A Collaborative Approach for Metadata Management for Internet of Things

Linking Micro Tasks with Physical Objects

Umair ul Hassan, Murilo Bassora, Ali H. Vahid, Sean O'Riain, Edward Curry

Digital Enterprise Research Institute
National University of Ireland
Galway, Ireland
{umair.ul.hassan, murilo.bassora, ali.vahid, sean.oriain, ed.curry}@deri.org

Abstract—There has been considerable efforts in modelling the semantics of Internet of Things and their specific context. Acquiring and managing metadata related to the physical devices and their surrounding environment becomes challenging due to the dynamic nature of environment. This paper focuses on managing metadata for Internet of Things with the help of crowds. Specifically, the paper proposes a collaborative approach for collecting and maintaining metadata through micro tasks that can be performed using variety of platforms e.g. mobiles, laptops, kiosks, etc. The approach allows non-experts to contribute towards metadata management through micro tasks, therefore resulting in reduced cost and time. Applicability of the proposed approach is demonstrated through a use case implementation for managing sensor metadata for energy management in small buildings.

Keywords—sensors, metadata, micro-task, collaboration

I. Introduction

One primary goal of the Internet of Things (IoT) is to enable humans and machines to better understand their surrounding environment, using the information generated by a variety of sensing devices [1]. This understanding is based on three types of data; the data generated by the devices, the data describing the devices, and the data describing the environment [2]. Low cost deployment of sensing devices has become relatively easy with the technical advancements in sensor technologies. However, descriptions of sensing devices still remain a challenge, especially in dynamic environments such as Smart Buildings and Smart Cities. The expertise required for maintaining IoT based systems can become a challenging factor in their successful deployment.

Typically, IoT devices are semantically described in terms of their sensing capabilities. The semantics of the environment are described according to the domain of application [3]. For instance, in the case of Smart Buildings the semantic description may contain data about floors, rooms, windows, etc. Consequently, decision support models are built using the metadata that describes both devices and their environment. Current research efforts have mainly focused on the challenges of interoperability, scalability, and integration for IoT [3]. The *dynamism* challenge of metadata in IoT has been largely overlooked. With this in mind, we argue that availability of high quality metadata is fundamental to effective decision making within dynamic IoT based environments.

Autonomous management of sensing devices [4] and the self-adaptation of IoT based systems is considered a key research problem in the dynamic environments [5]. In this regard, we propose a collaborative approach for metadata acquisition and management for IoT. In the proposed approach humans located near or around the sensing devices are involved in the management process by outsourcing micro tasks in a crowd sourcing manner [6].

Human involvement in IoT has been highlighted in research fields such as participatory sensing [7]. The underlying objective is to leverage humans along with sensing devices to monitor an environment and deliver personalized services. The main contributions of this paper are as follows

- A collaborative approach for metadata management through non-experts for IoT based systems.

- A location and availability based approach of task assignment in metadata management for IoT.

- A use case implementation of the proposed approach in the Smart Buildings use case for Energy Management.

The rest of the paper is organized as follows. Section II provides an overview of the proposed collaborative approach for metadata management in IoT systems. Section III describes the location and availability based technique for assignment of tasks to appropriate humans. Section IV details the use case implementation of metadata management system in Smart Building. Section V discusses the implications and limitations of the approach. Section VI summarizes the related research work. Section VII concludes with summary of research findings and future directions.

II. Collaborative Metadata Management

In following we highlight the main components of a system that implements collaborative management of metadata for IoT systems. We propose that the community of users can contribute towards the management process by performing micro tasks. Fig. 1 illustrates the main components of the system and
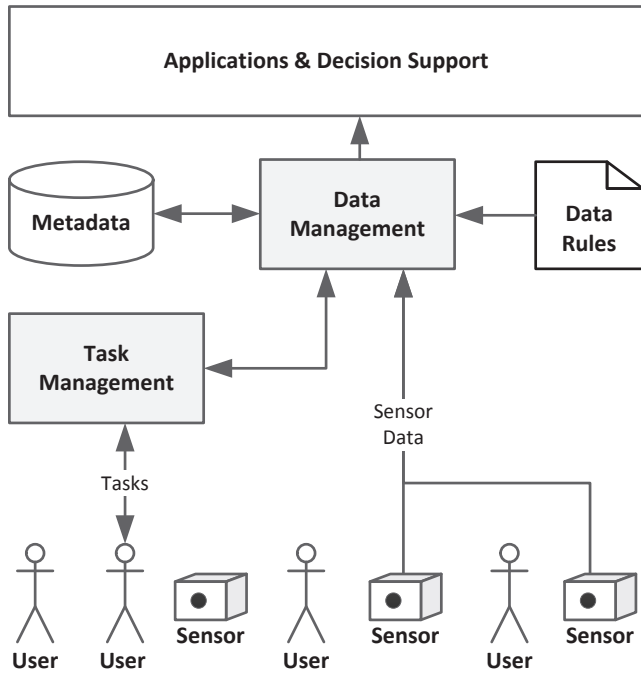
Figure 1. System overview of collaborative metadata management for Internet of Things.
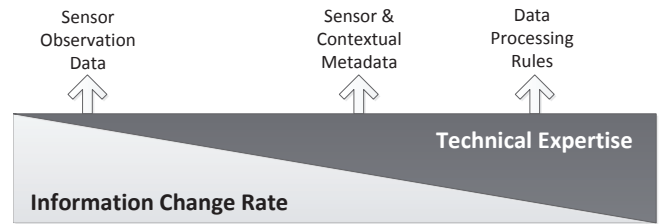

Figure 2. The continuum of rate of change for sensor data, metadata and rules in dynamic environments against the technical expertise required from users of the IoT based systems.

their communication paths. The rest of this section provides further details of two major components: *Data Management* and *Task Management*.

A. Data Management for IoT

The main objective of the *Data Management* component is to present a clean and consistent view of an IoT environment. This view is generated based three types of information [2], as described below:

- **Sensor observation data** is the actual measurements generated by the sensors. For example, a temperature sensor produces measurement of heat levels.

- **Sensor metadata** is the model that describes the sensors and their capabilities. For instance, the Sensor Modeling Language [8] and the Semantic Sensor Networks Ontology [9] are set of standards developed for representing sensor information in XML[1] and RDF[2] formats, respectively. **Contextual metadata** describes the environmental context in which sensors are deployed. For instance, a temperature sensor might be deployed in a room of an office building.

- **Data processing rules** are specified to help standardize the managed data for applications and decision support. These rules can be guidelines for data formatting, mappings between instances, or dependencies between attributes.

Sensor observation data is in the form of streams and events that keep changing at a high rate. The metadata and rules may require frequent updates in a dynamic environment, for instance due to installation of additional heater in a room. The rate of change is highest for the sensor data and lowest for the rules, as shown in Fig. 2. Conversely, the technical expertise required to preform relevant updates is lowest for sensor data and highest for rules. We propose that people with relatively low expertise can be involved in the metadata management process through collaboration. We assume that people in the monitored environment are willing to perform micro tasks voluntarily. Therefore, the *Data Management* component generates tasks based on the validity of metadata according to the data processing rules. These tasks are then submitted to *Task management* component for execution.

B. Task Management of IoT

The primary purpose of *Task Management* component is to provide a middleware for access to the users in the monitored environment. Furthermore, the *Task Management* is decoupled from the *Data Management* for the purpose of encapsulation. The three core functions of the *Task Management* are:

- **Task assignment** is concerned with matching between tasks and users of the monitored environment [10]. The matching can be based on characteristics of tasks or the specific requirements of tasks in terms of human capabilities of users [11].

- **Task presentation** involves rendering of task contents according to the platform used by human participants. A task should be displayed according to the delivery channel used such as mobiles, laptops, kiosks, etc.

- **Quality assurance** includes ensuring truthful and correct responses of tasks. This can be achieve by assigning same task multiple times or comparing task responses against gold truth data.

In the next section we discuss the specifics of task assignment from the perspective of IoT based environments.

III. TASK ASSIGNMENT

Effective assignment of tasks to appropriate users at right time is critical to dynamic IoT environments. Therefore information about a user's location and availability are required to assign tasks related to devices and "Things" around them.

---

[1]Extensible Markup Language
[2]Resource Description Framework

Figure 3. Example for a sensor and its QR code. The QR when scanned returns a URL for the tasks associated with the sensor.

There are a variety of methods for sourcing the location and availability information about users. In this work we propose a mixture of pull and push methods for sourcing participant's location for making assignment decisions.

1) Task pull: The linkage between physical sensing devices and tasks is made with the help of *Quick Response* (QR) codes. Fig. 3 illustrates an example of a QR code attached to a sensor. Each QR code represents an encoded URL for the sensor tasks. Resolving the URL through a browser renders the tasks associated with the sensor. Pull based assignment is suited for the situations where the information about sensors and users is not available.

2) Task Push: The data management tasks can be actively pushed to users in situations where the location and availability information of users is available. To achieve this the *Task Management* components keeps track of the sensor location information submitted by users and then pushes tasks that require description of sensor surroundings to nearby users.

Whether the pull or push method is used for assignment depends on the data management objective of the task. Therefore *data management* component must specify the assignment method for tasks, as well as the requirements in terms of human capabilities [11] .

IV. IMPLEMENTATION

We implemented the proposed collaborative approach to demonstrate its applicability in IoT based systems. The implementation was targeted at energy management systems for small commercial buildings [12]. The purpose of an energy management system is gather data from multiple sensors that are deployed in the building to monitor the energy consumption [13]. The energy consumption patterns are further analyzed to take suitable decision for energy optimization.

Our collaborative metadata management implementation was developed to support the energy management system that was deployed in the Digital Enterprise Research Institute (DERI) in National University of Ireland, Galway. The DERI

building has approximately 130 occupants including research students and staff. The building comprises 22 office units, 160 open plan workspaces, 6 conference and meeting rooms, 3 kitchens, a data center, a sensor network laboratory, and a café. In total 14 Constrained Application Protocol (CoAP) sensors were deployed in DERI at various locations, with their respective QR codes attached. Fig. 4 shows our collaborative metadata management system implemented for energy management in a small building.

The rest of this section details a linked dataspace, building energy data and micro tasks for energy management, along with results of an experiment to validate our approach.

A. Linked Dataspace for Energy Management

An enterprise energy management system have been developed in DERI based on the principles of *Dataspaces*, *Linked data* and *Complex Event Processing* to enable real-time holistic enterprise energy management [14]. A Dataspace for energy management using supports the coexistence of heterogeneous data sources with an incremental approach for integration. The Linked data enable interoperability between data sources using web-based standards. The inclusion of Complex Event Processing with sensor networks provides support for real-time processing of dynamic and background data for energy situation awareness. These technologies have been combined within a single architecture called the Linked dataspace for Energy Intelligence (LEI) that can be used to build energy management applications. Our collaborative approach based on micro tasks supports the incremental improvement of LEI through collaboration.

B. Building Energy Data

The sensor observation data in this scenario is mainly temperature measurement produced by the CoAP sensors. This data enables optimization of energy usage in the building. However, the decision support models, that use sensor observation data, require high quality metadata in terms of sensor locations and surrounding environment. The following data shows a sensor observation in JSON format:

```
{
  "sensorId": "coap_sensor_001",
  "timestamp": "2013-07-14 19:43:37",
  "reading": "25",
  "type": "temperature",
  "unit": "C"
}
```

In the scenario required sensor metadata includes entities and their attributes describing sensors and rooms. Each sensor entity was described in terms of its sensing capabilities and location. Each room entity was described in terms of devices such as bulbs, heaters and windows. The states of room devices influenced the energy optimizations decisions. Data processing rules were specified for sensor and room entities to ensure that there were no missing values. According to the specified data rules, human tasks were created on the task management engine described next. The following data describes a sensor and a room in JSON format:
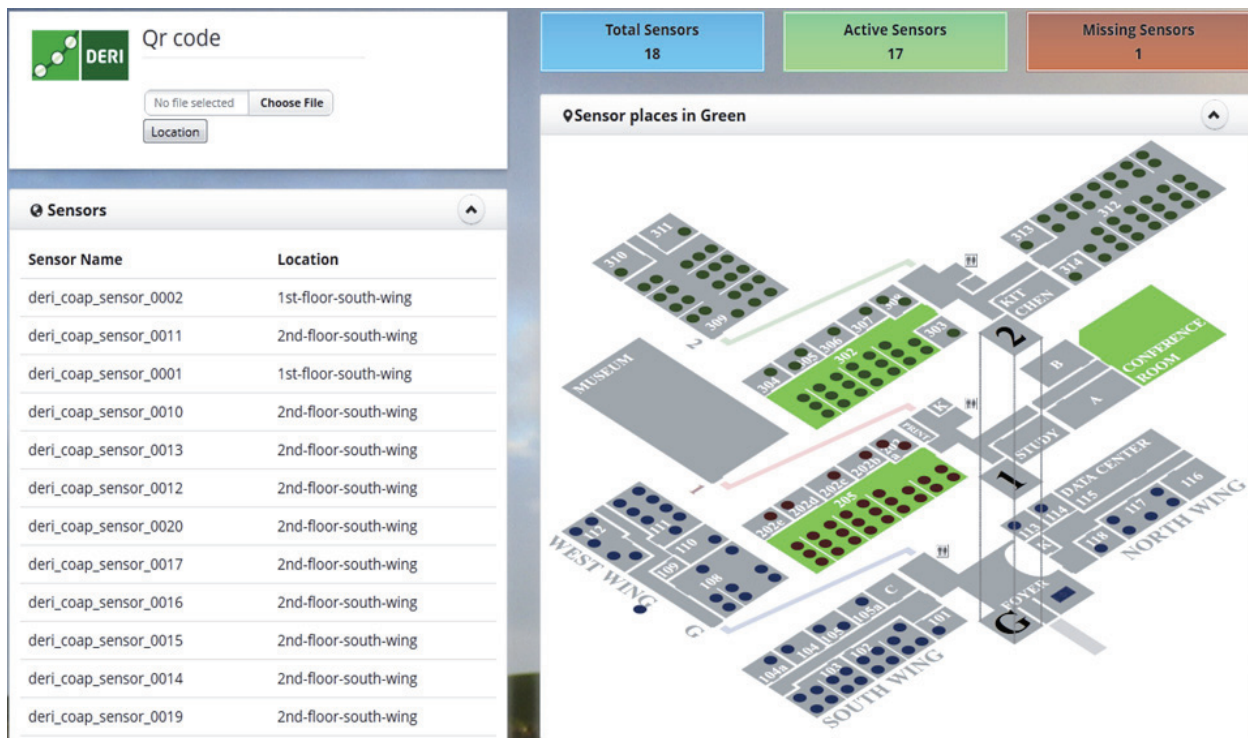
Figure 4. Screenshot of the collaborative metadata management system for energy management in a small building. The green color on the map shows the sensors located with the help of building occupants.

```
{
  "sensorId": "coap_sensor_001",
  "sensorName": "DERI CoAP Sensor 1",
  "location": "r112"
  "platform": "desk"
}
{
  "roomId": "r112",
  "wing": "2nd-floor-south-wing",
  "lights": 5,
  "windows": 2,
  "heaters": 2
}
```

TABLE I. LIST OF TASKS DEFINED IN THE TASK MANAGEMENT ENGINE FOR ACQUISITION OF METADATA FOR SENSORS AND ROOMS USING THE HUMAN PARTICIPANTS

Task	Description	Assignment Method
Sensor Location	This task requires participants to specify the location of sensor	Task pull based on QR code
Room Lights	This task asks participants to specify number of fluorescent light installed in the room	Task push based on person location
Room Heaters	This task asks participants to specify the number for heaters in the room	Task push based on person location
Room Windows	This task asks participants to specify the number of windows in the room	Task push based on person location

C. Micro Tasks

We implemented a task management engine that allowed human participants to access tasks from mobiles devices. Table I list the tasks defined in the task management engine. One task collects location information for the sensor with the help of QR code that is attached to the sensor. In this case, participants pull tasks by scanning a QR code and resolving the associated URL. Therefore, QR codes enable the linkage between human tasks and physical sensors.

Once the location of the sensor is submitted by participants, further tasks are pushed to them to collect further metadata about the surrounding environment. There are three tasks that collect information about fluorescent lights, heaters, and windows in the room. Provided that the metadata about sensors and rooms along with the sensor measurements is complete,

the energy management system makes recommendation about temperature control in the monitored building.

D. Experiment & Results

We evaluated the implemented system in terms of its utility for collecting high quality metadata about sensors and rooms. The metadata about sensor locations and room characteristics was not available at start of the experiment. This situation simulates the case when its difficult to gather all metadata upfront or the metadata become invalid due to changes in the environment. The objective of the experiment was to use the tasks defined beforehand for collection of metadata through building occupants. For this purpose, the occupants of the building were contacted through email to participate in the experiment. The occupants were asked to look for sensors

TABLE II. ACCURACY OF THE METADATA CONTRIBUTED BY
BUILDING OCCUPANTS AGAINST THE GOLD STANDARD DATA

Task	Accuracy
Sensor Location	85.71%
Room Lights	100%
Room Heaters	83.33%
Room Windows	100%

around them and scan the QR code on the sensors using mobiles phones. Then resolve the URL associated with QR code in a web browser and perform the available task.

The evaluation is based on the comparison of occupant contributed metadata versus gold standard data. The gold standard data was create manually. Table II shows the accuracy of data submitted by occupants of the building, with 5 hours of sending the email for participation. The reported accuracy is based on the data submitted by the first participants for each sensor and room. As can be seen, our approach achieved more than 80% accuracy in describing the sensors and rooms within 5 hours. These results provide preliminary evidence of the applicability of proposed approach. We discuss the limitation of the experiment in next section.

V. DISCUSSION

The main strength of the proposed approach that was observed in the use case implementation as its simplicity. A simple scan of QR code enabled building occupants to make contributions of metadata . In this sense the QR codes and their associated URLs enable the linkage between sensors, metadata tasks, and humans. In addition, the sensor location, submitted by a person, was used to assign further tasks related to the room in which the sensor was located. This enabled delivery of tasks to the right person at the right location. Since some rooms contained more than one sensor the room descriptions were correct although some sensor location were incorrect.

The results indicate the positive potential of collaborative metadata management. However, the suitability and scalability of the approach is yet to be established within different application domains. We observed that the relatively straightforward user interface and short completion time of tasks is important for non-expert contribution. The motivation of aspect of the approach deserves further investigation.

VI. RELATED WORK

The related literature is categorized into three main areas of research: *sensor data management*, *participatory sensing*, and *context awareness for IoT*.

While the core subject of data management in IoT has the data streams generated by sensors [15] there have been some efforts addressing the issue of metadata management. For instance, Dawes et al. [16] developed a wiki based metadata and provenance management system for sensors. Jeung et al. [17] proposed a real-time approach for fusion and management of sensor stream data and metadata. Fortuna et al. [18] compared embedded approaches against middleware approaches for metadata management for IoT. Most of these approaches are expert oriented as compared to our approach that focuses on the inclusion of non-experts through collaboration in metadata management.

Participatory sensing [7], crowd sensing [6], and citizen sensing [19], [20] are similar concepts describing the process of data collection through people. The collected data is generated by people or the mobile devices owned by them. In this case people are considered as the sensing devices. The main objective of these approaches is to utilize humans for sensing the monitored environment. By comparison, our approach utilizes humans for managing the metadata hence it is complementary to the participatory sensing.

Context awareness has been recognized as the major feature of next generation middleware for Internet of Things [21]. Ganz et al. [22] proposed an automated approach of inferring sensor context through association and registration mechanisms. Carrera et al. [23] utilize mobile based information to enable context awareness within crowd sensing. By contrast, our approach proposes context generation and management with the help of people.

VII. CONCLUSION AND FUTURE WORK

This paper addresses the problem of metadata management for IoT based systems in dynamic environments. A collaborative approach is proposed where simple tasks are outsourced to crowds for the management of metadata for sensors and their environment. Specifically, we propose a capability based approach for assignment of tasks to humans based on their location and availability. The applicability of the approach is demonstrated in use case implementation based on energy management system in small office building.

As part of future work we plan to expand the experiment the examine the long term metadata management of IoT. In this regard, the main challenges to be addressed include user interface design, human motivation, task redundancy, etc. A generic human task platforms for variety of tasks in cyberphysical systems is another interesting direction of the research work. For example, if a room is considered to be abnormally cold even when the heating is on then the person nearest to the room can be requested to check if the windows are open.

REFERENCES

[1] J. Gubbi, R. Buyya, S. Marusic, and M. Palaniswami, "Internet of things (iot): A vision, architectural elements, and future directions," *Future Generation Computer Systems*, vol. 29, no. 7, pp. 1645–1660, 2013.

[2] P. Barnaghi, W. Wang, C. Henson, and K. Taylor, "Semantics for the internet of things: Early progress and back to the future," *International Journal on Semantic Web & Information Systems*, vol. 8, no. 1, pp. 1–21, 2012.

[3] C. C. Aggarwal, N. Ashish, and A. Sheth, "The internet of things: A survey from the data-centric perspective," in *Managing and Mining Sensor Data*. Springer, 2013, pp. 383–428.

[4] D. Pfisterer, K. Romer, D. Bimschas, O. Kleine, R. Mietz, C. Truong, H. Hasemann, A. Kroller, M. Pagel, M. Hauswirth *et al.*, "Spitfire: Toward a semantic web of things," *Communications Magazine, IEEE*, vol. 49, no. 11, pp. 40–48, 2011.

[5] H.-D. Ma, "Internet of things: Objectives and scientific challenges," *Journal of Computer science and Technology*, vol. 26, no. 6, pp. 919–924, 2011.

[6] R. K. Ganti, F. Ye, and H. Lei, "Mobile crowdsensing: Current state and future challenges," *Communications Magazine, IEEE*, vol. 49, no. 11, pp. 32–39, 2011.

[7] D. Estrin, "Participatory sensing: applications and architecture [internet predictions]," *Internet Computing, IEEE*, vol. 14, no. 1, pp. 12–42, 2010.

[8] M. Botts and A. Robin, "Opengis sensor model language (sensorml) implementation specification," *OpenGIS Implementation Specification OGC*, pp. 07–000, 2007.

[9] M. Compton, P. Barnaghi, L. Bermudez, R. García-Castro, O. Corcho, S. Cox, J. Graybeal, M. Hauswirth, C. Henson, A. Herzog *et al.*, "The ssn ontology of the w3c semantic sensor network incubator group," *Web Semantics: Science, Services and Agents on the World Wide Web*, vol. 17, pp. 25–32, 2012.

[10] U. Ul Hassan, S. O'Riain, and E. Curry, "Slua: Towards semantic linking of users with actions in crowdsourcing," in *Proceedings of 1st International Workshop on Crowdsourcing the Semantic Web*, 2013.

[11] U. Ul Hassan, , and E. Curry, "A capability requirements approach for predicting worker performance in crowdsourcing," in *9th IEEE International Conference on Collaborative Computing: Networking, Applications and Worksharing*, 2013.

[12] E. Curry, J. O'Donnell, E. Corry, S. Hasan, M. Keane, and S. O'Riain, "Linking building data in the cloud: Integrating cross-domain building data using linked data," *Advanced Engineering Informatics*, vol. 27, no. 2, pp. 206–219, 2013.

[13] E. Curry and B. Donnellan, "Sustainable information systems and green metrics," *Harnessing Green It: Principles and Practices*, pp. 167–198.

[14] E. Curry, S. Hasan, and S. O'Riain, "Enterprise energy management using a linked dataspace for energy intelligence," in *Sustainable Internet and ICT for Sustainability (SustainIT), 2012*. IEEE, 2012, pp. 1–6.

[15] D. Carney, U. Çetintemel, M. Cherniack, C. Convey, S. Lee, G. Seidman, M. Stonebraker, N. Tatbul, and S. Zdonik, "Monitoring streams: a new class of data management applications," in *Proceedings of the 28th international conference on Very Large Data Bases*. VLDB Endowment, 2002, pp. 215–226.

[16] N. Dawes, K. A. Kumar, S. Michel, K. Aberer, and M. Lehning, "Sensor metadata management and its application in collaborative environmental research," in *Proceedings of the 2008 Fourth IEEE International Conference on eScience*. IEEE Computer Society, 2008, pp. 143–150.

[17] H. Jeung, S. Sarni, I. Paparrizos, S. Sathe, K. Aberer, N. Dawes, T. G. Papaioannou, and M. Lehning, "Effective metadata management in federated sensor networks," in *Proceedings of the 2010 IEEE International Conference on Sensor Networks, Ubiquitous, and Trustworthy Computing*. IEEE Computer Society, 2010, pp. 107–114.

[18] C. Fortuna, P. Oniga, Z. Padrah, M. Mohorcic, and A. Moraru, "Metadata management for the web of things: a practical perspective," in *Proceedings of the Third International Workshop on the Web of Things*. ACM, 2012.

[19] A. Sheth, "Citizen sensing, social signals, and enriching human experience," *Internet Computing, IEEE*, vol. 13, no. 4, pp. 87–92, 2009.

[20] D. N. Crowley, E. Curry, and J. G. Breslin, "Closing the loop-from citizen sensing to citizen actuation," *7th IEEE International Conference on Digital Ecosystem Technologies*, 2013.

[21] C. Perera, D. Georgakopoulos, A. Zaslavsky, and P. Christen, "Ca4iot: Context awareness for internet of things," in *IEEE International Conference on Green Computing and Communications*. IEEE, 2012, pp. 775–782.

[22] F. Ganz, P. Barnaghi, F. Carrez, and K. Moessner, "Context-aware management for sensor networks," in *Proceedings of the 5th International Conference on Communication System Software and Middleware*. ACM, 2011.

[23] I. Carreras, D. Miorandi, A. Tamilin, E. R. Ssebaggala, and N. Conci, "Crowd-sensing: why context matters."