MULTIMEDIA, WEB 2.0 and THE SEMANTIC WEB: A STRATEGY FOR SYNERGY

Lyndon J.B. Nixon

Networked Information Systems,
Free University Berlin
Takustrasse 9,
14195 Berlin, Germany
E-mail: nixon@inf.fu-berlin.de

Abstract.

Multimedia plays a role on the Web which is guaranteed to grow in significance as individuals and organizations increasingly generate digital content. Parallel to the development of richer media online is the discovery of "Web as platform", termed in some quarters as Web 2.0. Innovative websites like Flickr, del.icio.us and Yahoo Maps encourage social networking, tagging and "mashups". As rich media and Web 2.0 converge, new requirements arise which need new solutions. In this paper we consider the role that Semantic Web formulisms (ontologies) could play in realising a Web 2.0/multimedia synergy. We present the ontology-based SWeMPs multimedia presentation system and discuss how it can integrate with existing Web 2.0 content. As a result we identify necessary future developments for a truly multimedia rich Web 2.0 – and how the Semantic Web can play a role in that.

1 Background: Web 2.0 and Multimedia

A new term has been introduced to the Web community: "Web 2.0" [1]. While its significance as the name for a truly new generation of Web sites or role solely as a 'hype' word is debated, the very coining of the new term is symptomatic of the changes taking place in how Web sites interact with users and users with Web sites.

The core characteristic of Web 2.0 is that a website is no longer a static page to be viewed in the browser but is a dynamic platform upon which users can generate their own experience. The richness of this experience is powered by the implicit threads of knowledge that can be derived from the content supplied by users and how they interact with the site. Another aspect of this Web as platform is sites which provide users with access to their data through well defined APIs and hence encourage new uses of that data, e.g. through its integration with other data sources. Some well-known Web 2.0 platforms include Flickr, wikipedia and Yahoo Maps.

Flickr (http://www.flickr.com) changed the face of online photo storage by adding social features, most significantly **tagging.** Users can tag photos with keywords that describe their content and search over those keywords.

Wikipedia turned the concept of the Web providing content to the user on its head, creating a platform where anybody could add knowledge to the site.

Yahoo Maps is significant for providing a simple yet powerful API to access the map service, enabling developers to integrate those maps with other data on their own websites, termed **mashups** [3].

These platforms are exemplary of the Web 2.0 trend. However we must also turn to another developing trend on the Web – the growth in multimedia content. Technological progress has meant that we have never had access to so much media content as now. Future challenges for the Web will be the meaningful organization of this huge amount of online media content as well as its meaningful delivery to the user. Current trends – photo organization on Flickr, the audio files known as podcasts, video marketplaces like Google Video – are indicators of how Web 2.0 will have to also learn to provide the same appealing dynamic platforms with multimedia content.

2 The Rich Interactive Media Web, Presentation and Semantics

As the current trends develop we expect to experience a future Web which will be media rich, highly interactive and user oriented. The value of this Web will lie not only in the massive amount of information that will be stored within it but the ability of Web technologies to organize, interpret and bring this information to the user.

Media presentation is a key challenge for the emerging media-rich Web platforms. The growth of rich media on the Web together with the accessibility to rich data sources through Web platform APIs holds the promise for discovering new value through the synergy of both media and content streams. "Mashups" based on mapping services are an indicator of the possibilities. However, the present state of the art of media and Web technologies prevents richer integration. Laying data on top of a map requires being able to position the labels that represent that data correctly (i.e. having the knowledge of its co-ordinates). To integrate data with other images or media types the prerequisite knowledge is typically lacking and difficult to extract, e.g. knowing where a particular object occurs in an image or a video stream. These challenges hang primarily upon the generation and interpretation of appropriate knowledge.

The challenge of enabling computer systems to make better use of Web data by making that data machine-processable has been taken up by the Semantic Web effort, which proposes formal knowledge structures to represent concepts and their relations in a domain of discourse. These structures are known as ontologies and the W3C has recommended two standards, the simpler RDF[1] and the more expressive OWL[2]. While the Semantic Web to date exists primarily within research labs, others have noted that the terminological structures being generated in a bottom-up fashion by Web 2.0 site users through tagging of content forms a loose taxonomy of the domains covered by that content. This has led to the coining of the term "folksonomy" [4].

[1] http://w3.org/RDF

[2] http://w3.org/2004/OWL

The challenge to automated content generation systems – for example, for a Web application that is to select, interpret or present multimedia content - is that folksonomies are too semantically loose to be able to guarantee sufficient accuracy. Bridging the gap between the emerging folksonomies of Web 2.0 and the formal semantics of Semantic Web ontologies would benefit the Semantic Web community with being able to leverage the content and knowledge that Web 2.0 is already generating from its users and making available over standardized APIs. This applies even more in the multimedia community, where e.g. collaborative user-contributed media annotation on a Web scale is an attractive (compromised) solution to the problem of extracting knowledge out of large multimedia data stores. In recognition of this, we have chosen a Web 2.0 based scenario for our SWeMPs ontology-based multimedia presentation system. Firstly, it is necessary to introduce SWeMPs and its ontology.

3 SWeMPs: Semantic Web enabled Multimedia Presentation

3.1 SWeMPs Architecture

SWeMPs (see also [5]) is a research prototype of an automated multimedia presentation generation system based on the use of Semantic Web technologies and the integration with distributed ontologies and metadata on the future Semantic Web. Full information about SWeMPs, as well as its open source code, is online[3].

The SWeMPS architecture is influenced by the Reference Model for an Intelligent Multimedia Presentation System (IMMPS) [6]. Figure 1 below illustrates the SWeMPs architecture in UML. As well as some UIs providing the initial presentation goal and displaying the presentation result[4], the architecture contains these six components: Rulebase, Query interpreter, Reasoner, Service planner, Multimedia modeler and Presentation formatter. The core of SWeMPs is the rulebase, i.e. the internal application logic that realises the execution of the multimedia generation process, and the conceptual model, i.e. the knowledge base for the multimedia generation process expressed using the concepts formally given by the SWeMPs ontology (see 3.2). This process is based particularly on the layers of multimedia generation specified in the Reference Model of the IMMPS. It is however updated for the aims of SWeMPs, i.e. to support and interoperate with the Semantic Web:

[3] http://swemps.ag-nbi.de
[4] The two UIs may be, but are not necessarily, the same application.

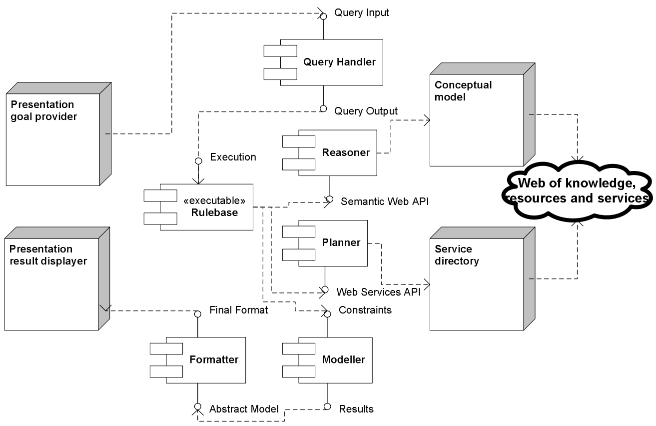

Fig. 1. SWeMPs architecture

- Activity 1, using the query interpreter: read in the information request, set up any necessary preconditions for the execution, break down the request into knowledge requirements for the individual multimedia generation task;
- Activity 2, using the reasoner: iteratively resolve the knowledge requirements by making requests upon the conceptual model through the reasoner. The reasoner will handle inferences upon the available knowledge in order to come to required conclusions. To support the inference, the activity may also make requests for acquiring new knowledge from identified resources.
- Activity 3, using the service planner: iteratively determine media representations for the knowledge conclusions derived from the previous activity. This determination is made by using available services to resolve knowledge deadlocks and to derive content and its presentation from that knowledge. As knowledge deadlocks we can consider knowledge gaps or ontology mismatches. Once any deadlocks are resolved, the next aim is knowledge to content conversion. This means resource acquisition (finding media representing a concept or relationship between concepts) and adaptation (resolving media characteristics to fit aspects of the information request).
- Activity 4, using the multimedia modeller: iteratively model the multimedia presentation by inserting the media found and adapted by the previous activity and determining appropriate constraints for the media items within the model. These constraints are based upon the semantics of the inserted media items and the semantic relationships between the concepts being represented by those media items.
- Activity 5, using the presentation formatter: when all prior activities are complete, take the multimedia model and format it into a final multimedia representation that is passed to an application that handles its correct display.

3.2 SWeMPs Ontology

Through the ontology we express multimedia generation processes in a formal, explicit, declarative, interoperable and domain-independant way. Any multimedia generation task can be modeled by the appropriate instantiation of SWeMPs concepts (e.g. using any ontology editor) and realized by plugging the resulting knowledge base into the SWeMPs system.

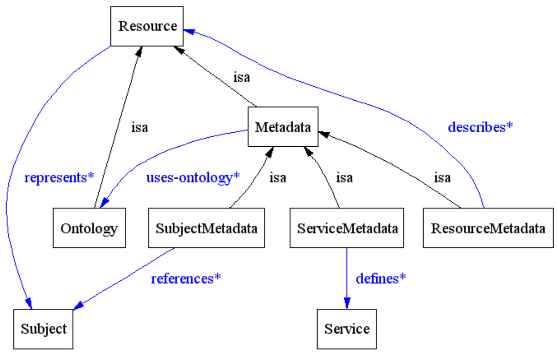

Fig. 2. SWeMPs Ontology, subsection

In Figure 2, the partial ontology is shown. In SWeMPs the multimedia generation process is modelled as involving three core types of object: a **Subject** – a topic which is to be represented in the presentation, a **Resource** – a unit of digital content which can be presented within the presentation and a **Service** – an application that can be executed through a call over the Internet and performs some specified action. Each can be associated to Metadata about it which is in turn associated to an Ontology which it uses. Both Metadata and Ontologies have Namespaces (the URIs from which their concepts are drawn). Resources and Services also have Occurrences (the URLs where representations of that resource/service can be found) and Media Types (e.g. MIME type identifying the data format used). The ontology enables us to make important distinctions between resources in the Web architecture sense (i.e. something which can be found at an URL) and their meaning within the multimedia generation process. For example, an Amazon web service could be:

- A subject in a presentation about web services on the Internet,
- A resource in that the web service URL returns a digital resource which describes the interface of the web service,
- A service which shall be used internally by the multimedia generation process to retrieve images of book covers for display in a presentation.

As in this example, it is clear that an URL alone would not suffice to enable a computer system to distinguish its different possible uses within a multimedia generation process. Likewise, we are able to explicitly represent the relationship between concepts (whether a subject, a resource or a service in the SWeMPs understanding of those terms) and metadata that describes it (in the Semantic Web architecture, there is no formal requirement on how metadata is related to what it describes).

4 Integration with the current Web 2.0

Given the lack of formal semantics on the present Web, we seek a looser integration with Web 2.0 content in order to study the potential of formally generating multimedia presentations with the currently available Web data. There are some interesting initiatives to generate useful semantic annotations from Web based content such as text and images e.g. through screen scraping or XSLT [7,8]. In our scenario we consider two other approaches using wikipedia and Flickr content respectively:

- Adding semantic markup to the open source Wikipedia system so that Wiki entries can be annotated;
- Generating RDF metadata for photos that are published online by Flickr.

There is some discussion at present on introducing semantic markup to Wikipedia so that articles could also contain properties and relationships that are exportable as RDF/OWL[5], since the present MediaWiki system is purely syntactic and uses its own markup. We choose to generate annotations for selection of relevant articles and their usage at different granularities of text and media. As a basis, we use the structure of the provided templates at Wikitravel[6]. We create articles following the appropriate semantic MediaWiki markup in order to express RDF concepts and properties. For example, in the envisaged semantic MediaWiki, stating that Coral Gables is located in South Florida could look like this:

```
Coral Gables is a city in [[is located in::South Flor-
ida]].
```

The article name after the :: (property relation) would be recognized and an article link would be created. A RDF crawler could extract from this a RDF triple like (throughout the paper we use this simple triple notation of subject predicate object):

```
wiki:Coral_Gables wiki:is_located_in
wiki:South_Florida
```

For Flickr, a Flickr2RDF service[7] is available that generates RDF annotations from Flickr photos. Tags are given URIs by appending the string to the namespace http://www.flickr.com/photos/tags/ and related to the photo by the `foaf:topic` property. Searches by tag can be made over the Flickr API which returns XML results. The matching images' URLs are then fed to the Flickr2RDF service to generate their metadata. A key aspect for generating suitable resource metadata from Flickr is mapping the (free text) tags to (Semantic Web) URIs. We handle this in that we assume combinations of tags are sufficient to uniquely identify what would otherwise possibly be ambiguous subjects, e.g. the tag "*venetianpool*" combined with the tag "*miami*" allows us to conclude that the image depicts the Venetian Pool in Coral Gables, FL as opposed to some pool in Venice, Italy. This disambiguation is the main requirement of the currently manual work in extracting RDF from Flickr. For example, we extract matching image data from http://www.flickr.com/photos/search/

[5] http://meta.wikimedia.org/wiki/Semantic_MediaWiki
[6] http://wikitravel.org/en/Wikitravel:Article_templates
[7] http://purl.org/net/kanzaki/flickr2rdf

tags:venetianpool,miami/tagmode:all and get 3 images[8] with the tags 'venetianpool' and 'miami'. Each image's metadata is extracted as RDF through Flickr2RDF. Given such a rule (using e.g. SWRL [9]):

```
_x swemps:represents
http://en.wikipedia.org/wiki/Venetian_Pool ←

_x foaf:topic
http://www.flickr.com/photos/tags/venetianpool,

_x foaf:topic http://www.flickr.com/photos/tags/miami
```

we have a basis to map between Semantic Web concepts instantiated in terms of an URI using the Wikipedia namespace and Flickr tags from the Flickr2RDF service. This could be automated where a Web service takes a String value of a chosen concept (extracted from typical metadata properties such as `rdfs:label` or `dc:title`[9]), retrieves the Flickr search results for photos with that tag and converts the metadata of the image set to RDF. SWeMPs then repeats its query for resources that represent the chosen concept and will retrieve relevant Flickr photos provided there are rules mapping between a combination of Flickr tags and the (domain of the) chosen concept.

Finally, we can "mashup" on the basis of this semantic information with the Yahoo Maps API. Using the HTTP POST method, this involves calling an URL with the syntax http://api.maps.yahoo.com/Maps/V1/annotatedMaps?appid={*service_id*}& xmlsrc={*location_of_xml_file*}. The XML source needs to be generated by the system. Fundamentally it contains the address of each object to be shown on the map, together with an optional label string.

Geotagging[10] is the addition of location metadata (e.g. longitude and latitude coordinates) to media. For example, one could use tags of the following form with Flickr photos of places (where xxx would be a decimal value, as used in GPS):

```
geotagged geo:lat=xxx geo:long=xxx
```

Using such metadata, locations could be unambiguously identified and we could integrate geotagged data onto maps automatically.

5 Realization of an interactive Web-based media platform

Many Web (2.0) sites publish APIs through which applications can access their content. In this scenario we have considered two sites which make an API available: Yahoo Maps and Flickr. Proposed Semantic Wikis will allow to insert RDF knowledge within Wiki articles which systems will be able to extract.

---

[8] Searched on 10 February 2006

[9] Naturally this is not perfect; for example ‚Crandon Park Beach' (also in South Florida) finds no photos, but ‚crandon' with 'park' would return one image, and 'crandon' with 'beach' returns seven images (search of 29 September 2005)

[10] See http://en.wikipedia.org/wiki/GeoTagging

The API calls or Wiki retrieval is encapsulated into an external Web service which is made available to the SWeMPs framework. We note that the Web 2.0 APIs use non-semantic content as both input and output and hence the Web service call must be composed with other services which mediate between the service input and output and the correct input and output for SWeMPs.

The principal aim of knowledge adaptation is to provide mappings between different ontologies. While this can be a simple case of stating equivalences, which is possible directly in OWL, semantic matching can also be less trivial and will require dedicated algorithms which could be wrapped as Web Services and made available to Semantic Web applications such as SWeMPs. In our scenario, we have one case where a service is necessary to map between Flickr and Wikipedia URIs. For a small set of concepts, this is done manually but semi-automatic methodologies become necessary for larger scale mappings, which is an area of further research.

A dedicated case of ontology mapping for SWeMPs regards the resource metadata. A common description is required internally by the SWeMPs framework to be able to reason over resource characteristics and determine constraints for the multimedia modeller. We have defined a common vocabulary (see [5]) drawn from the SWeMPs and ZyX ontologies (the latter is based on the ZyX document model [10]). Typical resource metadata vocabularies are mapped to these ontologies, and this mapping encapsulated in OWL/SWRL and made available to the multimedia generation process.

Presentation rules are written in SWRL and relate domain specific properties to sets of constraints between media objects. Authors can use the set of constraints (e.g. left-of, start-parallel-to, appears-on-click) available in SWeMPs, but are also able to write their own if they wish. Abstractions are implemented internally as Java methods which use the Cassowary constraints library [11]. A properties file is used by the multimedia modeller component to map a constraint (identified by URI) to the Java method to be called internally by the system.

Given the existence of the chosen sources of knowledge, content, (Web) services and (presentation) rules, an implementation of the SWeMPs framework for a particular task is realised in that the SWeMPs ontology is populated with instances of Ontology, Metadata and Service (together with XMLNamespace and Occurrence instances, as necessary, to reference namespaces used and URLs from where data can be retrieved). Currently, we use Protégé for setting up the SWeMPs conceptual model and writing the presentation rules. The instances used in this Web 2.0 scenario are enumerated below:

Ontology ID	Title	Service ID	Name
o_1	Gastronomy ontology	s_1a	Restaurant XML
o_2	Tourism ontology	s_1b	Yahoo Maps
o_3	MPEG-7 ontology	s_2a	Flickr XML
o_4	Yahoo Travel	s_2b	Flickr2RDF
o_5	Wikipedia Travel	s_3	Wikipedia text
o_6	Flickr	s_4	Wiki Travel
		s_5	Image transcode
		s_6	MPEG7 mapping
		s_7	Name extraction

Metadata ID	Title	Ontology used	Namespace used[11]
m_1	Coral Gables Restaurants	o_1	o_4
m_2	Coral Gables Sights	o_2	o_4
m_3	Tourism Video Annotation	o_3	o_5
m_4	Wikipedia/Yahoo concept mapping	OWL	o_4,o_5
m_5	Flickr/SWeMPs Mapping	OWL	o_6,swemps
m_6	Wikipedia/SWeMPs Mapping	OWL	o_5,swemps
m_7	Flickr/Wikipedia concept mapping	SWRL	o_5,o_6
m_8	Scenario Presentation Rules	SWRL	

The metadata, ontologies and services will be referred back to in the text by their ID given here.

5.1 Tourist information scenario

SWeMPs has two ways to acquire content:

- directly by reference in the SWeMPs conceptual model
- indirectly by invocation of a Web service which returns retrieved content

Note that SWeMPs does not directly reference content (Resources in the ontology) but rather knowledge about it (ResourceMetadata) which is used for the semantic selection from a range of potential content.

The only ResourceMetadata available at the execution of the multimedia generation task is the annotation of the video with a modified version of MPEG-7 (which can be mapped to a SWeMPs representation). We can consider how a multimedia generation process finds other content (from other sources) automatically. Their selection is based on a common understanding of their description (based on SWeMPs) and their integration into the presentation through the relevant presentation rules.

Given the query within the system (using an abstract form – upper case letters represent unbound variables, lower case letters bound variables, { } encloses specified resources and [] types the variable with a RDF class):

```
X {swemps:represents} y [t:Sehenswuerdigkeit]
```

which means „what are all things which represent a specific instance of a touristic sight?", we want to look at the retrieval of content. This takes place through answering the query through a list of concepts, and for those concepts then determining resources that can be used to communicate them to a user. There are no `swemps:represents` statements in the current knowledge base so first the system checks for services for knowledge look-up and then – as no services are present – attempts to map to known ontologies and in this case finds mappings to both o_3 and

[11] Meaning here the namespace in which instances of ontological concepts exist

o_6. So we can also extract the following mappings to SWeMPs (using Venetian Pools as an example sight) from the relevant sub-set of our (here, MPEG-7) metadata (mapping between the MPEG-7 and the SWeMPs ontology using Web service s_6):

```
…#vs0246456 [zyx:Video]

swemps:represents
http://en.wikipedia.org/wiki/Venetian_Pool

swemps:has-occurrence http://...../video.mp4

zyx:start  30000

zyx:duration 17000
```

In seeking further knowledge about this touristic sight (and hence, potentially, further content) we find two (composable) services which output ResourceMetadata: s_7 composed with s_2 can extract string keywords for a subject (based on key properties like `dc:title`) and use these keywords to acquire Flickr metadata, while s_4 will take a subject of type `w:TravelObject` to acquire SemanticWiki metadata. In the first case, the string "Venetian Pool" leads to a Flickr search on the tag venetianpool. An image such as the one below will be identified as relevant first after the further mapping to Wikipedia URIs using the available rules (m_7).

```
…#image2625 [foaf:Image]

foaf:topic  http://www.flickr.com/photos/tags/venetian-
pool

foaf:topic  http://www.flickr.com/photos/tags/miami

dc:source  http://photos22.flickr.com/2625_b64ad1_o.jpg
```

In the second case, RDF is extracted from the SemanticWiki entry for Venetian Pool, which is made possible by mapping m_4. The mapping rules are vital in that they establish the equivalence between the concept of type `w:TravelObject` (in the Wikipedia namespace) with the concept of type `t:Sehenswuerdigkeit` used in the original query. Other mappings (m_5, m_6) provide how the Flickr and Wikipedia resource metadata relate to the SWeMPs vocabulary (e.g. the `represents` property which ties Resources to the Subjects they can be used to represent).

To determine how this content is presented in relation to one another we apply a number of presentation rules (m_8, here in pseudo-SWRL syntax):

X before Y ← X zyx:start x Y zyx:start y x < y.	Y during X ← X swemps:is-of-type zyx:Video Y swemps:is-of-type (NOT zyx:Video) X swemps:represents C Y swemps:represents C
Z appears-on-click Y; X pause-on-click Y; Y top-left-corner X ←	R contains-h-layout Z ← Z appears-on-click Y

Y swemps:is-of-type zyx:Text Y during X Z during X Y swemps:represents C C rdfs:label Text1 Y swemps:has-occurrence O O swemps:text Text1	

These rules state:

- The video segments are ordered in parallel according to their respective start times (we assume there are no overlapping segments)
- During any video segment, the other resources which represent the same concept are also included in the presentation
- Of those resources, the Text resource which is the label of the concept is displayed in the top left corner of the video. Whenever it is selected by the user, the video is paused and the other resources are displayed
- The display of the resources that appear when the label is clicked is laid out horizontally. The constraint "contains-h-layout" defines a complex media element that acts as a container for the other resources and enforces a certain layout upon them. The container may not be able to display at one time all of the selected resources. One can use different types of overflow strategy to solve this – e.g. see the discussion in [12,13].

As a result, the scenario realises a summary of a tourism video in which segments relating to a particular sight are selected and presented chronologically, and during their presentation the name of the respective sight is displayed in the top left corner of the video. If the user interacts with this name (the form of the interaction would be set in the final formatting stage, dependent as it is on the target device and presentation format), the video is paused and media relating to the concept are presented, here some images extracted from Flickr and a text from the related Wikipedia article.

As a result, we can note that the process has taken a single concept (here, the Venetian Pools in Coral Gables, Florida) and through the use of metadata and services it could find resources from different sources (a tourist video, Flickr photos, Wikipedia text) which represent that concept, thus demonstrating **dynamic data integration**. We have also shown the **mediation between different knowledge representations**, at the ontological level both through a simple one-to-one mapping using OWL (m_5, m_6) as well as the less trivial generation of RDF statements from a XML data model (s_6), and at the instance level again both through a simple one-to-one equivalence (m_4) as well as the less trivial rules-based mapping (m_7).

In the scenario, initially only metadata for the tourism video and the Yahoo Travel-scraped knowledge about restaurants and sights is available. The Flickr and Wikipedia information demonstrates **dynamic knowledge retrieval**, where adding services to the model of the multimedia generation process that enable the retrieval of metadata from these sites combined with mappings from the metadata formats they produce and instances that they use has permitted the system to acquire and integrate additional resources into the presentation. These additional resources are linked to the segments of the video that deal with the same concepts, demonstrating **coherent presentation**.

5.2 Scenario involving personalization and adaptation

Let us consider another scenario using the same multimedia generation model but generating a different presentation on the basis of a different initial query:

```
X [r:Restaurant] {r:typeOfCuisine} y [r:Cuisine]
```

This query draws on the Gastronomy ontology o_1, and as the Cuisine instance also exists within the namespace of o_1 the metadata m_1 is included into the conceptual model. In resolving the query for a particular cuisine, a number of restaurants match from the Yahoo Travel metadata (scraped from the website). The system must now resolve resources that represent those restaurants for presentation to the user. Here, service s_1 composes a means to extract a Yahoo Map from the restaurant description, using the properties `r:name` and `r:address`.

Again we turn to the presentation rules to order these resources in a multimedia presentation, drawing upon their subject and resource metadata:

X title Y; Y subtitle X ← X swemps:is-of-type zyx:Text Y swemps:is-of-type zyx:Text X swemps:represents C Y swemps:represents C C r:name Text1 X swemps:has-occurrence O1 O1 swemps:text Text1 C r:address Text2 Y swemps:has-occurrence O2 O2 swemps:text Text2	X close-to Y ← X swemps:is-of-type zyx:Text Y swemps:is-of-type zyx:Image X swemps:represents C Y swemps:represents C X before Y ← X swemps:represents C1 Y swemps:represents C2 NOT (Y before X)

Here we see that the text is constrained as being either a 'title' or a 'subtitle', through which the relative style of the text is defined and its general spatial positioning is specified. Both text resources are constrained to being placed close to the image that they relate to conceptually. Finally, the resources for each distinct concept (a restaurant) are ordered temporally through a "before" rule. Note the use of the negation in the rule to ensure that once a set of resources are placed temporally before another set, the system can not make a contradictory inference (negation is not expressable in SWRL, but will hopefully be included in the W3C RIF[12] effort).

The result of this particular query is a slideshow effect, in which maps (from Yahoo) are displayed together with the restaurant name and address. Again, we see the combination of resources from different sources (Yahoo Maps and Yahoo Travel), and their presentation based on their relationships (an Image is titled and subtitled by resources representing a restaurant name and address, respectively). However, we will take this scenario a step further to introduce the idea of adaptation to context.

Let us add to this scenario the context that the user is using a mobile device and is located presently within the city about which the tourism video is relating. The user is hence not only interested in finding a restaurant, and that it serves a cuisine which he or she likes, but also that the restaurant is located close to his or her present location.

---

[12] http://www.w3.org/2005/rules/

Note that here the user should be presented with the relevant information with the minimum of interaction – e.g. the multimedia generation process is perhaps launched by a single click representing “Give me restaurants I like close to here”. Both the user’s preferred cuisine and current co-ordinates (as provided by GPS for example) are inserted into the multimedia generation process by adding them to the presentation constraints sent as input to the system, along with the device characteristics.

Additionally, we introduce a service into the conceptual model which takes as input the user location as (latitude, longitude) values and a set of RDF metadata using the Gastronomy ontology. The `r:address` and `r:city` properties are used with a geolocation service to derive their (latitude, longitude) values and distance between user and restaurant are computed. The service returns metadata in its own ‘proximity’ vocabulary which uses a `prox:distance-from-user` property which takes a literal (numerical) value.

Hence this last scenario introduces a few more ontologies, a mapping file and a service into the conceptual model (the other metadata is placed into the presentation constraints passed as input to the generation process, rather than into the conceptual model):

Ontology ID	Title	Service ID	Name
o_7	FOAF User Profile	s_8	Proximity Service
o_8	GeoRDF		
o_9	Proximity		
o_10	CC/PPx Device Description		

Metadata ID	Title	Ontology	Namespace
m_9	CC/PPx to SWeMPs mapping	OWL	o_10

Here the query expressed to the system is a join between three queries:

```
1. z [foaf:Person]   {foaf:likes} Y [r:Cuisine]

2. X [r:Restaurant]  {r:typeOfCuisine} y [r:Cuisine]

3. x [r:Restaurant]  {prox:distance-from-user}
d [xsd:decimal] <= 2.0
```

In other words, select those restaurants whose type of cuisine is liked by the current user and that are within 2 km of the user’s location. Note that in this scenario not only user but also device and the proximity service play a role in the selection and presentation of the resources. Firstly, the addition of the selection according to the user’s preferences filters the available restaurants. As in the previous scenario, we can retrieve for each restaurant an Image (Yahoo Map) as well as some text (name and address). However, the mobile device can not display images so only the text is selected (test against the CC/PPx metadata using m_9 for interoperability). Finally, to test for the third query we find no metadata using the proximity vocabulary defined in o_9 so services are examined for the production of metadata using this ontology. The proximity service s_8 is selected, and metadata with `prox:distance-from-user` statements is generated, leading to matches for the third query.

As always, the selected resources are organized in a presentation according to the available presentation rules:

X title Y; Y subtitle X (as above)	Y above X ← Y subtitle S X title T Y swemps:represents C1 C1 prox:distance-from-user Value1 X swemps:represents C2 C2 prox:distance-from-user Value2 C1 < C2

The text resources are placed spatially and styled according to the same 'title' and 'subtitle' rules from the previous scenario. The only other rule that need apply in this scenario is a spatial positioning of the restaurants based on their relative distance from the user. A rule checks their `prox:distance-from-user` and places distances with a lower value (i.e. closer) higher spatially in the presentation.

In this scenario, we not only see the re-use of presentation rules and how a presentation can be extended to provide new functionality through the identification of relevant ontologies, metadata and services, but also how the requirement of **adaptation to context** is met, here through three distinct contexts: User preferences, User location and Device characteristics.

6 The multimedia friendly Semantic Web

On the present Web, the lack of semantics prevents computer systems from being able to interpret Web information automatically. For a task such as Web-based multimedia generation and delivery, which could underlie a future rich media interactive Web, the Semantic Web plays a vital role in introducing such formal knowledge models onto the Web. While the Semantic Web is still largely within research groups, Web 2.0 has brought a significant amount of informal knowledge onto the Web, based around users as content providers, tagging to form loose "folksonomies" and open APIs to allow re-use of data in different settings. This trend is paralleled also by the growing ubiquity of digital media content, whose organization, interpretation and presentation requires suitable annotation and systems able to use that annotation.

This paper has described how the research prototype SWeMPs was integrated with the looser knowledge structures of the current Web 2.0 as an examination of the current possibilities for leveraging that user-provided content in the multimedia generation process. We find that much is already possible but still requires a level of manual preparation that will not scale up to open Web-based application, as a machine can not guess what a human has meant by textual annotations such as tags. Rather, we must hope that as users and developers continue to discover the added value of providing and using less ambiguous knowledge and media annotation, that this can push the Semantic Web within the Web 2.0 field and that the next paradigm shift on the Web can be both from multiple media to *multimedia* and from data to *knowledge*.

Acknowledgements

My PhD research was previously supported by the DAAD. Currently, my research is supported by the EU Network of Excellence KnowledgeWeb (FP6-507482).

References

1. Tim O'Reilly, "What is Web 2.0: Design Patterns and Business Models for the Next Generation of Software". Online http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html
2. Alan Smeaton, „Indexing, Browsing and Searching of Digital Video and Digital Audio Information". Lecture Notes in Computer Science, Volume 1980, Jan 2001, p93
3. "Mix, Match and Mutate". Online http://www.businessweek.com/magazine/content/05_30/b3944108_mz063.htm
4. Adam Mathes, "Folksonomies – Cooperative Classification and Communication Through Shared Metadata". In Doctoral Symposium, Computer Mediated Communication, University of Illinois Urbana-Champaign, December 2004. Online http://www.adammathes.com/academic/computer-mediated-communication/folksonomies.html
5. Lyndon J B Nixon, "Integrating knowledge, semantics and digital media into the multimedia generation process". In 2nd European Workshop on the Integration of Knowledge, Semantics and Digital Media (EWIMT '05), London, November 2005.
6. M Bordegoni, G Faconti, M Maybury, T Rist, S Ruggieri, P Trahanias and M Wilson, "A Standard Reference Model for Intelligent Multimedia Presentation Systems". Computer Standards & Interfaces, 18(6-7):477-496, December 1997
7. David Huynh, Stefano Mazzocchi and David Karger, "Piggy Bank: Experience the Semantic Web inside your Web browser". LNCS Volume 3729, Oct 2005, Pages 413-430
8. Dominique Hazaël-Massieux and Dan Connolly, "Gleaning Resource Descriptions from Dialects of Languages (GRDDL)". W3C Tech Report, 2004.
9. I Horrocks, PF Patel-Schneider, H Boley, S Tabet, B Grosof and M Dean, "SWRL: A Semantic Web Rule Language combining OWL and RuleML". W3C Member Submission, May 2004. http://www.w3.org/Submission/SWRL
10. Susanne Boll and Wolfgang Klas, "ZyX – A Multimedia Document Model for Reuse and Adaptation of Multimedia Content". IEEE Transactions on Knowledge and Data Engineering, Vol 13 No 3, May/June 2001, Pages 361-382.
11. GJ Badros, A Borning and PJ Stuckey, "The Cassowary linear arithmetic constraint solving algorithm". ACM Transactions on Computer-Human Interaction, 2001.
12. J van Ossenbruggen, J Geurts, F Cornelissen, L Hardman and L Rutledge, "Towards Second and Third Generation Web-based Multimedia". WWW 2001 Conference, Hong Kong, 2001.
13. J van Ossenbruggen, J Geurts, L Rutledge and L Hardman, "Towards a Multimedia Formatting Vocabulary". WWW 2003 Conference, Budapest, 2003.