Introducing SAM.F: The Semantic Ambient Media Framework

David Bouck-Standen
Kingsbridge Research Center
Hamburg, Germany
email: dbs@kingsbridge.eu

***Abstract*—In our digital society, any user can become a producer of media. With the heterogeneity of devices, which vary in their capabilities or hardware resources, the question of accessing these media in a meaningful and usable context grows more important, as devices and users are more and more interconnected through digital services. To encounter the various challenges these observations present, with the Semantic Ambient Media Framework, this article proposes a framework, in which media, devices, and services are extended and interconnected through semantic models for various contexts. Providing a Web-based API, the framework allows applications and devices to access media, which are provisioned through the framework's services. These services are automatically tailoring the media, depending on the semantic information on the context they are used in, their semantic interconnection with other media, and the specific application, device, and context they are accessed from. This contribution illustrates the concept and system architecture of the Semantic Ambient Media Framework from a developer's perspective and describes a practical scenario, in which the framework is already utilized, concluding with an outline of future work.**

Keywords-Semantic Media; Semantic Repository; Cross-Platform Media Provisioning.

I. INTRODUCTION

Today, interconnected and feature-rich multimedia systems allow users to produce high amounts of user-generated content. The contexts technology is used in also shift towards mobile and ubiquitous computing. The users utilize their personal mobile devices, such as smartphones, tablets, and other devices, to connect to interconnect with other systems through the Internet [1], [2].

As each multimedia system uses technologies with different interaction paradigms, they offer different capabilities for presentation, processing, and storing information in their own content repositories [3].

Focusing a vision of a convergence of personal or social information, at least the interconnection of multimedia systems, or at best a single multi-purpose multimedia repository system would be required [4]. The latter observation would also solve the problem of media being isolated for use in a single application or on a single device.

These challenges have been researched in various context-specific domains, as related work (cf. Section 2) indicates. With the *Semantic Ambient Media Framework* (SAM.F), this contribution presents a general context-independent approach. SAM.F is a framework that semantically interconnects (a) *media*, (b) *devices and applications*, and (c) *services*, which are enriched by digital properties in the form of semantic annotations. For both client application development, as well as the extension of framework functionality, SAM.F offers interfaces for developers.

In SAM.F, media consists of, e.g., text, photos, audio, videos, animations, or 3D objects. These are extended by digital properties, e.g., by classifying the media's content in the internal model of SAM.F.

Digital properties also include Meta data from the original file, such as Meta information on MIME type or encoding. For devices, in SAM.F, we model digital properties reflecting, e.g., the devices' capabilities', location, capacity, screen size, or screen resolution.

All digital properties are utilized by the services in SAM.F. Client applications running on users' devices access the services of SAM.F through Web-based interfaces. Each service serves a dedicated purpose, interconnecting devices and applications through the shared use of devices and media.

To be able to interconnect services and devices through media, SAM.F features an extendable service-based architecture, providing developers with dedicated interfaces and the means to develop new modularized services for SAM.F, as described in detail in this article. SAM.F is accessible for devices and applications through a Web-based API.

To illustrate the use of SAM.F, Section 2 outlines a practical scenario from a developer's perspective, referencing actual work [5]. In Section 3, related work is regarded. In Section 4, Semantic Media used in SAM.F, the system's architecture, as well as the modular service-based structure is illustrated in detail, followed by a summary and outlook in Section 5.

II. RELATED WORK

Semantic media comprises the integration of data, information and knowledge. This relates to the Semantic Web [6] and aims at allowing computer systems as well as humans to make sense of data found on the Web. This research field is of core interest since it yields naturally structured data about the world in a well-defined, reusable, and contextualized manner. The field of metadata-driven digital media repositories is related to this work [7] as well. Apart from the goals of delivering improved search results with the help of Meta information or even a semantic schema, SAM.F distinguishes itself from a pure repository by containing and using multiple repositories as internal components, as illustrated below. As Sikos [8] observes, semantic annotations feature unstructured, semi-structured, or structured media correlations. Sikos outlines the lack of structured annotation software, in particular with regard to generating semantic annotations for video clips automatically. SAM.F offers means for both structured and semi-structured semantic annotations. Through an interface, the functionality of SAM.F can be extended to, e.g., automatically annotate media as outlined below, but is not limited to video clips. By these means, SAM.F delivers even more sophisticated features.

In general, SAM.F facilitates collecting, consuming and structuring information through device-independent interaction with semantically annotated media, whereas the linked data research targets sharing and connecting data, information and knowledge on the Web [9]. The concept originally developed by the author [10] was already used in different contexts, e.g., the automatic reconstruction of 3D objects from photo and video footage based on semantically compiled sets of media [11]. However, with SAM.F, in this contribution the original concepts of the author [10] are presented in their revised version in order to expurgate over-weight services previously tailored for a narrow project-specific and less transferable use. Thus, SAM.F does not share code with or reuse code from any related work.

Blumenstein et al. [12] outline a technical concept in museum context, that relies on a server-based architecture to provide museum content in a multi-device ecology. SAM.F could be used in similar contexts, but is not limited to the use in museums.

Ambient systems can provide a platform for displaying

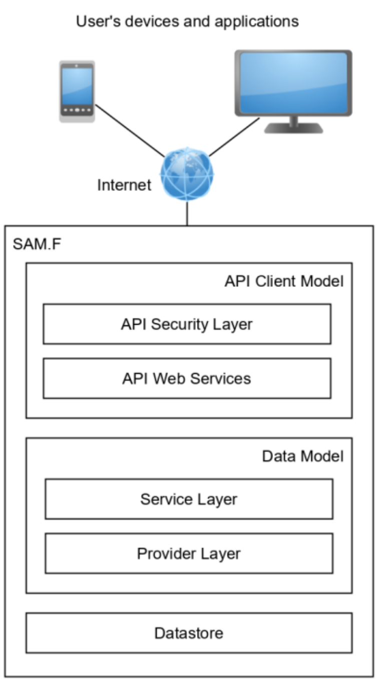

Figure 1: Layered architecture of SAM.F.

of and interaction with media [13]. In this context, the delivery of content on different devices is an important issue in SAM.F, e.g., with respect to the devices' capabilities or their context of use, and SAM.F addresses this challenge by provisioning media depending on applications and devices specifications or capabilities. SAM.F also addresses the issue of limited bandwidth of mobile devices.

The Social Web is related to this work, as it makes it easy for people to publish media online. Yadav et al. [14] propose a framework interconnecting Social Web and Semantic Web by semantically annotating and structuring information people share. SAM.F could be used in this way, but focuses on semantically enriched or described instances of media, devices and services. Semantic frameworks are used in various contexts, such as multimodal representation learning, as proposed by Wang et al. [15]. In their approach, Wang et al. use a deep neural framework to capture the high-level semantic correlations across modalities, which distinguishes this approach from SAM.F.

III. System Concept and Architecture

SAM.F is a smart media environment, which provides a device-independent access to and interaction with media through devices and applications.

The system's architecture of SAM.F is based on a system concept following these three considerations:

1. Web-based access provides platform-independent use of the services and access to media inside SAM.F and its repositories from the users' devices and applications.
2. a service-based modular architecture features extendibility, which provides developers with a framework to develop their own applications, which can be based on or reference to existing services within SAM.F.
3. the concept of Semantic Media regards media independently of their encoding or modality and automatically transcodes or converts media, where necessary and possible, to meet contexts, applications, and devices specifications or criteria.

In the following sections, we focus on the concept of Semantic Media fundamental to SAM.F. We illustrate the system's architecture and the service concept of SAM.F. Following, the application and device-specific media provisioning is outlined. In addition, technical details on the current implementation of SAM.F are given.

A. Semantic Media

In SAM.F, apart from services delivering media, media themselves are central. Semantic Media consist of plain media, such as text, audio, video, pictures, and 3D media, which are enriched by a dynamic set of semantic annotations. Together, plain media and semantic annotations form *Semantic Media* in SAM.F.

The dynamic set of semantic annotations stored in SAM.F for each media element consist of:

- the original Meta-data of the plain media file. For example, for photos taken with digital cameras, metadata usually contains information on the picture's location, and camera data such as camera make and model, or camera settings, such as camera capture settings. This data might be useful for SAM.F services and adding it to the set of annotations improves accessibility and performance when further processing media.

- data received from automated algorithms. Pictures for example are submitted to a Computer Vision algorithm by SAM.F automatically and in a background process in order to determine semantic annotations describing the media's content.
- data received from client applications. As the main user interaction with media through SAM.F is carried out through client applications, in which the context of use is known, this information is stored in additional semantic annotations. This information is collected automatically in a background process through the use of the SAM.F API Web Services, which implicitly reveal the context of use.
- data received from manual user interactions, such as manual annotations or correcting automatic annotations.

It should be noted that the semantic annotations of Semantic Media may not be complete or available for each media element at all times. This is, e.g., due to the context the media is created in, a foreign source the media is accessed from, or incomplete data entered by the user [16].

The set of annotations described above is not final and can be extended in context of client applications, devices or services.

In SAM.F, the complete set of semantic annotations are abstracted into the Data Model (cf. Figure 1) in order to be (i) accessible for all services running inside the framework and (ii) accessible independently of the underlying media repository in the Datastore layer (cf. Figure 1).

Not all annotations are made available for every client application or device through the API Web Services (cf. Figure 1), as the API Client Model only contains those properties that are required in the corresponding context. This way, overhead in the access of media through client applications is assumed to be significantly reduced. The effects on performance or bandwidth have however not been measured as part of this work.

It is one of the hypotheses of this work that the quality of semantic annotations as well as the interconnection of media will be a key issue for realizing appealing scenarios using SAM.F, as, e.g., described in the final Section of this article. An approach to achieve this is to gather additional sematic annotations through automated algorithms. As illustrated in Figure 2 and mentioned above, pictures, for example, are submitted to a Computer Vision algorithm. In the current implementation, SAM.F interfaces with Microsoft Cognitive Services. Thus, in the background, SAM.F computes additional semantic annotations, which are then stored in the internal Datastore (cf. Figure 1 and 2).

B. *System Architecture*

The architecture of SAM.F consists of a layer-based system concept, as illustrated in Figure 1. Client applications and devices utilized by users connect to the SAM.F *API Web Services* through the *API Security Layer* via the Internet in order to access media stored in SAM.F or interact with services in the *Service Layer* (cf. Figure 1).

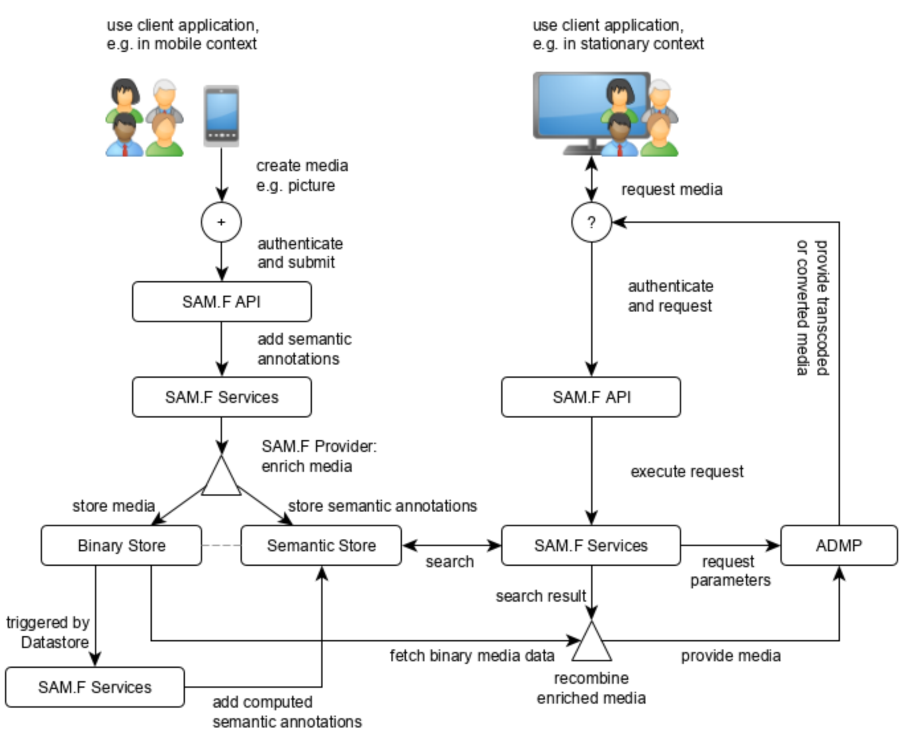

Figure 2: Media creation, enrichment through semantic annotations and retrieval. The Datastore consists of both Binary Store and Semantic Store.

When interacting with SAM.F, client applications as well as devices exchange information with the framework (cf. Figure 2) using a defined data model. Thus, for any context, the *API Client Model* can be extended to exactly match the needs of the application, device, or context, if necessary. API Web Services offer access to dedicated services provided by SAM.F, as the scenario described above outlines. Internally, SAM.F works with a dedicated *Data Model*, as illustrated in Figure 1. Any data is mapped from the *Datastore*, which includes external (semantic) databases as well as binary data stores, to the internal Data Model, which applies a homogenous model to potentially heterogenic data. Thus, SAM.F features the integration of different repositories and provides a combined access to Semantic Media. For simplification purposes, and in order to reduce the learning curve when implementing client applications accessing SAM.F, the internal Data Model is only used in the *Provider Layer*, which contains, e.g., authentication or data providers to be accessed by the upper Service Layer, and in the *Service Layer*, as shown in Figure 1. Any Semantic Media, together with semantic annotations, provided by a service to a client is mapped to the specific API Client Data Model, as outlined above, and being served through the API Web Services and the API Security Layer to the client application (cf. Figure 2).

With the Data Model only used internally, SAM.F accommodates different models used when storing media in digital repositories. A museum database for example differs significantly from, e.g., the DbPedia's semantic database. To be able to use heterogenous sources simultaneously, different data models are homogenized though the Data Model in SAM.F: by applying the data mapping techniques, the framework uses its own model internally, into which all other models are mapped. Applying data mapping in SAM.F produces constant overhead. However, services and applications, as well as their developers, benefit from only working with data models that are specific to the requirements of the services' or applications' context. This also reduces overhead when loading large sets of Semantic Media.

The range of functions of SAM.F is defined by the functionality provided by services residing in the Service Layer, as illustrated in Figure 1. In the scenario outlined below the developers extend SAM.F by implementing a custom service in order to realize the desired functionality. Thus, in the next section, the SAM.F services are regarded.

C. SAM.F Services

Following the implementation principles of SAM.F, a service features a dedicated set of functions in order to provide a certain functionality, e.g., for a use-case or scenario, as outlined above.

Utilizing the Data Model, through the Provider Layer, any service might access Semantic Media from the repositories included in SAM.F's Datastore layer. As a result, services may interchange information in a well-defined context.

SAM.F comes with a set of services that are useful to the developer in a Web-based environment and for developing applications in context of mobile use and the use of Semantic Media, explained in more detail below. In this article, we focus on the basic features the SAM.F services consist of:

- an authentication service to identify and authenticate sessions of applications, devices and users.
- a general media service that allows to retrieve or modify Semantic Media elements for a given keyword in a given general context. Media is retrieved both from the internal datastore, as well as external semantic databases housed in the Datastore layer (cf. Figure 1) and made available through SAM.F.
- the *Application and Device-specific Media Provisioning* (ADMP) service, which transcodes media based on different settings on client retrieval, as outlined below.

In the scenario outlined above, the developers extend the Service Layer of SAM.F (cf. Figure 1) and add their service to authenticate users on public displays. This service utilizes the modularized architecture of SAM.F and interfaces with the adjacent upper and lower layers. It also makes use of the default user authentication service. Service execution may either be triggered (i) on demand per request, or (ii) internally. This allows services of SAM.F to automatically run in the background without the necessity of user interactions.

D. Application and Device-specific Media Provisioning

Semantic Media in SAM.F can contain various types of plain media. However, their use is determined by the client applications. The devices running these applications are usually limited in their capabilities.

To address these challenges, SAM.F offers an *Application and Device-specific Media Provisioning* (ADMP) for any Semantic Media element retrieved through the API Web Services layer (cf. Figure 1).

In general, ADMP transcodes or converts Semantic Media due to specifications given. Trivial examples are the conversion of large photos into thumbnails, including cutting and cropping, if necessary.

ADMP is designed to work in two ways:

- on a per-request basis, in which the application submits the desired parameters (e.g., format, encoding, size, resolution) with every request, or
- on an application or device capability basis. As devices and applications are also represented in the Data Model (cf. Figure 1) of SAM.F, their capabilities are known to SAM.F. Thus, using per-request parameters can be omitted, if application or device capabilities can be generally set or are valid for multiple requests.

Especially in context of the Web-use of SAM.F and the heterogeneity of devices potentially accessing SAM.F, ADMP's usefulness can be illustrated through these examples, in which the correct parameter settings are presupposed: A video can be retrieved in different encodings or in matching screen size for the device's resolution. For example, ADMP can provide just the audio track of the video or just the textual transcript. The transcript can also be used to subtitle the video. More challenging 3D objects, which may not be viewed on any device, can be retrieved as a video of the 3D object rotating around the y-axis, or just as a picture in the form of a screenshot of the 3D object.

Reviewing key event-based multimedia applications, Tzelepis et al. [17] observe an enormous potential for exploiting new information sources by, e.g., semantically encoding relationships of different informational modalities, such as visual-audio-text. SAM.F provides these means by transcoding and converting Semantic Media in the background by automated processes.

As a side-effect, using ADMP reduces the use of bandwidth, which is of special interest in mobile contexts.

As these examples indicate, this way of provisioning media though SAM.F provides the means for a vast amount of use-cases. However, the author admits that not all possibilities have been implemented. The ADMP module, which also extends the Service Layer (cf. Figure 1), can be expanded, as it features an interface with an extendable list of parameters.

IV. Scenario

In context of a research project related to public displays, Josefine Kipke is part of a team developing a solution that features user authentication on public displays [5].

The motivation of this project is to provide access to private and sensitive information or functionalities in cases, where, e.g., the limited capabilities of a smartphone need to be extended, information or functionality is not to be made available on the user's device, or simply the use of a public display is intended by nature of the context.

This project presents serious challenges, as the user authentication on public displays in general is subject to vulnerability with regard to different sorts of attacks.

The team develops an interaction pattern, in which the users first authenticate themselves on their smartphone, and then enter a graphical code on the public display shown on their smartphone. Afterwards, the users confirm the logon using their smartphone again. Once the session has been confirmed, the users can start using the public display in a private context. Up to this point, the interaction pattern described is just a theoretical approach, which, to the developers, seems to cover the challenges with regard to a secure authentication on public displays. A prototype has yet to be implemented, not to mention to be validated and evaluated, as Josefine ascertains.

This rather simple approach is made possible through the interconnection of the user's smartphone and the public displays through the Internet and, most importantly, SAM.F.

The technical challenge presents itself in the fact that a public display, in practice, although connected to the Internet, for security and other reasons cannot and should not be remotely controlled by a smartphone that belongs to any random user passing by. These foreign devices are also not accessing the same network as the public display, which implies that any direct connection between a smartphone and a single public display is generally prohibited.

However, the team observes that public displays receive the media displayed from a server or a media framework, such as SAM.F and therefore can connect to SAM.F. Thus, SAM.F is able to identify a single public display through its registered session.

In addition, the team observes that connecting a user's device to the services of SAM.F poses no additional security issue, as SAM.F has been designed to interconnect devices and applications through services and Semantic Media, as outlined below.

After taking the observations mentioned above into account, the team develops a service for SAM.F using the interfaces provided by the framework. This service handles all the necessary steps to implement the interaction pattern for authentication the team developed. To be able to trigger and manage the authentication from a smartphone, the team also develops an application for smartphones, in this case a Web-based application that accesses SAM.F and the newly implemented service.

The team validates the functionality of their new extension of SAM.F under laboratory conditions. In the next step, they plan to evaluate the system under real conditions and with real users.

In summary, Josefine and the team have utilized the architecture of SAM.F, which interconnects the user's smartphones and public displays through the Internet, in order to provide an authentication mechanism for public displays. They have successfully extended SAM.F with a new service by implementing the corresponding interfaces.

V. Realization and Discussion

A first prototype implementation of SAM.F has been realized at the *Kingsbridge Research Center* (KRC). On the basis of a Windows Server system and its Internet Information Services (IIS) Web server, SAM.F is implemented in C# and runs as IIS Web application. Web services are provided using the Active Server Method File (ASMX) technology. Semantic annotations used in SAM.F are represented as RDF triples. For performance reasons analyzed under laboratory conditions in experimental settings, SAM.F's internally used RDF data is stored in a NoSQL database for performance reasons, although quantitative performance measurements are future work. SAM.F is compatible to semantic media repositories, e.g. using SPARQL to execute queries. Additionally, other required annotations for external media are stored in the internal datastore of SAM.F. In these terms, external media are media that are made available through SAM.F, but are stored in semantic datastores that are not managed by, but connected to SAM.F.

The approach of combining the automated enhancement of semantic annotations for media and delivering media in a device- or context-specific modality or encoding presents a technical novelty and distinguishes SAM.F from other media frameworks or repositories.

The current prototype has been validated under laboratory conditions. Computations are implemented to be carried out in a complexity of $O(n)$. Together with our project partner, as outlined below in more detail, we will integrate SAM.F for use in context of research projects. This will provide the opportunity to evaluate the system under real conditions with regard to functionality and performance.

 ISBN: 978-1-61208-739-9

VI. Summary and Outlook

With the Semantic Ambient Media Framework (SAM.F), this contribution presents a framework that semantically interconnects (a) *media*, (b) *devices and applications*, and (c) *services*. The practical scenario illustrated describes the use of SAM.F to provide means of a secure method of authentication for public displays [5] by developers. SAM.F provides Web-based access for devices and applications and features a service-based architecture, which allows for interaction with media, such as, e.g., text, pictures, audio, video, or 3D objects. The concept of SAM.F regards Semantic Media independently of their encoding and automatically transcodes or converts media, where necessary and possible, to meet contexts, applications, and devices specifications or criteria. Using SAM.F also solves the problem of media being isolated for use in a single application or on a single device, as SAM.F interconnects users and their devices through its services and Semantic Media.

SAM.F can be used in any context where interaction with Semantic Media is intended. Through technological means, SAM.F especially supports mobile contexts, e.g., through the application and device-specific provisioning of Semantic Media. Thus, SAM.F offers an enormous potential for exploiting new information sources, e.g., by the relationships of different informational modalities encoded semantically. As mentioned above, we have already used SAM.F in context of providing a multi-factor authentication method for public displays [5].

In future work, together with our project partner, the *Society for Audiovisual Archive of German-language Literature* based in the Hanseatic City of Bremen, we will utilize SAM.F as technical foundation to digitally enrich a cultural center for German literature. In this research project, SAM.F will interconnect media from various archives or libraries focusing on German literature. At the cultural center, the physical space will be enriched with digital media served provided by SAM.F. It is our hypothesis that providing meaningful digital content in a body- and space related environment fosters mindful knowledge.

The Kingsbridge Research Center (KRC) is a non-profit research company based in Hamburg, Germany. With our research and the systems, we develop, we have the goal to strengthen the meaningful use of digital technology in public environments, among others. We achieve this through our scientific and project-oriented work in an interdisciplinary team, by additionally achieving research funds through business returns to fund our non-profit activities, and the development of new future-oriented projects. At a time when many are confronting digitization with skepticism and uncertainty, we are committed to communicating security in the mindful use of these technologies.

References

[1] A. Whitmore, A. Agarwal, and L. Da Xu, ‘The Internet of Things—A survey of topics and trends’, Inf. Syst. Front., vol. 17, no. 2, pp. 261–274, Apr. 2015.

[2] Eurostat, ‘Internet use by individuals’, 260/2016, Dec. 2016.

[3] S. Vrochidis, B. Huet, E. Y. Chang, and I. Kompatsiaris, Big Data Analytics for Large-Scale Multimedia Search. John Wiley & Sons Ltd, 2019.

[4] C. Vassilakis et al., ‘Interconnecting Objects, Visitors, Sites and (Hi)Stories Across Cultural and Historical Concepts: The CrossCult Project’, in Digital Heritage. Progress in Cultural Heritage: Documentation, Preservation, and Protection, Cham, 2016, pp. 501–510.

[5] D. Bouck-Standen and J. Kipke, ‘Multi-Factor Authentication for Public Displays using the Semantic Ambient Media Framework’, In Press, ADVCOMP’19, Porto, 2019.

[6] T. Berners-Lee, ‘The Semantic Web’, Sci. Am., pp. 30–37, 2001.

[7] F. Nack, ‘The future in digital media computing is meta’, IEEE Multimed., vol. 11, no. 2, pp. 10–13, 2004.

[8] L. F. Sikos, ‘RDF-powered semantic video annotation tools with concept mapping to Linked Data for next-generation video indexing: a comprehensive review’, Multimed. Tools Appl., vol. 76, no. 12, pp. 14437–14460, Jun. 2017.

[9] C. Bizer, T. Heath, and T. Berners-Lee, ‘Linked data - the story so far’, Int J Semantic Web Inf Syst, vol. 5, no. 3, pp. 1–22, 2009.

[10] D. Bouck-Standen, ‘Construction of an API connecting the Network Environment for Multimedia Objects with Ambient Learning Spaces’, Master Thesis, DOI: 10.13140/RG.2.2.12155.00804, University of Luebeck, Luebeck, Germany, 2016.

[11] D. Bouck-Standen et al., ‘Reconstruction and Web-based Editing of 3D Objects from Photo and Video Footage for Ambient Learning Spaces’, Int. J. Adv. Intell. Syst., vol. 11, Jul. 2018.

[12] K. Blumenstein et al., ‘Bringing Your Own Device into Multi-device Ecologies: A Technical Concept’, in Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces, New York, NY, USA, 2017, pp. 306–311.

[13] P. J. Denning, Ed., The Invisible Future: The Seamless Integration of Technology into Everyday Life. New York, NY, USA: McGraw-Hill, Inc., 2002.

[14] U. Yadav, G. S. Narula, N. Duhan, and B. K. Murthy, ‘An overview of social semantic web framework’, in 2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom), 2016, pp. 769–773.

[15] C. Wang, H. Yang, and C. Meinel, ‘A deep semantic framework for multimodal representation learning’, Multimed. Tools Appl., vol. 75, no. 15, pp. 9255–9276, Aug. 2016.

[16] P. Oliveira and P. Gomes, ‘Instance-based Probabilistic Reasoning in the Semantic Web’, in Proceedings of the 18th International Conference on World Wide Web, New York, NY, USA, 2009, pp. 1067–1068.

[17] C. Tzelepis et al., ‘Event-based media processing and analysis: A survey of the literature’, Image Vis. Comput., vol. 53, pp. 3–19, 2016.

 ISBN: 978-1-61208-739-9