Combining Semantics and Augmented Reality to Support the Human Mind

Florian Michahelles
Siemens Corporate Technology
Berkeley, CA 94704, USA
florian.michahelles
@siemens.com

Kimberly García
Siemens Corporate Technology
Berkeley, CA 94704, USA
kimberly.garcia@siemens.com

Andrei Ciortea
Siemens Corporate Technology
Berkeley, CA 94704, USA
andrei.ciortea@siemens.com

Markus Funk
Siemens Corporate Technology
Berkeley, CA 94704, USA
funk.markus@siemens.com

UbiComp/ISWC'17 Adjunct , September 11–15, 2017, Maui, HI, USA
ACM 978-1-4503-5190-4/17/09.
http://dx.doi.org/10.1145/3123024.3129270

Abstract
Humans and machines work closer together as never before. Whether it is about sensors to expand humans' sensorium, exo-skeletons augmenting physical capabilities, augmented and digital worlds breaking with physical boundaries, or curated digital memories: the value of all these technologies rises and falls with their ability to synchronize with the user's current situation, understand the needs and provide appropriate support. In this position paper we want to outline how semantic technologies can be applied to add more context and meaning to the user's role and task, and use Augmented Reality to present this information to the user. Instead of proposing yet another framework representing world knowledge we describe how to build upon existing standards, descriptions of procedures and routines, and regulations that become machine accessible. This way, machines and humans should be able to work in symbiosis. In the following we describe our motivation, list upcoming challenges and provide a first direction of how to proceed.

Author Keywords
Semantics; Augmented Reality; Cognitive Support;

ACM Classification Keywords
H.5.m [Information interfaces and presentation (e.g., HCI)]: Miscellaneous

Introduction and Background

Humans have always been creatures of habit who create routines over time. Thus, we can describe reoccurring tasks and break down complex tasks into manageable sub-tasks and work flows. Most work flows that are performed by humans everyday can be described in a structured way. This can be a complex assembly instruction or something fairly simple e.g. taking medication every evening before going to bed. However, no matter how structured these work flows are, they might cause undesired situations, such as having to remember various steps, which might lead to a human forgetting an important step of a work flow, or a change of context that might require taking a different action that has to be first formulated, recognized, and learned.

Over the years, technology has been heavily shaped to make human life more comfortable, by providing easy, and ubiquitous access to all kinds of information, and services. What once was a dream consisting in having computers integrated seamlessly in day to day life [26] has now become a reality. More than ever, we are used to interact with different devices to accomplish easy, and complex tasks. However, these interactions still imply a trade off between the benefits they provide, and the learning curve they require to surpass. Machines are not yet capable of providing a human like behavior that reasons based on individual experience and current context. Despite the efforts made in the research community to create contextual solutions, and experience-based knowledge bases [10], solutions still limited for mass consumption. Interaction with technology remains limited in most cases to the size and resolution of screens, and the underlying information systems are slowly starting to turn from just queryable repositories into meaningful knowledge bases.

Semantic Technologies focus on providing meaning to data by establishing a common understanding of the world through ontologies of interlinked terms, concepts, relationships, and entities [15]. The Semantic field became popular with the introduction of the Semantic Web [2], whose main idea was to give structure to World Wide Web content in order to create a common understanding for agents scouting data that could lead them to provide services. In recent years, among other domains, the Internet of Things has paid special attention to Semantics Technologies, in order to enable interoperability among machines producing unstructured and meaningless data [16, 25]. In an ideal semantic world, applications are not just producing data that needs to be collected and later interpreted; applications produce meaningful data from interactions with the real world, and other semantic applications. Hence, they can easily profit from information of related domains provided by such meaningful applications. Ontologies have also been used for resource discovery, since their aggregation nature makes them perfect to get a deeper understanding of a simple piece of information [11, 14]. Moreover, recent developments at the World Wide Web Consortium (W3C) explore the use of ontologies to describe interactions with services [1], which enables the development of loosely coupled semantic applications. The Semantic Web vision is rapidly turning into reality, providing a suitable platform for machines to support humans and augment their capabilities.

Additionally, we are currently experiencing the proliferation of Augmented Reality (AR) devices to the mainstream market. These AR devices can be used to provide cognitive support during different tasks, e.g. providing instructions during assembly tasks [8] or providing cognitive support

[1] https://w3c.github.io/wot-thing-description/, last accessed: 08/02/2017

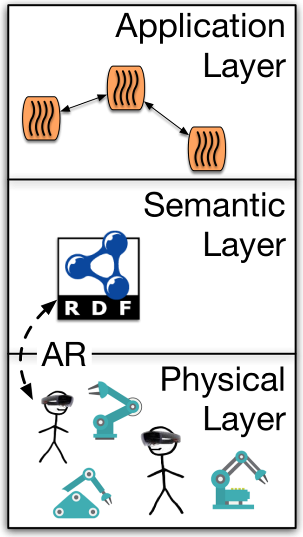

Figure 1: Integrating semantics and augmented reality (AR) to support the human mind.

while cooking [23]. Usually, AR is provided using three different categories of devices: hand-held screens that are carried by the user [13], in-situ projection [22], or using head-mounted displays (HMDs) [4]. All of these devices have been used to provide cognitive support e.g. for providing directions while walking [21], create spontaneous information displays [22], or providing AR assembly instructions. Especially when providing cognitive support for human workers through presenting assembly instructions, studies have shown that AR instructions can significantly reduce the errors that are made during object assembly tasks [24] by showing appropriate information to the user in real time. Especially individuals with limited cognitive capabilities (e.g. persons with cognitive disabilities) can highly benefit from AR support [7]. However, also workers which are new to a task can learn new work steps using AR instructions [6]. An overview about AR technologies as instruction systems is provided by Büttner et al. [3].

To sum up, using semantics is a great way of storing and accessing knowledge and AR is a great way in presenting information to a user. Previous work has applied this approach to learning of work related topics [12]. As semantics and AR are advancing, we believe that combining both technologies to constantly cognitively augment a user's mind is a promising approach for many everyday situations.

Concept: AR&Semantics

In our vision, bringing together semantics and AR is a promising approach to augment the human mind. Figure 1 illustrates this vision. Semantics provide the theoretical underpinning for enabling reliable cooperation between humans and machines: once the physical world is modeled in a given knowledge representation language, both humans and machines (via their application logic) are able to reliably interpret, reason, and act upon it. AR provides the means to seamlessly integrate the *physical* and *semantic layers* in order to support human-machine interaction: machines can communicate with humans by projecting graphical representations of semantic models into the physical world, but they can also receive input from humans by perceiving the physical world. In the following, we discuss in more detail each of these technologies.

The main benefit of current AR technology is that it includes sensing technology to recognize objects and position the HMD in the 3D space of the physical world. For example, the Microsoft HoloLens[2] contains depth sensors to build a spatial model of the surrounding environment that is updated while users are interacting with the device. Thus, cameras can use object recognition algorithms e.g. SURF [1] to recognize previously registered objects. This has been used in previous research for just finding lost objects [9], however we now can transfer the position and identity of recognized objects to a semantic backend.

On the semantic side the Open Semantic Framework (OSF) [17] provides an integrated solution for acquiring, enriching, maintaining, accessing, reasoning, and interacting with such machine-understandable knowledge. This effort enables subject matter experts, and final user applications to benefit from semantic knowledge, in a transparent way. OSF contains an extensible set of core ontologies (e.g., dimensions, and units) and Knowledge Packs (KP), which correspond to application specific ontologies . OSF provides access to knowledge stored in both the core ontologies, and the KPs through a controlled querying interface inside a REST API. This interface is based on prefabricated SPARQL query templates inside a KP. Thus, a KP determines the knowledge an application has access to, and the

[2] Microsoft HoloLens https://www.microsoft.com/en-us/hololens, last accessed: 08/02/2017

way it is accessed; preventing unwanted modifications to the knowledge models, and forbidding clients to indiscriminately extract all knowledge from the OSF.

We believe that this concept of combining AR with semantics can be used in various different application areas.

Use Cases

In the following we will provide two example use cases in order to illustrate our concept for combining AR and semantics: an industrial use case and a domestic use case.

In an industrial setting, production and factory managers are in charge of ensuring seamless and efficient operations with the help of their workforce. In order to do so, and to make the right decisions, managers need to have access to various information management systems that are typically independent from each other. Systems containing information such as workforce, assets, and logistics of those assets. Motivated by the difficulty that represents handling the aforementioned information systems; in [20] we proposed HoloInventory, an integrated information access platform based on semantics. HoloInventory provides role-based support for production or factory managers, and technicians wearing an AR device. Using voice commands, managers are able to quickly access data related to the part of the workforce they manage, and the status of assets that are under their responsibility. Managers can easily manipulate such information, which is then reflected in the semantic knowledge base. For technicians, HoloInventory provides functionalities that support them in the completion of their assigned tasks. Furthermore, in the future, we expect that the achievement of tasks in industrial settings will become a collective effort involving both humans and machines. In the UberManufacturing project [18], we show how a semantic layer can enable such cooperations. Machines are able to perceive humans via a Microsoft Kinect device. However, as future work, we envision further support for the human-machine symbiosis by projecting useful instructions for human workers in the environment.

As a domestic use case, we are envisioning a smart cooking scenario. A user is receiving AR instructions while cooking a meal (cf. [23]). The camera of the user's HMD could recognize the ingredients that are available in the user's kitchen and could suggest a recipe based on what is available. Further, through advanced computer vision techniques, the system could sense work steps that are being performed and send these logical steps to the semantic backend. There, the system can make sense of the actions that were performed during cooking and can send warnings in case something was done wrong or display next steps according to a selected recipe. Ultimately, we envision that a semantically enriched AR system is able to provide cognitive support in nearly any scenario.

Challenges

As using a novel technology usually comes with a couple challenges, in this position paper, we present the challenges that we identified when combining Semantics and Augmented Reality.

One of the main challenges that Semantic Technologies faces is their poor exposure. There is a lack of attractive outlets that let non-ontologists benefit from the information that a semantic model contains, and there is an even bigger lack of tools that let domain experts, not familiarized with ontologies, enrich existing knowledge models. Moreover, despite the great advantages that an interconnected model of concepts capable of reasoning, and providing comprehensive knowledge provides, there is still work to do regarding performance of the software that supports the creation, handling, and use of knowledge models. Thus, at this moment, the use of Semantic Technologies should be assessed, and customized for the type of application, the amount, and type of data that needs to be managed. Since a triple store will not perform well with big amounts of dynamic data, it would be necessary to combine a semantic model with a relational database management approach.

When using AR as an interface for humans to perceive additional information in the physical world, a few challenges become visible. The major challenge we currently see is the user acceptance. Especially for people wearing an HMD for a whole work day, this might be an issue as the users might feel patronized [6]. Another issue that is very important when designing AR experiences for cognitive support is that the user is not overwhelmed with too much information. Especially when using a semantic framework, the amount of information that can be shown to the user can be a lot. The designers of the human interface (which is the AR application in this scenario) should select carefully which information to present and which information not to present. Sometimes it can be better not to show a piece of information instead of showing too much. Lastly, the wearable sensors might be a privacy concern for bystanders as they also might be filmed when others are using the system. This could lead to a problem in social acceptability of this technology (cf. the social acceptance issues Google Glass had[3]).

Discussion

One crucial challenge mentioned above is how to collect domain knowledge and make it available in a semantic format. We have learned that earlier approaches of building up machine-readable knowledge models [19] have not really scaled due to their large and ambitious scope of representing general world knowledge. Instead, we propose to start with smaller and much more structured niches, such as industrial manufacturing, repair service or cooking recipes. Here, processes already are defined in handbooks, manuals, guidelines and processes. Thus, the remaining work is to translate those documents to machine-readable semantics. We believe that manuals, routines and standards would become much more versatile by releasing those as ontologies instead of PDF document targeting humans. Research [5] has shown that modeling standards can help to reduce inconsistencies and apply those as knowledge source for systems. We believe that AR can also be applied to provide a much more tangible access to these knowledge models, such that also non-semantic experts can be empowered to create, modify and curate ontologies as domain knowledge evolves over time. Obviously, we rely on the advances of usability for AR that are yet to be made in order to increase acceptance of long-term use in professional environments.

Conclusion and Outlook

We believe that if we can address all the challenges and apply the proposed concept to many different scenarios, the technology of combining semantics and Augmented Reality can truly become ubiquitous. In this position paper, we proposed a general concept of how this system architecture will look like and presented example use cases for using semantically enhanced AR applications.

In the near future, we are planning to apply our wearable prototype consisting of a HoloLens and a semantic backend to different specific industrial scenarios. When systems not only just take context into account, but also can make sense of it, in some years from now, we will look at the next

[3] https://www.scientificamerican.com/article/why-google-glass-is-creepy/, last accessed: 08/02/2017

generation of smart devices. We hope that this technology will be ubiquitously available for the mass market in a few years to cognitively support a huge number of persons on a daily basis.

REFERENCES

1. Herbert Bay, Tinne Tuytelaars, and Luc Van Gool. 2006. Surf: Speeded up robust features. *Computer vision–ECCV 2006* (2006), 404–417.

2. Tim Berners-Lee, James Hendler, Ora Lassila, and others. 2001. The semantic web. *Scientific american* 284, 5 (2001), 28–37.

3. Sebastian Büttner, Henrik Mucha, Markus Funk, Thomas Kosch, Mario Aehnelt, Sebastian Robert, and Carsten Röcker. 2017. The Design Space of Augmented and Virtual Reality Applications for Assistive Environments in Manufacturing: A Visual Approach. In *Proceedings of the 10th International Conference on PErvasive Technologies Related to Assistive Environments*. ACM, 433–440.

4. Thomas P Caudell and David W Mizell. 1992. Augmented reality: An application of heads-up display technology to manual manufacturing processes. In *System Sciences, 1992. Proceedings of the Twenty-Fifth Hawaii International Conference on*, Vol. 2. IEEE, 659–669.

5. Surbhi Dangi, Steven Ray, Ralph Hodgson, Gokhan Soydan, and Ankur Oberai. Using SPARQL/OWL for Validation of Smart Grid Standards.

6. Markus Funk, Andreas Bächler, Liane Bächler, Thomas Kosch, Thomas Heidenreich, and Albrecht Schmidt. 2017. Working with Augmented Reality? A Long-Term Analysis of In-Situ Instructions at the Assembly Workplace. In *Proceedings of the 10th ACM International Conference on PErvasive Technologies Related to Assistive Environments*. ACM, New York, NY, USA.

7. Markus Funk, Sven Mayer, and Albrecht Schmidt. 2015. Using in-situ projection to support cognitively impaired workers at the workplace. In *Proceedings of the 17th international ACM SIGACCESS conference on Computers & accessibility*. 185–192.

8. Markus Funk and Albrecht Schmidt. 2015. Cognitive assistance in the workplace. *IEEE Pervasive Computing* 14, 3 (2015), 53–55.

9. Markus Funk, Albrecht Schmidt, and Lars Erik Holmquist. 2013. Antonius: A mobile search engine for the physical world. In *Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication*. ACM, 179–182.

10. Kimberly García and Patrick Brézillon. 2017. From a Contextual Graph to a Tree Representation. In *Modeling and Using Context - 10th International and Interdisciplinary Conference, CONTEXT 2017, Paris, France, June 20-23, 2017, Proceedings*. 27–40.

11. Kimberly García, Sonia Mendoza, Dominique Decouchant, and Patrick Brézillon. 2016. Facilitating resource sharing and selection in ubiquitous multi-user environments. *Information Systems Frontiers* (2016), 1–21.

12. Chiara Ghidini, Viktoria Pammer, Peter Scheir, Luciano Serafini, and Stefanie Lindstaedt. 2007. APOSDLE: Learn@ work with semantic web technology. *Proc. of the IMEDIA and I-SEMANTICS* (2007), 262–269.

13. Mika Hakkarainen, Charles Woodward, and Mark Billinghurst. 2008. Augmented assembly using a mobile phone. In *Mixed and Augmented Reality, 2008. ISMAR 2008. 7th IEEE/ACM International Symposium on*. IEEE, 167–168.
14. Taekgyeong Han and Kwang Mong Sim. 2010. An ontology-enhanced cloud service discovery system. In *Proceedings of the International MultiConference of Engineers and Computer Scientists*, Vol. 1. 17–19.
15. Ian Horrocks. 2008. Ontologies and the semantic web. *Commun. ACM* 51, 12 (2008), 58–67.
16. Konstantinos Kotis and Artem Katasonov. 2012. An ontology for the automated deployment of applications in heterogeneous IoT environments. *Semantic Web Journal* (2012).
17. Simon Mayer, Jack Hodges, Dan Yu, Mareike Kritzler, and Florian Michahelles. 2017. An Open Semantic Framework for the Industrial Internet of Things. *IEEE Intelligent Systems* 32, 1 (2017), 96–101.
18. Simon Mayer, Dominic Plangger, Florian Michahelles, and Simon Rothfuss. 2016. UberManufacturing: A Goal-Driven Collaborative Industrial Manufacturing Marketplace. In *Proceedings of the 6th International Conference on the Internet of Things*. 111–119.
19. Marvin Minsky. 1974. A framework for representing knowledge. (1974).
20. Iori Mizutani, Mareike Kriztler, Kimberly García, and Florian Michahelles. 2017. Intuitive Interaction with Semantics using Augmented Reality - a Case Study of Workforce Management in an Industrial Setting. In *Proceedings of the 7th International Conference on the Internet of Things*.
21. Andreas Möller, Matthias Kranz, Robert Huitl, Stefan Diewald, and Luis Roalter. 2012. A mobile indoor navigation system interface adapted to vision-based localization. In *Proceedings of the 11th international conference on mobile and ubiquitous multimedia*. ACM, 4.
22. Claudio Pinhanez. 2001. The everywhere displays projector: A device to create ubiquitous graphical interfaces. In *Ubicomp 2001: Ubiquitous Computing*. Springer, 315–331.
23. Jürgen Scheible, Arnd Engeln, Michael Burmester, Gottfried Zimmermann, Tobias Keber, Uwe Schulz, Sabine Palm, Markus Funk, and Uwe Schaumann. 2016. SMARTKITCHEN Media Enhanced Cooking Environment. In *Proceedings of the 6th International Conference on the Internet of Things*. ACM, 169–170.
24. Arthur Tang, Charles Owen, Frank Biocca, and Weimin Mou. 2003. Comparative effectiveness of augmented reality in object assembly. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 73–80.
25. Wei Wang, Suparna De, Ralf Toenjes, Eike Reetz, and Klaus Moessner. 2012. A comprehensive ontology for knowledge representation in the internet of things. In *Trust, Security and Privacy in Computing and Communications (TrustCom), 2012 IEEE 11th International Conference on*. IEEE, 1793–1798.
26. Mark Weiser. 1999. The computer for the 21st century. *Mobile Computing and Communications Review* 3, 3 (1999), 3–11.