A Roadmap for Semantically-Enabled Human Device Interactions

Madhawa Perera[1,2], Armin Haller[1[0000−0003−3425−0780]], and Matt Adcock[2,1]

[1] Australian National University, Canberra ACT 2601, AU
{firstname.lastname}@anu.edu.au
https://cecs.anu.edu.au/
[2] CSIRO, Canberra ACT 2601, AU {firstname.lastname}@csiro.au
https://csiro.au/

Abstract. With the evolving Internet of Things (IoT), the number of smart devices we interact in our day-to-day life has significantly increased. The nature of human interaction with these devices must be perceived, because too much complexity could lead to the risk of IoT appearing unattractive for users or the system losing its efficiency, regardless of its potential. Therefore, it is important to address problems in Human Device Interactions to provide a high-quality User Experience (UX) in the emerging IoT.
This paper proposes a roadmap to address the complexities associated with human smart device (sensor or actuator) interaction through a methodology that incorporates context awareness in Augmented Reality (AR) using semantic Web technologies. Further, we analyse the use of Natural User Interfaces (NUI), such as hand gestures and gaze, to provide noninvasive and intuitive interaction in optimising the user experience.

Keywords: Semantic Web · Augmented Reality · Internet of Things · Smart Devices · Natural User Interfaces

1 Introduction

With the proliferation of IoT [63], the number of sensors deployed around the world started to grow at a rapid pace. With the increased demand created for smart devices (hereafter the term smart device includes sensors as well as actuators) [11], per unit prices of these are drastically decreasing which henceforth creates a high demand for building smart environments such as smart homes, automobiles, digital farming, modern smart hotel rooms, future cities etc. [45]. Statistics show that the total installed base of IoT connected devices is projected to increase to 75.44bn worldwide by 2025, a fivefold increase in ten years [63] and the number of network-connected smart devices per person around the world is predicted to grow from 0.08 to 6.58 [63].

Further, with this rapid evolvement of IoT [50,31], a human will often be confronted with smart devices which they are not familiar with. The number of smart devices that a human interacts in our day-to-day life has already significantly increased [45]. We face situations where either we have to interact with smart devices which we are not aware of the context and connections or have forgotten how to interact with due to the rapid growth of the number of devices that a modern human interacts with. Thus, a person must refer to alternate information sources to adduce context and details, such as user manuals or questions, or learning from an expert. This problem challenges the use of current systems that are built to establish HSI (in this paper Human Sensor Interaction (HSI) will be used synonymous with the phrase 'Human Device Interaction'). A key element of HSI is to provide access to the diverse information of smart devices that are relevant to a user in a given context, in the most convenient and usable way. Today, there are smart sensor hubs such as Amazon echo [13] and Google Home [23] which provide a HSI interface where users can use their voice to operate in a 'pre-configured' environment and where users know what sensors and actuators to interact with and how to. However, to interact with a specific sensor, a user must memorize and know the exact label that they have given to the sensor in order to establish an interaction through voice commands. In a real-world situation, these smart devices are built for different usages to different persons. Inability to provide user specific information is where the systems have failed to deliver a high-quality user experience. It remains as a challenge and an ongoing research area in IoT.

The above challenges can be further exemplified through a phenomenon in the domain of tourism and leisure. Currently, the world is rapidly moving towards automating customer experiences, especially in the hotel industry. Research and surveys are conducted to leverage the discomfort of being away from home and offer a better user experience in such situations [42,49]. As such, in the future, hotel rooms will consist of more sensors to provide better UX. In a modern hotel, even at the present day, the guest must try out several switches, remotes and potentially audio commands, read the hotel room manual or call the hotel reception to figure out how to operate smart doors, window blinds, adjust the thermostat, operate the audio system and the TV, etc. This is a very tedious task and sometimes a user may give up on some of the available features and will not gain the full experience that the environment offers. Key problems that ensue are how exactly are the users going to interact with those smart devices?

Therefore, to deliver a noninvasive interaction between human and smart devices, and to provide a better UX, we are aiming to utilize/consider a blend of semantic Web technologies, Augmented Reality technology along with Natural User Interfaces (NUIs) such as gaze, voice or hand gestures in HSI. In our research roadmap, we are investigating how one could use eye gaze to *detect*, and hand gestures or voice to *interact* with these feature-unfamiliar smart devices. That means we try to eliminate the burden of memorizing devices and their functionality in order to interact with them. Therefore, by looking at a smart device, the user should be able to read out and comprehend the capabilities of the device. Later these interactions should be made possible via hand gestures or voice.

Since this research is in its investigation phase we structured the paper as follows. Section 2 provides a discussion of the use of AR technology and we review and analyse Microsoft's HoloLens capabilities in regards to the identified challenges. Section 3 provides our definition of context, and why context is an important factor to consider. Section 4 discusses how we intend to incorporate semantic Web technologies to model the device capabilities and its context. Section 5 contains an analysis of how semantic Web technologies could blend with AR to provide better context-awareness. We propose a novel AR and semantic Web technology blend to address this problem and conclude the paper with a discussion section.

2 AR as a tool to enable Human-Device Interactions

2.1 Use of Augmented Reality

Augmented Reality (AR) is a field in computer science which uses computer vision-based techniques that enable superimposing interactive graphical content such as 2D and 3D multimedia content, on top of the view of real objects [59]. Therefore, AR could be used as a powerful visualization medium [59] to conveniently facilitate HSI. Its potential to blend real and virtual objects has opened up new opportunities for building interactive and engaging applications in multiple application domains [60]. AR has already been used in various domains to provide a better UX, for example, in Education [8,32,68], Marketing [70,16], Military [43,44], Medicine [12,34,52], Tourism [28,66], Entertainment [3,26,39] etc.

Looking at its widespread deployment in different use cases across multiple domains, we are aiming to investigate how successful it will be in addressing the aforementioned HSI challenges. In our research, we are aiming to utilize AR technology to extend a user's view in order to present them additional information regarding a smart device in the form of virtual content to then provide a mode of NUI to establish interaction between the smart device and the user.

Presently, AR technology has shown remarkable progress in building consumer-level hardware and use of it has spread rapidly in recent years [22,5,38]. It is estimated that by 2020, there will be 1 billion AR users where AR revenues will surpass Virtual Reality (VR) revenues [47]. Starting from Feiner et al's Head Mount Display (HMD) that was connected to a backpack containing a laptop and sensors such as GPS and gyroscopes [20], the technology has recently shrank to the size of handheld displays (HHD) such as mobile phones[1] [2] which have widened the access to AR experiences [56]. Research conducted by Billinghurst et al. utilizes both HMDs and HHD to provide better UX [15]. Advancement in processing and graphical performance of computing hardware and

[1] Google ARCore - https://developers.google.com/ar/discover/
[2] Apple ARKit - https://developer.apple.com/documentation/arkit

quickly growing bandwidth of mobile networks has been one of the key reasons behind this [60]. With these advances in the field of AR, the use of AR HMD such as Microsoft HoloLens (HL) is proposed in this research to track gaze actions and visualize content noninvasively to users. Further, this HMD is capable of identifying gaze points and recognizing hand gestures and voice. Thus, we will be utilizing these features in our proposed roadmap. The most critical functionalities of AR HMD are 1) Identifying the smart device a user is gazing at; 2) displaying relevant information to the user; 3) identifying user's hand gestures; 4) interpreting the hand gestures; and 5) communicating the interpreted information to the smart device. However, a key element of this research is how to present contextual information to users. Thus, in section 2.2, we investigate HoloLens' capability to identify smart devices located in physical environments with contextual information.

2.2 Analysis of AR HMD capabilities

Blending AR with natural user interaction methods like hand gestures, gaze and voice is a potential approach to make noninvasive and intuitive HSI. Therefore, in this section, we investigate the required capabilities of an AR HMD to potentially identify physical devices with their contextual knowledge.

If a user could get instructions on how to operate a smart device by looking at it (via visual instructions), we assume, that it saves time and maximizes the UX. Later, if the user can also interact via hand gestures or voice with the same device, in use cases such as smart hotel rooms, it would further improve the UX.

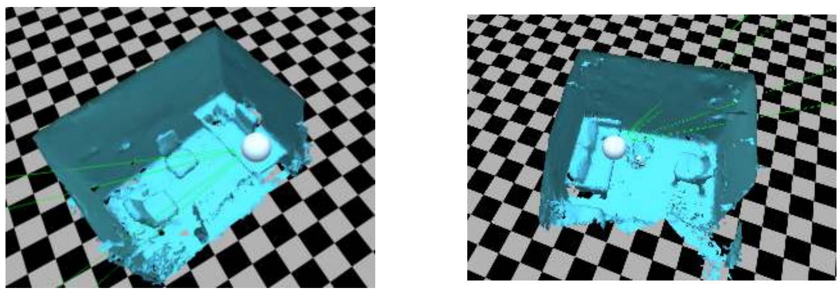
Yet, in each of these cases, the intermediate layer (which is the AR HMD) must be able to detect, and identify smart devices correctly. Then the HMD should be able to provide relevant information according to the user's context. For example, a frequent traveller to a specific hotel, does not need to know each information anew every time, thus the HMD need to be able to identify the user. A TV in a suite and a regular double room might look similar but might have different capabilities, thus the HMD needs to be able to identify its location. Likewise, depending on the contextual information the interaction will vary. Hence, looking at the problem holistically, we identified three main questions to be answered.

1. How could an AR HMD device detect and identify physical sensors and actuators within its context?
2. How could user preferences be extracted and generate contextually relevant information to a user?
3. How could this contextual information be added, maintained and altered without the help of experts in the AR domain?

As a contemporary example, we analysed Microsofts HL to see whether we can only use an AR HMD without integrating it with any other physical element to address these challenges. Looking at question 1, it seems it can be addressed with computer vision-based object detection techniques [64] along with an AR

HMD. However, contextual information could be absent from these systems without prior authoring (pre-configurations).

Automatic detection and segmentation of unknown objects in unknown environments is still work-in-progress for computer vision researchers [36,69]. Many existing object detection and segmentation methods assume prior knowledge about the object or human interference [37]. Even though techniques like segmentation, zero-shot learning and self-supervised learning can lead to developing the ability to predict an unknown object, predicting its context and associated details is still challenging.



(a) Room 1 spatial data representation (b) Room 2 spatial data representation

Fig. 1: Visual representation of spatial data (mesh) captured using Microsoft HoloLens

In an AR space, object recognition can be achieved either by using a 1)physical marker (or with the help of another physical element like a Bluetooth beacon), or by 2)direct object recognition (markerless AR). These markerless AR techniques use a combination of dedicated sensors, depth cameras, object recognition algorithms and environmental mapping algorithms to detect and map the real-world environment with objects [27]. However, Hammady et al. point to the possibility of even hybrid techniques. Further, the AR device has to know about its position in the world along with the awareness of its physical space [27]. For this purpose, AR devices use a technique called spatial mapping (also called 3D reconstruction) which maps the physical environment in order to blend the real and virtual worlds. This mapping helps the device to differentiate its physical locations and display virtual objects accordingly. It calculates this through the spatial relationship between itself and multiple key points. This process is called "Simultaneous Localization and Mapping (SLAM)" [6]. Figure 1 shows a spatial mapping mesh captured using Microsoft HL 1. It is also important to notice that spatial mapping provides a detailed representation of real-world surfaces which creates a 3D map of the environment [46] and as the user manoeuvres through space or objects move around, the mesh is updated to reflect the boundaries of the environment. Therefore, the device can understand and interact with the real world accordingly.

Spatial Mapping is identified as one of those capabilities that make the HL stand out from other AR HMDs. Yet, developers do not have direct access to the raw data of this mesh created by HL [41]. It is, however, possible to view the generated mesh and develop further on top of it and save it against a particular location as a visual representation [41]. However, HL only operates with a 3D mesh created by another HL and does not support the 3D meshes created by other devices or SLAM algorithms in general.

The following section analyzes the capabilities of using Microsoft HL for object detection with location awareness.

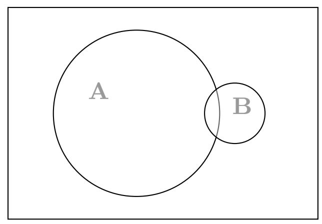


Fig. 2: Categorizations of devices and indoor environments

A:Known Devices, A':Unknown/New Devices (A' is everything that is not included in the set A), B:Mapped Environments, B':Unmapped/New Environments (B' is everything that is not included in the set B)

Our analysis was conducted under four major conditions. HL capabilities were analysed in detecting and interacting with a smart device when the HL is in 1)known environment with known smart devices. 2)Known environment with unknown devices 3)Unknown environment with known devices and 4)Unknown environment with unknown devices.

If the given environment is previously mapped (using an HL) and the smart devices are known (refer to scenario B ∩ A in Figure 2), then the HL itself can be used to interact with those objects (again) with contextual information such as location. With HL1 a gaze pointer can be used whereas with HL2 eye tracking will be available[3] to make the interaction more natural (noninvasive).

If the given environment is not mapped (using a HL) and the smart devices are known (scenario B' ∩ A), then the HL can use object detection algorithms (such as OpenCV [54], YOLO [57] etc.) without their context details. In this type of situation, objects could be labelled with physical markers to bind context-specific information. Additionally, beacons are another option which could be considered a solution to address this type of situation to provide context-awareness. In the AR domain beacons are commonly used to aide AR applications [24,53,55]

[3] HoloLens 2-Overview, Features, and Specs: Microsoft HoloLens (https://www.microsoft.com/en-us/hololens/hardware)

Even if the environment is mapped (using a HL) the smart devices could be unknown (that is object detection algorithms are not trained with a specific object's data previously), then it is not possible to interact with the smart devices by only using HL (scenario A' ∩ B). In this situation, again if the smart devices can be physically labelled (using markers or trackers) the contextual information could be attached to these markers.

Finally, if the given environment is not mapped (using a HL) and the smart devices are unknown (scenario A' ∩ B'), then it is not feasible to establish an interaction with a smart device with contextual information by only using HL. In such situations HL must pair with another physical element. Again, beacons and markers could be used to address the problem.

Table 1 summarizes the capabilities of Microsoft HL to interact with a smart device with contextual information, without the help of another physical element. Except for B ∩ A in which a prior authoring is conducted, all the other three scenarios need assistance of a beacon/physical marker along with a semantic description or visual recognition of the object and/or its position using semantic descriptions of either or both of these, i.e. the physical characteristics and/or its positional characteristics.

Table 1: Summary of HoloLens object recognition.

Scenario	Environment is mapped using a HL	Objects are known	HL itself can be used to		Has contextual awareness	Need the help of beacons	Need the help of physical markers
			Detect objects	Interact with objects			
B ∩ A	√	√	√	√	√		
B' ∩ A		√	√			√	√*
B ∩ A'	√				√		√
B' ∩ A'							√

* Optional. It is possible only with beacons, without markers.

Looking at question 1 again, it is clear that for HL to detect a smart device with its context, a pre-authoring process is essential. Therefore, either a physical marker creation or configuring a beacon with contextual data will require deep knowledge of AR and the suitable technology. This process includes collection, modeling, reasoning, and distribution of context-related information in relation to a smart device. This makes it a challenging and difficult task to build an AR application which will suit any given indoor environment in general. Thus, AR devices, applications and other relevant equipment needs to be authored to suit its application environment. This creates the necessity of an easy-to-use authoring tool.

The majority of currently available AR authoring tools, software, libraries and frameworks provide rich capabilities, but require advanced programming skills to use them [48]. There are a very few simple and easy-to-use authoring tools for non-technical users [48]. Yet, only very limited research has been conducted to building an easy-to-use AR authoring tool with the capability of adding contextual awareness.

3 Importance of Context

Humans can glance at objects and instantly identify or recognize them along with associated details, their location, and means and methods of interaction yet could struggle when the objects are unknown/unfamiliar or have a similar resemblance to another [65]. For instance, how many times one struggles to locate the exact room key among an unlabeled bunch of keys? Therefore, the contextual knowledge of an object is important to establish effective interaction with the same.

With the introduction of the term 'ubiquitous computing' by Mark Weiser in his seminal paper 'The Computer for the 21st Century' in 1991 [67], context-aware computing became a popular research area [51]. The term 'context-aware' was first used by Schilit et al. [51,61] in 1994. Thereafter many researchers attempted to address this concept in various applications and domains. According to Abowd et al. "Context is any information that can be used to characterise the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves" [1]. In HSI, the main concern is to present users with relevant and timely smart device data. Detection of sensors and then identifying and segmenting them according to the relevancy of the context is the best possible way of addressing this problem. To do this, we have to identify each smart device relevant to its context because the same device can be used for different purposes in indifferent contexts. Therefore the presentation technique (i.e. AR in our case) must identify the smart devices within their context.

Yet, we identified several challenges when addressing this problem such as; 1) how to identify an exact device that a user is gazing at; 2) how to distinguish similar-looking devices from each other; 3) how to model the device-specific information; and 4) how to change the interaction based on user preferences, location, time etc. Therefore, to present appropriate information related to each and every smart device, it is important to identify these smart devices within their context, which incorporates aspects such as indoor/outdoor location, user preferences, date and time, device capabilities etc. Detecting a physical device with its context is ongoing research in HSI and a complex interaction design challenge. Based on a literature review, we identified that the use of semantic Web technologies blended with AR provides a promising direction in solving this problem.

4 Incorporation of semantic Web technologies

According to J. Manyika et al., interoperability among IoT systems is required to capture 40% of the total potential value through the use of IoT [45]. According-ing to their research, there is a more than $4 trillion per year potential economic impact from IoT use in 2025, out of the total potential impact of $11.1 trillion predicted [45]. Recently, semantic Web technologies have been integrated in IoT with the aim of addressing interoperability challenges and reducing the heterogeneity in the domain [7,25].

The semantic Web is a term proposed by Tim Berners-Lee [10] "has been conceived as an extension of the World Wide Web that allows computers to intelligently search, combine and process Web content based on the meaning that this content has to humans" [29]. The semantic Web is aiming to provide a universal framework that allows data to be shared and reused across systems. It decouples applications from data through the use of an abstract model for knowledge representation [59]. Therefore, any application/system that understands the model can consume any data source that uses this model which in turn helps to address the problem of heterogeneity. By looking at the vast majority of manufacturers in the IoT domain and their heterogeneity, this type of knowledge representation model could provide a promising direction in providing a better HSI.

Using semantic Web technologies we can endow smart devices with their semantics (i.e. their intended use, capabilities and purpose). Thus, combining that information with contextual data allows to specify which conclusion should be drawn and then what information should be augmented and visualized (via an AR interface) to users to get an idea on what the smart device is for and how to interact with it.

There is a growing interest in the area of blending semantic with IoT and AR. Rumiski et al.'s findings suggest that an application of semantic Web techniques can be an efficient solution to search contextually described distributed resources constituting interactive AR presentations [60]. Further, Rumiski et al. have developed a semantic model for distributed AR services and built ubiquitous dynamic AR presentations based on semantically described AR services in a contextual manner. Yet, their work concerns integrating distributed services in AR, but does not focus on addressing the problem of maximizing the user experience in HSI, when humans are compelled to use multiple unfamiliar devices. However, the following researches addressed the user experience aspects in interactions. FarmAR by Katsaros et al. exploit AR technology to identify plants and to augment useful information to farmers. Their system is based on a knowledge base which consists of an ontology that describes information concerning the plant, such as its common scientific name and frequent plant diseases etc. [35]. Contreras et al. present a mobile application for searching places, people and events within a university campus. In their work they leverage semantic Web and AR to provide an application with a high degree of query expressiveness and an enhanced user experience [19]. In both Katsaros and Contreras approaches they have incorporated semantic Web technologies, yet their approaches do not consider contextual information. Further, they have used a handheld display (HHD) instead of AR HMDs which will create different UX in HSI.

L. Cheng et al.'s works shows that embedding semantic understanding with Mixed Reality (MR) can greatly enhance the user experience by helping to understand object-specific behaviours [17]. L. Cheng et. al demonstrate a framework for a material-aware prototype system for generating context-aware physical interactions between the real and the virtual objects. However, the focus of their research is on material understanding and its semantic fusion with the virtual scene in a MR environment hence not addressing HSI. Further looking at the context awareness, Hoque et al. have proposed a generic context model based on ontology and reasoning techniques in the smart home domain [30]. Zhu J. et al. have proposed a framework specifically designed for an assisted maintenance system in which they have incorporated context-aware AR in their research with semantic Web technologies to provide information that is more useful to the user [71]. Their main focus is towards a context aware AR authoring tool. Further, Flatt. H et al's proposed a framework towards a context-aware assistance system for maintenance applications in smart factories [21]. The central element of this approach is an ontology-based context-aware framework, which aggregates and processes data from different sources. Yet, their application is towards HHD AR and not addressing the HSI challenges. Thus, it is observed that, HSI and improving UX is not a concern in their work.

5 Blend of semantic Web technology with AR for context awareness

As per our literature review, the use of semantic Web technologies has provided a promising direction to add meaningful contextual information to AR presentations [59]. In this section, we explain the knowledge modeling approach.

Seydoux et al.s analysis of existing ontologies related to IoT, concludes that "some of the IoT ontologies cover most of the key concepts but none of them covers them all" [62]. Therefore, in our investigation we consider a combination of several ontologies like the DogOnt [14] which "aims at offering a uniform, extensible model for all devices being part of a local Internet of Things inside a smart environment", the Semantic Sensor Network (SSN) ontology [18], and the IoT-Lite ontology [9] which "is a lightweight ontology to represent Internet of Things (IoT) resources, entities and services. IoT-Lite is an instantiation of the SSN ontology", OneM2M [9] and IoT-O [62] which are widely used in IoT domain ontologies.

Preserving semantic Web best practices of reusability, instead of developing an ontology from scratch we analysed existing IoT ontologies to identify the suitability of using an existing knowledge model. Prior to this, as explained in section 2.2, we did an analysis of AR HMD to identify its limitations when handling contextual data and then to see whether we could use knowledge modeling to address those limitations. Our conclusion was, that except for B $\cap$ A (as described in Table 1) in all the other cases AR HMD itself could not resolve contextual data.

Therefore, firstly we identified the main contextual information that we required to blend with AR application in order to provide a better UX in HSI. As per our analysis, indoor/outdoor location, device capabilities and user information(users) are the high level concepts that are required. Secondly we investigate on suitable methods/ontologies to model these knowledge/concepts. Further, this knowledge model will be decoupled from the AR application which makes it customizable to different use cases/scenarios, without affecting the functionality of AR application.

In our roadmap, firstly we need to model a smart device which could be either a sensor or an actuator or both. Therefore, we require a combined and also a separated representation for sensors and actuators. Next, these device capabilities which could be either an observable property or an actuatable property, respectively need to be modelled. Presenting these capabilities to a user could vary based on the device location, user preferences and features of interest. Therefore these three are the next required modelling concepts. Further, in our investigation, we are researching on how to interact with a smart device using NUIs like gestures. Therefore, human gestures is another concept that we need to consider.

By looking at the conceptual requirements, we designed a high level concept overlapping diagram (See Figure 3) to identify the type of ontologies that we need to consider. This diagram depicts the cluster of concepts that are related to the domain of IoT. Each circle indicates the required concept and different colours depict the current representation level of these concepts within existing IoT ontologies. After identifying the required concepts we started analysing the existing IoT ontologies to see their adaptability. Table 2 below shows our analysis results.

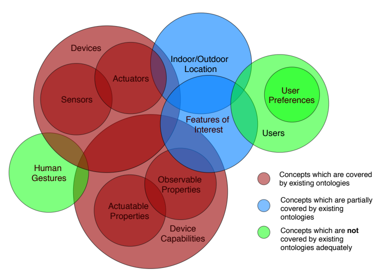


Fig. 3: Concept coverage in context aware IoT ontology development

Most of these ontologies are capable of modeling the knowledge specific to a device, sensor, actuator, and their capabilities and the location. However, user preferences and human gestures are not addressed directly in any of these ontologies. Even outside of the IoT domain, ontologies that define device users and potential interaction locations are rare. For example, Nazer et al. have defined a user's profile ontology, yet it is a use case ontology which aims at providing personalized food and nutrition recommendations [2]. Thus, there is a necessity for a global ontology to model user related and human hand gestures related knowledge. Table 2 summarized the fact that, we can reuse and merge some IoT ontologies to fulfil part of our conceptual requirements, yet user preferences and hand gestures modeling need to be further investigated to avoid redefinition as much as possible.

Table 2: Summary of existing IoT ontology evaluation.

	Device	Device Capabilities	Actuators	Actuator Capabilities	Sensors	Sensor Capabilities	Location	User	User Preferences	Features of Interest	Human Gestures
SSN	√	√	√	√	√	√				√	
IoT-Lite	√	√	√		√	√	√				
oneM2M	√	√									
IoT-O			√	√	√	√	√		√*	√	
DogOnt			√	√	√	√	√				

*Supported by an external ontology module.

6 Discussion

In this section we discuss the high-level process flow and highlight some of the potential challenges when establishing human device interactions in the IoT.

As the intention of the research is to maximize the UX by enabling an effective human - device interaction when a user wears an AR HMD, the visualized content has to be personalised to a specific user. For this, an AR device should be able to identify its user and associated information such as the user's preferences. A user study needs to assess natural interaction patterns and the results need to be encoded in an ontology. A study on how users naturally interact with unfamiliar devices and their behaviours. The intention is to identify/generalize ways to make human device interaction more intuitive and noninvasive. The user interaction data itself will then be captured and stored/updated accordingly as ontology instances. If it is a first-time user there will be no data recorded about previous interactions. Therefore, the AR application would not have the needed guidance on the user. The aim of studying user interactions is to reduce redundancy (by reusing previously stored interactions). When the information related to a user is processed, privacy and security is a concern. Proper authentication and authorization mechanisms need to be used when querying user specific data in the knowledge model. Therefore the specific requirements on privacy and security need to be further analysed in the long run.

Once the AR application is capable of identifying the user, the next concern is the device identification process. Based on the object recognition efficiency of the AR HMD, either a marker based or direct object recognition based or hybrid approach could be chosen. This needs to be further explored. If a marker based approach is selected, physical markers need to be pre-configured with a unique device identifier and their location information. This creates the necessity of an authoring task. In either cases, with a physical marker-based or direct object detection approach, there is the potential of facing processing delays. This could be due to the AR HMD's capability of recognizing a marker or an object as well as the size of information stored in the knowledge model (query time). Looking at the potential size of data sets, it is not feasible to store the data on the AR device. Thus, this creates the necessity of storing data in the cloud which brings the concern of network latency as well.

Once the smart device and user are identified, the relationships between the device and the user and previous interactions need to be identified. In this stage, an identified device could fall into one of the following categories.

1. Previously seen but not interacted with
2. Previously seen and successfully interacted with
3. Previously seen and unsuccessfully interacted with
4. Previously unseen and not interacted with

These information along with location details will be utilized when deciding which content to be displayed to the user.

Once the content is displayed, users will start interacting with the smart devices. Then the next concern is the human gesture interpretation process. Microsoft HL 1 has built-in functionality to recognize a restricted number of gestures whereas in Microsoft HL 2 this has been extended. [40]. Interpreting the meaning of hand gestures in accordance with a device-capability again requires to query the knowledge model related to human hand gestures. To the best of our knowledge, there is no study or ontology available that describes natural user behaviours when they are confronted with unfamiliar devices for the first time. These behaviours are most likely also culturally different, i.e. switches, for example, operate in opposite directions in different countries. Thus, human gestures could be changed based on personal preferences, geography and health conditions of a user. Therefore, the knowledge model is more important than a rigid mapping of device capabilities against fixed/defined hand gestures. Figure 4 shows a summary of overall process as flow in the roadmap.

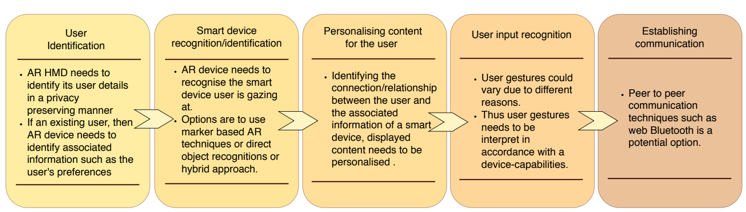


Fig. 4: High-level process flow of proposed roadmap

This AR based interaction technique needs to be evaluated against the commonly used, voice-based human sensor interaction methods such as Google Home or Amazon echo dot to assess whether users would like to wear a pair of AR glasses and to interact with smart devices.

It is important to know that, lighting conditions, distance between the physical marker/device and the AR HMD, can create delays when identifying a device. Yet, these devices are rapidly evolving and their capabilities are enhanced, addressing these limitations.

Hardware is also getting more user friendly and there are already wearables available that are similar to a pair of shades with AR features and functionality [58]. There are many ways of detecting hand gestures with the help of many industrial equipment such as Myo [4], the Leap motion [33], hand tracking gloves etc. These could be incorporated to reduce the invasiveness created by the hardware designs.

Finally, an easy-to-use authoring tool is additional benefit for this work. In our future research we are planning to investigate on how we can build such an authoring tool so that a general user can configure their environment. Further, merging real-time sensor data along with AR HMD sensor readings could help to provide real-time contextual data.

7 Conclusion

This paper presents a roadmap on how augmented reality could be used in combination with semantic Web technologies as a powerful interaction technique that yields a new types of user experience with the Internet-of-Things. The proposed methodology uses semantic Web technologies to produce context-aware interactions in AR presentations. Our key insight to build context awareness through ontologies is not only to enhance a user experience through device-specific behaviours but also to pave the way for solving complex interaction design challenges in HSI. We are planning to conduct quantitative and qualitative evaluations for the proposed methodology, and based on the results of these studies intend to show how this framework could further be enhanced to provide user friendly authoring interfaces for the purpose of creating context-aware AR presentations.

References

1. Abowd, G.D., Dey, A.K., Brown, P.J., Davies, N., Smith, M., Steggles, P.: Towards a better understanding of context and context-awareness. In: Proc. of International symposium on handheld and ubiquitous computing. pp. 304–307. Springer (1999)
2. Al-Nazer, A., Helmy, T., Al-Mulhem, M.: User's profile ontology-based semantic framework for personalized food and nutrition recommendation. Procedia Computer Science **32**, 101–108 (2014)
3. Alha, K., Koskinen, E., Paavilainen, J., Hamari, J.: Why do people play location-based augmented reality games: A study on pokémon go. Computers in Human Behavior **93**, 114–122 (2019)
4. Ali, S., Samad, M., Mehmood, F., Ayaz, Y., Qazi, W.M., Khan, M.J., Asgher, U.: Hand gesture based control of nao robot using myo armband. In: Proc. of 10th AHFE. pp. 449–457. Springer (2019)
5. Altinpulluk, H.: Determining the trends of using augmented reality in education between 2006-2016. Education and Information Technologies **24**(2), 1089–1114 (2019)
6. Andijakl: Basics of ar: Slam simultaneous localization and mapping (Sep 2018), https://www.andreasjakl.com/basics-of-ar-slam-simultaneous-localization-and-mapping/
7. Barnaghi, P., Wang, W., Henson, C., Taylor, K.: Semantics for the internet of things: early progress and back to the future. International Journal on Semantic Web and Information Systems (IJSWIS) **8**(1), 1–21 (2012)
8. Barrow, J., Forker, C., Sands, A., OHare, D., Hurst, W.: Augmented reality for enhancing life science education. In: Proc. of VISUAL 2019 (2019)

9. Bermudez-Edo, M., Elsaleh, T., Barnaghi, P., Taylor, K.: Iot-lite: a lightweight semantic model for the internet of things. In: Proc. of IEEE UIC/ATC/ScalCom/CBDCom/IoP/SmartWorld conference. pp. 90–97. IEEE (2016)
10. Berners-Lee, T., Hendler, J., Lassila, O., et al.: The semantic web. Scientific american **284**(5), 28–37 (2001)
11. Biederman, I.: Recognition-by-components: a theory of human image understanding. Psychological review **94**(2), 115 (1987)
12. Birkfellner, W., Figl, M., Huber, K., Watzinger, F., Wanschitz, F., Hummel, J., Hanel, R., Greimel, W., Homolka, P., Ewers, R., et al.: A head-mounted operating binocular for augmented reality visualization in medicine-design and initial evaluation. IEEE Transactions on Medical Imaging **21**(8), 991–997 (2002)
13. Black, M.: Your complete guide to amazon echo (Jun 2019), https://www.techadvisor.co.uk/new-product/audio/amazon-echo-3584881/
14. Bonino, D., Corno, F.: Dogont-ontology modeling for intelligent domotic environments. In: Proc. of ISWC 2008. pp. 790–803. Springer (2008)
15. Budhiraja, R., Lee, G.A., Billinghurst, M.: Using a hhd with a hmd for mobile ar interaction. In: Proc. of IEEE ISMAR. pp. 1–6. IEEE (2013)
16. Bulearca, M., Tamarjan, D.: Augmented reality: A sustainable marketing tool. Global business and management research: An international journal **2**(2), 237–252 (2010)
17. Chen, L., Tang, W., John, N., Wan, T.R., Zhang, J.J.: Context-aware mixed reality: A framework for ubiquitous interaction. arXiv preprint arXiv:1803.05541 (2018)
18. Compton, M., Barnaghi, P., Bermudez, L., GarcíA-Castro, R., Corcho, O., Cox, S., Graybeal, J., Hauswirth, M., Henson, C., Herzog, A., et al.: The ssn ontology of the w3c semantic sensor network incubator group. Web semantics: science, services and agents on the World Wide Web **17**, 25–32 (2012)
19. Contreras, P., Chimbo, D., Tello, A., Espinoza, M.: Semantic web and augmented reality for searching people, events and points of interest within of a university campus. In: Proc. of CLEI 2017. pp. 1–10. IEEE (2017)
20. Feiner, S., MacIntyre, B., Höllerer, T., Webster, A.: A touring machine: Prototyping 3d mobile augmented reality systems for exploring the urban environment. Personal Technologies **1**(4), 208–217 (1997)
21. Flatt, H., Koch, N., Röcker, C., Günter, A., Jasperneite, J.: A context-aware assistance system for maintenance applications in smart factories based on augmented reality and indoor localization. In: Proc. of 20th IEEE ETFA. pp. 1–4. IEEE (2015)
22. Garzón, J., Pavón, J., Baldiris, S.: Systematic review and meta-analysis of augmented reality in educational settings. Virtual Reality pp. 1–13 (2019)
23. Gebhart, A.: Everything you need to know about google home (May 2019), https://www.cnet.com/how-to/everything-you-need-to-know-about-google-home/
24. Guillama, N., Heath, C.: Personal augmented reality (Apr 25 2019), uS Patent App. 16/165,823
25. Gyrard, A., Serrano, M., Atemezing, G.A.: Semantic web methodologies, best practices and ontology engineering applied to internet of things. In: Proc. of 2nd IEEE WF-IoT. pp. 412–417. IEEE (2015)
26. Hamari, J., Malik, A., Koski, J., Johri, A.: Uses and gratifications of pokémon go: Why do people play mobile location-based augmented reality games? International Journal of Human–Computer Interaction **35**(9), 804–819 (2019)
27. Hammady, R., Ma, M., Powell, A.: User experience of markerless augmented reality applications in cultural heritage museums:museumeyeas a case study. In: Proc. of Salento AVR 2018. pp. 349–369. Springer (2018)
28. Han, D.I., Jung, T., Gibson, A.: Dublin ar: implementing augmented reality in tourism. In: Information and communication technologies in tourism 2014, pp. 511–523. Springer (2013)
29. Hitzler, P., Krotzsch, M., Rudolph, S.: Foundations of semantic web technologies. Chapman and Hall/CRC (2009)
30. Hoque, M.R., Kabir, M.H., Thapa, K., Yang, S.H.: Ontology-based context modeling to facilitate reasoning in a context-aware system: A case study for the smart home. International Journal of Smart Home **9**(9), 151–156 (2015)
31. Howard, P.N., Howard, P.N.: How big is the internet of things and how big will it get? (Jul 2016), https://www.brookings.edu/blog/techtank/2015/06/08/how-big-is-the-internet-of-things-and-how-big-will-it-get/
32. Ibáñez, M.B., Delgado-Kloos, C.: Augmented reality for stem learning: A systematic review. Computers & Education **123**, 109–123 (2018)
33. Jia, J., Tu, G., Deng, X., Zhao, C., Yi, W.: Real-time hand gestures system based on leap motion. Concurrency and Computation: Practice and Experience **31**(10), e4898 (2019)
34. Joda, T., Gallucci, G., Wismeijer, D., Zitzmann, N.: Augmented and virtual reality in dental medicine: A systematic review. Computers in biology and medicine (2019)
35. Katsaros, A., Keramopoulos, E.: Farmar, a farmer's augmented reality application based on semantic web. In: Proc. of SEEDA-CECNSM 2017. pp. 1–6. IEEE (2017)
36. Kikkawa, R., Sekiguchi, H., Tsuge, I., Saito, S., Bise, R.: Semi-supervised learning with structured knowledge for body hair detection in photoacoustic image. In: Proc. of 2019 IEEE 16th ISBI 2019. pp. 1411–1415. IEEE (2019)

37. Kootstra, G., Bergström, N., Kragic, D.: Fast and automatic detection and segmentation of unknown objects. In: Proc. of 10th IEEE-RAS. pp. 442–447. IEEE (2010)
38. Kotane, I., Znotina, D., Hushko, S.: Assessment of trends in the application of digital marketing. Scientific Journal of Polonia University **33**(2), 28–35 (2019)
39. Laine, T.H., Suk, H.: Designing educational mobile augmented reality games using motivators and disturbance factors. In: Augmented Reality Games II, pp. 33–56. Springer (2019)
40. Langston, J.: Hololens 2 gives microsoft the edge in next generation of computing (Jul 2019), https://news.microsoft.com/innovation-stories/hololens-2/
41. Legiedz, R.: A thorough look into spatial mapping with hololens (2017), https://solidbrain.com/2017/08/07/a-thorough-look-into-spatial-mapping-with-hololens/
42. Leonidis, A., Korozi, M., Margetis, G., Grammenos, D., Stephanidis, C.: An intelligent hotel room. In: Proc. of International Joint Conference on Ambient Intelligence. pp. 241–246. Springer (2013)
43. Livingston, M.A., Rosenblum, L.J., Brown, D.G., Schmidt, G.S., Julier, S.J., Baillot, Y., Swan, J.E., Ai, Z., Maassel, P.: Military applications of augmented reality. In: Handbook of augmented reality, pp. 671–706. Springer (2011)
44. Livingston, M.A., Rosenblum, L.J., Julier, S.J., Brown, D., Baillot, Y., Swan, I., Gabbard, J.L., Hix, D., et al.: An augmented reality system for military operations in urban terrain. Tech. rep., Naval Research Lab Washington DC Advanced Information Technology Branch (2002)
45. Manyika, J.: The Internet of Things: Mapping the value beyond the hype. McKinsey Global Institute (2015)
46. Microsoft: Spatial mapping - mixed reality, https://docs.microsoft.com/en-us/windows/mixed-reality/spatial-mapping
47. Moss, A.: 20 augmented reality stats to keep you sharp in 2019 (Jul 2019), https://techjury.net/stats-about/augmented-reality/
48. Nebeling, M., Speicher, M.: The trouble with augmented reality/virtual reality authoring tools. In: Proc. of IEEE ISMAR-Adjunct. pp. 333–337. IEEE (2018)
49. Oracle: Hotel 2025 emerging technologies destined to reshape our business (2017), https://www.oracle.com/webfolder/s/delivery_production/docs/FY16h1/doc31/Hotels-2025-v5a.pdf
50. Panetta, K.: Gartner top strategic predictions for 2018 and beyond, https://www.gartner.com/smarterwithgartner/gartner-top-strategic-predictions-for-2018-and-beyond/
51. Perera, C., Zaslavsky, A., Christen, P., Georgakopoulos, D.: Context aware computing for the internet of things: A survey. IEEE communications surveys & tutorials **16**(1), 414–454 (2013)
52. Peters, T.M.: Overview of mixed and augmented reality in medicine. In: Mixed and Augmented Reality in Medicine, pp. 1–13. CRC Press (2018)
53. Plescia, M., Hui, L.: Augmented reality background for use in live-action motion picture filming (Jun 6 2019), uS Patent App. 16/210,951
54. Pulli, K., Baksheev, A., Kornyakov, K., Eruhimov, V.: Real-time computer vision with opencv. Communications of the ACM **55**(6), 61–69 (2012)
55. Rajeev, S., Wan, Q., Yau, K., Panetta, K., Agaian, S.S.: Augmented reality-based vision-aid indoor navigation system in gps denied environment. In: Mobile Multimedia/Image Processing, Security, and Applications 2019. vol. 10993, p. 109930P. International Society for Optics and Photonics (2019)
56. Rauschnabel, P.A., Felix, R., Hinsch, C.: Augmented reality marketing: How mobile ar-apps can improve brands through inspiration. Journal of Retailing and Consumer Services **49**, 43–53 (2019)
57. Redmon, J., Divvala, S., Girshick, R., Farhadi, A.: You only look once: Unified, real-time object detection. In: Proc. of the IEEE CVPR. pp. 779–788 (2016)
58. Robertson, A.: It's 2019 - where are our smart glasses? (Jun 2019), https://www.theverge.com/2019/6/28/18761633/augmented-reality-smart-glasses-google-glass-real-world-big-picture
59. Rumiński, D., Walczak, K.: Semantic model for distributed augmented reality services. In: Proc. of the 22nd Web3D Conference. p. 13. ACM (2017)
60. Rumiński, D., Walczak, K.: Large-scale distributed semantic augmented reality services–a performance evaluation. Graphical Models p. 101027 (2019)
61. Schilit, B.N., Theimer, M.M.: Disseminating active mop infonncition to mobile hosts. IEEE network (1994)
62. Seydoux, N., Drira, K., Hernandez, N., Monteil, T.: Iot-o, a core-domain iot ontology to represent connected devices networks. In: European Knowledge Acquisition Workshop. pp. 561–576. Springer (2016)
63. statista.com: Iot: number of connected devices worldwide 2012-2025, https://www.statista.com/statistics/471264/iot-number-of-connected-devices-worldwide/
64. Svensson, J., Atles, J.: Object detection in augmented reality. Masters Theses in Mathematical Sciences (2018)
65. Trafton, A., Office, M.N.: How the brain recognizes objects (Oct 2015), http://news.mit.edu/2015/how-brain-recognizes-objects-1005
66. Wei, W.: Research progress on virtual reality (vr) and augmented reality (ar) in tourism and hospitality: A critical review of publications from 2000 to 2018. Journal of Hospitality and Tourism Technology (2019)

67. Weiser, M.: The computer for the 21st century. Scientific American **265**(3), 66–75 (1991), https://dl.acm.org/citation.cfm?doid=329124.329126
68. Wojciechowski, R., Cellary, W.: Evaluation of learners attitude toward learning in aries augmented reality environments. Computers & Education **68**, 570–585 (2013)
69. Zhang, D., Han, J., Zhao, L., Meng, D.: Leveraging prior-knowledge for weakly supervised object detection under a collaborative self-paced curriculum learning framework. International Journal of Computer Vision **127**(4), 363–380 (2019)
70. Zhang, X., Navab, N., Liou, S.P.: E-commerce direct marketing using augmented reality. In: Proc. of IEEE ICME 2000. vol. 1, pp. 88–91. IEEE (2000)
71. Zhu, J., Ong, S.K., Nee, A.Y.: A context-aware augmented reality assisted maintenance system. International Journal of Computer Integrated Manufacturing **28**(2), 213–225 (2015)