CAM4Home: A Generic Ontology for a Rich Multimedia Experience
Samir Amir, Ioan Marius Bilasco, Mika Rautiainen

To cite this version:

HAL Id: hal-00837766
https://hal.archives-ouvertes.fr/hal-00837766
Submitted on 29 Sep 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
CAM4Home: A generic ontology for a rich multimedia experience

Samir Amir
University of Lille 1, France
samir.amir@lifl.fr

Ioan Marius Bilasco
University of Lille 1, France
marius.bilasco@lifl.fr

Mika Rautiainen
University of Oulu, Finland
Mika.Rautiainen@ee.oulu.fi

ABSTRACT

This paper describes a new generic metadata model, called CAM Metamodel, that merges altogether information about content, services, physical and technical environment in order to enable homogenous delivery and consumption of content. We introduce a metadata model that covers all these aspects and which can be easily extended so as to absorb new types of models and standards. We ensure this flexibility by introducing an abstract metamodel, which defines structured archetypes for metadata and metadata containers. The metamodel is the foundation for the technical metadata specification. We also introduce new structures in the abstract and core metamodels supporting the management of distributed community created metadata.

General Terms:
semantic web, knowledge engineering

Keywords:
metadata; multimedia; ontology, rdf

1. INTRODUCTION

Digital home systems are a reality. The possibility of switching between access devices while accessing online and multimedia contents is becoming common usage [9][15]. Deploying the content regardless of the heterogeneity of devices capable of accessing multimedia contents is a key factor for producers. Once created, the contents should be deployable on various configurations without subsequent transformation performed manually by content creators, or completely redesigned by authors to make them deployable on new access devices. With regard to the current state of the art, the variety of devices as well as deployed technologies, it is naive to consider that a content can be transmitted anywhere as it is. Automatic transformations are necessary to modify the encoding, the structure or the content itself to make it accessible in constraint contexts. Among these transformations, we can mention: transcoding (i.e. changing the bitrate of the content), transmoding (i.e. changing the modality to convey the content), transcoding (i.e. changing the encoding format). These three operations are closely linked to the encoding of the content. Basically, they are characterized by a set of parameters describing, respectively, the expected output rate, the output format (video, audio etc.), and the output codec. They do not need any explicit knowledge about the content itself. The obtained output is a degraded version of the initial content.

More intelligent transformations like summarization, filtering or recombination of the content, require specific resources and information. Such techniques need a deep access to the knowledge embedded within the content and to the characteristics of the broadcasting context. The broadcasting context is composed of access devices, networks, environmental configurations (indoor/outdoor, noisiness, visibility, etc.), user and community interests and preferences, as well as available adaptation services deployed on intermediate proxies. The description of all this content-related and context-related knowledge must be made following a standard format accepted by all the devices and services involved in the delivery and adaptation process. Due to multimedia popularity and the wideness of its application domain, the amount of information describing the multimedia content and context has become essential. Information serves to describe all entities involved in multimedia systems in order to facilitate multimedia delivery and consumption. It starts with the description of the content itself (size, type, etc.), its semantics (objects appearing in a picture, place where a picture was taken, etc.), the characteristics of the devices transmitting or consuming the content (TV, networks, etc.) and finally the consumer profile (preferences, interests, etc.).

Currently, several standards just cover some part of the information. MPEG-7 [7] and MPEG-21 DIA [5], for instance, deal with the description of the structure and semantics of a media object. MPEG-21 DIA [5], CC/PP [16] or Device Independent Activity offer tools for modeling the utilization context. WSDL [8] and OWLS [6] standards deal with the characterization of Web Services that might be involved in some adaptation process on the network. However, up to now, no integrated solution has clearly emerged. The MPEG-21 [7] set of tools (namely DID and DIA ) seems the most prolific candidate as it proposes tools to describe the content, the context and the transformation to be applied to a given media by using BSDL [1]. But still, the difficulty of acquiring a deep knowledge on these tools might not encourage the designers, who try to find friendly solutions for specific applications. We underline here that, as far as we know today, there is no off-the-shelf solutions hiding to the designers the complexity of standards (in particular in terms of encoding and retrieving information) like the one proposed above. In this chapter, we report on the creation of a new metadata framework that embeds information related to the content, context and adaptation services. This is part of the CAM4Home ITEA2 project2. A group of twenty multimedia academic and industrial practitioners from TV, 3G and Internet application fields defined a
restricted set of metadata requirements in order to support the con-
vergence of multimedia content in Digital Home environments. A
unified model, called CAM Metadata model [4] whose role is to
merge information about content, services and environment in or-
der to enable homogenous delivery and consumption of content.
We ensure flexibility by introducing an abstract metamodel, which
defines structured archetypes for metadata and metadata contain-
ers. We introduce structures supporting the management of dis-
tributed community created metadata. The paper is organized as
follows. First we present some current approaches that deal with
the aggregation of multimedia experiences and multimedia object
characterization. Then, we introduce the Collaborative Aggregated
Multimedia for Digital Home (CAM4Home)\(^1\), EUREKA-ITEA2
funded project, which lays basis on the content delivery platform
and gives framework for our metadata specification. Section 4 de-
fines the abstract metadata metamodel (called CAM Metamodel) in
detail and illustrate an instantiation of this metamodel. In Section
5 we present the technical realization, i.e. the format for metadata
encoding. Finally, we enumerate some directions for future work.

2. STUDY OF EXISTING MULTIMEDIA
STANDARDS

The ubiquitous presence of multimedia data requires an extensive
use of metadata for multimedia content retrieval, filtering and adap-
tation. The main metadata specifications efforts are focused on
the description of the multimedia content and context [2]. The MPEG-7
standard represents the most successful result in this field. It stan-
dardizes the description format (syntax and semantics) and decoding
of media features and aggregation but does not cover other infor-
mation type such as user created metadata (comment) or networks
characteristics.

Tv-anyTime standard [17] encloses specifications for the controlled
delivery of multimedia content to a user’s digital video recorder. It
seeks to exploit the evolution in convenient, high capacity storage
of digital information to provide consumers with a highly personal-
ized TV experience. Users will have access to content from a wide
variety of sources, tailored to their needs and personal preferences.
Tv-anyTime does not cover all required information related to con-
tent and to context description. For instance, information describ-
ing device, service, network characteristics are not covered by Tv-
anyTime. METS is a metadata standard designed to encode meta-
data for electronic texts, still images, digitized video, sound files
and other digital materials within electronic library collections. In
doing so, it attempts to address the lack of standardization in digital
library metadata practices which is currently inhibiting the growth
of coherent digital collections. METS offers a coherent overall
structure for encoding all relevant types of metadata (descriptive,
administrative, and structural). However, this standard does not of-er a description of the use context. Besides this lack, the semantic
description of this standard is very limited.

Tv-anyTime standard deals with most of the issues that we have
enumerated above. Still, it only concerns multimedia content and
context description but does not allow the description of several
other information. For instance, it does not provide native support
for user created metadata (such as comments), for the aggregation
of content and services, etc. Still, it can be extended using its DIDL
part.

3. METADATA FOR HOMOGENEOUS
MULTIMEDIA DELIVERY IN CAM4HOME

The objective of the CAM4Home project is to create a metadata ena-
bled content delivery framework to allow end users and commercial
content providers to create and deliver rich multimedia experiences.
These multimedia experiences are based on a novel concept of
collaborative aggregated multimedia (CAM). The Collaborative
Aggregated Multimedia (CAM) refers to aggregation and compo-
sition of individual multimedia contents (called objects) into a con-
tent bundle. The project develops one common metadata frame-
work for CAM content that can be applied for both personal and
commercial applications and is interoperable with relevant standard
metadata and content representation formats.

Tv-anyTime standard encloses specifications for the controlled

delivery of multimedia content to a user’s digital video recorder. It
seeks to exploit the evolution in convenient, high capacity storage
of digital information to provide consumers with a highly personal-
ized TV experience. Users will have access to content from a wide
variety of sources, tailored to their needs and personal preferences.
Tv-anyTime does not cover all required information related to con-
tent and to context description. For instance, information describ-
ing device, service, network characteristics are not covered by Tv-
anyTime. METS is a metadata standard designed to encode meta-
data for electronic texts, still images, digitized video, sound files
and other digital materials within electronic library collections. In
doing so, it attempts to address the lack of standardization in digital
library metadata practices which is currently inhibiting the growth
of coherent digital collections. METS offers a coherent overall
structure for encoding all relevant types of metadata (descriptive,
administrative, and structural). However, this standard does not of-er a description of the use context. Besides this lack, the semantic
description of this standard is very limited.

Tv-anyTime standard deals with most of the issues that we have
enumerated above. Still, it only concerns multimedia content and
context description but does not allow the description of several
other information. For instance, it does not provide native support
for user created metadata (such as comments), for the aggregation
of content and services, etc. Still, it can be extended using its DIDL
part.

\(^1\)http://www.cam4home-itea.org/
such metadata. In the following we focus on the lifecycle of content and aggregated content and its impact on the metadata framework. The description of the service platform is out of the scope of this paper.

In addition to the content itself, the platform or bundle adaptation services and the device and/or the network are not capable of directly consuming the content, in its current format, is taken. If the consumption of the content, in its current format, is taken. If the consumption context is composed of user and community interests and preferences, device physical, software and hardware properties, network properties and available adaptation services.

The availability of these pieces of information related to the identified elements either as metadata or as annotations is then required. We organize them within a metadata metamodel that covers all these aspects. We introduce a metamodel, as in our vision it is impossible to construct a model that satisfies all the future needs of the platform related to specific applications. On the basis of the structures and entities available in the metamodel, requirements of specific implementations of the CAM4Home platform are easily build. The following section is dedicated to the presentation of the CAM4Home project metadata metamodel.

4. CAM METAMODEL FOR METADATA

CAM Metamodel consists of structures and rules needed to build the metadata for describing the content and its eventual aggregation into CAM Bundles. CAM Metamodel provides the core concepts and the required metadata level information for collaborative distribution of multimedia and software content. CAM Metamodel can be partially or fully instantiated, as metadata are used in various systems targeting the intelligent delivery of aggregated content over a wide range of medium and terminals. In addition, CAM Metamodel is designed to allow easy encapsulation of existing external metadata formats (WSDL for services or MPEG-7 for content) into the structures of the instantiated metadata. Furthermore, the most important design goal of CAM Metamodel is to create extensible models that allow the definition of new structures and associations that a system might need in its operation. Extended CAM Metamodel provides added-value for the systems that understand the extensive structures, but can also be used by systems conforming to original CAM Metamodel. CAM Metamodel is illustrated in Figure 3.

![Fig. 2. CAM Metamodel.](image-url)

![Fig. 3. CAM Metamodel.](image-url)
—CAM Core Metamodel represents the core CAM entities (Bundles and Objects). It defines the core structures and associations that are related to the distributed content and its basic metadata. CAM Core Metamodel supports the representation of a wide variety of multimedia content and multimedia services as CAM Objects, e.g. downloadable applications, software services, images, video, etc. Specific metadata is attached to different types of multimedia and service entities.

—CAM Supplementary Metamodel represents application dependent entities such as devices and users. It specifies metadata that is required to enable interoperability of the platform services and enriches the descriptions of the content with information such as targeted users and communities or devices. Structures for profiles of users, communities, devices, networks and platform services were defined. For each one of these entities several profiles can be associated in order to support time-related (e.g. in the morning, in the afternoon) and usage-related (e.g. at home, at work) characteristics.

—CAM External Metamodel acts as an interface towards external metadata established standards (i.e. SMIL, MPEG7). It defines the constructors and descriptors into CAM Metamodel that acts as an interface towards the external metadata formats and encapsulates them into CAM Metamodel. We distinguish between external metadata that addresses core aspects (for instance MPEG-7 content descriptors) and supplementary aspects (MPEG-21 DIA user preference descriptors).

4.1 Abstract Metamodel overview
CAM Metamodel is based on a higher level description, i.e. CAM Abstract Metamodel. CAM Abstract Metamodel acts as a connecting element between different categories of the CAM Metamodel enabling the extension of the model with new structures by providing the basic constructs and associations of all CAM Metamodel entities. CAM Abstract Metamodel defines a generic categorization of concrete metadata entities and associations between them on an abstract level. CAM Abstract Metamodel enables the following features: generic description, reusability, shared knowledge and explicit rule definition.

CAM Abstract Metamodel provides a common backbone for intelligent distribution, delivery and consumption of CAM content that can be extended and specialized to cater more specific subdomains without losing system compatibility. When the common concepts and their relations in a domain are specified, the domain model can be efficiently reused for defining metadata models for other subdomains. As the domain concepts are well-defined, the abstract model allows interpretation (e.g. searches) of the extended subdomain models. Explicit restrictions on the model level are propagated to the more specific metadata entities. Thus, extending CAM Abstract Metamodel prevents the required system level policies from prohibited instantiations of the model or association of model elements.

CAM Abstract Metamodel introduces core, supplementary and external generic concepts. Abstract Core Metamodel provides generic associations for metadata entities related to CAM Objects, CAM Object metadata, CAM Bundles and CAM Bundle metadata. Abstract Supplementary Metamodel introduces the high level description for devices, services, users and communities related to the CAM4Home domain. Abstract External Metamodel categorizes existing external metadata formats that can be associated with the content of CAM Bundle when a specific metadata description of content is required.

External metadata can be associated with core metadata or supplementary metadata in order to enrich existing description or to reuse descriptions available in other standards. Supplementary metadata is associated to core metadata in order to precise the context (user, community, device, network) for which the CAM object or CAM bundle was initially created.

4.1.1 Abstract core metamodel. The abstract core metamodel defines metadata and metadata containers (see Figure 4). A metadata container can either contain structured metadata (ContentMetadata) or simple metadata (described by literals). For reinforcing the types associated with simple metadata, we have adopted the simple types defined by XML Schema language. Those types are introduced into the CAM Metamodel RDF schemas using rdfs:Datatype constructs (e.g. `<rdfs:Datatype rdf:about="xsd:date">` introduces the `xsd:date` type). This choice is motivated by the fact that most of the existing XML parsers can manage and interpret at least simple XSD types. In this way, the implementation of the CAM4Home metadata framework can benefit from existing technologies.

The CoreMetadata class is the parent class of all core related metadata. The CoreMetadata is specialized in structured metadata ContentMetadata class or in metadata container class. A container regroups structured ContentMetadata through the hasStructuredMetadata association or simple textual or numeric metadata (e.g. title, description, etc.). For example, AppearingConcept is a structured metadata as it embeds information about the description of the concept, the location of the concept in the media, etc. These classes are further specialized with regard to the type of metadata they address.

Hence we have identified ContentAggregativeMetadata and ContentAggregativeMetadataContainer for describing the aggregation of content. Concrete classes extending these abstract classes are Relationship and CAMBundleMetadata respectively. The Relationship class describes how CAM objects are related within a given CAM bundle. The CAMBundleMetadata regroups information about the aggregation of CAM objects within CAM bundle. ContentEssenceMetadata and its container are used to describe the actual content essence such as the URI of the content. CommunityCreatedMetadata and its container are used for describing metadata datatypes that are used to represent community created metadata (comments, ratings, social tags, etc.).

At abstract level those classes are further specialized in order to distinguish between user created metadata stored locally CCLocalMetadataContainer and shared user created metadata CCCSharedMetadataContainer. Since the aggregation of content bundles may happen chronologically by the collaborative effort of user communities, it has been taken into account in the design of core metadata. The purpose of this metadata is to make community contributions independent from the typical clientserver structure that exists in
current multimedia content sharing services in the web. User comments, user ratings and social tags are some examples of metadata created in the user community.

Since community created metadata is treated similarly to other traditional content essence description data, such as title and description text, it can be consumed and delivered in hybrid network configurations. In other words, community created metadata is part of the content flow. ContentFeatureMetadata and its container are used for describing features of the content itself and not the containing media object or media subject (author, description, target community, appearing concepts). ContentFeatureMetadata class is specialized in order to represent amongst others, concepts that are present in a CAM object (AppearingConcept). The ContentFeatureMetadata Container is the parent class of CAMElementMetadata and CAMBundleMetadata classes regrouping respectively feature information about a single CAM object or an aggregated CAM bundle.

4.1.2 Abstract Supplementary Metamodel. This part of the model permits interoperability among platform services and enhances the manipulation of Core Metadata. This metamodel concerns the metadata related to domains outside the CAM Bundle concept e.g. users, devices, networks, communities and services with its respective environment descriptions.

The metamodel (partially illustrated Figure 5) consists of entities and profiles. The entities (such as user, community, device, etc) can be linked to different metadata metadata profile (such as user profile, community profile, device profile, etc). The supplementary metadata provides detailed information about the entities related to personal information and interests for the user or hardware and software information for the device. The entities are linked between each other by means of references (see Figure 6). A user entity can belong to a community entity. A user entity uses a given device entity on a given network entity. A device entity uses, at a given time, a network entity.

4.1.3 Abstract External Metamodel. The abstract external model provides basic structures for integrating existing standards into the CAM4Home framework. The ability to include external metadata in CAM4Home metadata framework is a key feature as it permits practitioners of CAM4Home to benefit from existing metadata standards.

Two categories of external metadata descriptions are considered: external core metadata and external supplementary metadata. The external core metadata descriptions are related to core aspects of a CAM object or a CAM bundle. The external supplementary metadata descriptions are related to the characterization of supplementary entities. In order to underline the type of external metadata classes that can be considered for integration into the CAM4Home metadata framework, we have defined several subclasses for each type of external metadata. Hence, we have at abstract level, core-related and supplementary-related external metadata classes that support the integration of existing content-related or context-related standard descriptions.

5. TECHNICAL REALIZATION

Previous section introduced an abstract metamodel for describing structured content and context entities. For the realization of this model we have chosen an RDF Schema representation. The classes presented in the metamodel are implemented using rdf:Class constructs and the relations are represented as rdf:property. All the properties related to a resource (CAM object or CAM bundle) are encoded as XML sub-elements of the main container representing the content. Descriptions of content are made by means of RDF statements.

Before presenting the effective implementation of the model using RDF Schema we discuss the technological alternatives that we have explored. We have considered the extension of existing comprehensive multimedia description standards such as MPEG-21. The MPEG-21 covers similar topics around content and context environment descriptions. However, it does not natively describe information about services or community created metadata. A mixed usage of MPEG-21 and service description standards such as WSDL, OWL-S etc would have been an option. The main inconvenient of this solution is that the client would need to be capable of interpreting a very large set of encoding format and structures (specific to MPEG-21, specific to OWL-S [13], WSDL [11], etc). Our goal has been to construct simple yet extensible metadata solution deployable also on thin clients that are typical in a home network environment. We have defined a compact set of descriptors and description structures that covers a set of information needed for consuming content in heterogeneous environments. External metadata complements the core metadata by allowing the inclusion of other encoding schemes that can be interpreted by clients with extended decoding capabilities without confusion with the core information designed for basic clients. Our metadata model assumes the fact the metadata is provided collaboratively by several entities (content creators, aggregation, creators, consumers). While using a plain XML solution, all these entities need to have direct access to the document including the whole metadata descriptions related to a given content. This implies that a regular (not specifically authorized) user can directly modify information previously defined by the creator.
We have preferred a semantic approach based on RDF, as services and users can enrich the information base by forming statement about a given CAM object or CAM bundle. We did not utilize OWL concepts as we wanted to keep the metadata encoding as simple as possible. The power offered by the RDF Schema language satisfies the needs of CAM Metamodel.

5.0.4 Relationship encoding conventions. Since RDF Schema language does not provide means to differentiate between aggregation and composition structures are available in UML for modelling relationships, we have adopted the following encoding conventions:

(1) compositions are encoded using nesting or rdf:resource references (see Figure 7). Both solutions have the same semantics as they result in similar sets of RDF statements

(2) aggregations are encoded using specific rdf:property constructs that contain string references representing unique CAM4Home identifiers (UID) (e.g. targetDevice Reference property contains the UID of the device entity the CAM bundle was aggregated for see Figure 8). This allows existence of unique objects in the system and allows creating dynamic references between objects.

Editing such rules ensure that the model remains consistent when removing or adding entities or relationships in the system: all members of a composition must be deleted when a parent item is deleted; items with aggregation type relationship are not affected when any of them is deleted.

In the following subsections we give an exemplary set of abstract classes and properties in RDF and then introduce the specialization of the abstract constructs by giving examples of instantiated classes and properties.

5.0.5 Class and properties encoding. First, we have defined the rdf:Class and rdf:property corresponding to the abstract part of the model. In Figure 9 we introduce the RDF Schema definition ContentFeatureMetadata dataContainer classes. The figure illustrates the definition of properties (hasFeatureMetadata) linking the structured feature metadata to the feature metadata container. The rdf:subClassOf and the rdf:subPropertyOf constructs are being used to link core class and property definitions to corresponding abstract concepts. An example of this separation between the classes of properties is shown in Figure 10. The example illustrates the generic templates for introducing simple and structured core metadata constructs to specialize the abstract concepts into concrete metadata elements:

![Fig. 7. Example of a composition encoding.](image-url)

![Fig. 8. Example of an aggregation encoding.](image-url)

---

**Fig. 7.** Example of a composition encoding.

**Fig. 8.** Example of an aggregation encoding.

In the following subsections we give an exemplary set of abstract classes and properties in RDF and then introduce the specialization of the abstract constructs by giving examples of instantiated classes and properties.

---

**Fig. 9.** RDF-S definition of abstract classes and properties.

**Fig. 10.** RDF-S definitions of concrete classes and properties.

---

**Fig. 11.** RDF examples of CAM Objects and Bundles.

—any simple metadata used for content feature description is directly associated with a property which specializes the simpleFeatureProperty rdf:property.

—any structured metadata used for content feature description extends the ContentFeatureMetadata class and it is linked to its container by a specialization of hasStructuredMetadata rdf:property.

---

In the example shown in Figure 10, hasAppearingConcept is presenting a complex property in that it uses another class definition in the metamodel (the AppearingConcept class).

Figure 11 gives examples of the CAM Object and Bundle instances as seen in Figure 2. The figure describes a CAM Object instance which is further associated with the CAM Bundle. The CAM Ob-
ject references to the original video essence file using the metadata field "isMetadataOf". CAM Object description contains a visual concept that appears in the video, a fish. The CAM Bundle contains also community created metadata in the form of shared social tags.

6. DISCUSSION AND CONCLUSION

CAM Metamodel, the metadata model proposed in this paper serves the homogeneous deployment of content and services aggregated within bundles. The richness of the metadata model provides practitioners with an access to the whole range of information they need in order to deliver the same content to a wide range of devices and in heterogeneous deployment context.

The core part of the model encloses information about the content conveyed by CAM bundles and CAM objects. The supplementary part offers information about the context for which the content was initially conceived as well as information about the current deployment context so that adaptation engines can exploit this information in order to apply adequate adaptations. The proposed solution is highly extensible and can be widely used outside the CAM4Home usage context as external information, which is not primarily supported by CAM4Home, can be seamlessly added to the metamodel. In [3], we have explored possibilities of natively interpreting these external metadata without requiring any preknowledge about the external standard, by extracting semi-automatic mappings between CAM Metamodel constructs and structures presented in external standards.

The collaborative annotation of content is supported by the community created and shared metadata. Basic metadata such as comments, ranks and tags can be created by the community. The abstract level that we have designed ensures the extensibility of this kind of metadata. The generic containers for community created metadata are easily extensible in order to support new types of metadata.

Our proposed metadata model contains several important design elements that address the challenges that distributed service platform development are facing for multimedia content delivery. First, the separation of core, supplementary and external domains split metadata between the bloodstream of content delivery; facilities for supporting the distributed delivery and consumption; and interfaces that make the metadata encoding compatible with the available encodings in other multimedia delivery and management systems. Second, the proposed metadata model encompasses both multimedia services and content into a unified framework that aims at enabling collaborative aggregation of multimedia content over heterogeneous networks and devices throughout the content lifecycle. Third, the proposed model contains specifications to support distributed and incremental aggregation of value-added information that acknowledges recent trends in consumption of networked content, such as user collaboration in creation, manipulation and consumption of meaningful content as well as accumulation of information that helps to locate, filter, access, interpret and communicate it with other communities of users.

7. ACKNOWLEDGEMENTS

This work was supported by the ITEA2 (Silver) CAM4HOME European project.

8. REFERENCES


