Interoperability Requirements for a Sustainable Component to Support Management and Sharing of Digital Resources

Martin Memmel

German Research Center for Artificial Intelligence DFKI GmbH
Knowledge Management Department
Trippstadter Straße 122
Kaiserslautern, Germany

martin.memmel@dfki.de

ABSTRACT

Platforms such as YouTube, Flickr or Delicious that allow users to manage and share different kinds of digital resources belong to the most popular applications in what is usually subsumed under the umbrella term *Web 2.0*. In the context of PLEs, the ability to manage and share digital resources used within a learning process is also one of the most important features. This paper gives a coarse overview of key aspects to consider when aiming to provide a sustainable, adaptable component for resource management and sharing that can be integrated into different, heterogeneous digital environments. The ALOE (http://aloe-project.de) system will be presented as an example for the realisation of a respective component meeting the presented demands.

Keywords:

Resource sharing, Metadata, Interoperability, Social media

1. INTRODUCTION

The ability to support management and sharing of digital resources is a key feature in any PLE. A variety of tools and platforms exist that support such features for different kinds of contents (e.g., music, video, photos) and application scenarios. Yet, most of them only offer few possibilities to integrate with other tools, and each of the platforms usually has to be accessed separately in order to add, annotate, manage and search for contents. Social bookmarking systems such as Delicious[1] or Diigo[2] are a means to annotate and store information about resources from different sources. However, the vast majority of these systems only provides very basic means to organise own contributions and is neither instantiable, nor can be adapted to the specific needs of a scenario. Consequently, it is doubtful that future and not yet anticipatable scenarios can be supported by means of these tools.

What is needed in order to ensure sustainability is a comprehensive approach and framework that allows contributing, managing, and sharing arbitrary digital resources, that allows to exchange information with components in potentially any kind of digital environment, and that allows to adapt to the specifics of different scenarios.

In the following, we will first provide a coarse overview of interoperability requirements that have to be met when aiming at such a sustainable approach for resource management and sharing. The ALOE system will then be presented as an example for the realisation of a respective approach meeting these demands.

2. INTEROPERABILITY REQUIREMENTS

The IEEE[3] defines interoperability as follows [3]:

Interoperability is the ability of two or more systems or components to exchange information and to use the information that has been exchanged.

For the design of an application to manage and share digital resources, this has to be considered for

- the selection of supported application scenarios,
- the resource types to be supported,
- the metadata to be used, and
- the interfaces offered to users and other systems.

In the following, we will briefly discuss each of these aspects.

2.1 Supported Application Scenarios

As a first step in the design process, one has to decide for which scenarios support should be provided. Concentrating on a very specific scenario (e.g., “knowledge workers in a research department”) in the system's design can provide the benefit of a customised solution that takes into account the very specific characteristics of this scenario and the needs of the involved users. Yet, such a very targeted approach inevitably has several downsides:

- A huge modelling effort is required, e.g., for specifying and generating complex and tailored structures such as ontologies.

[1] see http://delicious.com
[2] see http://www.diigo.com
[3] see http://www.ieee.org

- A created model can always only be a snapshot – yet, people and organisations evolve. Thus, maintenance is required, which is usually a very complex and time-consuming task. Furthermore, no model is able to anticipate all possible needs and scenarios.
- The restriction to a very specific scenario and model hinders interoperability with other components (e.g., tools, technologies, and other data sources) that might be used in such a scenario. Although the adaptation of such components is sometimes possible, this is once again a usually complex and time-consuming task.

Instead of focusing on specific scenarios and defining prerequisites that have to be met for infrastructures, domains or user types, a sustainable component should follow a generic approach that can potentially be applied to support access to digital resources wherever this support is needed.

2.2 Resources

As any kind of resource can be part of a learning process, it should be possible to incorporate any type of digital resource. This includes arbitrary types of multimedia resources (e.g., HTML, PDF, MPEG), but also services or even physical resources just represented by a URI in an information system. "Incorporate" here means:

- When a digital resource is newly created or not yet accessible in the respective environment, it should be possible to contribute this digital resource, and to make it accessible. A system that offers this realises a *repository*.
- For digital resources that are already accessible in the respective environment, it should be possible to integrate them into the system without having to physically copy them. Otherwise, the following problems are very likely to arise:

  Maintenance issues: When digital resources are copied from a source where new contents are added, or existing contents are deleted and modified, the system will have to react to these changes. This is usually an expensive and time-consuming task.

  Memory requirements: Every digital resource that is copied will require some memory capacity. For large collections or certain resource types such as videos, this can result in very high memory requirements.

  Legal concerns: Sometimes it is simply forbidden to physically copy existing digital resources and to provide them in a different system.

  A system that offers this realises a *referatory*.

2.3 Metadata

Before discussing interoperability aspects for metadata elements and representation formats, we will first briefly elaborate on the need to take into account subjectivity and diversity.

Subjectivity and Diversity

We always have to consider that metadata is created for certain purposes in certain contexts, and that it is impossible to anticipate for whom and for what reasons a resource might be considered as relevant in the future. We have to accept and to embrace the fact that there is no "single and correct" way to describe a resource. As a consequence, we should allow *subjectivity*, and also *diversity* in the metadata about resources, instead of a metadata monoculture[4]. The need to support diversity is also motivated by the fact that we aim to support the access to digital resources in a variety of application scenarios, especially with heterogeneous components and most likely also heterogeneous metadata formats used for resources. These requirements are also supported by Nilsson et al. in [6], where the authors identified the needs for Semantic Web architecture, concluding that it should be:

- Evolving, supporting a dynamic metadata eco-system
- Extensible, allowing introduction of new vocabulary with new semantics
- Distributed, supporting descriptions by anyone about anything, anywhere
- Flexible, supporting unforeseen uses of resources
- Conceptual, supporting the evolution of human knowledge

It is clear that a one-size-fits-all solution for metadata about resources will not fit these needs. Instead, an ideal infrastructure would be generic in a way that allows for the generation of adequate resource descriptions for different users in different scenarios. Therefore, potentially any existing metadata might be incorporated. Different approaches to generate metadata can only be applied successfully in certain scenarios and for certain types of digital resources and metadata, and each of them has its benefits and limitations. Ideally, a digital environment should allow in each scenario to combine the benefits of each of the metadata generation approaches and to avoid the limitations. To allow for subjectivity and diversity, human generated metadata is most important, as only humans can contribute with different views and opinions. The need for diversity demands a non-authoritarian approach, supporting different views of the same resource. Thus, social metadata (i.e., metadata generated in social media environments) is most likely to meet these requirements, because it allows any user to contribute metadata about a resource.

As a consequence, we should be able to make use of potentially any metadata existing in the environment where our component is introduced. Moreover, it should be possible to contribute a variety of different metadata for digital resources. Such metadata can immediately be helpful for end users (e.g., bibliographic information about a resource), and it can also be an important source

[4] The term "metadata monoculture" was coined by Randy Goebel in 2008

for several functionalities (e.g., search or recommendations).

Metadata Interoperability

Concerning metadata elements and representation formats, drawing upon standards is required. Duval et al. provided the following fundamental principles for interoperability [2] that where enhanced by Nilsson et al. [5] who added the principle "Machine-processability".

- *Extensibility:* The ability to create structural additions to a metadata standard for specific needs of a domain, community or application
- *Modularity:* The ability to combine different, heterogeneous metadata fragments
- *Refinements:* The ability to create more fine-grained descriptions compatible with more coarse-grained metadata, and to translate a fine-grained into a more coarse-grained description
- *Multilingualism:* The ability to express, process, and display metadata in a number of different linguistic and cultural circumstances
- *Machine-processability:* The ability to automate processing of different aspects of the metadata specifications (e.g., to handle extensions, or understand refinements

2.4 Interfaces

Our component will of course have to provide interaction means for other systems as well as users. The way this is realised also has a significant impact on the desired interoperability.

Access by Systems

In order to allow the usage of an approach in as many scenarios as possible, and to foster the adoption of as many users as possible, the following aims should be followed:

- *Low technical barriers for system usage:* Users should be able to use functionalities with minimal efforts. This means as few restrictions as possible concerning the technical environments in which the hub can be used, as well as minimal installation efforts.
- *Low conceptional barriers for system usage:* Conceptional prerequisites for system usage such as the use of certain metadata formats should be kept to a minimum, while still allowing to provide added value for as many scenarios as possible.

As we want to enable the integration in existing environments with different systems and components, we need more than "just" an adequate user interface. Interfaces allowing an easy creation of mash-ups and complex functionalities using information from our component are required. Thus, interoperability is a very important aspect, and we should offer access to potentially any data and functionalities, regarding privacy aspects at the same time.

User Interfaces

Of course we need to provide interaction possibilities in an adequate way so that users are encouraged to make use of them. It is thus important to provide a user interface following principles such as simplicity [4] and joy-of-use [7]. Furthermore, mechanisms that attract and motivate users (e.g., by using reward mechanisms or game-based approaches) can be offered.

In order to allow decentralised contributions of digital resources and metadata in a way that fosters interoperability, users should be offered the possibility to use functionalities of our component in their usual contexts and applications. This can of course be realised if the persons in charge integrate functionalities into the respective applications. A more lightweight approach that allows integrating information or functionalities including user interfaces is to use *widgets*[5]. A widget is an element of a graphical user interface providing information and/or interaction possibilities [8], and that can be embedded into existing environments (e.g., a lot of widgets exist that can be embedded in HTML pages).

Furthermore, as we aim at a generic approach that can be used in a variety of scenarios, the user interfaces should be adaptable in a way that allows to address specific needs of a scenario (e.g., concerning a corporate identity or a certain terminology).

3. THE ALOE SYSTEM

The ALOE is a web-based social resource sharing platform developed at the Knowledge Management group of DFKI. It allows contributing, managing and sharing arbitrary types of digital resources such as text documents, music, or video files. Users are able to either upload resources (using the system as a repository) or by referencing a URL (using the system as a referatory). Users can tag, rate, and comment on resources, they can maintain resource portfolios, join and initiate groups, etc. Furthermore, arbitrary additional metadata can be associated with resources. Further system features are, among others:

- Group management for open/closed/invisible groups.
- Publish as private, public, or only for a certain group.
- Find resources with different types of search filters (title, description, tags, ...).
- Rank search results according to different criteria (most viewed, best rated, most recent, most bookmarked...).
- Advanced search with different filter criteria (filter by mime type, filter by license, filter by date, ...)
- Feed support (Atom) and email reports for different topics (e.g., activities in groups, activities on resources).

[5] The term *widget* is an abbreviation of *window gadget*

- Automatic metadata generation based on the Aperture[6] framework.
- Embedded player for various resource types (e.g., flash, mp3).
- Thumbnail generation for all common multimedia formats.
- Optional parallel uploads or status updated in other platforms (e.g., Delicious, Diigo, and Twitter).
- Export of own resources, search results and group resources as Netscape Bookmark File (importable in all common browser and bookmarking platforms).
- Functionalities are also offered as services (SOAP/REST API). This allows for an easy integration in other contexts and (existing) components.
- Easy adaptation of design, menus and texts for new scenarios.
- Arbitrary metadata can be integrated into the system and associated with resources thus, the integration of existing data is easy to realise.

3.1 System Design

To allow the usage of ALOE in as many scenarios as possible, and to foster the adoption of as many users as possible, ALOE was designed as a server-based application where information is exchanged via HTTP. On the one hand, the system's functionalities are offered via a graphical user interface that can be accessed with any common web browser that can connect to the ALOE server. On the other hand, a Web Service API is offered that allows accessing the ALOE functionalities. For these purposes, SOAP was chosen as a standard and platform-independent, XML-based protocol.

To foster interoperability, ALOE uses several standards for content representation and delivery:

- SOAP (Document/Literal) is used to pull/push data from the MACE frontend.
- An OAI target allows the harvesting of social metadata.
- A CAM service for usage metadata is provided.
- ALOE metadata uses DC elements wherever possible (dc:contributor, dc:date and dc:format (all created automatically when contributing a resource), dc:creator, dc:description, dc:rights and dc:title).

ALOE can be used as a stand-alone component, but also realises a social backbone that allows introducing social media paradigms in existing (heterogeneous) infrastructures. The system comprises the following components:

- *AloeFeeds and AloeInfoMail:* To create feeds and email reports about a variety of system activities, these components directly access the ALOE database as shown in Figure 1.
- *AloeMultimediaServlet:* The AloeMultimediaServlet is responsible for the provision of all resources stored in the ALOE database (e.g., buddy icons or file resources that were uploaded).
- *AloeThumbnailer:* This component is requested when preview images of uploaded files shall be generated.
- *AloeView:* The AloeView realises the Web Interface as already presented.
- *AloeWebService:* This is the main component of ALOE that offers access to a variety of more than 150 methods to access, contribute, and manipulate user data, resources, collections, and groups.
- *ApertureWebService:* This service uses the Aperture framework to extract metadata about resources. It can be used, e.g., to provide recommendations when resources are contributed.

3.2 Sample Use Cases

ALOE was developed in a way that allows to access its functionalities in arbitrary contexts and environments. Furthermore, the AloeView can easily be adapted to the needs of a specific scenario. Consequently, several instances of ALOE are used in different scenarios and projects, among others:

- *ALOE-public*[7] is an ALOE instance that is publicly available since 2008. It is used in several real-world scenarios (e.g., by the Institut Henri Tudor in Luxemburg), but also as a simple playground.
- *Mindpool* is DFKI's internal social media suite for all DFKI employees (in Berlin, Bremen, Kaiserslautern and Saarbrücken). Mindpool consists of two components: mindpool hints is a microblogging tool based on the Open Source microblogging service status.net, and mindpool treasures is a social resource sharing platform based on ALOE.
- *MACE:* The objective of the European Project MACE[8] (Metadata for Architectural Contents in Europe) is to create a common infrastructure for enriching and retrieving educational contents about architecture in Europe. It was co-funded by the EU eContentPlus program from 09/2006 until 10/2009. All community features in MACE are realised using ALOE as a social backbone.
- *C-LINK:* The aim of C-LINK (Conference Link) was the development of a web based tool to support conference attendees. With C-LINK, users can share papers and presentations, generate individual conference schedules, get personalized recommendations to find interesting events and attendees, etc. C-LINK is based on ALOE and was used during the KI 2008 (the Annual German Conference on Artificial Intelligence) in Kaiserslautern and the ICDAR 2009 (the

[6] see http://aperture.sourceforge.net

[7] see http://aloe-project.de/AloeView

[8] see http://www.mace-project.eu

International Conference on Document Analysis and Recognition) in Barcelona.

- *RADAR:* The aim of the project RADAR[9] (Resource Annotation and Delivery for Mobile Augmented Reality Services) is the development of an ALOE-based infrastructure to contribute, organize and annotate multimedia resources that can be used within mobile augmented reality services. Besides adapters for existing services like Layar or Wikitude, a new personalized and location-based mobile augmented reality service will also be developed.

4. ACKNOWLEDGMENTS

ALOE was partly developed within the project *CoMet* funded by the Stiftung Rheinland-Pfalz für Innovation. RADAR is sponsored by the Stiftung Rheinland-Pfalz für Innovation from 03/2010 until 02/2011.

5. REFERENCES

[1] Anderson, R. E. Social impacts of computing: Codes of professional ethics. *Social Science Computing Review 10*, 2 (Winter 1992), 453-469.

[2] Duval, E., Hodgins, W., Sutton, S. and Weibel, S. L. Metadata principles and practicalities. *D-lib Magazine, 8,4* (April 2002),:1–16.

[3] Geraci, A. IEEE Standard Computer Dictionary: Compilation of IEEE Standard Computer Glossaries. *Institute of Electrical and Electronics Engineers* (1991).

[4] Nielsen, J. Designing Web Usability: The Practice of Simplicity. *New Riders Publishing* (2000).

[5] Nilsson, M., Naeve, A., Duval, E., Johnston, P. and Massart, D. Harmonization of metadata standards. *Technical report, ProLearn: Network of Excellence in Professional Learning* (February 2008).

[6] Nilsson, M., Palmér, M. and Naeve, A. Semantic web metadata for e-learning - some architectural guidelines. *Proceedings of the 11th World Wide Web Conference* (2002).

[7] Reeps, I. E. Joy-of-use: eine neue Qualität für interaktive Produkte. *Master's thesis, University of Konstanz* (2004).

[8] Swick, R. R. and Ackerman, M. S. The x toolkit: More bricks for building user-interfaces or widgets for hire. *USENIX Winter* (1988), 221–228.

[9] see http://radar-project.de