Conclusion

Our research set out to advance the state of the art in multimedia generation systems by introducing a greater level of semantic integration than in other current approaches. The work of the thesis is summarized through identification of the results of this research. We see how SWeMPs has brought IMMPS closer to its original vision. A consideration of the future development of the Web and the relevance of SWeMPs to these developments, raising new research directions, will conclude the thesis.

7.1 Results of this work

In formalizing SWeMPs as a conceptual framework and in the realization of the proof of concept system and related scenarios, much has been learnt about the practicalities and impracticalities of multimedia generation, particularly in terms of modeling it in a computer system with the aims of automation and adaptation. This research has produced a number of results which can be specified here:

- Formalization and prototyping of an IMMPS which, as has been argued in the previous chapter, fulfils the original design rationale of the Standard Reference Model and hence has met the requirements for an automated and adaptive multimedia generation process.

- The use of the emerging knowledge frameworks of the Semantic Web; not only in annotating resources and services for their use in automated systems but also to re-use knowledge about subject domains in a multimedia environment, integrating it with the resource and service annotations.

- The SWeMPs ontology is proposed as a core ontology for modelling the multimedia generation process, providing key conceptual distinctions (between subjects, resources and services; between these concepts and their metadata and occurrences). These results have been included in a new initiative to ground a Multimedia Ontology Framework[96] (refer also to section 2.4.4).

- The SWeMPs and ZyX ontologies together provide a common basis for interpreting and understanding resource descriptions, the abstract multimedia model, presentation constraints (such as device capability) and their relation to one another. A core set of required metadata properties for annotating media for a multimedia presentation generation process and their mappings to other media metadata schemes has been extracted from these ontologies and published under the name Multimedia Annotation Vocabulary (MAV) [Nixon,2006].

- The use of SWRL together with a constraints vocabulary in modeling design rules is proposed as a means to enable knowledge about multimedia layout at a conceptual level to be expressed and shared on the Semantic Web.

[96] http://www.acemedia.org/aceMedia/reference/multimedia_ontology/index.html

- The multimedia generation process is modeled on the basis of ECA rules (and implemented in a LP paradigm) which call external components and is guided fundamentally by semantic query on a conceptual model. Thus we make more concrete the multimedia generation process compared to that of the Standard Reference Model for IMMPS, defining all the components (unlike the "expert modules" in the SRM), their APIs and the knowledge representation formats used and the framework in which knowledge is processed.

In order to focus the research community more on this specific area of semantics and multimedia presentation systems, we are now co-organising a workshop which aims to further deepen the research in the field by bringing together other researchers whose work can contribute to the overall SWeMPs vision. The workshop 'Semantic-enabled Multimedia Presentation Systems (SEMPS)'[97] will take place during the Semantic and Digital Media Technologies Conference in December 2006.

7.2 SWeMPs and IMMPS: a comparison

We return to the Standard Reference Model for IMMPS (section 2.5) at this stage. In many ways, this was the first comprehensive attempt to lay down commonalities in implementation for a multimedia presentation system. As seen in Chapter 3, the model, on paper at least, did achieve the requirements that were specified for automated and adaptive multimedia presentation generation. However, actual implementations have failed to meet all of these requirements, and none can be considered a full concrete realization of the Standard Reference Model. Hence we turn again to the design decisions that guided the creation of the Standard Reference Model, which we have also chosen to integrate into the conceptual framework of SWeMPs, and ask first if SWeMPs can be considered a concrete realization of an IMMPS.

The underlying design rationale for the Standard Reference Model [Bordegoni,1997] has been expressed thus[98]:

- The adequate modularization of a generic process for multimedia presentation generation which breaks down the process into logically distinct and computationally feasible subtasks.
- The appropriate abstraction of a concrete implementation that can reflect the unique characteristics of multimedia presentation generation yet is general enough to model the whole class of presentation generation tasks.
- The identification and classification of knowledge sources which are required for the knowledge intensive task of multimedia presentation generation. The model should also make clear how processes and knowledge sources are related to each other.
- The modeling of shared resources in the client-server paradigm, so that 'expert modules' can serve requests from multiple clients.
- Openness to other standards as multimedia generation comprises subtasks which are being treated in other disciplines.

97 http://mmit.informatik.uni-oldenburg.de/SEMPS2006/
98 http://www.dfki.uni-sb.de/imedia/lidos/papers/csi97/node5.html

How does SWeMPs compare to this rationale?

Adequate modularization of the process

The process is realized generically in a logic-based rulebase and made concrete through the interaction with other components which handle specific subtasks of the process such as reasoning about available knowledge, planning Web services and building a consistent abstract multimedia model. The logic-based aspect of the application ensures backtracking, i.e. if an interaction with a component fails, the rulebase examines the working memory and attempts an alternative interaction.

Abstraction of a concrete implementation

The multimedia generation task has been traditionally coded into the system ensuring that systems could handle narrow ranges of tasks very well but were not adaptable to other domains or types of generation task. SWeMPs rather models the task in a formal and explicit manner in a conceptual model based on an ontology, abstracting details of a specific multimedia generation task from the generic process of multimedia generation coded inside the SWeMPs rulebase.

Identification and classification of knowledge sources

The SWeMPs conceptual model acts as the core model of the multimedia generation task which is queried by the rulebase to realize an instance of that task. However, the conceptual model does not include within itself all of the necessary knowledge, rather it acts as an upper ontology which points to the instances of knowledge which the system can use to resolve the multimedia generation task. The SWeMPs ontology allows conceptually differentiating between subjects, resources and services, which play distinct roles within the multimedia generation process, and their respective metadata. It also points to the ontologies which define the domains used by metadata and to occurrences which identify retrievable instances of individuals. From this model, the system can determine relevant knowledge and interpret it in the context of the user's information wish and valid presentation constraints. The process modeled in the rulebase and the knowledge modeled in the conceptual model are related by semantic queries that are handled by the reasoner component.

Modeling of shared resources

Both components (code) and data (the metadata of and referenced from the conceptual model) can be shared. The component based architecture of SWeMPs doesn't exclude the possibility of parallel multimedia generation tasks sharing the same components, allowing for resolution of concurrency (e.g. as the Prova rulebase can be executed within a Java program, one could use Java synchronization constructs). Likewise, as data is conceptualized as an instance in the conceptual model, the instance, uniquely identified by an URI, can be shared across conceptual models for different multimedia generation tasks. As a result, subjects, resources and services can all be re-used.

Openness to other standards

The component based architecture and the abstraction of data and processes (services) in the conceptual model allows for the use of different standards within the multimedia generation process without breaking the core process execution (modeled in the rulebase). Rather components can be changed with only the dedicated rule in the rulebase which interacts with that component needing to be updated (i.e. in terms of the new components API). Resources and services in the conceptual model can be related to a particular MediaType, and ServiceMetadata can describe services for mediating or transforming between different MediaTypes. The conceptual framework has been specified deliberately in an abstract fashion, even the knowledge representation of the conceptual model could be changed so long a reasoner which supports that new formalism is introduced. Hence as progress is made in the areas of Semantic Web, Semantic Web Services and multimedia modeling research, SWeMPs can be revised without negating the conceptual framework upon which it is based.

From this discussion, it appears to be clear that SWeMPs has met the original design rationale of the Standard Reference Model. We can conclude that SWeMPs is a form of IMMPS, closer to the original model than previous implementations and as illustrated in the evaluation that it meets the requirements of automated and adaptive multimedia generation which were specified in this thesis.

This lies in that other multimedia generation systems have not considered the Semantic Web to the extent that we have done so in modeling and implementing SWeMPs, where the Semantic Web has been considered from the very beginning in all the design decisions that were made.

7.3 Further remarks on SWeMPs and possible extensions

We add a few other comments relating to the SWeMPs framework as a further differentiation from past and current IMMPS efforts, including how SWeMPs could be further developed in line with Semantic Web and Semantic Web Service efforts.

As the application code (the rule-set) is generic, it does not need to be altered in the development of the different multimedia services. The implementer concentrates on identifying the relevant content, knowledge and services for the scenario at hand and instantiating the knowledge base for the scenario using the conceptual model vocabulary. The specific multimedia generation process is realized through the applications interaction with the knowledge base and the results of the subsequent operations through the system components. We have seen that as all necessary domain-specific content is referenced through concepts in the conceptual model, it is abstracted from the SWeMPs framework, maintaining a domain independent multimedia generation process. While both knowledge and (annotated) resources are expected to be increasingly available on the Web (or at least extractable through common methodologies), services and presentation rules were in particular needful of manual preparation in the scenarios.

However, the services used by the scenario are also semantic, i.e. they offer semantic information in their interfaces which can permit their dynamic selection and appropriate mediation for usage in an automated system such as SWeMPs. Hence we see that the emerging Semantic Web Service efforts, particularly in Web-based publication and discovery, can also be used by SWeMPs in dynamically finding suitable services for sub-tasks of the multimedia generation process. Semantic Web specific services such as the mediation between certain ontologies should emerge as a requirement of Semantic Web Service communication and can be re-used by SWeMPs. Web based services developed for SWeMPs can be advertised and hence re-used in different scenarios, not only by other SWeMPs-based systems but other implementations which would benefit from typical sub-tasks, e.g. extracting resources and RDF annotations of those resources from the comprehensive content available at sites like Yahoo Travel, Wikipedia and Flickr. Presently, SWeMPs has a local service directory (see section 4.3.3) which contains the descriptions of the services instantiated in the SWeMPs conceptual model and discovery is kept simple, e.g. the service planner selects services on the basis of dedicated SWeMPs properties on the service instances in the knowledge base (section 5.4). However, as Semantic Web Service efforts mature and a global Web directory of services becomes viable, it may be possible to direct SWeMPs to use a set of dedicated Web-based service directories and extend the service planner component to implement richer Web Service discovery, composition and invocation (as is indeed the research intention of Semantic Web Service efforts). While on one hand some issues would have to be resolved in dynamically selecting and using Web Services (such as security, trust and quality of service), this could mean that an implementer may not need to specify services at all when using SWeMPs, and rather rely on their dynamic discovery during execution. However this depends still a lot on the performance of Semantic Web Service infrastructure at a Web scale.

For the presentation rules, by taking the approach of modelling communicative abstractions in an exchangeable Semantic Web format (SWRL) and permitting them to be defined as Semantic Web concepts (by using URI identification), we leave open the possibility of presentation knowledge also being formulated by different users and distributed on the Semantic Web where it can be discovered and re-used. This will require some description scheme for these presentation 'packages' where they can be found and evaluated for use in a specific multimedia generation process, much as today Web site developers would seek to find and evaluate different Web page designs packaged as e.g. CSS files.

Finally, in the light of the shift in computing from software to service, we note that SWeMPs itself could be remodelled from a component-based architecture to a service-oriented architecture (SOA) and hence its own components distributed on the Web, re-used by different implementations and be dynamically selected at execution (e.g. an OWL-S service planner and a WSMF service planner switched between depending on the formulism used by a Semantic Web Service).

7.4 Future developments: multimedia and the Web 2.0

At the time of writing, a new buzzword has arisen in the Web community. Web 2.0 [O'Reilly,2005] refers to a new breed of website which uses an Ajax framework, i.e. asynchronous Javascript and XML. The fundamental paradigm shift which has provoked talk of a new generation of Web is that websites built in this way act more like applications than web pages, in which users can control aspects of their content in a very dynamic way (without HTTP reloads of the entire page). A well known example of such sites at present is Yahoo / Google Maps, which not only present maps to a user but can tag them with items of interest or allow to zoom in or out or scroll to another part of the map. Typical examples of Web 2.0 have some interesting features:

- They give users the chance to “tag” content on the site (e.g. Flickr, Technorati), with self-chosen strings which are used to group related content together. While the use of self-chosen strings would suggest that the typical problems of ambiguity and inaccuracy would exist, and they do, the expectation is that the community would regulate itself (i.e. each user wants to make content accessible to those who seek it through “tagging” and hence would seek to use understood and accurate tags). This has lead to so-called “folksonomies”, i.e. community built taxonomies.
- They open their content to users through defined public APIs so that others can write application code to retrieve content and display it. As well as dedicated applications built on these APIs, the application code can also be used on a Web server to create Web 2.0 type pages in which new things are done with the content of these sites. A popular trend here is “mushing”, which involves mixing content feeds from two or more sites to offer some new functionality, e.g. accommodation listings from Craigs List with Google Maps (to see where the accommodation is).

Provided that location information is clearly marked, mixing content with the mapping services of Yahoo or Google has become an interesting new trend in the Web. Yet in such an effort, the limitations of the present Web still come through: screen-scraping with Javascript, the need for consistency in the underlying HTML presentation, the lack of semantics so that a computer system can interpret Web information automatically. Likewise, while the mixing of text, images and maps can be called multimedia, it is a presentation which is controlled by the map service API (through which the items are added). Access to individual resources so that the user can decide on their layout and present them in a truly multimedia fashion (e.g. using SMIL) is often not possible, and while it is clear that a content provider may always want the final control over presentation of their content, the popularity of the Web 2.0 approach demonstrates the interest that Web users have to have more control over the data on the Web. The copyright issues and politics of the use of Web content aside, the trend is clearly towards more dynamic, interactive, multimedia-based use of Web content. What this thesis is working (or wishing) towards is perhaps “Web 3.0”, in which Web content is semantic, Web resources are semantically marked up and Web content presentation is dynamic, interactive, user-controlled and multimedia-based. This trend is encouraged also by the growing ubiquity of digital content and its access over the Internet through Web devices, which could be broadband (e.g. television in its future, watch-whatever-whenever, form) or low power mobile wireless – both require a new Web-based paradigm for accessing content beyond what is presently the case that proves to be automatable and adaptive.

7.5 Future impact and research directions

In this thesis, SWeMPs was motivated from a vision of a future intelligent, multimedia Web which offers richer delivery of information to users in the form of mixed media

Web content in a coherent, synchronized presentation. The vision is yet to be widely realized, and there is a lot of progress to be made on the public Web (as has been seen in Chapter 6). Semantics are not widely available on the Web and even the evaluation work done in this thesis had to deal with the limitations of the current state of the art (e.g. Flickr string tagging, Yahoo Map interaction limited by the range of the public API, screen-scraping information off Yahoo Travel). We used small sets of semantics and attempted reasonably simple tasks in order to demonstrate the feasibility of the approach. The implementation (at the time of writing) suffices as a proof of concept, but needs a wider development and lacks a user friendly interface. Issues of efficiency and scalability would be part of a further implementation. The Semantic Web is still maturing, in terms of tools and methodologies. Semantic Web Services are at an even earlier stage. However, it is our opinion that for any vision a first step should be made and we feel that this thesis is a strong motivation for further work, stepping out in the name of this vision to encourage the Semantic Web to become reality (with Web 2.0 as the possible bridge), to get real semantics out onto the World Wide Web, to see the framework realized in which SWeMPs, or whatever SWeMPs may evolve into, can realize automated and adaptive multimedia generation processes using these semantics, and that the Web - as the means to ubiquitous access to knowledge through Internet-enabled devices - will evolve to be the user-centred, dynamic, interactive environment that it has the potential to become.

To close this thesis then, we offer a few new research directions which arise from the experiences of this work and which contribute to realising the SWeMPs vision of a future multimedia and semantic Web:

- *Richer high level annotation of Web resources*. Coherent multimedia presentation is based on the conceptual relationships between Web resources, and the property resource-represents-concept is too simplistic to model more complex situations. A richer model for the conceptual meaning of a resource, e.g. MPEG-7's Semantic DS, must be complemented by tools to allow for instantiating such models and ideally in a (semi-)automatic fashion.

- *Rule languages for the Semantic Web*. While Prolog combined with Jena sufficed in the prototypical implementation of SWeMPs, we acknowledged that the logical basis differs from that of Semantic Web languages. The final logical basis for a Semantic Web rule language remains an open issue in current Semantic Web research: on one hand, the rule language needs to work consistently with RDF and OWL data, on the other hand aspects of logic programming such as negation are required in many scenarios. At present, reasoning in the rulebase upon the working memory and in the conceptual model using a Semantic Web reasoner is conceptually separated and a subset of knowledge in the conceptual model is re-modelled as Prolog facts in the rulebase working memory. In the future, we should be able to express the SWeMPs rulebase in a Semantic Web rule language and work directly with the SWeMPs conceptual model, provided issues such as Open vs. Closed World are adequately resolved.

- *Ontology mapping at the instance level*. The heterogeneity of ontologies is a recognised problem in Semantic Web research and there is a bulk of literature focusing on ontology alignment, matching, mapping and merging. However, the focus has been on the TBox, i.e. the classes and properties of an ontology, and the issue of heterogeneity at the ABox, i.e. the instances of a class, is unresolved. We found that in SWeMPs determining the relationship between two instances is an important requirement as Web resources are typically annotated as representing particular instances of a concept rather than the concept (ontological class) itself. As more metadata about instances becomes available, methodologies and tools should arise that allow for determining instance equivalence (and other types of relation).

- *Semantic Web Services for multimedia generation*. While the use of services in business contexts such as e-commerce is being widely considered, the role of services in multimedia contexts has not yet received much attention. As identified in section 5.4 in the context of the multimedia generation process services are required for resource adaptation and conversion. SWeMPs provides a vocabulary for describing such services. A scheme was also arbitrarily chosen for typical instances which need to be referred to in the service description, e.g. MIME Types. As a result, we see the need for agreement on description of Semantic Web Services for multimedia data. Furthermore, the description formulism (probably as an ontology) will need to be enriched to support describing the service's actions in order to support tasks such as the composition of such services to carry out specific multimedia analysis tasks. This requires process modelling and definition of multimedia analysis sub-tasks.

- *Multimedia presentation vocabularies*. The choice of SWeMPs and ZyX vocabularies together with mappings between them and other common schemas may be a step towards standardizing how media is annotated for the presentation generation process. However, we found that the presentation rules also required a vocabulary to express in a more abstract way than ZyX how media resources should be constrained in the presentation with respect to one another. A basic set of spatial, temporal and interactivity constraints was implemented and "named" as instances of a Constraint class in SWeMPs, however further work could be undertaken in a complete vocabulary to express relationships between media in a presentation in an abstract fashion as well as tools to interpret those relationships concretely in a multimedia presentation.

- *Multimedia end format for the Web*. The Web browser is effectively a hypertext system and multimedia is typically represented in-page through a different application, e.g. Flash or Quicktime. As a result the content of the multimedia object is not integrated with the content of the page. It proves a limitation to the multimedia Web which should offer rich, integrated and synchronized presentations of Web content. SMIL has been promoted by the W3C as a multimedia presentation standard for the Web, though lacking the textual layout support of HTML it will fail to replace the Web browser for a means for navigating the Web. An end format which covers XHTML+SMIL+SVG might be the solution for a new multimedia platform for surfing a multimedia Web.