Automatic Discovery and Composition of Multimedia Adaptation Services

Jean-Claude Moissinac
Telecom ParisTech, LTCI UMR 4141
46 rue Barrault
F-75 634 Paris cedex
jean-claude.moissinac @ telecom-paristech.fr

***Abstract*—In this paper, after presenting the context, which indicates a considerable increase in the need for the adaptation of multimedia documents, we show that these results can be obtained by the composition of basic services. Nevertheless, this requires the availability of semantic descriptions of services, for which a shared vocabulary and good practices still need to be defined. We identify a series of works that can contribute to this process and offer basic guidelines to establish these descriptions. This article especially highlights the importance of the development of semantic descriptions of several important families of multimedia processing and proposes a structure that is used to build and organize such descriptions.**

Keywords-multimedia; semantic web services; adaptation; service composition.

I. Introduction

Our environment is enriched every day by a greater number of communicating devices and multimedia document providers. From a user point-of-view, each of us today takes advantage of a finite number of devices, usually personal: a telephone, television, laptop, tablet, portable media player. The great variety in the features offered by each of these devices requires services returned to the terminals that are adapted to them. Tomorrow we'll probably be able to benefit from the functions offered by equipment found in the different places we move to [29].

From a provider of multimedia content point-of-view, this growing complexity is a problem. A provider is often obliged to offer the same multimedia content in several formats and presentations. The current methods of adaptation are not sufficient to cope with the variability of situations that must be taken into account: preferences or needs of users, equipment available, and context of use.

In this paper, we show why this situation makes it necessary to implement adaptation processes that are widely configurable and propose a methodology to do this.

Given the variety of multimedia documents that users are exchanging, it is difficult to require a producer of multimedia content to provide as many versions of a document as possible contexts of use. It is necessary to identify ways to adapt a variety of documents to different contexts, either known at the time the content is put online or unknown until the time of the consultation.

We consider it desirable to offer to Internet players the ability to provide processing resources for the adaptation of multimedia documents. We must define the methodology and establish the prerequisites to allow such operations.

Section 2 presents two usage scenarios that illustrate the need for the dynamic processing of service compositions for multimedia. Section 3 presents a set of technologies and works which can contribute, or have contributed to, the proposal of this article. Section 4 describes a general architecture for adaptation of multimedia documents. Section 5 provides guidance for the descriptions of processing services which focus on our work. Finally, Section 6 presents the next steps as we see them.

II. Example scenarios

To light the way, we present two usage scenarios, one inspired by [30] and [14] as an extension of work published in 2004 which is centered on the user, the other responding to the needs of multimedia providers.

A. Campus scenario

We assume that we are on the campus of an international university. Some courses are available as multimedia documents.

There are different situations in which the content is used: during a classroom course, to follow and annotate the current presentation; at home to learn; or, later, when using the knowledge acquired during the course.

Users access to that content in various ways as well. For example, a user preferring English might be using a terminal with a small screen (5cm x 5cm) and a good resolution (800x 600) with Wi-Fi access while another will be on a wired network with a large screen, and prefers Spanish. One user might be in a location where he can activate the sound, another not. Disabled users can be taken into account; for example, the text will be displayed larger for the visually impaired or will be converted by a Text-To-Speech utility if the context permits.

Finally, the emergence of new devices, tablets, media players with new features for restitution of the media, but also the ability to interact, requires taking into account these new modes of access.

In this scenario, it appears necessary to have a system that dynamically configures itself to provide the best adaptation of a multimedia document in a context only known at the time of the request. The system cannot be

 ISBN: 978-1-61208-195-3

limited to a fixed set of adaptations. It must be able to be configured dynamically and be extensible.

B. Broadcast ecosystem scenario

Another scenario can be found in the broadcast production industry, which needs to adapt a lot of content to many different user contexts.

The media industry has many new opportunities to exploit its productions and archives: mobile multimedia, on-demand content, new products built on archives, etc. To do that, the media industry must do a lot of various processing, dependent on the target.

To achieve this aim, the media industry must be able to provide different sort of processing, depending on the targeted user context; such modern media production facilities must to function enable to compose processes from a rich list of elementary processes such as transfer and storage, capture, transform, etc. This scenario is illustrated in the current effort of standardization of FIMS [8].

Our contribution focuses on the development of semantic descriptions of basic adaptation services, based on ontologies. These descriptions help to meet the need mentioned above, but may have many other uses in applications of Semantic Web Services. In the next section, we will discuss a series of works that contribute to, and complement our approach.

III. Related Works

In this section, we will discuss a set of works that may contribute to the approach presented in this document.

First, we assume that basic services will be accessed via the Internet. We include them in the generic class of Web services, either REST [7] or SOAP services [20] or other technologies to make services available on the Internet.

In order to achieve automated operation of these services, they must have a description formally usable by IT processes. A minimum concerns the description of each service interface; for this, the most common technology is WSDL [1]. We will see below that this is not enough.

We want to use a set of basic services that will work together to create a more complex service. For this, many works concern the composition of services. They generally focus on the fact that a developer creates a process by calling a set of Web Services. Major efforts are focused on this type of software production process applied to 'business' in business. A language emerged to describe the workflow created to oversee the services called: WS-BPEL[13]. 'runtimes' are able to supervise the execution of a process defined with WS-BPEL (e.g. we are working with ODE [23]).

Developers can read a service specification written in plain text to understand its role or do a search in a warehouse of services such as UDDI [23]- to find a service that meets their expectations.

To create an automatic dialing service, the WSDL description is not enough to describe the fact that it takes as input an image given by an URL to access it in the Internet; or to understand what transformation is applied to the image –the transformation is only known by its name. We need to have the role and effects of each service described: which is the role of a semantic description of services. Several techniques for semantic description of services have been proposed, including: SAWSDL [6], OWL-S [19], WSML [3]. The use of OWL-S to describe media adapters, for example, has been proposed as part of MPEG-21 [22]. The need arose to describe some effects of a service using rules. In common parlance, such a rule can define a part of the effect of an operation to resize an image 'if the object has a media width and that the service is applied, then the width of the media object will be changed'. The SWRL language [10] was proposed to represent such rules.

Planning an automated composition of services has so far resulted in only a few works. As for multimedia, the proposed solutions are, for example, heuristics [15], a systematic exploration of possibilities [16] or more complex methods based on rules describing a form of expertise [35][12][28]. An interesting solution was proposed by [8]. And it concluded, however, with the idea that ontology for multimedia adaptation services could help to solve the problems left open by the proposed solution. The search for such ontology and how to use it is at the heart of our proposal. More recently, [23] proposes a way to describe services with the goal to automatically build mashups. This work focuses on problems of automatically composing services with heterogeneous descriptions in heterogeneous domains and gives ideas on how to solve that important problem. Our work focuses on getting good enough descriptions in one domain, multimedia, to establish either widely used standard descriptions or to easily make a match from our description to another. In [34], we find an in-depth analysis of a composition process in the aim of performing various semantic analyses on multimedia content.

Very significant work was carried out around these concepts in the context of the European Initiative ESSI: WSML [3] is a language defined to formalize the modeling of web services offered by WSMO [35]; WSMX [36] defines a runtime environment and set of services.

Numerous studies have focused over the last decade on the adaptation of multimedia documents. For examples and significant elements of state of the art in this work; see [23] [32]. Pellan [23] proposed a method to directly choose an appropriate service, knowing an initial application and context. The proposal focuses on choosing a service tailored to a context; it has a real contribution in the way to obtain the appropriate service itself in a space of predefined variant of the service.

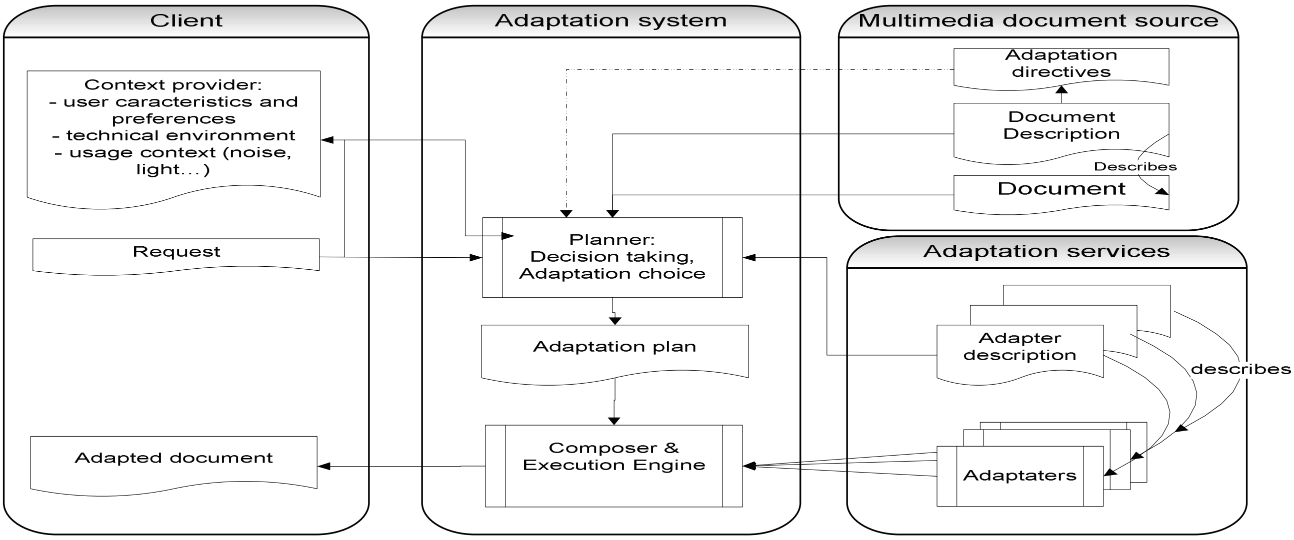

Figure 1. General architecture for distributed adaptation

MPEG-21 DIA [22] defines the desired adaptations to the media composing a document. This approach is insufficient because there is no possibility of describing dependencies between the different media adaptations. For example, if I describe that below a certain screen size, I no longer display a certain image, I have freed the space that can be used for text. But this dependence on the adaptation of the text based on that of the image cannot be described by DIA.

Even today, in many cases, adaptation is performed either at the server level by responding to a request with a different answer depending on what is known about the context of issuance of the request, or at the client -for example, the script by exploiting what is known locally in the terminal and about its user. The notion of proxy adaptation was introduced and is used by the network industry [14][23].

A very thorough study of the semantic description of multimedia services has been led by Bernhard Reiterer; the results are mainly available in German [28]. Very little research focuses on automatic composition applied to multimedia services; Derdour et al. [5] proposes a methodology for assembling basic services that provides services via mediation in order to make the entries of some services compatible with the output of other services.

IV. General Architecture for Distributed Adaptation

Fig. 1 shows the general principle of a distributed adaptation system as envisioned in this article.

A user requests a multimedia document or service. His request passes through an adaptation system. It is accompanied by an explicit or implicit context of use. We have presented in [21] an extended description of the following concepts. The main parts of an adaptation system are:

- the planner: it takes as input a description of the multimedia document and a description of a context and produces an adaptation graph, which describes the composition of a set of elementary steps, possibly subject to conditions, performed in parallel or in sequence, resulting in the appropriate document;
- context provider: as many works deal with the collection and provision of context, we leave this question out of our field of research and consider that a 'black box' is available and provides a context; there is a dependency between the context provider and the planner: the planner must be able to understand and use the context model of the context send by the context provider the context provider gives a context on demand or send an event each time a change occurs in the context;
- the source of multimedia documents: a component must manage the access to the multimedia document and its metadata description (source, nature...) which is to be adapted
- the composer searches for the needed services, while the runtime executes the plan and supervises the execution of selected services; at the end, it provides the result or a link to the result,
- elementary adapters: these components provide a specific adaptation for a part of the document.

The general principle is as follows. A consumer initiates an adaptation cycle. He/she sends a request to the planner; part of this request consists of a reference to a multimedia document and part of a reference to a context. The planner

 ISBN: 978-1-61208-195-3

uses this information to apply its adaptation algorithms and find a plan. It decides what will be the sequence of adaptation operations. The adaptation plan is sent to the composer who seeks elementary adapters to compose the ready to execute representation of the plan. The execution of the sequence is supervised by the runtime and returns a reference to access the resulting document.

V. DESCRIPTION OF ADAPTERS

A. Adaptaters for basic media

Most of the adapters we want to consider perform an elementary operation on a media category.

We first need to list the media types which must be taken into account. Beyond an obvious list (text, image, audio, video), we believe it is necessary to consider other media. Two examples are:

- A musical score, which is not a text or an image or sound, although it may be transformed into these three forms,
- A map, which is neither text nor an image, but an object which is much more complex.

In the present state of our descriptions, we have also introduced: 2D graphics, animated 2D graphics, 3D graphics and animated 3D graphics.

It is necessary to establish a methodology that allows some extensions, to define a new media that is useful, for example, in a specific activity, or some specializations, e.g. to distinguish speech from music as two kinds of audio documents.

We also see that some media have evolve in time; we are not yet sure of the best classification to be adopted. (Is a 2D graphic a degenerate case of an animated 2D graphics? Is an animated 2D graphic a 2D graphic extended by a description of a temporal process?)

Each media type must be associated with a list of characteristics that define it. Many ontologies have attempted to define the most common media and their representative characteristics. This situation is due to the fact that each ontology has its relevance in a given application. The W3C has taken steps to ensure correspondence among the models whenever possible [33]. In a preliminary study in 2007 [10], about two dozen models for describing media types, at least some media types, were identified. We found more during our work. In the work on WSMX, the concept of 'Data Mediation' [1] was introduced and can be a way to cope in an automated way with this problem. We find a similar concept in [5] in the work on the adaptation of input/output in a UML diagram representing the processing of a multimedia document.

B. Adapters

In [21], we described the top-level categories - transmoder, transformer, transcoder, extractor, composer- that we use as the basis for the definition of service classes.

- transmoder: changes a media from one modality to another –like creating an image of a text,
- transcoder: changes the format used to code a media without changing any other parameter –like transcoding an image from Jpeg to PNG)
- transformer: changes one or several intrinsic parameters of a media –like changing the size of an image,
- extractor: extracts each media and rules of composition from a composed multimedia document,
- composer: creates a composed multimedia document from a set of media and some rules to compose them.

These categories are then refined according to the media they take as input, the output they provide and the changes they perform on the media. We undertook a systematic description of adapters and have already identified about forty relevant types of adapters.

For example:

- text to speech is a text to audio transmoder whose input is mainly a text and output is an audio sequence,
- scaling of an image is a transformer that goes from one image to another image by changing certain characteristics.

We can see the existing services as being instanciated from the semantic description of some classes of services:

- class 'transformer/scaling', applicable to several media types: image, video, 2D graphics, animated 2D graphics,
- class 'transformer/summary' applicable to text, video, audio...

Whenever possible an adapter will be in one of the main classes. One last class has been defined to contain all adapters that are not clearly an instance of one of the previous classes. This last class is to be avoided because it conveys the poorest semantic. This class will include such additional adapters specific to a particular treatment on a type or a specific document format, for example, an adapter for PowerPoint documents or any document type specific to a specific activity domain.

Each adapter must have a basic WSDL description to conform to the call mechanisms of SOA services. But, as is now well identified, WSDL only provides technical information on how the call is made and no information on the meaning of the parameters, the nature of the result, the preconditions for the call or the effect of service execution on the surrounding world. To some extent, this information will be inherited from the ontology. However, each service may need a specific description not defined by the ontology, for example:

- the type of a parameter does not have an exact correspondence in the ontology and we need to define the mapping between the types and provide the type expected by the service (e.g., string versus URL),

- constraints on a parameter, for example, the width and height to resize an image can be limited to be homothetic,
- technical constraints imposed by the instance of the process, for example, the size of data transmitted is limited,
- pre-conditions of a nature that have nothing to do with the functions of the service (access control, security, etc.) can be attached to a service; these conditions involve concepts that are beyond the functional area under consideration -multimedia, and other concepts are necessary: other descriptors, other ontologies.

Several technical solutions have been proposed to complete the WSDL description for the semantic enrichment. We have begun experimenting with various solutions (OWL-S, WSML, SAWSDL and proprietary descriptions by extending WSDL). Apart from these experiments, we are trying to define the necessary descriptors, independently from a description language.

We believe that conceptually it is not the media that is provided to the adapter, but the access method of the media. In practice, the adapters receive as input a URI to access the media.

All descriptions deemed necessary in the context of Semantic Web Services (SWS) are often referred to by the acronym IOPE, Inputs Outputs-Preconditions Effects.

In the case of media processing, the minimum is to determine which characteristics of the media were changed and which descriptors are useful for the result of the transformation.

Following the work of [28], we consider different versions of the same multimedia document as variants of this document. We can describe the result of an adaptation, not exhaustively, which would not be possible, but only through changes made to formally described characteristics of the original.

As a general principle, we will consider that all the attributes of a transformed media remain the same, with the exception of those whose transformation is described. This has an advantage if a descriptor is added to a future version of a transformation: the adapters that were based on the current ontology work without that descriptor, by default, as if it were granted that they do not change the descriptor; this hypothesis seems relevant because, if it were not the case, it would mean that when we had done the initial description of the adapter, we forgot an important part of the description.

Consider two services to reduce the size of a picture, cropping and scaling. In both cases, the result is an image, which is a variant of the original image. In both cases, the width and height characteristics of the images are changed by the transformation. What differentiates the two transformations is that in the case of a crop, the image portion resulting from the processing is extracted from the original image while in the case of a change of scale, the resulting image is the result of the processing of all the image data. Most of the other features remain unchanged and can be skipped from the description. The amount of data used to represent the image is –generally- changed; we must mention that fact in the description.

Through this example, we see how difficult it can be to describe the adapters, but also the richness of the approach to build a large catalog of such descriptions. We undertook this work, which is being refined gradually; we are aware of similar work, for example at the University of Klagenfurt, but it seems that all the research projects we have identified are currently stopped.

The scientific community on multimedia adaptation and media processing and the one on Semantic Web Services will benefit from progress on these types of descriptions. Collective work will be necessary to achieve the goal of establishing shared concepts and vocabulary, to design a formal representation and to create the tools to facilitate the specification of new services based on the proposed model.

C. *Adaptation of a multimedia service*

In the work of Pellan [23], three dimensions of an adaptation process of a multimedia service are to be taken into account: a spatial adaptation, a temporal adaptation, and an adaptation of interactions.

We have begun to take into account the depth of all three dimensions for all media types and all categories of adaptation, but this work must still be completed.

We retain the assumption of Pellan: useful results can be obtained by considering that these three dimensions can be treated independently and that a composition of adaptations selected along each axis can be chosen.

On one hand, works such as those of [15] propose a method to adapt the layout of a document (spatial adaptation). On the other hand, we are exploring the possibilities of abstract representations of interactions [3][33][35], which could then allow concrete instantiations adapted to each situation.

VI. Conclusion and Future Works

We believe that the proposed approach will, in the future, be followed by other research. Indeed, it responds to the need to move from distributed storage on the Web to distributed processing. This approach benefits from unused software resources and from available bandwidth and processing capabilities on the Internet, usable in a decentralized manner and dynamically reconfigurable. These features could be a major asset for the spread of pervasive computing.

We think that describing all the known categories of multimedia services is possible; we have identified more than 50 categories and, probably, the categories added in the future will remain under a total of 100. Our main results are in the structuration of the categories, the principles of the description of each category and a first description of a group of categories.

Our future work is to complete the list of categories, clearly describing them all and, most importantly, to publish and share these descriptions to encourage their adoption.

ACKNOWLEDGMENT

This work was supported by the SOA2M project in the UBIMEDIA LAB, common with Alcatel/Lucent Bell Labs.

REFERENCES

[1] Christensen E., Curbera F., Meredith G., and Weerawarana S., "Web Services Description Language (WSDL) 1.1" http://www.w3.org/TR/wsdl [retrieved: April, 2012].

[2] Cimpian E., "Process Mediation in Semantic Web Services" 4th European Semantic Web Conference, June 2007, pp. 16-20, Innsbruck, Austria.

[3] Coutaz J., "PAC: an object oriented model for implementing userinterfaces" SIGCHI Bulletin Oct. 1987 Vol. 19 Num. 2

[4] Bruijn J. de, et al. "Web Service Modeling Language" http://www.w3.org/Submission/WSML [retrieved: April, 2012].

[5] Derdour M., Zine Ghoualmi N., Roose P., and Dalmau M., "UML Profile for Multimedia Software Architectures" Multimedia International Journal of Intelligence and Security, pp. 209-231, 2010

[6] Farrell J., et al. "Semantic Annotations for WSDL and XML Schema" http://www.w3.org/TR/sawsdl/, [retrieved: April, 2012].

[7] Fielding R., "Architectural Styles and the Design of Network-based SoftwareArchitectures" PhD Dissertation, University of California, Irvine

[8] "FIMS Media SOA Framework 1.0", September 2011, AMWA-EBU Specification

[9] Girma B, Brunie L, and Pierson J. "Content Adaptation in Distributed Multimedia Systems" Journal of Digital Information Management, special issue on Distributed Data Management, Vol. 3 No. 2, June 2005

[10] Hausenblas M. et al., "Multimedia Vocabularies on the SemanticWeb" www.w3.org/2005/Incubator/mmsem/XGR-vocabularies-20070724/ [retrieved: April, 2012].

[11] Horrocks I., Patel-Schneider P.F., Boley H., Tabet S., Grosof B., and Dean M., "SWRL: A Semantic Web Rule Language", http://www.w3.org/Submission/SWRL/ [retrieved: April, 2012].

[12] Jannach D., Leopold K., Timmerer C., and Hellwagner H., "A knowledge-based framework for multimedia adaptation", Applied Intelligence, 24 (2), pp. 109-125.

[13] Jordan D., et al. "Web Services Business Process Execution Language Version 2.0" http://docs.oasis-open.org/wsbpel/2.0/OS/wsbpel-v2.0-OS.html [retrieved: April, 2012].

[14] Kazi-Aoul Z., Demeure I. and Moissinac J.C., "PAAM: a Web Services Oriented Architecture for the Adaptation of Multimedia Composed Documents", "Parallel and Distributed Computing and Networks (NCSP)", Innsbruck, Austria.

[15] Kimiaei-Asadi M. and Dufourd J.C., "Context-Aware Semantic Adaptation of Multimedia Presentations", Proc.IEEE International Conference on Multimedia & Expo (ICME 2005), Amsterdam, pp. 362-365, July 2005

[16] Lardon J., Gravier C., and Fayolle J. (2010). "DOM tree estimation and computation: overview of a new web content adaptation system". In Proceedings of the 2nd ACM SIGCHI symposium on Engineering interactive computing *systems* (EICS '10). ACM, New York, NY, USA, pp. 357-360.

[17] Lee W. "Ontology for Media Resource", http://www.w3.org/TR/mediaont-10/ (retrieved 02/28/2011)

[18] Lopez-Velasco C., Villanoca-Oliver M. Gensel J., and Martin H. « Adaptabilité à l'utilisateur dans le contexte de services Web, Extraction des connaissances : état et perspectives » Revue Nouvelle des Technologies de l'Information, RNTI-E, ed. Cépaduès, 2005, pp. 153-158

[19] Martin D., et al. "OWL-S: Semantic Markup for Web Services" http://www.w3.org/Submission/OWL-S/ [retrieved: April, 2012].

[20] Mitra N. et al., « SOAP Version 1.2 Part 0: Primer » (link check: 03.03.2011) http://www.w3.org/TR/2003/REC-soap12-part0-20030624/ [retrieved: April, 2012].

[21] Moissinac J.C., Demeure I., and Kazi-Aoul Z., « Services d'adaptation de contenus multimédia, composition de services et pair-à-pair », CRIMES 09, St-Denis de la Réunion

[22] "MPEG-21 ISO / IEC JTC1/SC29/WG11/N6168, MPEG-21 Part7: Digital Item Adaptation", March 2007

[23] OASIS, http://www.oasis-open.org/specs/index.php#uddiv [retrieved: April, 2012].

[24] ODE, http://ode.apache.org/ [retrieved: April, 2012].

[25] Pellan B. and Concolato C., "Scalable Multimedia Authoring of Documents", Multimedia Tools and Applications, vol. 43, No. 3, pp. 225-252.

[26] Pietschmann S., Radeck C., and Meißner K. "Semantics-Based Discovery, Selection and Mediation for Presentation-Oriented Mashups", in Proc. Of Mashups 2011, September 2011; Lugano, Switzerland

[27] Rabin, J. "Guidelines for 1.0 Web Content Transformation Proxies" http://www.w3.org/TR/ct-guidelines/ [retrieved: April, 2012].

[28] Reiterer B., "Beschreibung von Multimedia-Adaptierungsoperationen als Semantic Web Services" PhD dissertation, Klagenfurt University

[29] Rodriguez B.H., Moissinac J.C., and Demeure I., "Multimodal pervasive services for the semantic web", UBIMOB 2010, Lyon, France

[30] Saathoff C. and Scherp A. "Unlocking the Semantics of Multimedia Presentations in the Web with the Multimedia Metadata Ontology "

[31] Salomoni P., Mirri S., Ferretti S., and Roccette M. "E-Learning Galore!Providing Quality Educational Experiences Across a Universe of Individuals with Special Needs Through Distributed Content Adaptation", LILW 2007

[32] Scherp A. "A Component Framework for Personalized Multimedia Applications" PhD Dissertation, University of Oldenburg

[33] Vella G. "XPL, a Presentation Language based on User Interface Design Pattern" in Proc. of 6th International Conf. on Computer and Information Science, 2007, pp. 285-290.

[34] Verborgh R., Van Deursen D., Mannens E., Poppe C. and Van de Walle R. "Enabling context-aware multimedia annotation by a novel generic semantic problem-solving platform" Multimed Tools and Applications, 2011

[35] WSMO, http://www.wsmo.org [retrieved: April, 2012] .

[36] WSMX, http://www.w3.org/Submission/WSMX [retrieved: April, 2012] .

[37] Yanagida T., Nonaka H., and Kurihara M. "User-Preferred Interface Design with Abstract Interaction Description Language" in IEEE International Conference on Systems, Man, and Cybernetics, Taipei, vol. 3, pp. 2458-2463.

[38] Zeng L., Benatallah B., Dumas M., Kalagnanam J., and Sheng Q.Z., "Quality Driven Web Services Composition", Proc.International WWW Conference, 2003, Budapest., pp. 411-421