A Semiotic Approach to EUD for the Internet of Things

Barbara Rita Barricelli, Stefano Valtolina

Department of Computer Science, Università degli Studi di Milano
Via Comelico 39/41, 20135 Milano, Italy
{barricelli, valtolina}@di.unimi.it

Abstract. With this paper we describe the ongoing research we are developing in the Internet of Things (IoT) domain. The description of IoT as an ecosystem of objects and services highlights the central role of the end user in extracting, merging, analyzing, visualizing, and sharing data. We also provide a discussion of data under different perspectives to underline how promising the IoT field is in terms of possibilities for End-User Development activities with a specific focus on the semiotic approach we use in our research.

Keywords: Internet of Things, EUD, semiotics.

1 Introduction

Internet of Things (IoT) [1] is the evolution of the 1970s lifelogging, initially conceived as a 24/7 broadcasting of self-videos. IoT has become today a wide spreading phenomenon strictly related to the so-called quantified-self movement, that allows people to keep track of their habits, health conditions, physiological data, and behavior, and to monitor conditions and quality of the environments in which they work and live. The spread of such technology has become possible thanks to the evolution of the devices, both small/wearable and unmovable, and designed for home or office use: they are now easily affordable to the masses and can be easily interconnected by means of broadband Internet connections. Today, IoT is successfully adopted in several application domains and it is estimated that in 2015 the number of objects connected will be around 12 billion, while in 2020 it will be 50 billion [2, 3]. IoT allows the end users to manage physical devices, interactive systems, and lifelogging data by deciding how to create new usage scenarios and this empowers them more than ever, making them evolve from passive end users to active end users and to some extent also to end-user developers [4]. However, lifelogging tools produce a very large quantity of data especially in the long term and when shared with other people, leading to a very challenging information overload. The collected data have to be organized and aggregated in order to enhance the knowledge of the user and also the other persons involved. In particular, End-User Development (EUD) systems able to filter and visualize large quantity of data, making them available on mobile devices with reduced display size are needed; it must be left to the user to decide the level of detail that it is needed to show and to have control on the use of the data granting a good level of privacy and security.

Proc. of Third International Workshop on Cultures of Participation in the Digital Age - CoPDA 2015
Madrid (Spain), May 26th, 2015 (published at http://ceur-ws.org).

2 Ecosystem of IoT

IoT concept is based on an ecosystem of elements (hardware and software) that need to be combined and connected through the Internet to exchange data, and to act and react according to events, and/or users' preferences, rules, or decisions [5].

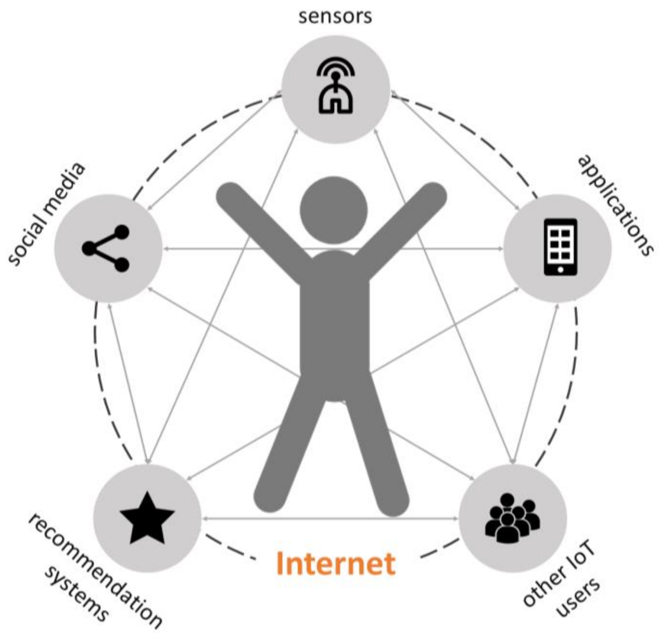

Fig. 1. Ecosystem of the Internet of Things.

Designing for IoT means to put at the center of this ecosystem the end user, the one who actually generates the data, manages the IoT elements in the ecosystem, and personalizes the IoT environment defining the interactions among the elements and the elements' behavior. The elements of the IoT ecosystem (depicted in Figure 1) can be categorized into five groups:

1. **Sensors**: typically built-in components in electronic devices aimed at collecting data of various nature. Examples are sensors embedded in weather stations, activity tracking armbands, or Wi-Fi body scales. For their nature, IoT devices can be mobile – meant to follow the user everywhere (e.g. activity trackers) – or unmovable – designed for being placed in a specific place and not moved around (e.g. weather stations). The data collected through the sensors can be treated mainly in two ways: they can be sent directly by the sensors or they can be gathered by the user when needed. Sensors typically come with proprietary applications that enable the access to both settings and data through an ad-hoc interface.

Proc. of Third International Workshop on Cultures of Participation in the Digital Age - CoPDA 2015
Madrid (Spain), May 26th, 2015 (published at http://ceur-ws.org).

2. **Applications**: the access points for the end users to access the data of the IoT devices. They can be bundled with specific devices or compatible with several devices. Applications are usually designed to be mobile-compliant giving the end user the chance of interacting with the devices on remote setting.
3. **Social media**: the communication channels that today are chosen by end users to share the data over the Internet. Most of the IoT devices and applications are ready to send over social media and share data with virtual or real communities.
4. **Recommendation systems (RSs)**: intelligent systems aimed at suggesting aggregation, integration and distribution ways for the gathered IoT data. The suggestions are based on end users profile, needs, and habits, so as on the behavior of the communities they belong. However, the use of automatic suggestions might be not appreciated by the end user that may feel frustrated in using the recommendation features, whenever they appear to be inappropriate. To deal with this problem, some RSs offer solutions able at exploiting end user's social relationships for improving the service quality.
5. **Other IoT users**: the communities' population that typically share with the end user some particular interests, life choices, or other characteristics. It is the end user who chooses the people to be connected with on the basis of personal searches or suggestions made by the applications (also thanks to the recommendation systems).

3 End-User Development for IoT

As an extension of Jennings et al. [6] research on human-agent collectives (HACs), we suggest to analyze IoT sensors/devices and the EUD activities that can be implemented on them, considering the main characteristics of the data that are gather and conveyed: type, usage, elaboration, and presentation.

Data type. With data type, we mean to describe the kind of data that are gathered and deployed by sensors embedded in IoT devices. According to the type of device, we can first distinguish between personal data (originated by personal sensors) and ambient data (streaming from ambient sensors). It is important to consider that sometimes the users need to fill some gaps in the data collection by manually input data by themselves. This can happen when they are using devices that are not provided with Wi-Fi connections or when the data are related to subjective states of the user (e.g. mood). The data collected through IoT devices and related applications can be classified in two main groups: quantitative and qualitative data. Quantitative data can be further categorized in two classes: automatic and manual. Automatic data are collected through electronic devices either with or without intentional user actions. In case of unintentional user actions, the device is able to recognize autonomously if the user is starting a specific activity, e.g. run, walk, sleep. Intentional user actions, on the other hand, are those performed by the users to tell the device the kind of activity they are going to perform in a specific moment. This happens in those devices that are not designed to detect in real time the change of the user activities. For example, since sleeping is quite complex to being detected, the majority of lifelogging devices need the user to instruct the device

Proc. of Third International Workshop on Cultures of Participation in the Digital Age - CoPDA 2015
Madrid (Spain), May 26th, 2015 (published at http://ceur-ws.org).

at the beginning of such activity. The data that the user needs to input to an application because the device is not directly connectible to the Web are named manual quantitative data. Examples of such devices are non-IoT devices or devices that are not compatible with specific applications. Furthermore, there are some data that cannot be gathered by electronic devices and for which the human intervention is fundamental (e.g., the kind of food that the user eats, the type of exercises that the user practices in the gym). Qualitative data need to be classified in two main groups: objective and subjective. Objective data are originated from the observation of facts. An example of qualitative objective data is the state of sleep: electronic devices can detect the different sleep stages during the night and offer an interpretation of them that lead to results like “light sleep”, “deep sleep”, and “awake”. Subjective data, on the other hand, are qualitative data that cannot be interpreted by an electronic device and its applications. For example, many applications ask the user to tell their mood and most of them use an emoticon scale to express it. While subjective data is completely in the hands of the user, objective data is gathered by a device but is not meant to be always trustful: it is in fact important to give the user the chance to correct the objective data with their own interpretation. If, for example, a user knows for sure that she was not asleep at a specific time, she should be able to rectify the data produced by the device. From the data type perspective, the EUD activity that can be enabled regards the configuration of data collection, giving the end user the possibility of deciding about the type of data to collect and the way of collection.

Data usage. IoT devices and related applications allow to use data in different ways, each one having different levels of complexity. Complexity of data use can be measured by studying the level of interaction that is expected by the users for exploiting the IoT ecosystem that they use. The more the user is asked to interact with a device and/or an application, the more the complexity grows but with it grows also its potentials. With the analysis of the state of the art of IoT devices that we performed in the last months of this year (2015), we identified four main classes of data usage depending on the level of human intervention in the process: non-interactive/automatic, low interaction/semi-automatic, feedback-based/recommendation-based, and highly-interactive/EUD. Non-interactive or automatic applications and related devices, after a first configuration for which the user intervention is needed, are able at collecting, aggregating and elaborating data into information in automatic and non-interactive ways. There is no need for the user to make any decisions and the IoT’s role here is to silently monitor the user and to present the results of this monitoring via the application whenever the user accesses it. Low interaction or semi-automatic devices and applications need a minimal human intervention for working correctly. As an example, there are those devices that need the user to specify that a specific activity is starting. The user role is fundamental for the correct initialization of monitoring but the required interaction is very low. Feedback-based or recommendation-based applications and devices that work on feedbacks and recommendations are strictly connected with the behavior of their user and the member of the communities with whom they share their data. In this context of use, the devices’ behavior changes accordingly to the user(s) behavior and the social dimension of IoT becomes central. Highly interactive or EUD class of devices and applications

Proc. of Third International Workshop on Cultures of Participation in the Digital Age - CoPDA 2015
Madrid (Spain), May 26th, 2015 (published at http://ceur-ws.org).

allow the user to actively change the behavior or the single device/application or of the network of devices and applications by becoming in some extent the developer of their own IoT system. A very high interaction is needed for the users to influence the behavior of the IoT system but there is no need for them to learn a programming language. Lifelogging brings important implications on the sociotechnical side: configuring and even programming IoT devices is becoming suitable for all thanks to the design of interfaces for supporting customization, personalization and to some extent also EUD. Today, this is mainly implemented in two ways: a) through rules definition-wizards that rely on the states of sensors/devices, or b) by providing a visual pipeline generator for letting the end user creating aggregation, filtering, and porting of data originated by sources.

Data elaboration. Proper detection, aggregation, integration, filter and fusion of data coming from different sources are required to present significant information to the users according to their context, profile, and needs. Some of the most advanced IoT devices and applications offer solutions based on artificial intelligence and expert systems for avoiding to prompt users too often and risking to bother them with too many questions. The idea to make devices and applications able to take decisions on behalf of the users aims at not disturbing and overwhelming people in their everyday lives. In lifelogging devices, streams of data can be produced at different spatial and temporal granularity concerning the fact that the time frame and the location associated to an event can be envisaged at several levels of detail (e.g., hour, day, month, house, district, city). For handling the information produced by a single sensor or a set of data stream sources, the user needs to use specific ETL (Extract-Transform-load) operations. Basic operations concern simple projections or filtering of the data for determining subsets of them or for pruning irrelevant values. Other operations aim at transforming data: (1) for changing the unit of measure (e.g. from yards to meters) or geographical coordinates (from one standard to another one); (2) for introducing new attributes relying on the values assumed by other attributes; (3) for checking that data conforms to given validation rules or (4) for splitting a value in multiple values. Other complex data elaborations concern two classes of operations used to integrate data streams coming from different sensors: conversion and fusion. With conversion, different streams can be generated at different spatiotemporal granularities and in order to properly compare and analyze conversion operations should be supported. In this activity, the relationships among different granularities are fundamental for providing transformations that are sound and meaningful by applying some conversions to the streams. With fusion, when streams are produced from two or more sensors the need arises to combine or fuse their data. The temporal and spatial granularities of the different sources should be identical in order to obtain meaningful information. Moreover, the issue of the identification of a common temporal starting point should be faced in order to combine or fuse together data that are temporally aligned. Once the temporal and spatial granularities of the data streams are aligned, we need to establish the operations to apply on the single values.

Proc. of Third International Workshop on Cultures of Participation in the Digital Age - CoPDA 2015
Madrid (Spain), May 26th, 2015 (published at http://ceur-ws.org).

Data presentation. Data streams are characterized by a form and a spatio-temporal nature. They could be distributed in a raw form (e.g. flow of numerical signals) or in a structured form by producing texts, images, video, or sounds. Moreover, various sources provide real-time data with spatial characteristics – such as geographical locations and spatial extents. Beyond stationary sensors, data is also continuously generated by moving objects thanks to advances in GPS and wireless technology. Therefore, with data representation we mean the description of the nature of the data both in terms of the media used to provide them and for explaining their spatiotemporal characteristics. At a presentation level, data coming for sensors need to be shown on the screen of the proper device for enabling the user's consultation. Data can have multimedia form – i.e. multimedia applications are needed to present and distribute it – or it can have the form of a flow of signals (numeric or alphanumeric) representing the output of the sensors. In the latter case, we need a transformation of the raw data into a more meaningful and abstract representation by applying different kinds of operations. For example, it is possible to summarize, at day level, the heartbeats per hour, it is possible to point out particular outliers or deviations in presence of an anomaly or warning situation, and so on.

4 Ongoing Research on Semiotic Approach to EUD for IoT

As presented in [5] we defined a new EUD paradigm and language for IoT that extends the current trends (IF-THIS-THEN-THAT and WHEN-TRIGGER-THEN-ACTION) by supporting the end user in composing space/time-based rules related to events captured by IoT sensors/devices. On top of this, we are now formalizing an internationalized visual language, localizable in terms of language and culture able to express the end user defined rules leaving the end users free to express themselves in a more natural way and not being forced by English-biased syntax and semantics. At a methodological level, we are adapting the semiotic model to participatory knowledge-management design [7] and to EUD [8] that we defined in the past years and that lies on SSW methodology [9] and on Tondl [10] and De Souza [11] work on semiotics and semiotic engineering. Such an approach is aimed at studying how eventual communication breakdown between the end user and the designer can affect the successful EUD activity in IoT.

5 References

1. Ashton, K.: That 'Internet of Things' Thing. RFID Journal, June (2009). Available online: http://www.rfidjournal.com/articles/view?4986 (accessed on January 19th, 2015).
2. Connections Counter: The Internet of Everything in Motion. Available online: http://newsroom.cisco.com/feature-content?type=webcontent&articleId=1208342 (accessed on June 30th, 2015).
3. Evans, D.: The Internet of Things. How the Next Evolution of the Internet is Changing Everything. Cisco Internet Business Solutions Group – White Paper. (2011). Available online:

Proc. of Third International Workshop on Cultures of Participation in the Digital Age - CoPDA 2015
Madrid (Spain), May 26th, 2015 (published at http://ceur-ws.org).

http://www.cisco.com/web/about/ac79/docs/innov/IoT_IBSG_0411FINAL.pdf (accessed on June 30th, 2015).

4. Lieberman, H., Paternò, F., Klann, M., Wulf, V.: End-User Development: An Emerging Paradigm. In: H. Lieberman, F. Paternò, & V. Wulf (Eds.) End-User Development, pp. 1--8). Springer (2006).
5. Barricelli, B.R., Valtolina, S.: Designing for End-User Development in the Internet of Things. End-User Development. Lecture Notes in Computer Science, 9083, pp. 9-24 (2015).
6. Jennings, N. R., Moreau, L., Nicholson, D., Ramchurn, S., Roberts, S., Rodden, T., Rogers, A.: Human-Agent Collectives. Communications of the ACM, 57(12), pp. 80--88 (2014).
7. Valtolina, S., Barricelli, B.R., Dittrich, Y.: Participatory Knowledge-Management Design: a Semiotic Approach. Journal of Visual Languages and Computing (JVLC), 23(2), pp. 103-115 (2011).
8. Barricelli, B.R.: An architecture for End-User Development supporting global communities. PhD dissertation. Advisor: Prof. Ernesto Damiani. Co-advisor: Dr. Stefano Valtolina. Università degli Studi di Milano, Italy (2011).
9. Costabile, M.F., Fogli, D., Mussio, P., Piccinno, A.: End-user development: the software shaping workshop approach. pp. 183–205, Springer (2006).
10. Tondl, L.: Problems of Semantics. ReidelPublishing (1981).
11. de Souza, C.S.:The Semiotic Engineering of Human–Computer Interaction, The MIT Press (2005).

Proc. of Third International Workshop on Cultures of Participation in the Digital Age - CoPDA 2015
Madrid (Spain), May 26th, 2015 (published at http://ceur-ws.org).