The Internet of Speaking Things and Its Applications to Cultural Heritage

Fiammetta Marulli[1], Remo Pareschi[2] and Daniele Baldacci[3]
[1]*DIETI, University of Naples Federico II, Via Claudio 21, Naples, Italy*
[2]*Department of BioScience and Territory, University of Molise, Pesche (IS), Italy*
[3]*Olosproject Ltd, 7 St. John's Road, Harrows, Middlesex, U.K.*

Keywords: Internet of Things, Cultural Heritage, Multimedia Content Processing, Advanced Human Computer Interaction, Smart Environments, Holographic Simulations, Speech Recognition, Natural Language Processing.

Abstract: The initial driver for the development of an Internet of Things (IoT) was to provide an infrastructure capable of turning anything into a sensor that acquires and pours data into the cloud, where they can be aggregated with other data and analysed to extract decision-supportive information. The validity of this initial motivation still stands. However, going in the opposite direction is at least as useful and exciting, by exploiting Internet to make things communicate and speak, thus complementing their capability to sense and listen. In this work we present applications of IoT aimed to support the Cultural Heritage environments, but also suitable for Tourism and Smart Urban environments, that advance the available user-experience based on smart devices via the interaction with speaking things. In the first place we describe a system architecture for speaking things, comprehensive of the basic communication protocols for carrying information to the user as well as of higher-level functionalities for content generation and dialogue management. We then show how this architecture is applied to make artworks speak to people. Finally, we introduce speaking holograms as a yet more advanced and interactive application.

1 INTRODUCTION

Internet of Things (IoT) has to date been viewed as an infrastructure and a project essentially aimed to massively strengthen the already impressive capabilities of information acquisition of the Internet.

Let everything, in the broadest sense of the term - from the household refrigerator to the probe roaming the immensity of space, from the car that juggles its way in the metropolitan traffic to the hare that races carefree through woods and fields - turn into a sensor capable of acquiring data and of pouring them into the cloud, where they can be aggregated with other data and analysed in order to extract decision-supportive information: such has been so far the mainstream of IoT evolution.

This approach to IoT fits neatly with the trend of Big Data, of which it multiplies the capabilities and potential. Its validity stands on very solid grounds, yet there is just as much validity in traveling the road carved by IoT in reverse: namely, given that data and information are now generated, acquired and made available in ever larger amounts, can we use IoT as a way to serve them to users according to modalities radically more innovative and more accessible with respect to those in place?

The benefit of doing so is clear and evident, in that it maximizes the usefulness of the available information. Equally clear is that there are powerful and mature technologies that can make the IoT data outflow fully user accessible, just as proven methods and algorithms of big data analytics can make manageable and exploitable the IoT data inflow.

Indeed, stepping beyond dedicated interfaces like desktops, laptops, tablets, and smartphones, which are rooted in information technology and telecommunications, IoT, coupled with existing capabilities of speech processing and of management of dialogues in natural language, can create the basic prerequisites for an ecosystem of speaking things.

This will enable a direct and natural connection between human users and the information cloud, thereby making effectively "smart" not only terminals like computers and mobile phones but also physical

Marulli, F., Pareschi, R. and Baldacci, D.
The Internet of Speaking Things and Its Applications to Cultural Heritage.
DOI: 10.5220/0005877701070117
In *Proceedings of the International Conference on Internet of Things and Big Data (IoTBD 2016)*, pages 107-117
ISBN: 978-989-758-183-0

environments such as cities, homes, offices, shops and museums.

We illustrate this possibility by showing how the Internet of Speaking Things can be applied to the domain of Cultural Heritage, which fits perfectly well with our vision: we often say metaphorically that the past speaks to us, and we translate metaphor into reality by making it literally speak!

We thus describe an architecture for the Internet of Speaking Things in this domain and provide a specific instance of its application through a case study of “speaking vessels”. We also show how to go beyond cultural heritage by applying the same techniques to a fully animated “thing”: the holographic human being, which originates from the coupling of the ancient tradition of performing arts with the most recent advances of 3D cinematic photography.

2 BACKGROUND AND RELATED WORKS

Internet of Things (IoT) and Internet of Services (IoS) are currently paving the way towards unified ICT solutions that support variety of applications composing smart environments.

The feasibility of IoT derives from the level of maturity reached by several enabling technologies such as wireless sensor networks (WSN), Bluetooth Low Energy (BLE) and Near Field Communication (NFC).

A parallel trend to IoT is Big Data, where data coming at high velocity in high volumes and variety are increasingly collected, stored and analyzed by modern organizations. A large bulk of such data typically describe different dimensions of the daily social life and are the heart of a knowledge society, where the understanding of complex social phenomena is sustained by the knowledge extracted from the miners of big data across the various social dimensions by using mining technologies (Zafarani, 2014). Mining Big Data is an emerging and specific field where there are more problems than ready solutions, among which the data retrieval.

Remarkable opportunities for acquiring these data are surely provided by IoT infrastructures, capable as they are of turning anything into a sensor that acquires and pours data into the cloud, where they can be aggregated with other data and analysed to discover interesting knowledge and to extract decision-supportive information.

Going in the opposite direction is, however, at least as useful and exciting, by exploiting IoT to make things communicate and speak, thus complementing their capability to sense and listen.

From this perspective, IoT perfectly fits with the Information-Centric Network (ICN) approach (Xylomenos, 2014) that has the potential to provide very natural and efficient solutions for many of today’s important communication applications including but not limited to highly scalable and efficient content distribution (for example, of web pages, videos, documents, or other pieces of information), thus motivating the development of novel Internet architectures based on named data objects (NDOs). The approach of these architectures, commonly called ICN, can be contrasted with host-centric networks where communication is based on named hosts, for example, web servers, PCs, laptops, mobile handsets, and other devices (Ahlgren, 2012). The ICN architectures leverage in-network storage for caching, multiparty communication through replication, and interaction models that decouple senders and receivers. The common goal is to achieve efficient and reliable distribution of content by providing a general platform for communication services.

As discussed at length in (Ahlgren, 2012), communication is driven by receivers requesting NDOs. Senders make NDOs available to receivers by publishing the objects. From a networking perspective, these objects can be viewed as named data chunks without semantics, but some ICN designs have an information abstraction model including multiple representations, i.e., unique bit patterns, for the same object.

However, to make the synergy between Big Data and ICN real, smart systems have to be provided with information logic and semantics so as to turn a service infrastructure into an effectively intelligent system.

In (Singh, 2014), authors show how billions of devices could emerge into a single system and how raw data can be interpreted through meaningful inferences. They view IoT as the biggest promise of technology today, but they evidence that it still lacks higher-level semantic capabilities.

Cultural Heritage (CH) and Assisted Tourism (AT) have turned out to be examples of suitable domains in which such achievements can be profitably exploited, since they require addressing a variety of interdependent aspects: sustainability of interventions, people enjoyment, personalized and customized interactions, promotion and safeguard of spaces and cultural objects.

In recent years, several ubiquitous and multimedia systems for Cultural Heritage that leverage the IoT paradigm have been proposed and discussed, moving from the premise that IoT (Atzori, 2010) leads to full maturity the concept of intelligent environment with numerous applicative effects in everyday life.

Smart cultural sites (Amato, 2012) provide a relevant application of IoT into CH, aiming to involve visitors with more amazing and personalized experiences in living culture. “Talking” museums (Amato, 2013) exploit a novel approach in the story telling of an art exhibition. More generally, cultural objects and sites (sculptures, drawings, buildings, etc..) can be enabled to support the visitors in the contextualizing of their exploration and interactions with cultural objects; if supported by intelligent infrastructures, objects like statues, paintings, jewels or bowls, can tell about themselves, their stories, why and for whom they are born, how they stand in relation with the other surrounding objects, the “geographical” place and the historical epoch of origin.

Nowadays, a very common and widespread example of smart environment implementation, basing on IoT infrastructures and services, is represented by the smart guides and the multimedia content delivery services.

In (Chianese, 2013) the authors describe a Location Based Service System whose main components are an indoor positioning system, a multimedia contents repository server and a smart app to guide users during an art exhibition visit and help them in the acquisition of contents. In (Marulli, 2013) a more advanced location based application, enhanced by advanced data features extraction and multimedia contents recommendation strategies is presented to support cultural events and spaces.

In (Bordoni, 2013), a perspective on the support of Artificial Intelligence to CH is given, introducing an ontology based approach to improve the effectiveness of recommendation systems. In (Semeraro, 2012), folksonomies are introduced as the base of a strategy to enhance content based filtering techniques of multimedia objects, by the exploitation of semantic annotation (tags) released by users on cultural digital objects during their web navigation.

In (Valente, 2013), a workflow based approach, exploiting a semantic enrichment of the cultural contents system via the integration of Social Networks pulses as a further knowledge source, is proposed.

The authors of (Kang, 2012) proposed a network based ticket reservation system which is a localization-based smartphone application with augmented reality. In (Husain, 2012), a Personalized Location-based Recommender System provides personalized tourism information to its users.

Most of the constraints to be taken into account, when designing a system providing pervasive information, are given by the actual domain where they are deployed. In this scenario, a not trivial issue concerns the selection and organization of knowledge which has to be delivered to users.

Users, coming from diverse cultural and social background, and of diverse age and sensitivity, have to be approached in different ways, in order to reach an effective engagement with the context they are experimenting.

In (Marulli, 2015), the author proposes an authoring platform for automatic generation of tailored and personalized types of profiled textual artworks biographies (fables for schoolchildren). Users profiled textual artworks descriptions are employed to feed a mobile app, as part of an IoT smart infrastructure, supporting users during a real-life “talking” sculpture exhibition.

As a matter of fact, IoT enables things to communicate with each other, but what they should be able to tell and how they could communicate with their human interlocutors cannot be taken for granted. Nowadays, the most recent progresses in IoT technologies, as evidenced by the newest integrated solutions like the Smart Beacon (Smart Beacon, 2015) platforms, further simplify and reduce the effort in the design and implementation of smart environments. Therefore, effective bottlenecks of IoT based solutions are no longer located in the communication infrastructure and its deployment, but in the quality and the type of provided services, and in the interaction paradigm.

To better understand the motivations behind our work, it is important to analyze the relationship that exists between cultural exhibitions and their visitors: the purpose of the visitor remains, as it has always been, to see and learn more, hence the deployed technology should stay a means rather than become an end. For these reasons and to better promote an exhibition or a museum heritage, it appears preferable to provide exhibition objects with the capability of telling their story, rather than let users require in first person (through multimedia guides) more information about an object as has happened so far in the majority of smart museum scenarios.

As a matter of fact, visitors are not infrequently bored by spending their time to look at their devices to know more about what they are admiring. Studies performed by logs computation from a multimedia-enhanced sculpture exhibition (Chianese, 2015) evidenced that users do not linger to dismiss their smart devices, and proceed in their visit just as in the old days by admiring statues or, at most, by reading flyers and paper catalogues. In (Benedusi, 2015), a further evidence of the still insufficient involvement of people in cultural topics and events is assessed by mining Twitter messages produced by users' during their participation to cultural events.

These phenomena evidenced that interaction must be made more interesting and engaging to win users' appreciation. It's a pity, as Michelangelo would say, that such statues or artworks can't really say a word! In such a perspective, a more direct and natural interaction between humans and things could reverse the situation, by providing a sort of natural human dialogue between the parts.

Nowadays, an evolution towards more proactive scenarios is evidenced by the availability of many applicative and infrastructural solutions, so as to support more engaging interactions, mainly when the amount of exhibited objects is very large. A current limitation of existing smart cultural sites service infrastructures lies, on the other hand, in the partial or total lack of support for natural language interactions, and for the effective dynamic composition of personalized services.

Indeed, the easiest way for a user to acquire information and to express needs and preferences regarding a desired service or condition, is to adopt natural language. In what is called a "smart environment" users expect to use natural speaking processes to drive their interactions with the surrounding environment and its elements.

Service customization and composition (contents tailoring and delivery, visit paths recommendations, etc..) mean to create new services by composing existent ones. In traditional approaches, this composition is done by a human expert because the composition task requires an understanding about the service semantics.

The main challenge lies in the fact that natural languages incomplete and ambiguous, while the service composition process should lead to valid services. In (Cremene, 2009) authors propose a natural language service assemblage method based on composition templates (patterns).

The criteria which has driven the design and development of the proposed service architecture derives from the basic concept of trying to shorten the distance between users and cultural objects, going a step further than simple smart device support. A process driven by natural language, so as to make objects and holograms interact naturally, is needed to let visitors engage, in front of an artwork, into new exciting experiences of knowledge.

3 THE APPROACH

In order to provide elements for an effective comprehension of the proposed approach, we can consider the case of a tourist visiting an Archeological Museum exhibiting ancient Greek vessels. We assume that this environment offers a wireless sensors network (Bluetooth technology), enabling visitors to interact with the cultural things exposed in the exhibition area, by exploiting their personal smart device and an appropriate application.

We can image a situation in which the visitor is walking within a given exhibition room through several ancient objects (bowls, statues, drawings, weapons, jewels, utensils, etc.) and when he/she is particularly close to one of them, its mobile devices detected by the sensor placed on the basement of such an object.

Once the user mobile device has been detected thanks to the WSN technology, the cultural object is animated and is enabled to talk about itself, its author, its story and its status. In this way, a "static" art exhibition is able to transform itself into a "living" one and the speaking artworks surround the visitor in a place that is now both "populated" and "animated".

Figure 1 shows, at a glance, the described environment of speaking things.

Figure 1: The Speaking Things System Environment.

Our approach builds on an existing paradigm of pro-active interaction but at the same time forces a significant paradigm shift, where interactions are effectively driven by natural language.

In fact, in the current, well assessed paradigm, contents are delivered using multimedia facilities and taking into account user preferences (e.g., audio, text and images data). In this case, the object "speaks" by means of contents delivered via a smart application to the users' devices; the interaction is mediated by the massive usage of personal devices. Narrations are provided as textual documents, audio files and browsable multimedia contents, hence they have to be managed and driven by user actions on the smart device.

In the second and novel approach, when a visitor is located as close to any art object, a push notification appears on the screen of the user's smartphone, asking to submit a request or a question to the object in natural language, to meet a possible user's desire to know more about the object. Figure 2 depicts, at a glance, this type of scenario. Users speaking and submit a question among a set of proposed ones, that are suggested as tips on the screen of his device.

If the visitor accepts the proposed interaction with the art object, the latter starts to speak, telling about its history, suggesting other objects categorized as similar, according to such criteria as art style, historical period, shared artist etc. The visitor can thus interact with the object, interrupting and conversing with it as in a real human conversation.

Some bookmarks to skip parts of the art object story telling are suggested to users in a "magic words" list, provided on the App display.

Figure 2: Speaking Things and Humans Interaction.

Users submits their requests for contents or explanations, by using the microphone of their personal device. At the time of the first request, if correctly delivered to the system (e.g. no noisy contexts or bad spelling, or unrecognized or unsupported languages), through the smart application communicating with sensors equipping the art objects, the system automatically recognizes the user's language request.

In the current implementation of this system, Italian and English are supported. In the next future, other three languages as German, French and Spanish will be integrated in the system. As default choice, the first interaction is proposed in English, but users can immediately request to obtain the questions list in other supported or preferred languages. If a request is correctly delivered to the backend (Multimedia Contents and Service Provider Tier) of the Speaking Things system, it is processed according to Speech to Text and Text Analysis computation processes, and a matching answer, if existing, is delivered to user as a vocal message, through a Text to Speech process.

The matching process, as it will be detailed in the next section, is based on a Question & Answer engine, and selects the corresponding answer, given a well identified and recognized question or significant part of it. If any question can't be recognized by the system (too noisy environment, bad spelling, unsupported languages or questions), the system produces a message asking the users to re-submit his/her request.

Otherwise, if a question is correctly recognised and a matching answer is available, the latter is emitted and played through a human natural voice, output either by the speakers equipping the basement of the art objects or the speakers of the user's personal device.

4 SYSTEM ARCHITECTURE

A functional overview of the system architecture supporting the Speaking Things environment is shown in Figure 3.

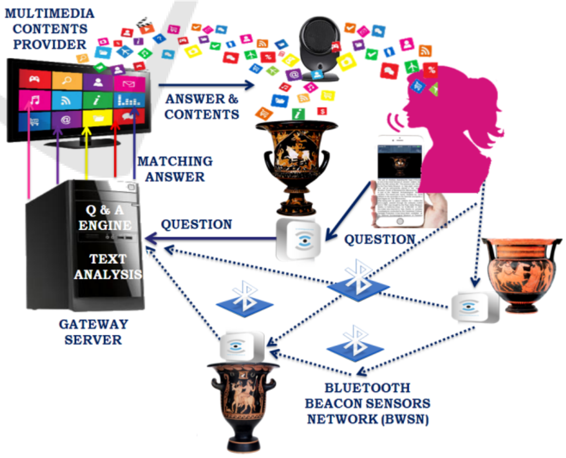

Figure 3: Service System Architecture.

The most advanced Bluetooth 4.0 smart and low energy consumption technologies have been exploited to build a robust communication and service infrastructure. Smart Beacon (Smart Beacon, 2015) devices were employed to identify users' location and deliver specific contents, according to the art object they were closest to.

The architecture is built from the following main components:

- A Wireless Sensor Network (WSN) composed by a set of beacons, each one deployed near a cultural artefact (e.g. a vessel, a sculpture or any other object) and communicating with a Gateway Server (an advanced base station). Such a component scans, using the Bluetooth protocol, the areas surrounding sensors in order to detect possible users that could be interested in the observed object.

When a device is recognized, the device is effectively connected to the data network.

- A Gateway Server (GS) hosting a set of services able to filter and process information coming from the WSN and implementing the core computations underlying the Speaking Things system. It mainly consists of a Natural Language Processing Service (concerning Speech To Text and Text To Speech Translations and Advanced Text Analysis) and a Question-Answer Engine. It is responsible for users' question analysis, matching answers selection, triggering a Multimedia Contents and Services Provider for delivering selected answers and suitable multimedia contents, to enrich the experience of users. All information about interactions occurring between users and objects are properly stored in dedicated logs for further analysis and system refinement.

- The Multimedia Contents and Services Provider (MCSP) accepts a request for extracting contents from the GS (answers and multimedia collections), so delivering them to users' devices App. The MCSP manages a multimedia repository and exploits proper multimedia delivery techniques to propose users other object of interest arranged in the shape of multimedia stories.

- A Smart Assistant App (SAA), leading and managing connection and communication between users' personal smart devices and Smart Beacons. App layout and other preferences can be customized via a dedicated environment for service composition. It enables users to submit the system requests in natural language, and to acquire multimedia contents, thus playing the provided voice answers.

4.1 The Art Things Speaking Process

Figure 4 describes the core process for the natural language driven interaction. A Question-Answer (Q-A) engine is employed to find matching answer for well-formed request.

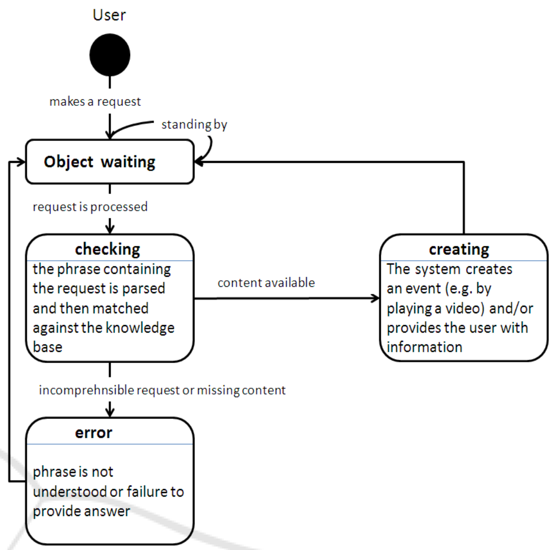

Figure 4: The Natural Language Based Q&A Process.

At a glance, the processing flow is composed of the following four steps:

1. *Speech to Text Conversion*. An instance service is listening on a communication bus, waiting for a user's request. Speech Recognition and translation into Textual form was implemented employing Google Speech Recognition API (for the web application version) and Microsoft DotNet 4.0 Framework Solution (ASR and TTS SDKs, for the desktop solution version).
2. *Text analysis and Request Categorization*: typical text analysis (parsing, word sense disambiguation, lexical features extraction, etc.) and categorization strategies are applied to incoming textual requests. The output is represented by lists of relevant terms and categories useful to match suitable answers and to select multimedia contents.
3. *Question and Answer Matching:* Categories and terms summarizing the request are matched against a Knowledge base (KB). Cultural Heritage Experts provided specific domain ontology, vocabularies and answers to populate effectively the underlying KB.
4. *Text to Natural Language Speech Generation*: If an answer matches against the submitted

request, an event consisting in a voice explanation or in playing an audio or video, is created and such event is delivered to be delivered and notified to requesting users.

5 CASE STUDIES

In this section two real case studies are described. The first one concerns an archeological exhibition held in an Italian Museum, in which smart technologies deliver desired multimedia contents driven by a natural language communication. The second one deals with an installation of holographic totem at the Expo 2015, a wide multicultural fair exhibition held in Italy.

5.1 The Speaking Vessels

As one of the effective examples for our application, we considered the case of a tourists and local people visiting The National Archaeological Museum of Sannio Caudino, within the Bourbon Tower of Montesarchio (Benevento, Italy). In such a place, one of the most beautiful vessels from Ancient Greece, "*the Crater of Assteas*", is exposed as the main attraction for cultural exhibitions and events.

For long times, the bowl was far from its original location in the Sannio, after being sold by grave robbers and sold in antiques market.

Only in 2007 the crater was returned to Italy, thus increasing the attention and expectations for visiting and looking at it. Other 50 bowls and vessels were exposed in the same exhibition and provided with the beacons in order to speak about their story.

In 2015, this cultural site was set up, during an exhibition organized to celebrate the return back to its original location of the precious vessel after many years, with a temporary beacon sensors network, so enabling the possibility for visitors to interact with the cultural things arranged in the exhibition area. Every time a user was close to a vessel, it tried to catch visitors' attention starting to tell about their story and curiosities.

A rich multimedia collection containing textual documents describing the charming story of the precious vessels, digital reproductions of the complex mythological scenes depicted on the vase, narrated by educational videos, audio guides, and hypermedia documents for a deeper comprehension, was proposed to tourists and visitors, driven by natural language requests and replies. Two different languages, Italian and English, were available for natural language interactions and contents explanations.

5.1.1 Implementation Details

In this section, we report some implementation details concerning the developed prototype.

As previously said, the WSN has been realized and tested using Smart Beacons, small devices using Bluetooth 4.0 smart and low energy consumption technologies for interacting with users. When placed in a physical space, such as a museum or an exhibition area, a smart beacon broadcasts tiny radio signals around itself at low power consumption, being able to interact with users' personal smart devices or any other device supporting Bluetooth (BLE) communication.

When a user walks through an area and his position is detected, s/he will receive custom notifications and invitations to interact with the surrounding objects. The micro-location system is able to locate mobile devices that are as close as four inches away, or as far as over 200 feet away and even more.

The Smart Beacon platform solution also offers a prebuilt but customizable App, useful for typical interactive museum scenarios. In our case of study for the National Archaeological Museum of Sannio Caudino, we have just implemented a customized application supporting Android and IOS platforms, in order to craft our specific kind of interactions (mainly, the Speaking Things Proactive Mode).

In the first place, each deployed smart beacon in the museum area was assigned a unique identifier (SB-UID), so as to maintain dedicated links between each beacon and its corresponding speaking artwork, and its location over the exhibition area.

On the other hand, when a user's device is detected for the first time in the smart beacons network, a unique user identifier (UUID) is assigned to it.

Each beacon can be combined with more than one art object. According to our experience, it is possible to combine up to 5 different objects with the same beacon, with a precision of about 90% as far the correct detection of the object is concerned.

This result can be guaranteed under the assumption that the exhibition layout is designed to fit such distances, so as to avoid overlapping regions of beacons inquiry.

When a beacon scans the surrounding area to detect other BLE devices, a localization algorithm detects and computes the existing distance (RSSI value estimation), thus inferring the proximity of the detected device to the right object.

Our choice can be further motivated by taking into account that such devices allow to control the Bluetooth inquiry area with seven different power levels in order to set different coverage zones from 5 to 50m that can be decreased to obtain higher level of precision in detecting users' devices in the artworks surrounding area. The distance of any detected device is calculated exploiting the RSSI value, received during the inquiry process slog with the UUID of the Bluetooth device.

In our case study, we were forced to design a hybrid solution, because of the unmovable position of the archaeological art objects. The room, among the six of the museum, hosting the ancient Greek vessels was arranged in a set of display cabinets very close to each other, each containing a different number of objects (from one to five), as evidenced in Figure 5. Therefore, we designed a network deploying just one smart beacon node covering a 2 metres space line (one beacon for each pair of cabinets); we also reduced the sensibility for a long range detection, in order to solve the overlapping situations.

Overlapping areas were also managed by taking into account the movement direction of the users. A user localized in an overlapping area covered by two different beacons and standing at the same distance from each of them, will be notified by the beacon that has been detected after in the timeline of the user's path. In other words, the winning beacon is the most recent one detected by the user's device.

Figure 5: The National Archaeological Museum of Sannio Caudino artworks layout.

From the user's device standpoint, the App is designed to be in a "listening" state, waiting for invitations from smart speaking objects.

The UUID is used by the App to figure out a course of actions and to track the user's path, for conflict resolution (overlapping areas) and deep post-processing (users' log behavioural analysis for contents and topics recommendation or user's profiling). In the future release, the App would check the phone's default language and automatically use it in what it displays or plays.

The GS component has been implemented via several JAVA libraries exploiting multi-threading facilities, able to inquiry the advanced NLP and Q&A features provided by the Intelligent Platform Cogito (Cogito, 2016), in order to identify the matching contents to be retrieved by MCPS and delivered to users' devices and to the output audio infrastructure.

In addition, a proper repository managed by the NoSQL DBMS MongoDB was employed to store interaction logs.

The MCSP component exploits ad-hoc JAVA libraries to build the multimedia story of an artwork on top of the multimedia collection managed by a PostgreSQL.

Finally, a user can interact with our system using at the moment an Android or IOS SAA, implemented by exploiting Smart Beacon SDK. The presentation logic is based on apposite widgets. The client requests are elaborated by JAVA Servlets and the contents are sent to the client in form of JSON data.

5.1.2 Usability, Enjoyment and Naturalness Estimation

A number of trials have been performed to assess the behaviour, the users' enjoyment and, consequently, the usability and the utility of the proposed application. A sample of about 100 visitors were logged during one of the events organized for celebrating the return to its original location for the Crater of Assteas.

These participants were engaged at the entrance of the exhibition, before starting the visit and were given a 10-minute presentation about the infrastructure.

According to the usability dimensions for a mobile application, as proposed by the literature in (Baharuddin, 2013), we investigated three of these dimensions to have an overall estimation for the proposed approach. We considered the following dimensions: simplicity (SIM), usefulness (USN) and enjoyment (satisfaction) (ENJ). For a better investigation, we added a further dimension, the naturalness of interaction (NAT).

Participants were asked to fill in a post-visit questionnaire. These questionnaires stimulated users to express their level of agreement with a set of statements, using a 10-point Likert scale, or to make choices between proposed options.

Table 1 summarizes results extracted from the users' answers, showing the most relevant questions related to the four dimensions of the usability considered and their average ratings.

The overall degree of satisfaction manifested by participants towards the proposed infrastructure was positive with an average rating of 8.86 (ENJ08).

Table 1: Post-Visit Questionnaires Results.

ID	Description	Value
SIM01	It was easy to interact with the exhibit artworks.	8.56
SIM02	It was easy to obtain useful multimedia contents.	7.81
SIM03	It was easy to navigate among the mobile App functionalities.	8.02
USN01	The infrastructure was overall useful during the visit.	7.83
USN02	Using the infrastructure was useful to gain knowledge about the exhibit artworks.	7.66
USN03	Using the infrastructure was useful to get a deeper insight on the museum themes.	7.89
ENJ01	I appreciate the mobile Assistant App GUI.	8.32
ENJ02	I appreciate the artworks detection metaphor.	8.45
ENJ03	I appreciate the image galleries.	7.44
ENJ04	I appreciate reading cultural information about exhibit artworks.	7.06
ENJ05	The quality of the sound was high.	7.52
ENJ06	Using the infrastructure contributed to increase my will to visit other art exhibitions.	8.09
ENJ07	Using the infrastructure positively contributed to the enjoyment of my visit.	8.87
ENJ08	I overall appreciated the infrastructure and the proposed approach.	8.86
NAT01	I appreciate listening cultural information about exhibit artworks.	8.98
NAT02	I appreciate the clearness of the spoken dialogue.	8.32
NAT03	The waiting time in the performing interaction attended my expectations.	7.89
NAT04	I appreciate the naturalness of the interaction with the environment	8.45

Furthermore, the overall degree of perceived naturalness in the proposed interaction modality (NAT04) and the expected waiting time in the performing interaction (NAT03) were positive with an average rating of 7.89 (NAT03) and 8.45 (NAT04), respectively.

Multimedia features such as image-galleries (ENJ03), texts (ENJ04) and the quality for audio responses (ENJ05), were rated 7.44, 7.06 and 7.52, respectively. As for the usefulness dimension, users agreed that the application was useful overall (USN01, 7.83), facilitating to a certain degree the acquisition of a better knowledge (USN02, 7.66) and a deeper insight (USN03, 7.89) on the artwork on display.

Additionally, the analysis of the ease of use dimension pointed out that participants found the information access about the artworks quite easy (SIM01, 8.56) as well as the multimedia content browsing (SIM02, 7.81).

5.2 The Holographic Human Being

Holographic human beings bring one step further the capabilities of interaction of speaking things, in that they can provide support for full-fledged dialogues in natural language. This patented basic technology allows capturing, in the form of holographic simulations, sequences of actions performed by human actors that can be matched with requests coming from human users through the integration with technologies for speech processing, natural language understanding, gesture recognition, linked data and knowledge representation.

Deployment is through standardized carriers such as ordinary totems optimized for 3D displays with very high resolutions (at least 4K UHD), always maintaining full size reproduction of the holographic human being so as to make the user experience totally natural and familiar.

Applications range widely, cultural heritage being one (as exemplified by the installation at EXPO 2015 where the historical characters of Teodolinda and Vergil tell the visitor about the history of Lombardy (https://www.youtube.com/watch?v=dB6wUGG9Oys&feature=youtu.be) but include also info-points, CRM, augmented shops, training, home automation and robotics (e.g. by providing the coordinating interfaces for teams of communicating home appliances and robotic agents).

A fundamental step in supporting interactivity is to make the holograms cloud-connected and capable of transferring information back and forth over the Internet.

In fact, this is a necessary condition for the holographic human beings to be able to answer the requests of the users providing them with the needed information. The IoT brokerage services are essential in this respect, being this a typical case of communication between the machines maintaining the information and the "animated thing", namely the holographic human being, that would provide them to the user. Pro-active and reactive event processing, audio mining, content optimization and context-aware recommendation can also be exploited effectively to turn holographic human beings into revolutionary user experiences and interfaces. A patent for this application was registered (Patent N. 001416412, June 11, 2015).

6 CONCLUSIONS

We have shown how an Internet of Speaking Things can become at least as impactful as an Internet of Sensing Things.

In fact, while one is already widening the perspectives and the applications of Big Data Analytics and computer-supported Decision Making, the other has the potential to open a radically new view on man-machine interfaces, where things of all kind bring to users the information residing on the cloud.

A novelty aspect of the proposed approach, when compared to the state of art in the smart applications supporting CH field, is the strong human driven communication strategy. Currently, prototype versions of the system are able to manage and automatically recognize two languages (Italian and English), but extensions for supporting more other European and Asiatic languages are ready to be integrated, evenly supported by linguistic specialists.

The adopted approach promises to be scalable and flexible enough to support extensions for other types of interfaces or application domains, when specific domain ontology and lexical resources are available to manage Natural Language driven interactions.

Open issues concern the robustness of the technological solutions supporting our approach, against environment or infrastructural faults. Refinements, supported by a massive testing action, have to be introduced in order to assure real-time interactions, when environment is too noisy or network latencies are over acceptable rates.

Another open issue, aim of future investigations, is the absence of a standard evaluation metrics to establish a human-machine interaction quality baseline.

In the case studies that we reported we focused on things full of meanings handed down from the past, such as speaking statues and interactive holograms that embody digital resurrection of historical characters and give them effective interactive capabilities.

But nothing prevents us from pursuing equally exciting applications with objects and situations that belong to everyday life, from the speaking fridge asking for instructions to shop to the holographic butler that coordinates appliances within an automated home. As in all new areas, the sky is the limit to the possibilities that open up, and there are so many things that have interesting stories to tell their users!

REFERENCES

Zafarani, R., Abbasi, M. A. and Liu, H., 2014. Social media mining: an introduction. Cambridge University Press.

Xylomenos, G., Ververidis, C. N., Siris, V. A., Fotiou, N., Tsilopoulos, C., Vasilakos, X., Katsaros, K.V. and Polyzos, G.C., 2014. A survey of information-centric networking research. Communications Surveys & Tutorials, IEEE, 16(2), pp.1024-1049.

Ahlgren, B., Dannewitz, C., Imbrenda, C., Kutscher, D. and Ohlman, B., 2012. A survey of information-centric networking. Communications Magazine, IEEE, 50(7), pp.26-36.

Singh, D., Tripathi, G., Jara, A. J., 2014. A survey of Internet-of-Things: Future Vision, architecture, challenges and services. In *Internet of Things (WF-IoT), 2014 IEEE World Forum on*, vol., no., pp. 287-292, March 6-8, 2014.

Atzori L., Iera A., and Morabito G., 2010. The Internet of Things: A survey. In *Computer Networks*, vol.54, no.15, pp. 2787-2805.

Amato, F., Chianese, A.,Moscato, V.,Picariello, A., Sperli, G., 2012. SNOPS: a smart environment for cultural heritage applications. In proceedings of the *12th International Workshops on Web Information and Data Management*., pp. 49-56.

Amato F., Chianese A., Mazzeo A., Moscato V., Picariello A., Piccialli F., 2013. The talking museum project. In *Procedia Computer Science*, Vol. 21, pp. 114-121.

Chianese, A., Marulli, F., Moscato, V., Piccialli, F., 2013. A smart multimedia guide for indoor contextual navigation in Cultural Heritage Applications. In proceedings of *2013 Workshop on Location-based services for Indoor Smart Environments (LISE2013)*, collocated in 2013 International Conference on Indoor Positioning and Indoor Navigation, IPIN2013, Montbeliard, France, October 28-31, 2013.

Marulli, F., Chianese, A., Moscato, V., Piccialli, F., 2013. "SmARTweet: A Location-based smart application for Exhibits and Museums". In proceedings of *Workshop on Cultural Information Systems (CIS2013)*, collocated with *the 9th International Conference on Signal Image Technology & Internet Based Systems (SITIS2013)*, Kyoto, Japan, December 2-5, 2013.

Bordoni, L., Ardissono, L., Barceló, J. A., Chella, A., de Gemmis, M., Gena, C., Iaquinta, L., Lops, P., Mele, F., Musto, C., Narducci, F., Semeraro, G., Sorgente, A. 2013. The contribution of AI to enhance understanding of Cultural Heritage. In *Intelligenza Artificiale*, vol. 7, no. 2, pp. 101-112, 2013.

Semeraro, G., Lops, P., d, M., Musto, C., Narducci, F, 2012. Folksonomy-based recommender system for personalized access to digital artworks. In *JOCCH*, vol. 5, issue 3, no. 11, 2012.

Valente, I., Chianese, A., Marulli, F., Piccialli, F., 2013. "A novel challenge into Multimedia Cultural Heritage: an integrated approach to support cultural information enrichment". In Proceedings of *the 9th International Conference on Signal Image Technology & Internet Based Systems (SITIS2013)*, Kyoto, Japan, December 2-5, 2013.

Kang, S. K., Kang, H. K., Kim, J. E., Lee, H., Lee, J. B., 2012. A study on the mobile communication network with smart phone for building of location based real time reservation system. In *IJMUE*, vol.7 no.2, pp 17-36.

Husain W., Dih, L. Y., 2012. A framework of a Personalized Location-based traveler recommendation system in mobile application. In *International Journal of Multimedia and Ubiquitous Engineering*, vol. 7, no.3, pp. 11-18.

Marulli, F., 2015. IoT to enhance understanding of Cultural Heritage: Fedro authoring platform, artworks telling their fables. In proceedings of 1st *EAI International Conference on Future Access Enablers of Ubiquitous and Intelligent Infrastructures (FABULOUS 2015)*, Ohrid, Republic of Macedony, September 23-25, 2015.

Cremene, M., Tigli, J. Y., Lavirotte, S., Pop, F., Riveill, M., Rey, G., 2009. Service Composition based on Natural Languages Requests. In Proceedings of *IEEE International Conference on Services Computing 2009 (SCC '09)*, pp. 486-489.

Chianese, A., Marulli, F., Piccialli, F., Benedusi, P., 2015. An Associative Engines Based Approach supporting Collaborative Analytics in the Internet of Cultural Things. In proceedings of the *3rd International Workshop on Cloud and Distributed System Application and he 10th International 3PGCIC-2015 Conference,* Krakow, Poland, November, 4-6, 2015.

Benedusi, P., Marulli, F., Racioppi, A., Ungaro, L., 2015. What's the matter with Cultural Heritage tweets? An Ontology–based approach for CH Sensitivity Estimation in Social Network Activities. In *Proceedings of 11th International Conference on Signal Image Technology and Internet Based System (SITIS2015),* Bangkok, Thailand, November 23-27, 2015.

Smart Beacon, 2015, http://www.smartbeacon.it/

Cogito, 2016. Cogito Intelligence Platform – Expert System, http://www.expertsystem.com/cogito/

Baharuddin, R., Singh, D. and Razali, R., 2013. Usability dimensions for mobile applications—A review. *Res. J. Appl. Sci. Eng. Technol,* 5, pp.2225-2231.