research problems with respect to CEP. The first one is the use of CEP engines in mobile devices, due to their limited memory resources. Note that some of the applications of this project require running CEP systems in combined environments (cloud and mobile devices) but others are solely based on mobile devices. The second requirement is the need to process, query and analyze information which, in our case, is not structured as sequences of events, but as graphs of interconnected data that evolve over time. Our group has started to work on this second problem, for which some results already exist [BBT+18].

1.4.4. Uncertainty

Uncertainty is the quality or state that involves imperfect and/or unknown information. It applies to predictions of future events, estimations, physical measurements, or unknown properties of a system. *Measurement uncertainty* is a kind of uncertainty that refers to the inability to know with complete precision the value of a quantity, an intrinsic aspect of any physical setting. Measurement uncertainty can be due to different causes, such as unreliable data sources and communication networks; tolerance in the measurement of the physical elements values; estimates due to the lack of accurate knowledge about certain parameters, or the inability to determine whether a particular event has actually happened or not. The explicit representation and management of measurement uncertainty is a crucial issue in any faithful model of a given physical system.

We have started working on the representation of measurement uncertainty in software models [BBM+18, BMB+18] and in CEP systems [MBB+18], using a probabilistic approach (instead of employing fuzzy logic or possibility theory). We plan to extend these results so that measurement uncertainty can be considered in the information managed by the DAs and in the patterns that determine their behavior, too, with the goal of enriching the models and the applications, which can now represent and process information in a more realistic way. Note that we will focus on measurement information (also known as Epistemic uncertainty) and in the confidence that we have on the data we handle, and on the rules that determine the behavior of the DAs (Aleatory uncertainty). Other kinds of uncertainties, such as application design uncertainties, are not part of this project.

1.4.5. Model-Based Software Engineering

Model-based Software Engineering (MBSE) [BCW17] is an approach to system design, architecting, analysis, development, operation and maintenance of systems in which models and model transformations play the primary roles. Models are used to represent systems and to capture the aspects of interest at the right level of abstraction, while model transformations are used to manipulate models in order to extract information out of them, convert them into other models (or into code), or analyze them.

MBSE technologies also provide support for manipulating models, such as querying, transforming, merging, relating and analyzing (including the execution of) models. MBSE approaches explicitly support modularity, separation of concerns and automated generation of major system artifacts from models (for example, test cases and code), using modeling languages to represent both the systems and the operations performed on them.

1.4.6. Analysis techniques

The high-level architectonical models of SC applications may be used for their formal analysis. Given the different alternatives for the location of computational nodes (e.g., server-centric vs. mobile-centric) plus the novelty of the different alternatives for the design of collaborative applications, we are particularly interested in the analysis of their privacy, performance and reliability properties. The use of probabilistic models for the specification of the DAs leads us to consider the use of probabilistic model checkers for the analysis of such properties. In particular, we plan to provide analysis mechanisms to be capable to determine whether a specific service (attending to their QoS requirements and the environment capabilities, i.e. mobile device resources, network performance, etc.) could be hosted in the mobile device or it would need to have part of their functionality residing in cloud/fog servers.

Among the variety of model checkers available, we plan to use VeStA, specifically one of its re-implementations PVeStA or MultiVeStA, based on simulations to perform statistical model checking. Other alternatives, such as PRISM or MRMC may also be considered. Although the formalisms supported by these tools is a priori less expressive, the probabilistic computation tree logic (PCTL) supported by PRISM is very attractive, and MRMC presents very good performance metrics (in terms of both computation time and memory).

We have an extensive experience in the development and use of formal tools. Specifically, we have used simulation-based analysis techniques for CEP models in [MBB+18], we have developed a CEP-based mediation platform for dynamic verification of service-oriented things [BCG+14], and used PVeStA for the quantitative analysis of e-Motions and BPMN behavioral models in [DMA16, DRS18].

1.5. Application Scenarios

To illustrate the scope and implications of this project, and to validate the project results, a series of scenarios have been selected, covering different aspects of the kinds of applications that our framework will support. Each one exhibits different features and specific characteristics. The motivation and most relevant features of each scenario are highlighted below.

1.5.1. Scenario 1: Elderly people in rural areas

María is an elderly person living in a small village in a rural and almost deserted area. After her husband died, and with her children living away, María started feeling lonely and considered moving to a nursing home. However, she finally decided to stay in her lifelong house while she was able to help herself properly. After all, she is completely autonomous, and only needs her sugar levels and blood pressure to be checked frequently, since she now suffers from diabetes.

In order to avoid recurrent and cumbersome trips to the primary healthcare center, which is located in another town, the regional health services have provided María with some smart gadgets that can be connected to the smartphone she normally uses to speak with her children, and to text with her grandson. A glucometer and a blood pressure monitor send their data by Bluetooth Low Energy (BLE) to her phone. The records of these measurements are stored there, together with relevant information about her daily activities, such as patterns of movements inside her house and within the village, and phone usage records (calls, app execution). Bluetooth is also used to detect proximity to other smartphones, particularly to those of her contacts, which enables detecting patterns of visits and social relationships between María and her friends, neighbors and relatives. In this way, both the healthcare center and María's children can be aware of her health conditions, well-being, and even of her mood. For example, abnormal sugar levels in her blood, or the fact that she is not visiting her neighbor one afternoon, as she always does on Tuesdays, can be easily detected. In case something seems to be wrong, the smartphone could ask María whether everything is ok, or raise an alert to the appropriate contact person or to the emergency services.

In this scenario, the smartphone plays a central role in capturing, processing and storing information about its owner and the readings from the sensors and devices that monitor Maria's context: GPS readings, movements detected by the phone accelerometer, BLE signals, usage of mobile apps, and IoT healthcare devices. The rules for processing all this information and detecting that any significant event has likely occurred are also stored in the phone, and managed by the DA. Note the importance of dealing with confidence levels (i.e., aleatory uncertainty) when making decisions, given that all sensed data is subjected to deviations and potential measurement errors (a sensor may stop working properly for a short period of time), as well as María's variations in their regular habits due to uncertain environmental conditions (e.g., bad weather) or unexpected situations (a surprise visit from a distant friend) that may cause changes in her daily habits but do not represent any challenge to her health.

1.5.2. Scenario 2: Collaborative transportation

Ana and Pedro are two university students that live in the same town, although they have never met. Every day, they travel separately to the same university, which is located in a city a few miles away from their hometown.

The DAs in their smartphones have been collecting information about their trips, including their frequency, typical schedules and routes. To detect the use of specific transportation means, their movements are cross-checked against different open data sets, including those provided by the city council, the regional bus company, or open map initiatives such as openstreetmap.org. Using this information, Pedro's phone can easily detect that he is using a particular suburban bus line, which at very regular times takes him to the university campus and returns him home. In contrast, Ana's smartphone infers that she is normally travelling by car, as well as the route she uses. The DA of her phone can ask her if she drives her own vehicle, and whether she would be willing to accept other passengers on her daily trips to the university.

Today, Pedro has finished his lessons earlier than usual and the next bus leaving directly from the campus is due in two hours. Thus, he decides to use his DA to launch a query in order to find alternative transportation means. From the answers received from those DAs that predict a similar trip by car this afternoon, Pedro select Ana's, since she is expected to return soon, and they both seem to live close to each other. Then, their DAs will ask each other if they are willing to make the trip together and share travelling expenses.

This scenario shows how several users' DAs can cooperate in achieving a common goal, such as a collaborative transportation. Habits of transportation would be autonomously inferred by the smartphones, and the transportation method and routes would be detected by crossing information against open data sources. Moreover, behavioral rules in the DAs govern both the collaboration between avatars and the interaction with their owners. There is no need of a third party centralizing information about trips offered and requested by users, and the system would solve this transportation problem only with direct collaboration between them. Data security is ensured since only the information that phone owners are willing to share with that particular application or group of users is disclosed.

1.5.3. Scenario 3: Theaters audience profile

The managers of a network of theaters in large cities are interested in characterizing the profile and behavioral habits of the people that attend the shows. In particular, they interested in the following information:

- Place of residence. The theater managers want to know if spectators come mainly from the same city in which a show is performing, or they also travel from other suburban areas, or even from more distant locations. This information will help planning the shows advertising campaigns, and where the publicity should be placed.
- Means of transportation. How does the audience travel to the see these shows? Do they mainly use their cars, or prefer public transportation? The answer to these questions would help planning public park slots and public transportation lines, allowing to predict their use according to the schedule of the shows programmed in the different theaters of the network.
- Collocated habits. What other activities do spectators perform before or after attending the show? Do they go back home right after the show? Do they have dinner in a nearby restaurant? By characterizing these kinds of habits, managers would be able to provide combined offers and other discounts to specific target audiences.

In order to address these and other similar questions, it won't be necessary to carry out expensive surveys, or to develop a specific mobile app for the network of theaters — most users are normally reluctant to download and install these kinds of apps in their smartphones. It won't even be necessary to conduct a general survey in order to identify the potential target users and to ask them about their habits. It will be enough asking the DAs in a given area whether their owners attend regularly to any of the theaters of the network. Only the DAs of the appropriate users would react to that request, provided of course that they are allowed by their owners to give away this information. That set of users will then be asked about their place of residence, places visited before or after the show, itineraries, habits, etc. All these queries can be directly managed by the DAs, without requiring the intervention of their owners. Obviously, theater managers may encourage users to grant them permission to ask for this information by offering discounts or any other benefits. The information provided by each DA can be aggregated for their analysis, building collective sociological profiles.

In this third scenario we also see how the information stored in the DAs can be offered as a service to third parties, and how the producers of the information can get benefits from its exploitation. Moreover, appropriate confidence levels can be assigned to these collective sociological profiles, to enable performing the corresponding risk analysis of the business decisions to be made. This scenario also aims at showing that the proposed framework is general enough to successfully design and develop server-based cloud applications, in addition to peer-to-peer systems.

1.6. Methodology

Our project will develop an architectural framework to provide all the necessary infrastructure and common services to design, develop, analyze and deploy the applications described in these three scenarios. The project work plan, schedule and methodology are described in this section.

1.6.1. Project Workplan

The project plan has been broken down into four technical work packages (WP1-WP4), plus one more (WP0) dedicated to project management, and other devoted to dissemination and exploitation (WP5). The description of these work packages follows, including goals, tasks, responsible persons, and deliverables for each of them.

WP 0: Project Management

Goals: This WP is responsible for coordinating the joint efforts of the researchers during the execution of the project, ensuring the smooth progress of the work plan by monitoring its progress towards milestones and deliverables. Periodical monitoring of each WP tasks advance towards the objectives will be carried out in the context of this WP. Also risk and quality assessment and the edition and presentation of the interim and final reports of the projects are assigned to this WP.

Tasks
- Task 0.1. Project progress monitoring: This task will be devoted to ensure that the project progress adequately, following the established methodology, and taking counter measures in the case of issues not considered in the project plan.

WP Leader: A. Vallecillo, C. Canal (Project PIs)

Deliverables:
- D0.1 Interim activity report (M12)
- D0.2 Interim activity report (M24)
- D0.3 Final project activity and management report (M36)

WP 1: Architectural framework

Goals: This WP is aligned with project objective O1, and will be devoted to develop an architectural framework and a set of common services to support the effective development of collaborative social computing applications, including the internal architecture of the DAs and skeleton implementations for them.

Tasks
- Task 1.1. Core concepts definition and requirement analysis: This task will be aimed to define the features that DAs (both at the structure and at the behavioral level) have to present in order to fulfil the requirements obtained from the typical use cases derived from the scenarios considered in WP4.
- Task 1.2. APIs to access and modify the information and behavior of DAs: This task focuses on defining an API to provide a common way to invoke services capable to deal with the information stored in a given DA, and to apply (but also modify) the behaviour associated to that DA.
- Task 1.3. Prototype implementation of basic services: An implementation will be provided in order to effectively use a minimum functionality to deal with DAs.

WP Leader: E. Pimentel

Deliverables:
- D1.1 Architectural description of DAs (M6)
- D1.2 API specification for service-oriented manipulation of DAs (M12)
- D1.3 Prototype implementation of a basic set of services (M24)

WP 2: Modeling languages

Goals: This WP will design and develop the appropriate modeling languages for the specification of virtual profiles and the information they store; for querying and modifying such information; as well as script languages for interacting with the DAs. It corresponds to the goals stated in objectives O2 (tasks 2.1 and 2.2), O3 (task 2.3) and O4 (task 2.4).

Tasks
- Task 2.1. Information modeling: This task will be devoted to the design and development of the appropriate modeling languages for the specification of virtual profiles and the information they need to store.
- Task 2.2. Interacting with DAs: This task focuses on the design of modelling languages for querying DAs and scripting languages for interacting with the DAs.
- Task 2.3. Data management: Extensions to CEP systems to deal with information structured not only as streams, but also as graphs of interconnected data, in contexts on reduced memory and low computing and network resources.
- Task 2.4 Uncertainty representation and management: mechanisms and tools for the representation, management and propagation of measurement uncertainty by the social computing applications, including both epistemic and aleatory uncertainty.

WP Leader: A. Vallecillo

Deliverables:

- D2.1 Information models for storing and managing DA structure and behavior (M12)
- D2.2 Query languages for interacting with DA (M12)
- D2.3 Extension to CEP systems (M24)
- D2.4 Libraries for modeling and implementing uncertainty in DAs (M12)

WP 3: Analyses

Goals: This WP will be in charge of defining the kinds of analyses that will be supported by our proposal, as well as the appropriate methods and tools for performing them. Given specifications of the architecture of an SC application and of its components (including services offered, resources used and interactions), a deployment plan, and models of the devices used (including smartphones and cloud/fog servers) and the provided infrastructure, our tools will allow the study of its performance, reliability and privacy. These properties are different in nature, and will require different information in the models for the analysis. We will adapt existing tools and methods on which the project team has an extensive experience for the development of the tools and methods tailored to the infrastructure and nature of SC applications built using the planned infrastructure. This WP is aligned with objective O5.

Tasks

- Task 3.1. Performance: We will use the statistical model checker VeStA for the analysis of the predictive analysis of SC applications. The main reason for choosing this tool for performing the analysis is the expressiveness of the language it uses for the specification of systems: Maude. Its implementation in MultiVeStA provides support for the parallel execution of the analysis, speeding up the analysis time. We will also consider the use of Simulizar, a Palladio plug-in for analyzing self-adaptive systems, PRISM and MRMC.
- Task 3.2 Reliability: The same Maude specifications of SC applications, the infrastructure and the devices will allow us to perform the analysis of their reliability using the wide range of tools available in the Maude formal environment, which includes, tools for reachability analysis and model checking. Simulation-based techniques and statistical model checking seem also appropriate in the context of uncertainty, as well as VeStA.
- Task 3.3. Privacy: In the context of collaborative SC applications the verification of security properties are in order. Specifically, we plan to mechanically prove privacy guarantees on the developed infrastructure, and guarantees on the use of people's information as granted by the virtual avatars through this infrastructure.

Task Leader: F. Durán

Deliverables:

- D3.1 Maude specification of the infrastructure and APIs (M12)
- D3.2 Methods and tools for the analysis of performance and reliability (M24)
- D3.3 Report on the privacy guarantees of the infrastructure (M36)

WP 4: Applications and Validation

Goals: In this WP we will validate the whole project approach, including methodology, tools and services by developing the three application scenarios described in Section 1.5, corresponding to objective O6. In this way, we will proof the feasibility and soundness of our proposal, as well as the results of the different technical tasks carried out the project. This WP will serve also as a proof of concept showing how the proposal can be successfully applied to the development of collaborative social computing applications.

Tasks

- Task 4.1. Requirement analysis: The purpose of this task is gathering and specifying the requirements of the three scenarios described in Section 1.5. This task will be performed in close collaboration with Tasks 1.1 and 1.2, providing and getting feedback from them. Additionally, we will explore different technological alternatives that make possible building applications as those described in the scenarios.
- Task 4.2. Application design and development: This task will take profit of the ongoing results of WPs 1 to 3, tailoring them for making real the three application scenarios. It includes the definition of particular aspects of the virtual profiles and interactive behavior of virtual avatars for each of the scenarios, as well as defining specific interfaces both to the user and to other systems.
- Task 4.3. Evaluation: Each scenario puts its focus on different aspects of the proposal: individuals, collaborative groups, third parties, interaction with IoT and other systems, etc.

Hence, evaluating the implementation of the scenarios will serve both as a validation of the ongoing results of the different technical tasks and WPs and as a proof of concept of the whole approach.

WP Leader: C. Canal

Deliverables:

- D4.1 Requirement specification of the application scenarios (M9)
- D4.2 Prototype version of each application scenario (M18)
- D4.3 Final versions of each application scenario (M30)
- D4.4. Evaluation of the results of new opportunities of research and development (M36)

WP 5: Dissemination and Exploitation

Goals: Although this is a research oriented project, it is expected that its results can be tested in real environments, and also have an impact on standards. Some of our industrial partners (Section 1.6.6) have expressed interest in the project, and will help with the non-academic dissemination inside their own companies and in possible projects in collaboration with our group. Technology transfer will be a main objective of the project.

Tasks

- Task 5.1. Academic Dissemination: The basic form of academic dissemination will be through high-quality journal and conference papers. Members of the project group have a previous record on publishing in top-notch journals. We also plan to organize international workshops about the topics of the project, using the relationship of this project with other projects and initiatives in which our group is also engaged. These events will allow us not only to disseminate the results of the project, but also to collect and integrate as much as possible information from external sources. The main achievements of the project and its progress will be documented in a dedicated project webpage.
- Task 5.2. Exploitation and relationship with industry: This task will be devoted to make the results of the project visible to the industry, and carried out in relationship with some industrial partners. We also plan to patent some of the results of the project, given its expected novelty and applicability. Our group has started working on this direction, and we plan to further exploit this way of technology transfer using the results of this project.
- Task 5.3. Standardization activities: Our group actively participates in different international standards organizations (such as ISO and OMG). It is important that the project efforts and results are fully aligned with existing and forthcoming international standards, and also that our results and findings can influence and contribute to new standards or the revision of existing ones. Participation and attendance to standardization meetings and venues is also an integral part of our dissemination plan for this project.

WP Leader: A. Vallecillo, C. Canal (Project PIs)

Deliverables:

- D5.1 Project Webpage (M06)
- D5.2 Final Exploitation Report (M36)
- D5.3 One international patent or participation in one international standard (M36)

1.6.2. Methodological approach

The project will follow an iterative and incremental approach when possible. This means that the WPs will overlap in order to allow for feedback and adaptation to the evolution of the project and the state of the art. We will use a fixed time span of six months for each iteration. A set of requirements will be signed to the iteration following a risk-oriented approach. Each iteration will have the following phases:

- Risk analysis and requirement selection. Before starting an iteration, the project scope is analyzed in order to select those work areas where focus should be put in that iteration. The work areas will be selected to minimize the project risks and maximize the added value of the effort (taking into account dissemination and exploitation).
- Iteration Planning. A detailed work plan for the time span will be developed, including only those work areas that can be developed during the time span. If some expected progress is not finally achieved, it will be delayed until the next time period.

- Development. A conventional analysis-design-implementation cycle will be followed during this phase. In the first iterations, the analysis and design phases will have greater importance, since the developments will be oriented to clarify requirements and to build proof of concept components.
- Validation. At the end of each iteration, a validation phase will be carried out. Validation and verification activities will be of paramount importance during the project development. Validation activities will run in parallel to the development in such a way that possible problems can be discovered as soon as possible.

1.6.3. Risks and Contingency plans

The development of the project will follow a risk-oriented methodology whenever it is possible. This means that the most critical activities have been planned to be carried out in the initial stages of the process, in such a way that the appropriate measures can be taken as soon as possible in the project planning. Anyway, the following set has been identified as primary risks of the project:

- *Rapid evolution of underlying technologies*. CEP, cloud/fog and mobile technologies are rapidly evolving. New systems, standards and tools will probably emerge during the next few years and they can affect the methodologies and tools developed in this project. The implication of some of the members of the project team in international standardization organizations and the close relationship of the group with leading industries in the area will try to minimize these effects.
- *Inability to define languages to represent and query the patterns.* The project incorporates leading-edge theoretical research, whose absolute success cannot be predicted. Although we strongly believe that is entirely feasible to find abstract languages for the kind of concepts and mechanisms that we want to express, the intrinsic nature of the research objectives makes it impossible to assure that such languages can be found in the context of the project. If this happens, the results of the project will be adequately reported to the scientific community so that the community can benefit from the project. In order to reduce risk, we will try to reuse and adapt existing languages in related domains, trying to avoid creating completely new languages from scratch.
- *Lack of tooling and technologies.* We may develop our own tools or extend already existing ones, which may represent a heavy overload for the project, due to the effort required to build tools. Same for some of the technologies of the project, in particular CEP engines and databases for mobile phones. We expect to be able to reuse some of the existing tools and technologies, extending them accordingly, but the lack of maturity of this field represents a challenge that cannot be neglected. The experience of our group in building tools is a valuable asset that we expect can mitigate this risk.

1.6.4. Coordination meetings and problem resolution

Given the physical proximity of the team members, there will be a continuous contact among them. In any case, and for coordination purposes, periodic follow-up meetings have been scheduled every three months. In these meetings each WP leader will report on the status and level of achievements of his objectives, and corrective actions will be agreed in case of potential delays.

Generally, all issues and problems found will be dealt with at WP level first, and the reported at the periodical coordination meetings if solved. WP leaders will also be in charge of coordinating with other WP leaders in case some problems are detected that may affect other WPs. The IP will always be informed of these issues, not only for coordination and project management purposes but also for recording all detected problems.

1.6.5. Relationship with industry and institutions

There are several companies and institutions that have expressed their interest in the results of the project. Although their interests vary depending on their domain, we have strong relationships with the following companies from past and on-going projects:

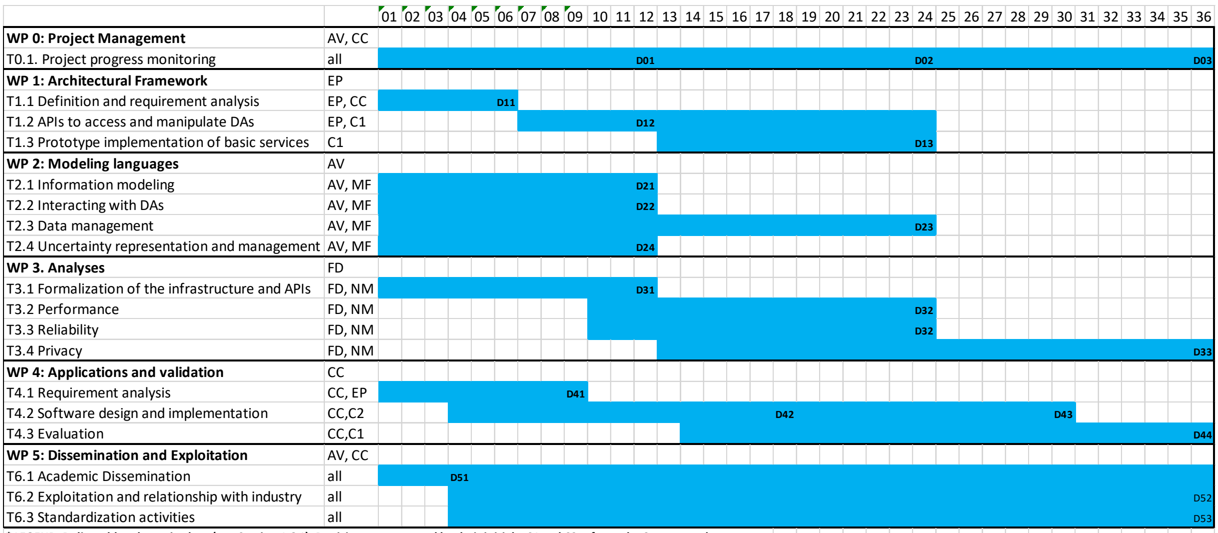
- **ATOS** is a large company whose business lines include, among others, the development of software applications and services. ATOS has been our industrial partner in several EU projects, and this project is useful for them in the realm of service-oriented applications in mobile-based and cloud applications.
- Software for Critical Systems S.L. (**SOFTCRITS**) is a technological start-up company from the University of Málaga with wide experience in sensor networks and critical systems. Its main lines of business are related to nuclear plant simulators, software for non-destructive Analysis in critical systems (aeronautics and nuclear), and critical infrastructure monitoring. They were our partners in our previous TIN2014 project, and they have shown their interest in our adaptation and use of CEP and database technologies in smartphone, and on the management of measurement uncertainty in the applications, which are two our main contributions.
- **CORDIS** is a leading company in the use of models for the development of simulation and analysis tools for Industry 4.0 applications. They have recently open a branch in the Málaga Technology Park, and they have already showed their interest in collaborating with our group, particularly in dealing with uncertainty in their models. Our results in these topics could be incorporated in some of their simulation tools.
- Finally, the **Málaga Town Hall**, through its Computing Center (**CEMI**), is actively working on implementing the Smart Cities initiative, including different kinds of services and applications. For instance, they have created an open data portal to provide useful information to Málaga citizens (https://datosabiertos.malaga.eu/), and are currently promoting the development of social computing applications that use open data. The present project will serve them as a perfect testbed for the kind of mobile-based applications they are considering, in order to study their feasibility, applicability, potential advantages and limitations, and associated problems and costs. We will be working in close cooperation with them during the whole project.

1.7. Labs and existing infrastructure

The group has a research laboratory in the premises of the ETSI Informática at the University of Málaga, and another laboratory in the premises of the recently inaugurated CAITI software institute. These laboratories would perfectly serve to host the new personnel that we are requesting in this proposal (2 contracts, 1 FPI). Most of the material needed to build and deploy the different CPS required in this project is already in place, acquired for previous and on-going projects. However, to carry on the tasks in the proposed project we will need to buy the new hardware and software licenses contemplated in the budget.

1.8. Project Chronogram

The proposed plan spans for 3 years. The chart below displays the expected schedule and duration for each WP and task, as well as the project participants that will work on each one.



*** LEGEND: Deliverables shown in chart (see Section 1.6.1). Participants are named by their initials. C1 and C2 refer to the 2 contracted persons.**

1.9. Personnel

The research team is composed of four researchers who exhibit a significant scientific productivity and have a recognized activity in different fields directly related with the proposal (Ernesto Pimentel, Francisco Durán, Carlos Canal and Antonio Vallecillo). In addition, we have incorporated two other researchers (Nathalie Moreno and Manuel Fernández Bertoa) who have experienced in the last years some difficulties to maintain their scientific productivity at the expected level. Although in recent years they have developed a research activity in topics closely related to the project, such as service computing and quality measurement, for different reasons this has not led to a better publication record. However, they have recently started producing good publications in these topics, and we hope that their participation in the project will help them improve their production record again.

The present proposal is very intensive in design and implementation of software. The development of the architectural framework and the common services; the design of the analysis tools; the implementation and testing of the case study applications, as well as their evaluation, will require a significant amount of effort that cannot be achieved without extra personnel. These activities require highly qualified staff, and preferably with experience in similar or related matters. This is why we have included in the project working plan two engineers (identified as C1 and C2 in the chart shown in Section 1.8) with similar profiles: Software Engineers, with previous experience in R&D activities, and in particular with one (or more) of the following topics: MBSE, real-time distributed computing, open data, and mobile applications. The plan is to contract one for 30 months, starting in month 6, and the second one for the last two years of the project. The expected tasks and activities of these engineers within the scope of the project, as well as those of the rest of the participants, are those indicated in Sect. 1.6, and shown in the chronogram displayed in Sect. 1.8.

1.10. Research group's previous experience

The members of our group have been successfully collaborating together for the last years (e.g. in EU projects such as SEACLOUDS or Smart-FI), joining our knowledge and experience in the design and development of distributed applications, service-oriented computing, MBSE and formal methods. This combination has proved to be essential for enriching the overall group's expertise and achievements, and has leveraged the potential impact of the results and the international visibility of the group. The research team is a joint effort by two subgroups (SCENIC and ATENEA) of the much broader GISUM team (Grupo de Ingeniería del Software de la Universidad de Málaga).

As mentioned above, when describing the project's supporting technologies, previous research activities of our group are closely related to those proposed in this project. In fact, this project builds on our results from previous ones, including SEACLOUDS, on service-oriented computing; Smart-FI, on smart cities and open data; DADIoT, which focused on the construction of users' virtual profiles; EU COST Action MPM4CPS, on effective modeling of physical systems; and POLYCIMS (TIN2014-52034-R), which introduced the use of CEP technologies for monitoring and analyzing streams of data coming from critical systems, and the representation and management of measurement uncertainty in that domain. They constitute the basis of this project.

For illustration purposes, our group produces **per year** an average of 5 (JCR) journal publications, 7 international conference papers, 9 workshops papers, and 6 national conference papers, all related to the topics of the project. We constantly participate in European and industrial projects, being very active both at national and international levels. At international level the group also maintains an active presence in conferences and journals, with participation in the editorial boards of journals (SoSym, JOT, STIOT, Computing, Services, Trans. on Internet and Inf. Systems), participation in the program committees of conferences, organization of international conferences (e.g. MODELS 2013, ESOCC 2013, ICSOC 2017, ECOOP 2017) and workshops, edition of special issues at various journals, etc. The complete list of publications (including journals, books, conference papers, book chapters, etc.) and projects can be obtained from the individual CVs and web pages of the project participants.

Another strong value of our group is the collaboration with other research groups, both at national and international levels, as demonstrated by the authorship of our joint publications with international researchers, and by our international projects. Our group is also an active member of the GEMOC initiative (http://gemoc.org), which explores the necessary breakthrough in software languages to support a global software engineering, defining the relationships between the languages and models involved in multi-paradigm modeling approaches. We also work in different standardization bodies, namely ISO/IEC, ITU-T and OMG. With ISO/IEC and ITU-T we have worked in the Reference Model for Open-Distributed Processing (RM-ODP) framework [LMTV12], on the representation using models of the framework languages, being editors of several international standards. In OMG we are part of the submission team on the new OMG standard on “Precise Semantics for Uncertainty Modeling” (PSUM, http://doc.omg.org/ad/2017-12-1). This project will be aligned with all these initiatives and standardization efforts, ensuring the corresponding technology transfer, dissemination and internationalization activities.

Similarly, we have a strong connection with industry and other institutions, as demonstrated by our continuous collaboration with them. In particular, we have been recently collaborating with the Málaga Town Hall in the development of mobile applications that make use of their open data portal, employing CEP techniques to analyze the data they provide and to extract useful information from it. We expect this project to make significant advances in our cooperation with them.

References

[AA16] A. Ahmed, E. Ahmed. A survey on mobile edge computing. In: 10th IEEE International Conference on Intelligent Systems and Control (ISCO), Proceedings, pp. 1-8, 2016.

[AWM14] A. Asadi, Q. Wang, V. Mancuso. A survey on device-to-device communication in cellular networks. IEEE Communications Surveys & Tutorials, 16(4):1801–1819, 2014.

[BBM+18] L. Burgueño, M. F. Bertoa, N. Moreno, A. Vallecillo. “Expressing Confidence in Models and in Model Transformation Elements”. In Proc. of MODELS 2018. LNCS, Springer, Oct. 2018.

[BBT+18] G. Barquero, L. Burgueño, J. Troya, A. Vallecillo. “Extending Complex Event Processing to Graph-structured Information”. In Proc. of MODELS 2018. Springer, Oct. 2018.

[BBV18] L. Burgueño, J. Boubeta-Puig, A. Vallecillo. “Formalizing Complex Event Processing Systems in Maude”. IEEE Access, 6(1):23222-23241, 2018

[BCF+12] P. Bellavista, A. Corradi, M. Fanelli, L. Foschini. “A survey of context data distribution for mobile ubiquitous systems.” ACM Comput. Surv. 44(4):1–45, 2012.

[BCW17] M. Bambrilla, J. Cabot, M. Wimmer. Model-Driven Software Engineering in Practice: Second Edition, Morgan & Claypool Publishers, 2017.

[BCG+14] A. Brogi, J. Cubo, L. González, E. Pimentel, R. Ruggia. Dynamic Verification of Mashups of Service-Oriented Things through a Mediation Platform, Journal of Universal Computer Science, 20(8): 1049-1070. 2014.

[BGC+16] J. Berrocal, J. Garcia-Alonso, C. Canal, J.M. Murillo. Situational-Context: A Unified View of Everything Involvedat a Particular Situation. In Proc. of ICWE 2016, LNCS 9671, pp. 476–483. Springer, 2016.

[BGV+17] J. Berrocal, J. Garcia-Alonso, C. Vicente-Chicote, J. Hernández, T. Mikkonen, C. Canal, J.M. Murillo. Early analysis of resource consumption patterns in mobile applications, Pervasive and Mobile Computing 35:32–50, 2017.

[BMB+18] M. F. Bertoa, N. Moreno, G. Barquero, L. Burgueño, J. Troya and A. Vallecillo “Expressing Measurement Uncertainty in OCL/UML Datatypes”. In Proc. of ECMFA 2018. LNCS 10890, pp. 46-62. Toulouse, France, Springer, 2018.

[CDE+07] M. Clavel, F. Durán, S. Eker, P. Lincoln, N. Martí-Oliet, J. Meseguer, and C. Talcott: All About Maude – A High-Performance Logical Framework. Volume 4350 of LNCS. Springer, Heidelberg, Germany, 2007.

[CFL06] C. Charron, J. Favier, J., C. Li. Social Computing: how networks erode institutional power, and what to do about it, Forrester Customer Report, 2006.

[CM12] G. Cugola and A. Margara. 2012. Processing flows of information: From data stream to complex event processing. ACM Comput. Surv. 44(3)15:1–15:62, 2012.

[Coh12] Cohen, B. What exactly is a smart city? Co. Exist 19, 2012.

[DMA16] F. Durán, A. Moreno-Delgado, J. Álvarez-Palomo: Statistical Model Checking of e-Motions Domain-Specific Modeling Languages. In Proc. of FASE 2016, pp. 305-322, Springer, 2016

[DRS18] F. Durán, C. Rocha, and G. Salaün: Stochastic analysis of BPMN with time in rewriting logic. Science of Computer Programming 168:1-17, 2018.

[EN10] O. Etzion and P. Niblett. 2010. Event Processing in Action. Manning Publications.

[GA05] R. Gross, Acquisti, A. Information Revelation and Privacy in Online Social Networks, Proc. ACM Workshop Privacy in the Electronic Society (WPES05), 71–80, ACM, 2005.

[GMB+14] J. Guillén, Miranda, J., Berrocal, J., Garcia-Alonso, J., Murillo, J.M., Canal, C. People as a Service: A Mobile-Centric Model for Providing Collective Sociological Profiles, IEEE Software, 31(2):48–53, 2014

[HSK09] J. Hong, Suh, E., Kim, S.J. Context-aware systems: a literature review and classification, Expert Systems with Applications, 36(4):8509–8522, Elsevier, 2009.

[IJR16] S. Inzunza, Juárez-Ramírez, R., Ramírez-Noriega, A. User and Context Information in Context-Aware Recommender Systems: A Systematic Literature Review, In: New Advances in Information Systems and Technologies, pp. 649–658, Springer, 2016.

[LMTV12] P. F. Linington, Z. Milosevic, A. Tanaka, A. Vallecillo. “Building Enterprise Systems with ODP. An Introduction to Open Distributed Processing” Chapman & Hall/CRC Press. 2012. ISBN: 978-1-4398-6625-2

[LRU14] J. Leskovec, Rajaraman, A., Ullman, J. D. Mining of Massive Datasets, Cambridge University Press, 2014.

[MBB+18] N. Moreno, M. F. Bertoa, G. Barquero, L. Burgueño, J. Troya, A. García-López and A. Vallecillo. “Managing Uncertain Complex Events in Web of Things Applications”. Proc. of ICWE’18. LNCS 10845, 349-357. Springer, 2018.

[MCC+14] C. Manville, Cochrane, G., Cave, J., Millard, J. Pederson, J. K. Thaarup, R. K., Liebe, A., Wissner, M. Massnink, R. Kotterink, B. Mapping Smart Cities in the EU. European Parliament, 2014. http://www.europarl.europa.eu.

[MMG+15] J. Miranda, Mäkitalo, N., Garcia-Alonso, J., Berrocal, J., Mikkonen, T., Canal, C, Murillo, J. M. From the Internet of Things to the Internet of People, Internet Computing Magazine, IEEE, 19(2):40–47, 2015.

[MSS13] P. Makris, Skoutas, D.N., Skianis, C. A survey on context-aware mobile and wireless networking: On networking and computing environments' integration. IEEE Communications Surveys and Tutorials 15(1):362–386, 2013.

[NCD+18] S. Nastic, J. Cubo, M. Donato, S. Dustdar, Ö. Guthu, M. Jonsson, Ö. Özdemir, E. Pimentel, M. S. Yümlü. “SMART-FI: Exploiting Open IoT Data from Smart Cities in the Future Internet Society”. In “Internet of Everything: Algorithms, Methodologies, Technologies and Perspectives”, pp. 153-173, Springer, 2018.

[NP11] T. Nam, Pardo, T. A. Conceptualizing Smart City with Dimensions of Technology, People and Institutions. In: Proceedings of the 12th Annual International Digital Government Research Conference: Digital Government Innovation in Challenging Times. 282–291, 2011.

[RMM+12] M. Raatikainen, Mikkonen, T., Myllarniemi, V., Makitalo, N., Mannisto, T., Savolainen, J. Mobile Content as a Service: A Blueprint for a Vendor-Neutral Cloud of Mobile Devices, IEEE Software 29(4):28–32, 2012.

[ROA09] M. Raento, Oulasvirta, A., Eagle, N. Smartphones: An Emerging Tool for Social Scientists. Sociological Methods Research 37(3):426–454, 2009.

[Rus11] M. Russell, Mining the Social Web: Analyzing Data from Facebook, Twitter, LinkedIn, and Other Social Media Sites, O’Reilly Media, 2011.

[SW14] I. Stojmenovic, Wen, S. The fog computing paradigm: scenarios and security issues. In: IEEE Federated Conf. on Computer Science and Information Systems, pp. 1–8, 2014.

[WYJ+14] Y. Wang, Vasilakos, A.V., Jin, Q., Ma, J. Survey on mobile social networking in proximity (MSNP): approaches, challenges and architecture. Wireless Networks, 20(6):1295–1311, 2014.

[XJZ+15] C. Xu, Jia, S., Zhong, L., Muntean, G.M. Socially-aware mobile peer-to-peer communications for community multimedia streaming services. IEEE Communications Magazine, 53(10):150–156, 2015.

[YH13] W.S. Yang, Hwang, S.Y. iTravel: A recommender system in mobile peer-to-peer environment. Journal of Systems and Software, 86(1):12–20, 2013.

2. IMPACTO ESPERADO DE LOS RESULTADOS - *EXPECTED RESULTS IMPACT*

The results of the project are aimed towards two main kinds of impacts. First, at **research** level it expects to produce significant advances in the state-of-the-art and state-of-the-practice in the domains of Social Computing and (Model-Based) Software Engineering, in particular in the design and development of collaborative applications using mobile-phones, which brings along quite stringent requirements of memory, battery lifetime, network bandwidth, and communications reliability, while at the same time introducing uncertainty as an essential element of all computations and decisions. Second, one of the most valuable contributions of the project is expected in terms of the definition of a new paradigm in the scope **social computing**, where the exploitation of people's information can be owned and controlled by the people themselves, instead of being used or commercialized by third parties — without neither transparent control nor clear benefits to the data owners. These results will also give the group a perfect opportunity to be more competitive and visible at international level.

The goals and expected results of the project are fully aligned with the challenges declared in the Spanish Scientific Research and Innovation Plan for the period 2017-2020. In particular, Challenge 7 (Digital Economy, Society and Culture) includes two specific strategic priorities in which this proposal will provide clear contributions, as described in Section 1.2. In fact, the project proposes solutions to exploit the "*Internet of Things and their applications*", which is explicitly mentioned as the strategic priority 7(i). The new scenarios opened with the proposed approach will provide new opportunities for the *"development, innovation and adoption of solutions and technologies bound to: cloud computing and massive processing technologies of data"*, which again is explicitly declared as part of the strategic priority 7(iv). In this context, the project results are expected to be very useful to some of our partners, and in particular to the Málaga Town Hall, for implementing new forms of Social Computing applications within their Smart City initiative. Our collaboration with them on this project is expected to have a relevant impact on their applications within the next 3-4 years. We also expect some collaborations in H2020 EU projects based on the results of this project, particularly with ATOS, since they have shown interest in our potential results.

Dissemination of the project activities will be aimed at promoting and publicizing the work achieved and the results obtained, with a direct influence on the final impact of the project. The basic form of academic dissemination will be through high-quality international journal and conference papers. Both national and international workshops about the topics of the project will also be organized, not only to publicize the project results and improve its visibility, but also to discuss and exchange information and experiences with other research groups working on related topics and projects, and to start collaborations that enable the development of joint European projects. Last, but not least, it is also very important in our case that the project results are fully aligned with existing and forthcoming international standards, and that our results can influence and contribute to new standards or the revision of existing ones. Thus, we will coordinate our efforts with those currently in place within ISO/IEC and the OMG on Smart Cities and IoT, and on Uncertainty Modeling, respectively.

3. CAPACIDAD FORMATIVA - *TRAINING CAPACITY*

The research-oriented nature of the group provides the natural environment for the development and supervision of PhD thesis. The following is the list of PhD thesis supervised by the group members in the last 10 years (on average, one thesis defended per year). We also include, for each one, where he or she is currently working (University or Industry):

1. Loli Burgueño. "On the properties on Model Transformations: Correctness and Performace". **2016** (Dir: A. Vallecillo) – Currently in academia (post-doc at Universitat Oberta de Catalunya, UOC)
2. Javier Troya. "On the Model-Driven Performance and Reliability Analysis of Dynamic Systems" **2013**. (Dir: A. Vallecillo) – Currently in academia (profesor ayudante doctor at U. Sevilla)
3. Nathalie Moreno "WEI: Integración de aplicaciones externas en sistemas Web mediante modelos" **2012**. (Dir: A. Vallecillo) – Currently in academia (profesor contratado doctor at U. Málaga)
4. José A. Martín Baena. "Secure Adaptation of Software Services" **2012** (Dir: E. Pimentel) – Currently in industry (VirusTotal, Google)
5. Meriem Ouederni. "On Checking the Compatibility of Service Interaction Protocols" **2011** (Dir: G. Salaün, E. Pimentel). – Currently in academia (Associate Professor at IRIT/INP Toulouse, Francia)
6. Manuel Roldán. "Estrategias para el guiado y la monitorización de la ejecución de sistemas en Maude" **2011** (Dir: F. Durán) – Currently in academia (profesor Titular de Universidad at U. Málaga)
7. José Luis Pastrana. "Uso de conectores para la coordinación, composición y adaptación de componentes" **2011** (Dir: E. Pimentel, M. Katrib) – Currently in academia (profesor Titular de Universidad at U. Málaga)
8. José Eduardo Rivera "On the semantics of Real Time Domain Specific Languages", **2010** (Dir: F. Durán y A. Vallecillo) – Currently in academia (Lecturer at Marbella International University, MIUC)
9. Javier Cubo. "DAMASCO: Discovery, Adaptation and Monitoring of Context-Aware Services and Components" **2010** (Dir: E. Pimentel, C. Canal) - Currently in industry and academia (profesor asociado at U. Internacional de La Rioja)
10. Javier Cámara. "Behavioural Adaptation of Components and Services" **2009** (Dir: C. Canal, G. Salaün) – currently in academia (Lecturer at U. York, UK)

Members of our group actively participate in the Malaga University PhD programs, both in the past and now in the newly accredited PhD Program of the University on "Tecnologías Informáticas" (RD99/2011). The following PhD thesis are currently in progress in our group:

– José Carrasco. "Trans-Cloud: CAMP-TOSCA-based management of cloud applications" ~Early 2019 (Dir. E. Pimentel & F. Durán)
– Patricia Araújo de Oliveira. "Predictive Analysis of Cloud Systems with e-Motions" ~Late 2019 (Dir. F. Durán & E. Pimentel)
– Alejandro Pérez Vereda. "New Models for Social Computing". ~Early 2020 (Dir. C. Canal).
– Gala Barquero Moreno, "Tratamiento y análisis de flujos de datos estructurados" ~Late 2020 (Dir. A. Vallecillo)

The present project proposal opens up new research lines that complement those previously tackled by the group, and require dedicated personnel. We believe, based on our previous records, that our group is able to successfully train one PhD and produce high-quality professionals (not only for research in academia, but also for industrial positions). Given that two of the theses in progress will be defended during the first months of 2019, the training capability of the group will be more than enough to host the new PhD student we apply for in this project proposal.

Research plan, method and activities. The research plan for that candidate assumes a three-years activity program, with a possible extension to four years if necessary. Design Science will be followed as research methodology, since our experience has shown that it provides an adequate scientific and systematic approach for conducting research in the areas of Software Engineering.

The research problem to be addressed is related to the analysis of emergent properties of complex applications that involve many different participating agents (humans, physical devices and communicating networks) with uncertain environmental conditions. The applications designed and developed in this project are of that nature, and hence its suitability and the opportunity to develop a PhD within this project.

Concerning the research plan for the candidate, apart from the inherent research activities of any PhD, other activities to be performed by the candidate include: training courses both in general research subjects (research methods, scientific writing skills, etc.) and in specific topics related to her research theme (formal methods, analysis tools); short stays in research centers to collaborate with other researchers; and at least one longer stay (4 months) to work in a topic related to the thesis in another lab or a top-notch research center, probably with one of our current international contacts.

4. IMPLICACIONES ÉTICAS Y/O DE BIOSEGURIDAD - *ETHICAL AND/OR BIOSAFETY IMPLICATIONS*

N/A