ACSIJ Advances in Computer Science: an International Journal, Vol. 4, Issue 4, No.16 , July 2015
ISSN : 2322-5157
www.ACSIJ.org

ACSIJ
WWW.ACSIJ.ORG

Multi-lingual and -modal Applications in the Semantic Web: the example of Ambient Assisted Living

Dimitra Anastasiou

**Media Informatics and Multimedia Systems, Department of Computing Science,
University of Oldenburg, 26121 Oldenburg, Germany**
dimitra.anastasiou@uni-oldenburg.de

Abstract

Applications of the Semantic Web (SW) are often related only to written text, neglecting other interaction modalities and a large portion of multimedia content that is available on the Web today. Processing and analysis of speech, hand and body gestures, gaze, and haptics have been the focus of research in human-human interactions and have started to gain ground in human-computer interaction in the last years. Web 4.0 or Intelligent Web, which follows Web 3.0, takes these modalities into account. This paper examines challenges that we currently face in developing multi-lingual and -modal applications and focuses on some current and future Web application domains, particularly on Ambient Assisted Living.

Keywords: *Ambient Assisted Living, Multimodality, Multilinguality, Ontologies, Semantic Web.*

1. Introduction

Ambient Assisted Living (AAL) promotes intelligent assistant systems for a better, healthier, and safer life in the preferred living environments through the use of Information and Communication Technologies (ICT). AAL systems aim to support elderly users in their everyday life using mobile, wearable, and pervasive technologies. However, a general problem of AAL is the digital divide: many senior citizens and people with physical and cognitive disabilities are not familiar with computers and accordingly the Web. In order to meet the needs of its target group, AAL systems require natural interaction through multilingual and multimodal applications. Already in 1991 Krüger [1] said that "natural interaction" means voice and gesture. Another current issue to bring AAL systems into the market is interoperability to integrate heterogeneous components from different vendors into assistance services. In this article, we will show that Semantic Web (SW) technologies can go beyond written text and can be applied to design intelligent smart devices or objects for AAL, like a TV or a wardrobe.

More than 10 years ago, Lu et al. [2] provided a review about web-services, agent-based distributed computing, semantics-based web search engines, and semantics-based digital libraries. The challenges of the SW at that time were the development of ontologies, formal semantics of SW languages, and trust and proof models. Zhong et al. [3] were in search of the "Wisdom Web" and Web Intelligence where "the next-generation Web will help people achieve better ways of living, working, playing, and learning." The challenges described in [2] have now been sufficiently addressed, whereas the vision presented in [3] has not yet gained ground. d' Aquin et al. [4] presented the long-term goal of developing the SW into a large-scale enabling infrastructure for both data integration and a new generation of intelligent applications with intelligent behavior. They added that some of the requirements of an application with large-scale semantics are to exploit heterogeneous knowledge sources and combine ontologies and resources. In our opinion, multimedia data belong to such heterogeneous sources. The intelligent behavior of next-generation applications can already be found in some new research fields, such as AAL and Internet of Things.

Many Web applications nowadays offer user interaction in different modalities (haptics, eye gaze, hand, arm and finger gestures, body posture, voice tone); few examples are presented here. Wachs et al. [5] developed GESTIX, a hand gesture tool for browsing medical images in an operating room. As for gesture recognition, Wachs et al. [6] pointed out that no single method for automatic hand gesture recognition is suitable for every application; each algorithm depends on each user's cultural background, application domain, and environment. For example, an entertainment system does not need the gesture-recognition accuracy required of a surgical system. An application based on eye gaze and head pose in an e-learning environment is developed by Asteriadis et al. [7]. Their system extracts the degree of interest and engagement of students reading documents on a computer screen. Asteriadis et al. [7] stated that eye gaze can also be used as an indicator of selection, e.g. of a particular exhibit in a museum, or a dress at a shop window, and may assist or replace mouse and keyboard interfaces in the presence of severe handicaps.

This survey paper presents related work on multilingual and multimodal applications within the field of Semantic

Copyright (c) 2015 Advances in Computer Science: an International Journal. All Rights Reserved.

Web. We discuss challenges of developing such applications, such as the Web accessibility by senior people. This paper is laid out as follows: in Sect. 2 we present how the multilingual and multimodal Web of Data is envisioned. Sect. 3 presents the challenges of developing multi-lingual and -modal applications. In Sect. 4 we look at some current innovative applications, including Wearable Computing, Internet of Things, and Pervasive Computing. The domain of AAL and its connection with the SW and Web 4.0 is presented in detail along with some scenarios in Sect. 5. Finally, we summarize the paper in Sect. 6.

2. Multi-linguality and -modality in the Semantic Web

Most SW applications are based on ontologies; regarding the multilingual support in ontologies, W3C recommends in the OWL Web Ontology Language Use Cases and Requirements [8] that the language should support the use of multilingual character sets. The impact of the Multilingual Semantic Web (MSW) is a multilingual "data network" where users can access information regardless of the natural language they speak or the natural language the information was originally published in (Gracia et al. [9]). Gracia et al. [9] envision the multilingual Web of Data as a layer of services and resources on top of the existing Linked Data infrastructure adding multilinguality in:

i) *linguistic information* for data and vocabularies in different languages (meaning labels in multiple languages and morphological information);
ii) *mappings* between data with labels in different languages (semantic relationships or translation between lexical entries);
iii) *services* to generate, localize, link, and access Linked Data in different languages.

Other principles, methods, and applications towards the MSW are presented by Buitelaar and Cimiano [10].

As far as multimodality is concerned, with the development of digital photography and social networks, it has become a standard practice to create and share multimedia digital content. Lu et al. [5] stated that this trend for multimedia digital libraries requires interdisciplinary research in the areas of image processing, computer vision, information retrieval, and database management. Traditional content-based multimedia retrieval techniques often describe images/videos based on low-level features (such as color, texture, and shape), but their retrieval is not satisfactory. Here the so-called *Semantic Gap* becomes relevant, defined by Smeulders et al. [11] as a "lack of coincidence between the information that one can extract from the visual data and the interpretation that the same data has for a user in a given situation." Besides, multimodality may refer to multimodal devices, like PC, mobile phone, PDA. In this paper, though, by "multimodality" we refer to multimodal input/output:

i) *Multimodal input* by human users (in)to Web applications, including modalities, like speech, body gestures, touch, eye gaze, etc.; for processing purposes, this input involves recognition of these modalities;

ii) *Multimodal output* by Web applications to human users; this involves face tracking, speech synthesis, and gesture generation. Multimodal output can be found in browser-based applications, e.g. gestures are performed by virtual animated agents, but it is even more realistic to be performed by pervasive applications, such as robots.

2.1 Breaking the digital divide: heterogeneous target group

Apart from the so-called "computer-literate" people, there are people who do not have the skills, the abilities, or the knowledge to use computers and accordingly the Web. The term "computer literacy" came into use in the mid-1970's and usually refers to basic keyboard skills, plus a working knowledge of how computer systems operate and of the general ways in which computers can be used [12]. The senior population was largely bypassed by the first wave of computer technology; however, they find it more and more necessary to be able to use computers (Seals et al. [13]). In addition to people with physical or cognitive disabilities, people with temporal impairments (e.g. having a broken arm) or young children often cannot use computers efficiently. All the above groups profit by the interaction with multimodal systems, where recognition of gesture, voice, eye gaze or a combination of modalities is implemented. For the "computer-literate" people, multimodality brings additional advantages, like naturalness, intuitiveness, and user-friendliness. To give some examples, senior people with Parkinson have difficulties controlling the mouse, so they prefer speech; deaf-mute people are dependent on gesture, specifically sign language. Sign language, as with any natural language, is based on a fully systematic and conventionalized language system. Moreover, the selection of the modality, e.g. speech or gesture, can also be context-dependent. In a domestic environment, when a person has a tray in their hand, (s)he might use speech to open the door. Thus, as the target group of the Web is very heterogeneous, the current and future applications should be context-sensitive, personalized, and adaptive to the target's skills and preferences.

Copyright (c) 2015 Advances in Computer Science: an International Journal. All Rights Reserved.

2.2 Multimodal applications in the Semantic Web

Historically, the first multimodal system was the "Put that there" technique developed by Bolt [14], which allowed the user to manipulate objects through speech and manual pointing. Oviatt et al. [15] stated that real multimodal applications range from map-based and virtual reality systems for simulation and training over field medic systems for mobile use in noisy environments, through to Web-based transactions and standard text-editing applications. One type of multimodal application is the multimodal dialogue system. They are applicable both in desktop and Web applications, but also in pervasive systems, such as in the car or at home (see AAL scenarios in 5.2.2). *Smartkom* [16] is such a system that features speech input with prosodic analysis, gesture input via infrared camera, recognition of facial expressions and emotional states. On the output side, the system features a gesturing and speaking life-like character together with displayed generated text and multimedia graphical output. *Smartkom* provides full "symmetric multimodality", defined by Wahlster [17] as the possibility that all input modes are also available for output, and vice versa. Another multimodal dialogue system is *VoiceApp* developed by Griol et al. [18]. All applications in this system can be accessed multimodally using traditional GUIs and/or by means of voice commands. Thus, the results are accessible to motor handicapped and visually impaired users and are easier to access by any user in small hand-held devices where GUIs are in some cases difficult to employ.

He et al. [19] developed a dialogue system called *Semantic Restaurant Finder* that is both multimodal and semantically rich. Users can interact through speech, typing, or mouse clicking and drawing to query restaurant information. SW services are used, so that restaurant information in different city/country/language are constructed, as ontologies allow the information to be sharable.

Apart from dialogue systems, many web-based systems are multimodal. In the assistive domain, a portal that offers access to products is EASTIN[1] [20]. It has a *multilingual* (users should forward information requests, and receive results, in their native language) and *multimodal* (offering a speech channel) front-end for end-users. Thurmair [20] tested the usability of the portal and found that most people preferred to use free text search.

3. Challenges in developing multi-lingual and -modal applications

In this section we discuss some challenges for multi-lingual and -modal applications from a development perspective. A basic challenge and requirement of the future Web is to provide Web accessibility to everybody, bearing in mind the heterogeneous target group. Web accessibility means to make the content of a website available to everyone, including the elderly and people with physical or cognitive disabilities. According to a United Nations report [21], 97% of websites fail to meet the most basic requirements for accessibility by using units of measurement (such as pixels instead of percentages), which restrict the flexibility of the page layout, the font size or both. Today worldwide 650 million people have a disability and approximately 46 million of these are located in the EU. By 2015 20% of the EU will be over 65 years of age, the number of people aged 60 or over will double in the next 30 years and the number aged 80 or over will increase by 10% by 2050. These statistics highlight the timeliness and importance of the need to make the Web accessible to more senior or impaired people. W3C has published a literature review [22] related to the use of the Web by older people to look for intersections and differences between the accessibility guidelines and recommendations for web design and development issues that will improve accessibility to older people. W3C has a Web Accessibility Initiative [23], which has released accessibility guidelines, categorized into:

i) *Web Content*: predictable and navigable content;
ii) *User Agents*: access to all content, user control of how content is rendered, and standard programming interfaces, to enable interaction with assistive technologies;
iii) *Authoring Tools*: HTML/XML editors, tools that produce multimedia, and blogs.

Benjamins et al. [24] stated that the major challenges of SW applications, in general, concern: (i) the availability of content, (ii) ontology availability, development and evolution, (iii) scalability, (iv) multilinguality, (v) visualization to reduce information overload, and (vi) stability of SW languages. As far as multilinguality is concerned, they state that any SW approach should provide facilities to access information in several languages, allowing the creation and access to SW content independently of the native language of content providers and users. Multilinguality plays an important role at various levels [24]:

i) *Ontologies*: WordNet, EuroWordnet etc., might be explored to support multilinguality;

[1] www.eastin.eu, 10/09/14

Copyright (c) 2015 Advances in Computer Science: an International Journal. All Rights Reserved.

ii) *Annotations:* proper support is needed that allows providers to annotate content in their native language;
iii) *User interface*: internationalization and localization techniques should make the Web content accessible in several languages.

As far as the challenges related to multimodal applications are concerned, He et al. [19] pointed out that the existing multimodal systems are highly domain-specific and do not allow information to be shared across different providers. In relation with the SW, Avrithis et al. [25] stated that there is a lot of literature on multimodality in the domains of entertainment, security, teaching or technical documentation, however the understanding of the semantics of such data sources is very limited. Regarding the combination of modalities, Potamianos & Perakakis [26], among other authors, stated that multimodal interfaces pose two fundamental challenges: the combination of multiple input modalities, known as the *fusion problem* and the combination of multiple presentation media, known as the *fission problem*. Atrey et al. [27] provided a survey about multimodal fusion for multimedia analysis. They made several classifications based on the fusion methodology and the level of fusion (feature, decision, and hybrid). One other challenge of multimodal systems is low recognition. Oviatt & Cohen [28], on comparing GUIs with multimodal systems, stated that, whereas input to GUIs is atomic and certain, machine perception of human input, such as speech and gesture, is uncertain; so any recognition-based system's interpretations are probabilistic. This means that events, such as object selection, which were formerly basic events in a GUI (point an object by touching it) are subject to misinterpretation in multimodal systems. They see that the challenge for system developers is to create robust new time-sensitive architectures that support human communication patterns and performance, including processing users' parallel input and managing the uncertainty of recognition-based technologies.
Apart from the above challenges, an additional challenge is twofold: i) develop multi-lingual and -modal applications in parallel and ii) tie them with a language-enhanced SW. Today there are not many applications that combine multiple modalities as input and/or output and support many natural languages at the same time. Cross [29] states that current multimodal applications typically provide user interaction in only a single language. When a software architect desires to provide user interaction in more than one language, they often write a multimodal application for each language separately and provide a menu interface to a user that permits the user to select the language that the user prefers. The drawback is that having multiple versions of the same multimodal application in various languages increases complexity, which leads to an increased error rate and additional costs.

4. Current domain applications

In the last years the usage of the Web has shifted from desktop applications and home offices to smart devices at home, in entertainment, the car, or in the medical domain. Some of the latest computing paradigms are the following:

- **Wearable computing** is concerned with miniature electronic devices that are worn on the body or woven into clothing and access the Web, resulting in intelligent clothing. A commercial product is the MYO armband by Thalmic Labs [1] with which users can control presentations, video, content, games, browse the Web, create music, edit videos, etc. MYO detects gestures and movements in two ways: 1) muscle activity and 2) motion sensing. The most recent Apple Watch [2] is designed around simple gestures, such as zooming and panning, but also senses force (Force Touch). Moreover, a heart rate sensor in Apple Watch can help improve overall calorie tracking.
- The **Internet of Things** (IoT) refers to uniquely identifiable objects and their virtual representations in an Internet structure. Atzori et al. [30] stressed that the IoT shall be the result of the convergence of three visions: things-oriented, Internet-oriented, and Semantic-oriented visions. Smart semantic middleware, reasoning over data, and semantic execution environments belong to the semantic-oriented visions. A recent survey of IoT from an industrial perspective is published by Perera et al. [31]. [31] stated that "despite the advances in HCI, most of the IoT solutions have only employed traditional computer screenbased techniques. Only a few IoT solutions really allow voice or object-based direct communications." They also see a trend from smart home products that it also increasingly uses touch-based interactions.
- **Pervasive context-aware systems**: Pervasive/ ubiquitous computing means that information processing is integrated into everyday objects and activities. Henricksen et al. [32] explored the characteristics of context in pervasive systems: it exhibits temporary characteristics, has many alternative representations, and is highly interrelated. Chen et al. [33] developed the *Context Broker Architecture*, a broker agent that maintains a shared model of context for all computing entities in the space and enforces the privacy policies defined by the users when sharing their contextual information. They believe that a requirement for realizing context-aware systems is the ability to understand their situational conditions. To achieve this, it requires contextual information to be represented in ways that are adequate for machine

[1] https://www.thalmic.com/myo/, 08/06/15
[2] https://www.apple.com/watch/, 08/06/15

Copyright (c) 2015 Advances in Computer Science: an International Journal. All Rights Reserved.

processing and reasoning. Chen et al. [33] believe that SW languages are well suited for this purpose for the following reasons: i) RDF and OWL have rich expressive power that are adequate for modeling various types of contextual information, ii) context ontologies have explicit representations of semantics; systems with the ability to reason about context can detect inconsistent context knowledge (result from imperfect sensing), iii) SW languages can be used as meta-languages to define other special purpose languages, such as communication languages for knowledge sharing.

- **Location-based services and positioning systems:** Positioning systems have a mechanism for determining the location of an object in space, from sub-millimeter to meter accuracy. Coronato et al. [34] developed a service to locate mobile entities (people/devices) at any time in order to provide sets of services and information with different modalities of presentation and interaction.
- **Semantic Sensor Web:** Sheth et al. [35] proposed the semantic sensor Web (SSW) where sensor data are annotated with semantic metadata that increase interoperability and provide contextual information essential for situational knowledge. The SSW is the answer to the lack of integration and communication between networks, which often isolates important data streams [35].
- **Ambient Assisted Living (AAL):** AAL is a research domain that promotes intelligent assistant systems for a better, healthier, and safer life in the preferred living environments through the use of Information and Communication Technologies (ICT). More information on AAL is provided in the next Section.

5. Ambient Assisted Living

The aging population phenomenon is the primary motivation of AAL research. From a commercial perspective, AAL is rich in terms of technology (from tele-health systems to robotics) but also in terms of stakeholders (from service providers to policy makers, including core technology or platform developers) (Jacquet et al. [36]). The program AAL JP [37] is a funding initiative that aims to create a better quality of life for older adults and to strengthen the industrial opportunities in Europe through the use of ICT. In the next sections we will discuss the connection between AAL and the SW and Web 4.0, the reason why ontologies play a role in AAL, and the way AAL is realized along with two scenarios

5.1 Ambient Assisted Living and Web 3.0 – Web 4.0

According to Eichelberg & Lipprandt [38], the typical features of AAL systems, as standard interactive systems, are adaptivity, individualization, self-configuration and learning aptitude. These features have been traditionally achieved with methods developed within the field of Artificial Intelligence (AI). However, they believe that the Internet has been the driving force for the further development of those methods and mentioned the problems of Web services: i) integration and evaluation of data sources, like ambient technical devices; ii) cooperation between services of different kinds, such as device services; and iii) interoperability of the above mentioned services. These problems are very similar to the interoperability issues arising between AAL system components, hence Eichelberg & Lipprandt [38] state that the success of AAL systems is tightly coupled with the progress of Semantic Web technologies. The goal of using SW technologies in AAL is to create interoperability between heterogeneous devices (products from different vendors) and/or IT services to promote cooperation between AAL systems and the emergence of innovative business models [38].

Web 4.0, the so-called Intelligent/Symbiotic Web, follows the Web 2.0 (Social Web) and Web 3.0 (Semantic Web) and is about knowledge-based linking of services and devices. It is about intelligent objects and environments, intelligent services, and intelligent products. Thus there is tight connection between Web 4.0 and AAL, since AAL is realized in smart and intelligent environments. In such environments, intelligent things and services are available, such as sensors that monitor the well being of users and transfer the data to caregivers, robots that drive users to their preferred destination, TVs that can be controlled through gestures, etc.

Aghaei et al. [39] points out that Web 4.0 will be about a linked Web that communicates with humans in a similar manner that humans communicate with each other, for example, taking the role of a personal assistant. They believe that it will be possible to build more powerful interfaces, such as mind-controlled interfaces. Murugesan [40] stated that Web 4.0 will harness the power of human and machine intelligence on a ubiquitous Web in which both people and computers not only interact but also reason and assist each other in smart ways.

Moreover, Web 4.0 is characterized by the so-called ambient findability. Google allows users to search the web and users' desktop and also extend this concept to the physical world. Some examples are to tag physical objects with the mobile phone, such as wallet, documents, but even people or animals. Users can use Google to see what objects have been tagged and Google can also locate the objects for the user. In this concept, RFID-like technology, GPS and mobile phone tricorders are needed. Also here the connection between findability and AAL is present, as smart objects with RFID are an important component of AAL and IoT (see scenarios in 5.2.2).

Copyright (c) 2015 Advances in Computer Science: an International Journal. All Rights Reserved.

To sum up, Web 4.0 can support AAL by linking intelligent things and services through Web technology keyed to sensors, like RFID and GPS. However, according to Eichelberg & Lipprandt [38], until the integration of AAL systems into the era of Web 4.0, there is still significant progress needed concerning the semantic technologies. For instance, development of tools for collaborative development of formal semantic knowledge representations; integration of domain experts and standardization; ontology matching, reasoning and evaluation as well as semantic sensor networks and "Semantic Enterprise" methods for the migration of IT-processes in linked systems. Information of how ontologies are related to AAL is given in 6.2.1. An example project which combines sensor networks is SHIP (Semantic Heterogeneous Integration of Processes) [41]. It combines separate devices, components and sensors to yield one coherent, intelligent and complete system. The key concept of SHIP is a semantic model, which brings together the data of the physical environment and the separate devices to be integrated. One application domain of SHIP is the Bremen Ambient Assisted Living Lab[1], where heterogeneous services and devices are combined in integrated assistants

5.2 Realization and evaluation of AAL

AAL is primarily realized in domestic environments, i.e. the houses of senior people. Homes equipped with AAL technology are called smart homes. Moreover, AAL systems can be applied in hospitals and nursing homes. Generally speaking, the objective of AAL systems is to increase the quality of life of the elderly, maintain their well-being and independence. However, achieving these outcomes requires the involvement of third parties (e.g. caregivers, family) through remote AAL services. Nehmer et al. [42] distinguished three types of remote AAL services: emergency treatment, autonomy enhancement, and comfort services.

The projects funded by the AAL JP programme cover solutions for prevention and management of chronic conditions of the elderly, advancement of social interaction, participation in the self-serve society, and advancement of mobility and (self-) management of daily life activities of the elderly at home. Thus AAL is multifaceted with specific sub-objectives depending on the kind of application to be developed.

Regarding the involvement of end users in AAL, the project *A2E2* [43] involves users in several phases of the project, including focus groups, pilots, and an effectiveness study. Three groups are used: eldery clients, care professionals, and care researchers. Users are interviewed to find out which requirements they have on the particular interface to be developed. In the project *CARE* [44], which develops a fall detector, more than 200 end users in Austria, Finland, Germany and Hungary were questioned regarding the need for a fall detector; they answered that the current fall detectors (wearable systems) are not satisfactory and do not have high acceptance in the independent living context. Thus generally speaking, end-users are involved in current AAL related projects either through answering questionnaires or participating in user studies. Their involvement includes analysis of technical achievements/ requirements of the developed product, acceptance and usability of the prototypes, and also often ergonomic, cognition, and psychological aspects.

As for adoption of AAL systems by end users, this depends on various aspects, such as an application's obtrusiveness and the willingness of users. In many AAL systems, bio-sensors (activity, blood pressure- and weight sensors) are employed to monitor the health conditions of the users. Sensors/cameras are placed at home, so that the seniors' activities are monitored and shared between informal carers, families and friends. The assisted have to decide whether their well being should be monitored in order to avoid undesired situations, but also to keep the technology as unobtrusive as possible, so that they preserve dignity and maintain privacy and confidentiality. Weber [45] stated that an adequate legal framework must take the technology of the IoT into account and would be established by an international legislator, which is supplemented by the private sector according to specific needs.

Sun et al. [46] referred to some other challenges of AAL systems: i) dynamic of service availability and ii) service mapping. The Service Oriented Architecture, which supports the connection of various services, tackles the dynamicity problem. For service mapping, ontology libraries are required to precisely describe the services. There should be a so-called "mutual assistance community" where a smart home is managed by a local coordinator to build up a safety environment around the assisted people and the elderly should find themselves with a more active living attitude [46].

5.2.1 Ontologies and AAL

AAL applications are trans-disciplinary, because they mix automatic control with modeling of user behavior. Thus, the ability to reuse knowledge and integrate several knowledge domains is particularly important [36]. Furthermore, AAL is a very open and changing field, so extensibility is key. In addition, an AAL environment requires a standard way of exchanging knowledge between software and hardware devices. Therefore [36] believe that ontologies are well adapted to these needs:

i) are extensible to take into account new applications;
ii) provide a standard infrastructure for sharing knowledge;

[1] www.baall.org, 05/09/13

Copyright (c) 2015 Advances in Computer Science: an International Journal. All Rights Reserved.

iii) semantic relationships, such as equivalence, may be expressed between various knowledge sources, thus permitting easy integration.

Jacquet et al. [36] presented a framework in which ontologies enable the expression of users' preferences in order to personalize the system behavior: it stores preferences and contains application-specific modules.
Another ontology-centered design is used in the SOPRANO Ambient Middleware (SAM) [47]. SAM receives user commands and sensor inputs, enriches them semantically and triggers appropriate reactions via actuators in a smart home. The ontology is used as a blueprint for the internal data models of the components, for communication between components, and for communication between the technical system and the typically non-technical user.
In AAL there is often a problem of disambiguation between complex situations and simple sensors events. For example, if the person does not react to a doorbell ring, it may indicate that they have a serious problem, or alternatively it may indicate that they are unavailable, e.g. taking a bath [48]. Therefore, Muñoz et al. [48] proposed an AAL system based on a multi-agent architecture responsible for analyzing the data produced by different types of sensors and inferring what contexts can be associated to the monitored person. SW ontologies are adopted to model sensor events and the person's context. The agents use rules defined on such ontologies to infer information about the current context. In the case that agents discover inconsistent contexts, argumentation techniques are used to disambiguate the situation by comparing the arguments that each agent creates. In their ontology, concepts represent rooms, home elements, and sensors along with relationships among them.
Furthermore, Hois [49] designed different modularized spatial ontologies applied to an AAL application. This application has different types of information to define: (1) architectural building elements (walls), (2) functional information of room types (kitchen) and assistive devices (temperature sensors), (3) types of user actions (cooking), (4) types of furniture or devices inside the apartment and their conditions (whether the stove is in use), and (5) requirements and constraints of the AAL system (temperature regulations). Hois [49] designed different, but related, ontologies to manage this heterogeneous information. Their interactions determine the system's characteristics and the way it identifies potential abnormal situations implemented as ontological query answering in order to monitor the situation in concrete contexts.
Last but not least, in the project OASIS, one of the challenges was to achieve interoperability spanning complex services in the areas of *Independent Living*, *Autonomous Mobility* and *Homes and Workplaces*, including AAL. Due to the diversity of types of services, Bateman et al. [50] suggested the support of cross-domain networked ontologies. They developed the OASIS Common Ontological Framework, a knowledge representation paradigm that provides: (i) methodological principles for developing interoperable ontologies, (ii) a 'hyper-ontology' that facilitates formal semantic interoperability across ontologies and (iii) an appropriate software infrastructure for supporting heterogeneous ontologies.

5.2.2 AAL Scenarios

Two AAL scenarios will now be presented that demonstrate how multi-lingual and -modal applications, coupled with SW and Web 4.0, can improve the quality of life of senior citizens:

- *Scenario 1:* John is 70 years old and lives alone in a smart home equipped with intelligent, height-adaptable devices. He just woke up and wants to put on his clothes. His wardrobe suggests to him to wear brown trousers and a blue pullover. Then he goes to the supermarket for his daily shopping. He comes back and puts his purchased products into the fridge. The fridge registers the products. Then he wants to take a rest and watch TV. He lies on the bed; the bed is set to his favourite position with the headrest and footrest being set slightly higher. While he was at the supermarket, his daughter called him. The TV informs him about this missed call. Then he wants to cook his favourite meal; he goes to the kitchen and the kitchen reminds him about the recipe going through all the steps. The next day, when he goes again to the supermarket, his mobile reminds him that he has to buy milk.
- *Scenario 2:* Svetlana is from Ukraine and lives together with Maria, 85 years old from England, at Maria's smart home. Svetlana is caring staff, i.e. she cooks, cleans, helps in shopping, etc. Svetlana does not speak English very well; thus she speaks in Ukrainian to Maria, but also to electronic devices (TV, oven, etc.) and Maria hears it back in English. Alternatively to the voice commands, they can control the devices through a GUI or through haptics on the devices that this is available.

The above scenarios have a lot of hardware and software requirements, some of which are currently under development in the project SyncReal[1] at the University of Bremen; we will study these scenarios in more detail in the subsequent paragraphs.

- **Intelligent devices:** these are the wardrobe, fridge, cupboards, bed, and TV. The clothes in the wardrobe are marked with an RFID tag (IoT – Web 4.0) and the wardrobe can suggest to the user what to wear through

[1] http://www.syncreal.de, 25/06/2013

Copyright (c) 2015 Advances in Computer Science: an International Journal. All Rights Reserved.

these RFIDs and motion sensors (Beins [51]). This is useful, among other benefits, for people with memory deficit or visual impairments. It can also remind people to wash clothes if there are not many clothes left in the wardrobe. Similarly, the fridge and the cupboards register all products that are placed in and taken out by storing them in a database. This information is then transferred to other devices, such as mobile phones, so that John could see the next day that there is no more milk in the fridge (Voigt [52]). The bed is set automatically to a specific height position every time that he wants to watch TV (context-sensitive).

- **Semantic integration of ambient assistance:** John could see the missed call of his daughter on the TV owing to formal semantic modeling and open standards; the semantic interoperability allows the integration of the telephone with the TV.

- **Speech-to-speech dialogue system:** the language barrier between Svetlana and Maria is not a problem due to the speech-to-speech technology that is implemented in the home system. It includes three technologies: i) speech recognition, ii) Machine Translation, iii) speech synthesis; advantages and drawbacks of all three technologies have to be balanced. The dialogue system is also multimodal giving the possibility to interact with either through GUI, voice commands or haptics. It can be applied not only in electronic appliances, but also in robots. Information about speech-to-speech translation in AAL can be found in Anastasiou [53].

6. Summary and Conclusion

In this paper we focused on multimodal applications of the SW and presented some challenges involved in the development of multi-lingual and -modal applications. We provided some examples of current and future application domains, focusing on AAL. As there are large individual differences in people's abilities and preferences to use different interaction modes, multi-lingual and -modal interfaces will increase the accessibility of information through ICT technology for users of different ages, skill levels, cognitive styles, sensory and motor impairments, or native languages. ICT and SW applications gain ground rapidly today in everyday life and are available to a broader range of everyday users and usage contexts. Thus the needs and preferences of many users should be taken into account in the development of future applications. High customization and personalization of applications is needed, both because the limitations of challenged people can vary significantly and change constantly and in order to minimize the learning effort and cognitive load.
AAL can efficiently combine multimodality and SW applications in the future to increase the quality of life of the elderly and challenged people. The AAL market is changing and is expected to boom in the next few years as a result of demographic developments and R&D investment by industries and stakeholders. Currently the ICT for AAL is very expensive; projects test AAL prototypes in living labs that can be applied in domestic environments in the future. The technology is still often obtrusive (motion sensors), although researchers are working towards a goal of *invisible technology*. In addition, often the data is "noisy", as it is based on fuzzy techniques, probabilistic systems, or Markov-based models. Generally speaking, in regards to the future of intelligent ambient technologies, not only intelligent devices (Web 4.0), but also semantic interoperability between devices and IT-services (Web 3.0) are necessary. In our opinion, as emphasized by the term "*semantic*", the SW should be context-sensitive, situation-adaptive, negotiating, clarifying, meaningful, and action-triggering. All these aspects are important both for SW-based dialogue systems and multimodal interfaces including various input and output modalities. We share the opinion of O Grady et al. [54], on their vision about evolutionary AAL systems, about the necessity for an adaptive (robust and adapting in real-time), open (not propriety AAL systems), scalable (integration of additional hardware sensors), and intuitive (support for many interaction modalities) software platform that incorporates autonomic and intelligent techniques.

References

[1] Krueger, M.W., Artificial Reality, Second Ed. Addison, Wesley, Redwood City CA, 1991.
[2] Lu, S., Dong, M., Fotouhi, F., "The Semantic Web: opportunities and challenges for next-generation Web applications", Information Research, 2002, 7 (4).
[3] Zhong, N., Liu, J., Yao, Y., In search of the wisdom web. Computer, 2002, 37 (11), pp. 27-31.
[4] D'Aquin, M.D., Motta, E., Sabou, M., Angeletou, S., Gridinoc, L., Lopez, V., Guidi, D., "Toward a New Generation of Semantic Web Applications", IEEE Intelligent Systems, 2008, pp. 20-28.
[5] Wachs, J., Stern, H., Edan, Y., Gillam, M., Feied, C., Smith, M., and Handler, J., "A hand-gesture sterile tool for browsing MRI images in the OR", Journal of the American Medical Informatics Association, 2008, 15, pp. 3321-3323.
[6] Wachs, J.P., Kölsch, M., Stern, H., Edan, Y., "Vision-based hand-gesture applications", Communications of the ACM, 2011, 54 (2), pp. 60-71.
[7] Asteriadis, S., Tzouveli, P., Karpouzis, K. Kollias, S., "Estimation of behavioral user state based on eye gaze and head pose-application in an e-learning environment.", Multimed Tools Appl, 2009, 41, pp. 469-493.
[8] OWL Web Ontology Language Use Cases and Requirements: http://www.w3.org/TR/webont-req/, June 2015
[9] Gracia, E. M. Ponsoda, P. Cimiano et al. "Challenges for the multilingual Web of Data", Journal of Web Semantics, 11: 2012, 63-71.

Copyright (c) 2015 Advances in Computer Science: an International Journal. All Rights Reserved.

[10] Buitelaar, P., Cimiano, P. (Eds.), Towards the Multilingual Semantic Web, Principles, Methods and Applications, Springer, 2014.
[11] Smeulders, A., Worring, M., Gupta, A., Jain, R., "Content-Based Image Retrieval at the End of the Early Years.", in: Proceedings of IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22 (12), pp. 1349-1380.
[12] Computerized Manufacturing Automation: Employment, Education, and the Workplace (Washington, D. C., U.S. Congress, Office of Technology Assessment, OTACIT-235, 1984.
[13] Seals, C.D., Clanton, K., Agarwal, R., Doswell, F., Thomas, C.M.: Lifelong Learning: Becoming Computer Savvy at a Later Age. Educational Gerontology, 2008, 34 (12), pp. 1055-1069.
[14] Bolt, R.A.: Put-that-there: Voice and gesture at the graphics interface. ACM Computer Graphics, 1980, 14 (3), pp. 262-270.
[15] Oviatt, S.L., Cohen, P.R., Wu, L. et al. "Designing the user interface for multimodal speech and gesture applications: State-of-the-art systems and research directions for 2000 and beyond", in Carroll, J. (Ed.), Human-Computer Interaction in the New Millennium, 2000, 15 (4), pp. 263-322.
[16] Wahlster, W., Reithinger, N., Blocher, A. "SmartKom: Multimodal communication with a life-like character", in Proceedings of the 7th European Conference on Speech Communication and Technology, 2001, pp. 1547-1550.
[17] Wahlster, W. "Towards Symmetric Multimodality: Fusion and Fission of Speech, Gesture and Facial Expression", in Günter, A., Kruse, R., Neumann, B. (Eds.): KI 2003: Advances in Artificial Intelligence. Proceedings of the 26th German Conference on Artificial Intelligence, 2003, pp. 1-18.
[18] Griol, D., Molina, J.M., Corrales, V., "The VoiceApp System: Speech Technologies to Access the Semantic Web", in Advances in Artificial Intelligence, 2011, pp. 393-402.
[19] He, Y.; Quan, T., Hui, S.C. "A multimodal restaurant finder for semantic web", in Proceedings of the 4th International Conference on Computing and Telecommunication Technologies, 2007.
[20] Thurmair, G. Searching *with* ontologies – searching *in* ontologies: Multilingual search in the Assistive Technology domain. Towards the Multilingual Semantic Web, 2013.
[21] United Nations Open Audit of Web Accessibility: http://www.un.org/esa/socdev/enable/documents/fnomensarep.pdf
[22] Web Accessibility for Older Users: A Literature Review: http://www.w3.org/TR/wai-age-literature/. Accessed 25 Aug 2013
[23] W3C Web Accessibility Initiative (WAI): http://www.w3.org/WAI/
[24] Benjamins, V.R., Contreras, J., Corcho, O., Gómez-Pérez, A. "Six Challenges for the Semantic Web", in KR2002 Semantic Web Workshop, 2002.
[25] Avrithis, Y., O'Connor, N.E., Staab, S., Troncy, R. "Introduction to the special issue on "semantic multimedia", in Multimedia Tools Applic, 2008, 39, pp. 143-147.
[26] Potamianos, A., Perakakis, "Human-computer interfaces to multimedia content: a review", in Maragos, P., Potamianos, A. Gros, P. (Eds), Multimodal Processing and Interaction: Audio, Video, Text, 2008, pp. 49-90.
[27] Atrey, P.K., Hossain, M.A, El Saddik, A., Kankanhalli, M.S. "Multimodal fusion for multimedia analysis: a survey" Multimedia Systems, 2010, 16, pp. 345-379.
[28] Oviatt, S., Cohen, P. "Multimodal interfaces that process what comes naturally", Communications of the ACM, 2000, 43 (3), pp. 45-52.
[29] Cross, C.W., Supporting multi-lingual user interaction with a multimodal application, Patent Application Publication, United States, Pub. No: US/2008/0235027, 2008.
[30] Atzori, L., Iera, A., Morabito, G. "The Internet of Things: A survey", in Computer Networks, 2010, 54, pp. 2787-2805.
[31] Perera, C., Liu, C.H., Jayawardena, S., Chen, M., "A Survey on Internet of Things From Industrial Market Perspective", in IEEE Access, 2015, 2, pp. 1660-1679.
[32] Henricksen, K., Indulska, A., Rakotonirainy, A. "Modeling context information in pervasive computing systems", in Proceedings of the 1st International Conference on Pervasive Computing, 2002, pp. 167-180.
[33] Chen, H., Finin, T., Joshi, A. "Semantic Web in the Context Broker Architecture", in Proceedings of the Second IEEE Annual Conference on Pervasive Computing and Communications, 2010, pp. 277-286.
[34] Coronato, A., Esposito, M., De Pietro, G. "A multimodal semantic location service for intelligent environments: an application for Smart Hospitals", in Personal and Ubiquitous Computing, 2009, 13 (7), pp. 527-538.
[35] Sheth, A., Henson, C., Sahoo, S. "Semantic Sensor Web" IEEE Internet Computing, 2008, pp- 78-83.
[36] Jacquet, C., Mohamed, A., Mateos, M. et al. "An Ambient Assisted Living Framework Supporting Personalization Based on Ontologies. Proceedings of *the 2nd International Conference on Ambient Computing, Applications, Services and Technologies*, 2012, pp. 12-18.
[37] Ambient Assisted Living Joint Programme (AAL JP): http://www.aal-europe.eu/. Accessed 5 Mai 2015
[38] Eichelberg, M., Lipprandt, M. (Eds.), Leitfaden interoperable Assistenzsysteme - vom Szenario zur Anforderung. Teil 2 der Publikationsreihe "Interoperabilität von AAL-Systemkomponenten". VDE-Verlag, 2013.
[39] Aghaei, S., Nematbakhsh, M.A., Farsani, H.K. "Evolution of the World Wide Web: From Web 1.0 to Web 4.0.", in International Journal of Web & Semantic Technology, 2010, 3 (1), pp. 1-10.
[40] Murugesan, S. "Web X.0: A Road Map. *Handbook of Research on Web 2.0, 3.0, and X.0: Technologies, Business, and Social Applications", in* Information Science Reference, 2010, pp. 1-11.
[41] Autexier, S., Hutter, D., Stahl, C. "An Implementation, Execution and Simulation Platform for Processes in Heterogeneous Smart Environments", in Proceedings of the 4th International Joint Conference on Ambient Intelligence, 2013.
[42] Nehmer, J., Becker, M., Karshmer, A., Lamm, R. "Living assistance systems: an ambient intelligence approach", in Proceedings of the 28th International Conference on Software Engineering, ACM, New York, NY, USA, 2006, pp. 43-50.

Copyright (c) 2015 Advances in Computer Science: an International Journal. All Rights Reserved.

[43] A2E2 project: *http://www.a2e2.eu/*. Accessed 11 June 2015
[44] CARE project*: http://care-aal.eu/en*. Accessed 11 June 2015
[45] Weber, R.H., "Internet of Things – New security and privacy challenges", Computer Law & Security Review, 2010, 26 (1), pp. 23-30.
[46] Sun, H., De Florio, V., Gui, N., Blondia, C. "Promises and Challenges of Ambient Assisted Living Systems", in Proceedings of the 6th International Conference on Information Technology: New Generations, 2009, pp. 1201-1207.
[47] Klein, M., Schmidt, A., Lauer, R. "Ontology-centred design of an ambient middleware for assisted living: The case of SOPRANO", in the 30th Annual German Conference on Artificial Intelligence, 2007.
[48] Muñoz, A., Augusto, J.C., Villa, A., Botia, J.A. "Design and evaluation of an ambient assisted living system based on an argumentative multi-agent system, Pers Ubiquit Comput 15: 377-387, (2011)
[49] Hois, J. "Modularizing Spatial Ontologies for Assisted Living Systems", in Proceedings of the 4th International Conference on Knowledge Science, Engineering and Management, 2010, 6291, pp. 424-435.
[50] Bateman, J., Castro, A., Normann, I., Pera, O., Garcia, L., Villaveces, J.M. OASIS common hyper-ontological framework, OASIS Project, Tech. Rep., 2009.
[51] Beins, S.: Konzeption der Verwaltung eines intelligenten Kleiderschranks, Bachelor Thesis, Fachbereich 3: Mathematik und Informatik, University of Bremen, 2013.
[52] Voigt, M. Entwicklung einer mittels Barcode-Lesegerätes automatisierten Einkaufsliste, Bachelor Thesis, Fachbereich 3: Mathematik und Informatik, University of Bremen, 2013.
[53] Anastasiou, D. "Speech-to-Speech Translation in Assisted Living. Proceedings of the 1st Workshop on Robotics in Assistive Environments", in the 4th International Conference on Pervasive technologies for Assistive Environments, 2011.
[54] O'Grady, M.J., Muldoon, C., Dragone, M., Tynan, R., O'Hare, G.M.P. "Towards evolutionary ambient assisted living systems", Journal of Ambient Intelligence and Humanized Computing, 2010, 1 (1), pp. 15-29.

Dr. Dimitra Anastasiou finished her PhD in 2010 within five years on the topic of "Machine Translation" at Saarland University, Germany. Then she worked for two years as a post-doc in the project "Centre for Next Generation Localisation" at the University of Limerick, Ireland. There she designed guidelines for localisation and internationalisation as well as file formats for metadata, leaded the CNGL-metadata group, and was a member of the XML Interchange File Format (XLIFF) Technical Committee. In the next two years she continued with the project "SFB/TR8 Spatial Cognition" at the University of Bremen, Germany. Her research focused on multimodal and multilingual assistive environments and improvement of dialogue systems with relation to assisted living environments. She run user studies in the "Bremen Ambient Assisted Living Lab (BAALL)" with participants interacting with intelligent devices and a wheelchair/robot and did a comparative analysis of cross-lingual spatial spoken and gesture commands. Currently she is working at the University of Oldenburg, Germany in the DFG project SOCIAL, which aims at facilitating spontaneous and informal communication in spatially distributed groups by exploiting smart environments and ambient intelligence.
In 2015 she has been awarded a Marie Curie-Individual Fellowship grant on the topic of Tangible User Interfaces. In the last years, she has supervised numerous BA, MA and PhD students. In total she has published a book (PhD version), 17 journal/magazines papers, 32 papers in conference and workshop proceedings, and she is editor of 6 workshop proceedings. In addition, she is a member of 13 programme committees for conferences and journals (such as Jounral of Information Science, Computer Standards and Interfaces journal). She has 6-year teaching experience mainly in the field of Computational Linguistics.

Copyright (c) 2015 Advances in Computer Science: an International Journal. All Rights Reserved.