Semantic Web, linked data and broadcasting

— more in common than you'd think!

Jean-Pierre Evain
EBU Technical

Tobias Bürger
Independent Expert

The purpose of this article is to raise awareness of Semantic Web technologies and on-going convergent developments between broadcasting and the Internet, i.e. in the W3C, the EBU, TV-Anytime and IPTC. It provides an overview of the genesis of the Semantic Web and its evolution towards Linked Data. Moreover, it introduces related technologies and W3C specifications.

After providing basic know-how on ontology modelling, the authors present in detail the W3C Media Annotation ontology, a metadata format for the description of media resources on the Web. This is followed by the identification of the main advantages and distinct features of RDF/OWL, and some disadvantages which are mainly due to a lack of market maturity of related technologies.

Finally the big picture around interoperability between the W3C Media Annotation Working Group (MAWG) and the main media metadata communities (EBUCore, MPEG-7, TV-Anytime and NewsML-G2) is drawn at the eve of the introduction of HTML5 and RDFa, two key enabling technologies.

Introduction to the Semantic Web and Linked Data

In 2001, Tim Berners-Lee et al. introduced their vision of the Semantic Web, as an extension of the current Web in which information has *"well-defined meaning, hence better enabling computers and people to work in cooperation"* [1]. The most essential part of this next-generation Web is content that is formally described via ontologies, metadata conforming to these ontologies, logic, and agents [2]. Many definitions of the term ontology exist. The most popular is by Gruber who defines an ontology as "an explicit specification of a conceptualization" [3]. This definition is further extended by Studer et al. to "Ontologies are a formal, explicit specification of a shared conceptualization" [4].

Conceptualization refers to an abstract model of some part of the world which identifies the relevant concepts and relations between these concepts. ***Explicit*** means that the type of concepts, the relations between the concepts, and the constraints on their usage, are explicitly defined. ***Formal*** refers to the fact that the ontology should be machine readable. Finally, ***shared*** means that the ontology should reflect the understanding of a community and should not be restricted to the comprehension of some individuals. By doing so, it captures consensual knowledge [5]. Ontologies occur in different degrees of formality, ranging from thesauri to richly-axiomatized structures [6].

A huge momentum has recently been building up in Semantic-Web research, due to the ongoing implementation of a vision of a *Web of Data* formulated by Tim-Berners Lee. In this vision, formerly-fragmented data is connected and interlinked with each other, based on the so-called Linked Data principles[1]. Over just a few years, the so-called Linked Open Data (LOD) cloud, which represents a huge interconnected data set, has been steadily growing *(see Fig. 1)*.

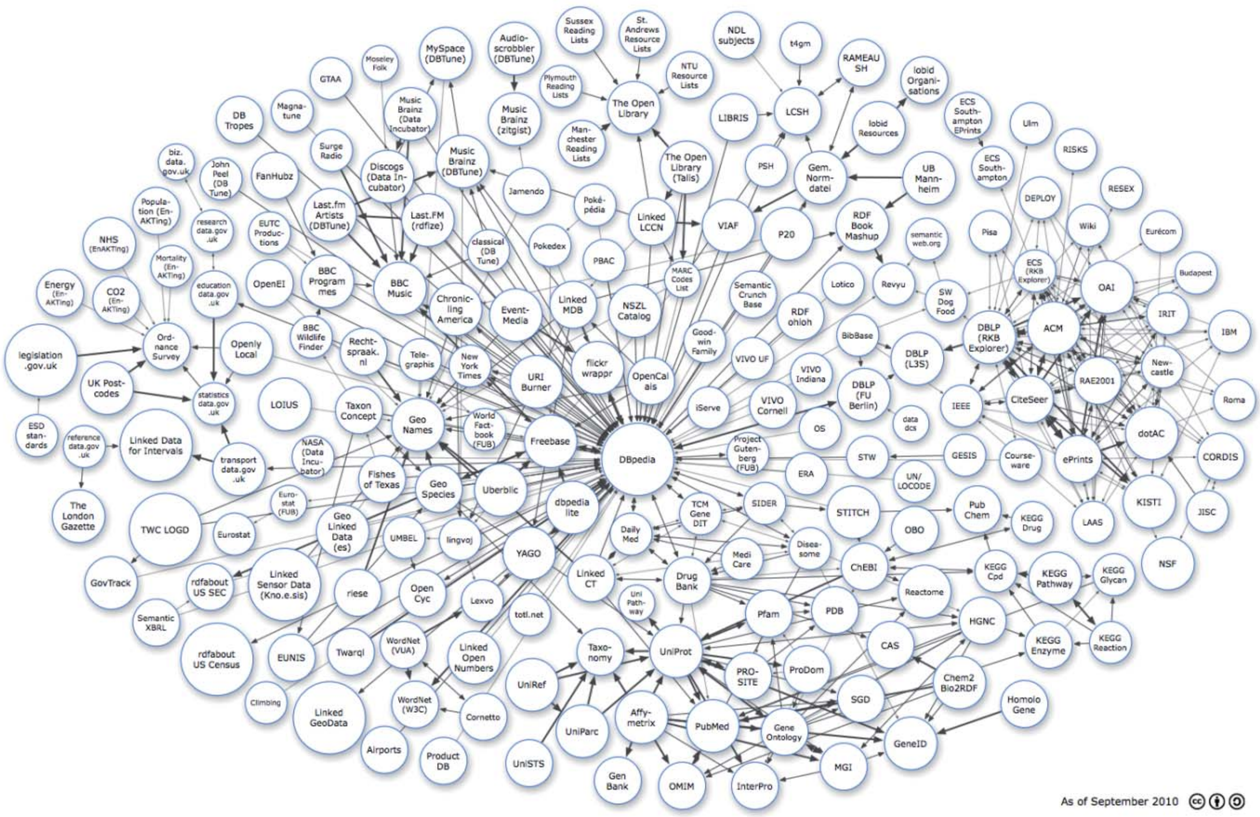


Figure 1
The linked data bubble

In early 2007, the LOD community project had been launched within the W3C Semantic Web Education and Outreach group. It bootstraps the *Web of Data* by publishing datasets using the Resource Description Framework (RDF), the metadata model primarily used on the Semantic Web. RDF enables automated software to store, exchange, and use machine-readable information distributed throughout the Web which, in turn, allows users to deal with the information with greater efficiency and certainty. Currently, the LOD project includes more than 200 different datasets *(Fig. 1)*, ranging from rather centralized ones, such as DBpedia[2], a structured version of WikiPedia, to those that are very distributed, for example the FOAF-o-sphere.

The current LOD cloud contains data from diverse domains such as people, companies, books, scientific publications, films, music, television and radio programmes, genes, online communities, statistical or scientific data [7]. Datasets were contributed both by researchers as well as by industry.

The *key success factor* of the LOD movement is the simplicity of its underlying principles:

1) all items should be identified using URIs;
2) all URIs should be dereferenceable;
3) when looking up a URI, it should lead to more (linked) data;
4) links to other URIs should be included in order to enable the discovery of more data.

1. Linked Data Principles: **http://www.w3.org/DesignIssues/LinkedData.html**
2. **http://dbpedia.org/About**

Multimedia and Linked Data – the need for a common descriptive framework

Despite recent efforts by companies such as the BBC (**http://www.bbc.co.uk/programmes**), or the European Commission's FP7 NoTube project (**http://notube.tv/**), we can observe that multimedia content in the *Web of Data* is still under-represented which, amongst others, is due to a lack of integrated means to describe multimedia content on the Web [8].

Research confirms that there is a need for going beyond current metadata standards to annotate and describe media assets [9]. One reason for this is because current XML-based standards, or proprietary schemas, are not ad hoc interoperable. The crux of metadata standards is the diversity of standards used: multimedia systems typically contain digital documents of mixed media types, which are indexed on the basis of strongly divergent metadata standards. This severely hampers the inter-operation of such systems because, for instance, different content description standards are used in different communities. This fact makes it hard to provide a generic search interface on top of these contents.

The automated integration of XML-based metadata standards is not feasible since (i) the semantics of the standards are often provided in natural language and (ii) named elements are sometimes used differently in two or more standards. Semantic technologies have proven to be a credible answer to this problem. They can help to overcome well-known drawbacks of XML-based languages, such as the vague notion of elements, attributes or nesting and the difficulties of dealing with graph-like structures which are pre-dominant on the Web and which are implicitly given by relations among deployed resources. The abundance and diversity of content represented in different formats on the Web (i.e. videos, images, audio, or 3D animations) directly translates to the need for a common framework for the description and representation of content.

Steps and criteria for developing an ontology

Key enablers for an integrated Web of Data, in which multimedia is a first-class citizen, are so-called ontologies as a formal means to describe their contents and technical characteristics. Typically, ontologies are built in a collaborative fashion according to ontology engineering methodologies which typically reflect the ontology lifecycle as introduced in [10][11].

Ontology engineering methodologies follow guidelines and major steps including domain analysis, conceptualization, implementation, and finally test and evaluation. For the sake of simplicity of the whole process, the design criteria that were proposed by [12] and the main ontology engineering steps proposed in [13] can be followed: the design criteria should guarantee the objectivity of the ontology. They include (i) clarity, i.e. the definition of terms should be clear without any ambiguity, (ii) coherence, i.e. the ontology must not have any contradictory statement, (iii) extensibility, i.e. the model and underlying classes should allow its extension and customization and (iv) minimal ontological commitment, which leads to a balance between a simple ontology and an over-axiomatized structure. In order to develop an ontology, the main ontology building steps as proposed by Noy et al. can be adopted [13]:

Abbreviations

API	Application Programming Interface
CCDM	Class Conceptual Data Model
EXIF	EXchangeable Image File format
FOAF	Friend Of A Friend
HTML	HyperText Markup Language
IPTC	International Press Telecommunications Council **http://www.iptc.org**
LOD	Linked Open Data
MAWG	(W3C) Media Annotation Working Group **http://www.w3.org/2008/WebVideo/Annotations/**
RDF	Resource Description Framework
OWL	Web Ontology Language *(instead of "WOL")*
URI	Uniform Resource Identifier
W3C	World Wide Web Consortium **http://www.w3.org/**
XML	eXtensible Markup Language

1) **Determine the domain and scope of the ontology**. In this step the coverage and usage of the ontology is determined next to its domain and scope. This phase typically involves the definition of so-called competency questions, i.e. for what types of questions the ontology should provide answers, that should capture the information-representation needs of the ontology. This step is typically carried out by domain experts.
2) **Consider re-using existing ontologies.** The next step considers reuse of ontologies that (at least partially) capture the domain and scope of the ontology. This step is important to ensure compatibility with other applications on the Web and facilitate the adoption of Linked Data.
3) **Enumerate the terms of the ontology.** This step involves the definition of terms which the ontology should make statements about.
4) **Class definition and Class hierarchy.** This step is due to the identification of classes and to arrange the classes hierarchically. In this step it is suggested that we apply the middle-out approach as proposed in [14]. In the middle-out approach, you first identify the core of basic classes and then specify and generalize them as needed. This step is typically carried out by ontology development experts. It is further recommended to consider the need for classes (and their instances) to be duly uniquely identified. Some classes can be naturally derived from database structures. The identification of relevant manageable classes is the more critical step in ontology design.
5) **Determine the datatype and object properties of classes.** This step takes the terms identified in step 3 that were not classified as classes in step 4. The terms not identified as classes are most probably object or datatype properties. This step is typically carried out by ontology development experts.
6) **Determine the restrictions of the datatype and object properties.** The intention of this step is to identify and specify restrictions describing the possible value types, the allowed values, etc. This step is typically carried out by ontology development experts.
7) **Creation of individuals.** In the last step, individuals are created.

The W3C Media Annotation ontology

The original purpose of the W3C Media Annotation Working Group (MAWG)[3] was to identify a minimum set of core properties necessary and sufficient to describe and retrieve information about media resources (video, audio, images) on the Web, the so-called *Ontology for Media Resources 1.0* [4]. These core properties have been mapped to more than 15 metadata formats from a wide range of communities such as ID3, EXIF, MPEG-7 and others. In the course of the definition of the mappings, the EBU provided the mappings from the core properties to EBUCore[5], TV-Anytime[6] and IPTC's NewsML-G2[7].

The aim of the group is to provide two main realizations of the Ontology for Media Resources 1.0 (ma-ont):

- A WebIDL[8]-based API[9] (with bindings to ECMAScript and Java) to datamine and query metadata repositories in different formats through their mapping with the ma-ont properties. A typical implementation is the development of Web services as query endpoints[10].

---

3. http://www.w3.org/2008/WebVideo/Annotations
4. http://www.w3.org/TR/2010/WD-mediaont-10-20100608/
5. http://tech.ebu.ch/docs/tech/tech3293v1_2.pdf
6. http://tech.ebu.ch/tvanytime
7. http://www.iptc.org/site/News_Exchange_Formats/NewsML-G2/
8. W3C WebIDL: http://www.w3.org/TR/WebIDL/
9. http://www.w3.org/TR/2010/WD-mediaont-api-1.0-20100608/

- A Semantic Web ontology implemented with RDF-OWL[11] for Semantic Web and Linked Data integration, which naturally relates to HTML5[12] and RDFa[13]. This is the main topic of this article.

As mentioned earlier, one key step when developing an ontology consists of defining a set of relevant classes assembled in a structured model. The ma-ont ontology has been built based on the list of properties defined in the W3C Ontology for Media Resource 1.0 which itself is based on common properties from various metadata standards and their mapping potential. Based on this phase, which is typically referred to as domain analysis, the actual conceptualization and implementation of a Semantic Web compatible ontology has been made. The identified class model is based on the EBU's CCDM (Class Conceptual Data Model) for distribution.

The ma-ont ontology defines a set of media- and non-media-specific classes:

- A set of media classes (*Figs 2* and *3*) which include:

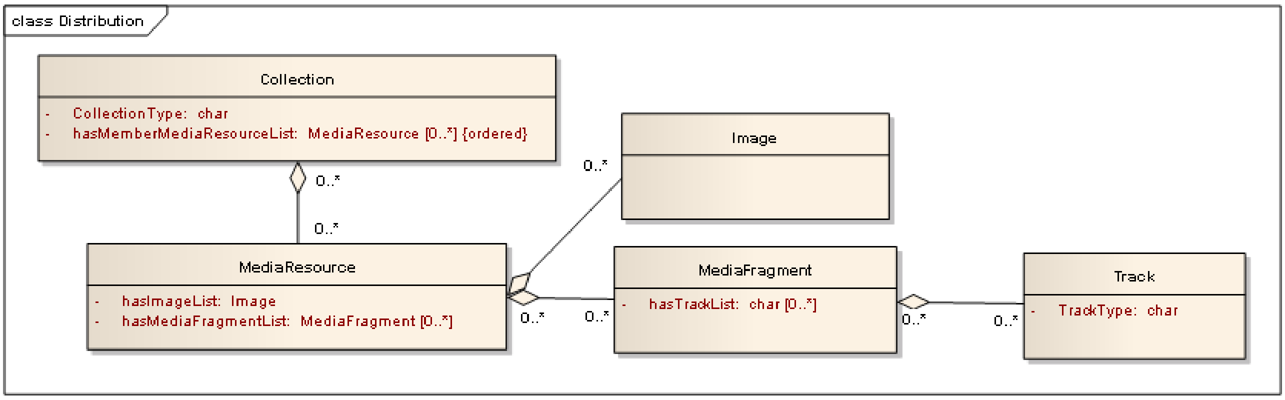


Figure 2
MAWG UML Class diagram

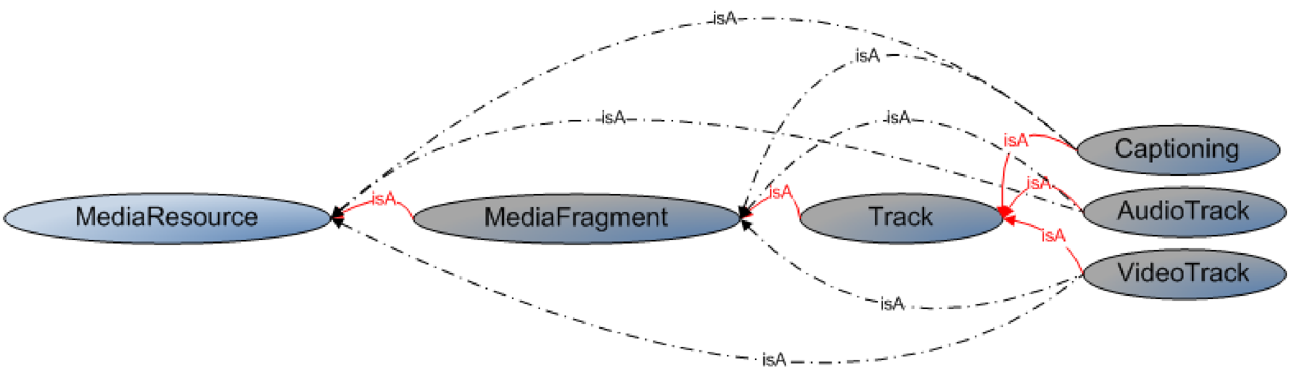


Figure 3
MAWG asserted and inferred axioms of the MediaResource class hierarchy

  - **`MediaResource`**, which identifies content e.g. a programme, which can be member of a group (like the episode of a series).
  - **`MediaFragment`**, which identifies a part / segment and other directly accessible items of a MediaResource (programme).

---

10. Media Annotations Working Group – WebServices Reference Implementation: **http://mawg.joanneum.at/web/mawg.html**
11. **http://dev.w3.org/2008/video/mediaann/mediaont-1.0/mediaont-1.0.html#ont-RDF**
12. **http://www.w3.org/TR/html5/**
13. **http://www.w3.org/2010/02/rdfa/sources/rdfa-primer/**

- **`Track`** and associated subclasses **`Audio`** (track), **`Video`** (track) and **`Captioning`** (track) to identify specific media components of a MediaResource.
- **`Image`** to declare, for instance, images related to the MediaResource such as keyframes.
- **`Collection`** which defines a group of MediaResources (e.g., a programme group or a series).

- A set of non-media classes which include:
  - **`Agent`**, which can be a Person or Organization associated with the resource.
  - **`Rating`**, which allows collecting user evaluations of the resource (ma-ont specific). The use of the class rating is justified by the possibility of an application to store in a database all ratings attributed by different users about different MediaResources.
  - **`TargetAudience`**, which is used to define the audience for which the media resource has been editorially defined. The different target audience schemes used in different countries or for different applications (e.g. television or games) can also be managed in relational databases.
  - **`Location`**, which allows associating location information with the resource, e.g. a shooting and/or fictional location depicted in the resource.

The classes defined in the ma-ont ontology have a number of characteristics and relations with each other which are captured in either objectProperties (relations between classes) and dataProperties (as a property of a class) The definition of properties were mainly based on the need for rich and coherent user queries capturing what the user is supposed to look for.

The following ma-ont ontology examples illustrate the use of properties by the MAWG:

- ObjectProperties define the relations between classes (e.g. “Collection *hasMember* MediaResource” or “MediaResource *hasFragment* MediaFragment”).
- ObjectProperties can be used to represent roles. For example, roles of contributors are defined by object properties such as “MediaResource *hasCreator* Agent”.
- ObjectProperties can be used to represent types. A type of fragment is defined by “MediaResource *hasChapter* MediaFragment” or “MediaResource *hasTrack* Captioning”.
- ObjectProperties can be used to link to external classification schemes and in particular to SKOS concepts. The EBU has specified many classification schemes available as Web resources in XML[14] and HTML[15]. Several of these classification schemes have been translated to SKOS (e.g. the EBU ‘genre’ and ‘role’ lists)[16].
- Inverse objectProperties enrich by inference (result of an analysis done by an RDF/OWL parser called a ‘reasoner’) the number of relations in the knowledge base on which queries (e.g. SPARQL[17]) can be done, thereby delivering richer results. Examples of MAWG inverse properties are: ‘*hasFragment’ & ‘isFragmentOf’* or *‘hasTrack’* & *‘isTrackOf’.*
- DataProperties define intrinsic characteristics of each class (e.g. “MediaResource *title* xsd:string”).
- In the same manner as for classes, subProperties can be specified or user-defined to specialise the properties. The property ‘*hasContributor*’ is an example of an objectProperty for which subProperties have been defined e.g. ‘*hasCreator*’. The property *‘date*’ is an example of a dataProperty for which subProperties have been defined such as *‘createDate’.* A query on *‘date’* or on *‘hasContributor’* will respectively list all dates and contributors (per role) associated with the MediaResource.

14. EBU XML classification schemes: http://www.ebu.ch/metadata/cs/
15. EBU HTML classification schemes: http://www.ebu.ch/metadata/cs/web/
16. EBU SKOS classification schemes: http://www.ebu.ch/metadata/ontologies/skos/
17. http://www.w3.org/TR/rdf-sparql-query/

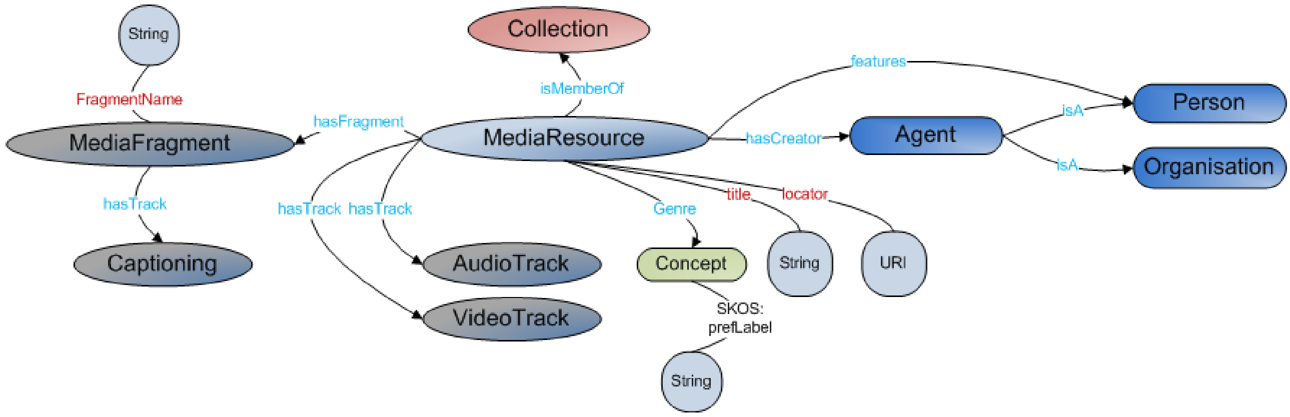


Figure 4
Example of a MediaResource RDF representation

Fig. 4 illustrates an example of a MediaResource description using several classes and properties of the ma-ont ontology. The next step would consist of instantiating/attributing values (URIs, strings, dates, etc.) to each relevant property.

Compared to XML, RDF/OWL presents significant advantages:

- **RDF/OWL instances are lists of flat statements** (so-called triples: subject + predicate + object) in opposition to XML structured data. Extensions simply mean adding new statements. Missing statements do not break the model.
- **Designed to link and integrate data**: It is the answer to the XML broken link with classification schemes. It is also the natural/native solution to link to other RDF/OWL models or other Web resources; the use of URIs within the RDF model means that information from various sources merges naturally. Furthermore information expressed using different schemata can be captured in one model.
- **Backward and forward compatibility**: New applications can fully exploit old and new classes and properties while old applications will continue to support descriptions and relations using the original model even if extended with new classes and properties. In RDF/XML, there are none of the compatibility issues that developers have to take care of in XML.
- **Agility**: New classes and properties can be added flexibly, either as an extension to a specification or as custom extensions by users (preferably but not necessarily within their own namespace). For example, the ma-ont ontology could be extended to further specialise classes e.g. declaring a subclass of '*AudioTrack*' called '*AudioDescription*' for accessibility applications, etc.
- **Formality**: The use of Semantic Web compatible ontologies implies the use of unique identifiers for conceptual elements to resolve ambiguities and the formal semantics of the Semantic Web languages such as RDFS or OWL supports the exclusion of unwanted interpretations of its elements and the automatic inferencing of implicit facts.
- **Rich queries**: Compared to a traditional database, storing RDF/OWL statements results in creating one table with one column and many rows all indexed via an 'inference graph' generated by a 'reasoner' (including the creation of new statements from inverse properties). Queries can be made on a combination of each class, property or value corresponding to statement elements (subject, verb, complement). SPARQL[18] is a new query language that has been developed to take maximum benefits from RDF/OWL ontologies and inference graphs.
  - Example 1: List all tracks at MediaResouce and MediaFragment level
    - SPARQL query: "MediaResourceX ma-ont:hasTrack ? Track"
  - Example 2: List all MediaResource with Captioning
    - SPARQL query: "? MediaResource ma-ont:hasTrack Captioning"

18. http://www.w3.org/TR/rdf-sparql-query/

- Examples 3: List all actors of a MediaResource
  - SPARQL query: “MediaResourceX ma-ont:features ? Person”
- **Web compatibility**: The use of Web standards such as URIs is compliant with the Web. A simple data model to express statements about resources (RDF) and a serialization of the model that seamlessly embeds into existing Web pages (RDFa) is a potential means to naturally embed metadata about multimedia resources in the Web graph.

Besides the advantages mentioned above, a number of disadvantages exist:

- Even with editors, writing a first coherent ontology follows a steep learning curve.
- RDF/OWL can be represented in different formats. Although RDF/XML will look more familiar to XML developers, others will prefer simpler Turtle annotations. There are on-line validation tools[19] although debugging messages are often obscure.
- Only a few editing tools support RDF and OWL such as ‘Protégé’ [open source][20] or TopBraid [commercial][21]. These tools can handle RDF/XML and Turtle, among others.
- RDF/OWL-compliant databases (sometimes referred to as ‘triple stores’) are not very common yet.
- OWL and OWL-2 present different profiles with varying degrees of implementation. MAWG has taken the option of minimizing complexity by seeking compliance with the OWL-DL[22] and OWL2-RL[23] profiles.
- SPARQL is on-going into a second phase of specification within W3C. Furthermore, writing SPARQL queries are not lambda-user friendly and must be hidden behind well-designed user interfaces. This is the end of the loop that started at the time of ontology design by asking what the user should or will be looking for!

The primary purpose of the Ontology for Media Resources was for datamining data repositories. A next promising field of exploitation of the ontology is publishing on and harvesting from the Web. There is a lot of attraction and promises around HTML5 with its video tag and large support for accessibility. This is accentuated by the demand of browsers for a language to expose RDF/OWL statements on Web pages.

What metadata format for HTML5? The example of RDFa

Two specifications are currently developed called RDFa[24] (initially developed for XHTML and extended to cover HTML5) and Microdata[25]. They share a common principle that consists of using HTML markup to insert structured RDF/OWL statements/triples in Web pages.

The following exemplary HTML5 excerpt shows metadata associated with a media resource which is exposed via the HTML5 video tag.

```
<html>
 <body>
  <video xmlns:ma-ont="http://www.w3.org/ns/ma-ont/"
  src="http://www.example.org/testv.ogg" typeof="ma-ont:MediaResource">
   <a rel="ma-ont:hasRelatedImage" href="http://www.example.org/test.png"/>
```

19. http://www.w3.org/2001/sw/wiki/OWL_Validator
20. http://protege.stanford.edu/
21. http://www.topquadrant.com/resources/community.html
22. http://www.w3.org/TR/2004/REC-owl-ref-20040210/#Sublanguages
23. http://www.w3.org/TR/owl2-primer/#OWL_2_RL
24. http://www.w3.org/2010/02/rdfa/sources/rdfa-primer/
25. http://dev.w3.org/html5/md/

```
   <img typeof="ma-ont:Image" src="http://www.example.org/test.png" />
   <span property="ma-ont:title" content="The title of the video." />
   <span property="ma-ont:duration" content="60"/>
   <span rel="ma-ont:creator" href="person1"/>
   <span about ="person1" typeof ="ma-ont:Person"
  xmlns:foaf="http://xmlns.com/foaf/0.1/"property ="foaf:name"content="The Creator"/>
  </video>
 </body>
</html>
```

Using the Turtle notation, the instance would read:

```
<http://www.example.org/testv.ogg> a ma-ont:MediaResource ;
  ma-ont:creator <person1> ;
  ma-ont:duration "60" ;
  ma-ont:hasRelatedImage <http://www.example.org/test.png> ;
  ma-ont:title "The title of the video." .
<http://www.example.org/test.png> a ma-ont:Image .
<person1> a ma-ont:Person ;
  foaf:name "The Creator" .
Using the RDF/XML notation, more familiar to XML literates, the instance would read:
<rdf:RDF>
 <ma-ont:MediaResource rdf:about="http://www.example.org/testv.ogg">
  <ma-ont:duration>60</ma-ont:duration>
  <ma-ont:creator>
   <ma-ont:Person rdf:about="person1">
    <foaf:name>The Creator</foaf:name>
   </ma-ont:Person>
  </ma-ont:creator>
  <ma-ont:title>The title of the video.</ma-ont:title>
  <ma-ont:hasRelatedImage>
   <ma-ont:Image rdf:about="http://www.example.org/test.png"/>
  </ma-ont:hasRelatedImage>
 </ma-ont:MediaResource>
</rdf:RDF>
```

Once harvested, RDF triples can be stored in databases, and inference graphs can be generated using reasoning on known ontologies such as ma-ont. The data is then ready to be queried through, e.g. SPARQL, the designated query language for RDF. Exposing RDF in Web pages via RDFa and thus exposing structured data about the entities in a Web page yields benefits, and supports Linked Data such as linking content to a licence or linking to further data about these entities in different datasets such as DBPedia or the LinkedMovieDatabase.

Convergence with audio-visual semantic and broadcasting

Several other communities had been investigating the RDF/OWL prospect before W3C MAWG was formed. Of these communities, the EBU, TV-Anytime, MPEG-7 and IPTC directly relate to broadcasting.

These four communities have all started from their XML specifications namely EBUCore, TV-Anytime, MPEG-7 MDS and NewsML-G2, with which MAWG wanted to interoperate by implementing mappings specified in its upcoming recommendation. MAWG didn't have to bother about any XML legacy schema when working on its RDF/OWL ontology. The convergence of these two approaches has proved constructive and models are converging, which is essential to future RDF/OWL semantic interoperability and richer linked data. In particular it has been once more demonstrated that direct XML to RDF/OWL conversion is not advisable[26,27] and work on modelling cannot be spared.

26. **http://tech.ebu.ch/docs/metadata/ibc2009_JPE_SemanticWeb.pdf**

27. **http://tech.ebu.ch/docs/techreview/trev_2009-Q3_SemanticWeb_Evain.pdf**

The cornerstone for success is undoubtedly the fact that the EBU, TV-Anytime, MPEG, IPTC and MAWG to a large extent share a common conceptual data model as illustrated in *Fig. 5*.

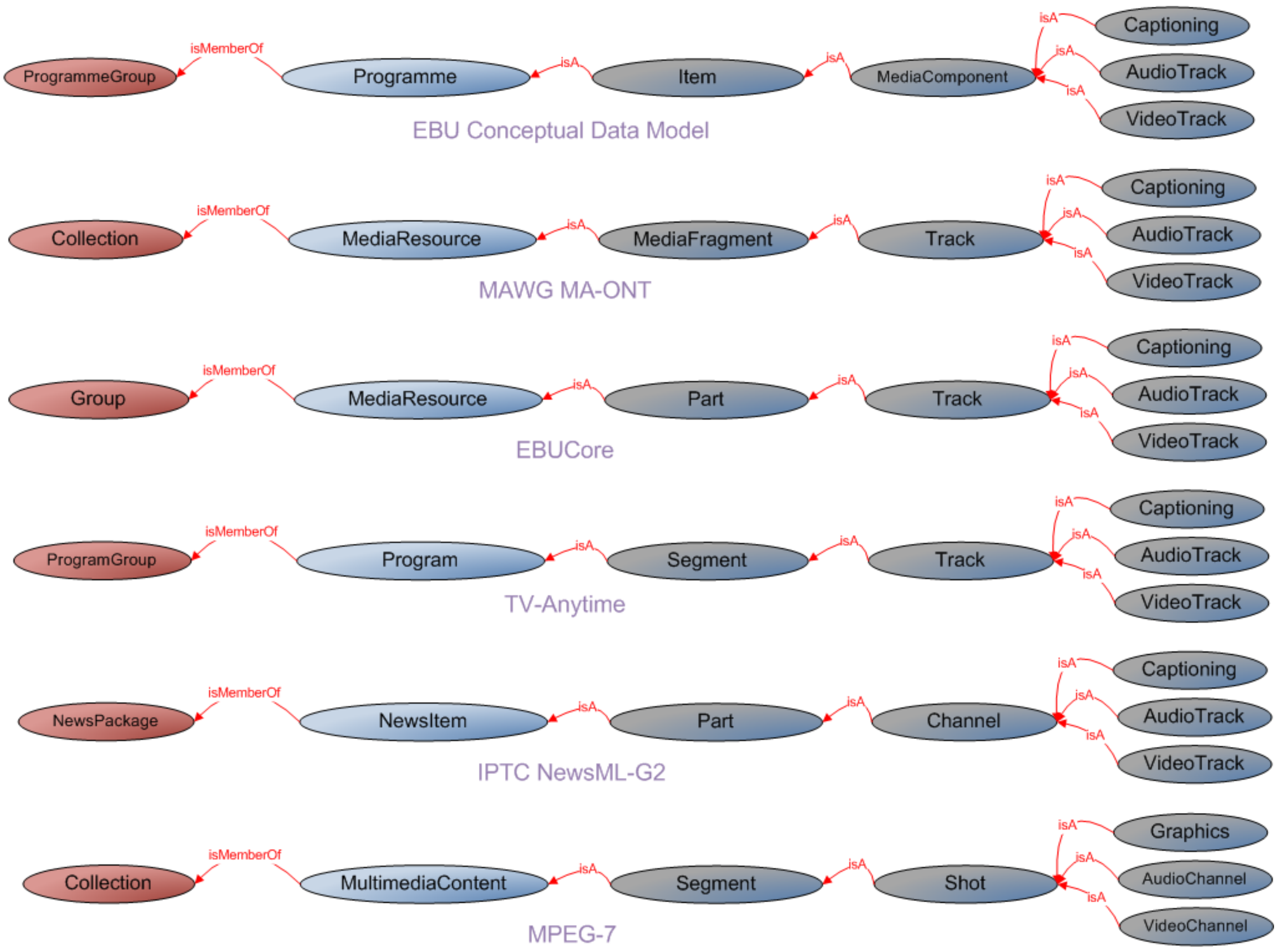


Figure 5
One conceptual data model for media ontologies

Commonalties shown on media classes in *Fig. 5* also apply to most non-media-specific classes of the different ontologies.

Due to the convergence of the core of the respective models, mappings can simply be realized by declaring equivalence between classes using the built-in objectProperty "owl:equivalentClass", once each community has published its RDF/OWL ontology.

The same diagram as in *Fig. 5* could be repeated for dataProperties, showing that each class has similar intrinsic characteristics. Mappings again can be done using the built-in "owl:equivalentProperty".

Finally, the same equivalence can be declared between objectProperties if the ontologies follow a common modelling approach. As mentioned earlier, direct translation from XML to RDF is not recommended in particular to harmonise modelling concepts.

- For example, Dublin Core declares at the same level 'contributor' and specialised roles such as 'creator' and 'publisher'. This confusion is often maintained in XML schemas. Option 1 consists of representing roles as classes (a class 'Contributor' and sub-classes 'Creator' and 'Publisher'). Option 2 use objectProperties ('hasContributor' with sub-objectProperties 'hasCreator' and 'hasPublisher'). In Option 1, the class 'Contributor' needs a 'role' property that is in conflict with its sub-classes (confusion in the role concept between the class and the role property). In Option 2, creating as many sub-objectProperties as required is sufficient. From an ontological perspective, Option 2 is preferable and will be promoted in all ontologies.

Jean-Pierre Evain (EBU) and Tobias Bürger, along with colleagues from the University Claude Bernard in Lyon, are the co-authors of the W3C MAWG ontology. The EBU has already produced drafts of the EBUCore and TV-Anytime ontologies, which will be adjusted to reflect the semantic choices finally made in the MAWG. The EBU has also submitted to IPTC a simplified draft ontology for NewsML-G2, compliant with the latest version of the specification and the profile used for Eurovision services. The EBU is also promoting an MPEG-7 standard for automatic extraction, which may lead to RDF/OWL modelling.

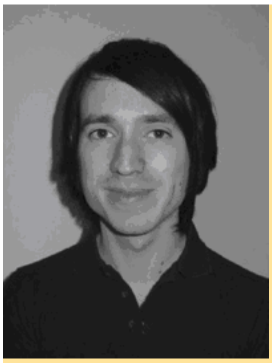
As soon as the ontologies are released, the EBU will collaborate with interested parties on an RDF/OWL ontology which specifies mappings and thus bridges MAWG, EBUCore, TV-Anytime, MPEG-7 and NewsML-G2. As a result, queries will be possible across domains. For example, provided that a search engine collects RDF/OWL MAWG, EBUCore, TVA, MPEG-7 and NewsML-G2 statements, a query on MAWG will generate hits from other domains.

Conclusions

The main motivation for each community has always been to provide better access to media content of interest to users. It is believed that this work will significantly enrich search capabilities across a wider range of metadata sources.

The implementation of this technology to develop search engines for professional centralised or distributed archives is being investigated within the EBU (e.g. for news).

Information on Semantic Web and Linked Data is regularly provided at **http://tech.ebu.ch/semanticweb_ebu**, hosted by 'EBU Technology and Development'.



Dr Tobias Bürger is a researcher in the field of multimedia and Semantic Web technologies with a high interest in Linked Data and multimedia on the (Semantic) Web. He holds a PhD in Computer Science from the University of Innsbruck and a Diploma in Computer Science from the University of Passau. He held positions as a senior researcher at the Semantic Technology Institute (STI), Innsbruck and Salzburg Research, and as a research assistant at DERI Innsbruck.

Dr Bürger has been involved in a number of European projects, including SALERO, ACTIVE, LIVE and INSEMTIVES. He has published more than 50 papers around the aforementioned topics and he actively participates in standardization activities such as the W3C Media Annotation Working Group.

Jean-Pierre Evain has worked for EBU Technical and Development since 1992 on new systems and services. He is in charge of coordinating all EBU metadata activities for production and distribution. He is involved in the conception and development of several EBU metadata specifications and, in particular, the EBUCore and the EBU Conceptual Class and Data Model CCDM.

Mr Evain is the main point of contact for TV-Anytime (ETSI maintenance, website, etc.). He is a member of the W3C Media Annotation Working Group where he is the co-author of the RDF/OWL ontology (Semantic Web and Linked Data ready) in relation to the EBUCore RDF/OWL ontology. After developing the audiovisual extension of NewsML-G2, he is now an active member of the IPTC SemWeb group developing an RDF/OWL ontology for News.

Jean-Pierre Evain represents the EBU in key European projects such as EUScreen and Europeana, respectively the European portal to audiovisual archives (using EBUCore) and the European Digital Library. He co-manages the EBU-AWMA FIMS project on SOA on behalf of EBU, and also represents the EBU in discussions with the industry on content management and copy protection.

Acknowledgement

The authors would like to thank the members of the W3C Media Annotation Working Group for their support, active review and constructive comments during the development of the RDF/OWL ontology.

Bibliography

[1] T. Berners-Lee, J. Hendler and O. Lassila: **The semantic web**
The Scientific American, 284(5):28{37, May 2001.

[2] G. Antoniou and F. van Harmelen: **A Semantic Web Primer**
MIT, 2004.

[3] T.R. Gruber: **A translation approach to portable ontology specifications**
Knowledge Acquisition, 5(2):199-220, June 1993.

[4] R. Studer, R. Benjamins and D. Fensel: **Knowledge engineering: Principles and methods**
Data & Knowledge Engineering, 25(1-2):161-198, March 1998.

[5] D. Fensel: **Ontologies: A Silver Bullet for Knowledge Management and Electronic Commerce**
Springer, Berlin, 2003.

[6] D.L. McGuinness. In D. Fensel, J. Hendler, H. Lieberman and W. Wahlster, editors: **Spinning the Semantic Web: Bringing the Worldwide Web to Its Full Potential**, chapter Ontologies Come of Age
MIT Press, 2003.

[7] C. Bizer, T. Heath and T. Berners-Lee: **Linked Data - The Story So Far**
Special Issue on Linked Data, International Journal on Semantic Web and Information Systems (IJSWIS), 2009. Eds: T. Heath, M. Hepp and C. Bizer.

[8] B. Schandl, B. Haslhofer, T. Bürger, A. Langegger and W. Halb: **Linked Data and Multimedia: The State of Affairs**
Accepted for publication in Multimedia Tools and Applications journal, Special issue 'Semantic Multimedia', Springer, 2010.

[9] T. Bürger and M. Hausenblas: **Why Real-World Multimedia Assets Fail to Enter the Semantic Web**
In: Proceedings of SAAKM 2007.

[10] A. Gomez-Perez, M. Fernandez-Lopez and O. Corcho: **Ontological Engineering**
Springer, Berlin, 2004.

[11] E. Simperl, C. Tempich and T. Bürger: **Methodologies for the creation of semantic data**
In: Miguel-Angel Sicilia (ed.) Handbook of Metadata, Semantics and Ontologies, World Scientific Publishing Co., 2011 (forthcoming).

[12] M. Uschold and M. Gruninger: **Ontologies: principles, methods and applications**
In: Knowledge Engineering Review, Vol. 11 / No 2, 1996, pp. 93-155.

[13] N. Noy and D.L. McGuinness: **Ontology Development 101: A Guide to Creating Your First Ontology**, 2001

[14] M. Uschold and M. King: **Towards a Methodology for Building Ontologies**
In: Proceedings of the Workshop on Basic Ontological Issues in Knowledge Sharing, held in conjunction with IJCAI-95, 1995.

This version: 20 January 2011

Published by the European Broadcasting Union, Geneva, Switzerland

ISSN: 1609-1469

Editeur Responsable: Lieven Vermaele

Editor: Mike Meyer

E-mail: tech@ebu.ch

The responsibility for views expressed in this article rests solely with the authors