A Flexible Architecture for Semantic Annotation and Automated Multimedia Presentation Generation

Payam M. Barnaghi[1] and Sameem Abdul Kareem[2]

[1] School of Computer Science and IT, the University of Nottingham (Malaysia Campus)
43500 Semenyih, Malaysia
payam.barnaghi@nottingham.edu.my
[2] Faculty of Computer Science and IT, the University of Malaya
50603 Kuala Lumpur, Malaysia
sameem@um.edu.my

Abstract. Multimedia information system design has been recently influenced by state-of-the-art technologies such as those provided by semantic web. This paper introduces an information search and retrieval methodology that employs semantic web technologies for its data representation and reasoning tasks. Creating a meaningful multimedia presentation is an attempt to answer a user's query using a knowledge representation structure. In this case, the system instead of providing a list of results, as happens in many typical information search and retrieval systems, collects a list of selected items based on the relevancy to the queried topic and also meaningful relationships between data objects. The collected information is used for the construction of an automated presentation to demonstrate the results to the user. The paper concentrates in particular on the content selection, narration structuring, and presentation design and generation processes. The implementation of an automatic presentation generation facility in an integrated system, called MANA, is described through the paper. MANA is designed to generate adaptive and automatic presentations based on users' perspectives and preferences on the queried topic.

Keywords: Multimedia Data Representation, Automated Generation of Multimedia Presentations, Semantic Associations Search, Adaptive Multimedia Presentation Generation.

1 Introduction

Current advances in Web technologies and information systems have led to the provision of rich media descriptions and enhanced reasoning techniques to interpret the contents and meaningful relationships of multimedia data. Although current advances of the Web and information presentation systems allow having rich media enabled information presentation in the form of hypermedia, most of the current information search and retrieval systems only demonstrate a list of the results to the user. Among the various research works in the information search and retrieval area, a part is focused on developing efficient methods to provide more relevant information and enhanced ranking mechanisms which are implemented in the algorithms such as PageRank [1], HITS [2], and Teoma [3]. On the other hand the user's main purpose of searching the web is frequently for obtaining knowledge, not a list of related documents. By employing an automated presentation generation mechanism, information retrieval systems would be able to present the query results to users in an improved form acting as knowledge retrieval systems instead of being only information retrieval systems.

Initial work in this area was carried out by introducing a standard reference model for intelligent multimedia presentation systems (SRM-IMMPS) [4]. The standard reference model describes higher-level design characteristics and plan-based approach to generate multimedia presentations. The proposed approach assumes that the required media attributes and specifications for multimedia data are available. The standard model emphasises the significance of knowledge representation and processing in producing adaptive multimedia presentations [4,5]. In the current research, we consider different levels of metadata and knowledge representations for multimedia data and discuss how different aspects of the automated presentation generation process are implemented in an integrated presentation generation system, called MANA. MANA is an automated hypermedia presentation generation engine which utilises semantic web technologies and reasoning methods in response to information queries.

This paper focuses on multimedia data annotation and multi levels of knowledge representation in its design to address different aspects of an adaptive and automatic multimedia presentation generation process. We describe how semantic support can be applied to multimedia data descriptions and discuss how implicit data hidden in meaningful relationships between the objects is used to extract explicit knowledge to acquire the relevant data. The main goal is in organising the objects in a presentation scenario in response to information queries.

2 The Presentation Generation Process

The presentation generation process attempts to develop a narrative structure to illustrate the relevant information regarding a main topic. The main goal is answering the user's query with a hypermedia presentation instead of providing a listing of results. The relationships between the participant objects of a presentation have to be identified and organised in order to represent a smooth and meaningful narrative to describe the retrieved results to the user. An automated presentation generation system facilitates discovery of knowledge from information resources, and presents this knowledge to the users, i.e. obtaining a knowledge-driven presentation tailored to the user's query and preferences. The proposed system enables searching, accessing, and presenting the information based on two main approaches: 1) *information discovery* to facilitate processing of enriched multimedia data. The information discovery is the data collection process. It refers to the content descriptions to address the relevant multimedia data to a query topic 2) *information presentation* to allow the user to view the relevant information through a presentation construction. The presentation structure is intended to be constructed in a flexible and adaptable form with respect to the user's browsing environment and device, and her/his contextual preferences on the queried topic.

We have developed a specific document annotation and description structure for multimedia data based on semantic web information representation structures. The semantic annotation and description model is incorporated with higher-level domain and design knowledge which allow reasoning methods to interpret the contextual information and meaningful relationships for the multimedia data. An automatic presentation generation system is then developed with respect to the data annotation and specification model. The development of the presentation generation in the MANA system consists of the following steps:

- implementation of the data annotation model and organisation of information structures based on content and media specific attributes of multimedia objects and meaningful relationships between the objects.
- adaptation of specific search and categorisation methods for content selection and collection purposes.
- development of intelligent information filters responsible for the discovery of meaningful relationships and addressing the relevant data based on the collected contents.
- construction of narration structures based on the collected data and user preferences to view the results.
- presentation generation to construct a structure based on data objects, and their relationships, and to deliver the results in a suitable format to the user.

All of these cooperatively provide an integrated architecture to retrieve information and obtain knowledge from the annotated multimedia objects to generate a presentation to answer information queries.

3 Concepts and Semantics in a Multi-Facet Environment

The system utilises an annotation structure to describe the multimedia data with specific reference to those aspects which are necessary to facilitate the presentation generation process. The descriptions should provide two main views to the objects: the content-specific data (e.g. what, where, who, is described in the object), and technical and media-dependent descriptions (e.g. size, width, height, etc). The implementation of this structure includes a query and retrieval mechanism which facilitates the collection of suitable objects from the repository for the presentation generation purpose. The system needs to access the annotations and to process the specifications in order to select the appropriate data items based on the submitted query. Although there are tools to retrieve and process the annotations, but when it comes to interpreting the content and extracting useful information, the capabilities of current software are still very limited. An alternative approach is to represent the data in a form that is more machine-accessible and interpretable. The semantic web technologies [6] and ontology-based knowledge representations provide machine-accessible and machine-interpretable information specifications which are used in the current architecture.

In order to create an annotation and specification structure capable of expressing different aspects of multimedia data, the following requirements have been considered.

- including content-specific features and also media dependent attributes to satisfy both the content selection and technical presentation generation requirements with regards to common annotation structures and vocabularies such as MPEG-7 [7] and Dublin Core [8].
- encapsulating the interdependencies amongst the multimedia data through sets of classes, relations, functions and constraints for the domain of discourse (i.e. semantic support for multimedia data and higher-level ontologies) and specification of interpretation rules (i.e. description logic).
- defining contextual attributes which define a generalised view to the concepts in a particular domain. The contextual specifications will allow the system to adapt the presentation based on a user's expectations of the query.
- expressing subject matter metadata which depicts the association of the multimedia content to the domain concepts represented in a knowledge-base. This allows the system to associate high-level concepts and low-level media dependent features. The difference between the low level feature descriptions provided by content analysis tools and the high level content descriptions required by the applications is often referred to, in the literature, as the "Semantic Gap" [9]. A higher-level explanation of multimedia data according to a standard vocabulary of domain concepts represented in domain ontology (i.e. knowledge-base) would allow information engineers to effectively address this issue.

While the specified requirements describe the main aspects of a data model scheme, a proposed model that satisfies these requirements also has to remain adequately simple and flexible to be applicable. The model also needs to remain sufficiently expressive for the majority of multimedia data types.

4 Multimedia Data Annotation Model

We define an annotation and description model in which the structure of a multimedia data representation consists of different distinct specifications described through the schemas shown in a UML diagram in Fig. 1.

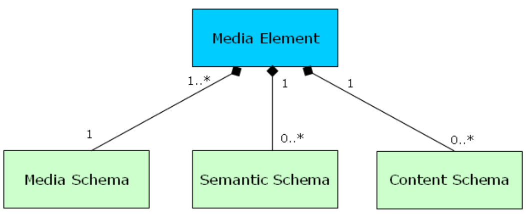


Fig. 1. The structure of multimedia annotation model

The annotation structure consists of content, semantic, and media schemas. The content schema describes the multimedia contents and explains their conceptual notions. The schema involves entities such as objects, abstract concepts and relationships, topic, and application domain related attributes. The media schema describes multimedia data from its media and type dependent viewpoint. This schema involves entities such as size, type, length, and other temporal and spatial features of multimedia data. The semantic schema defines a scheme that is used to describe the events, theme, contextual and internal relations of the multimedia data. We have chosen fine arts as the application domain and in particular painting as a sample to choose data sets and to describe our framework. For our sample collection, we use multimedia data taken from the Getty Museum [10]. To define the vocabularies to describe the annotation structure, we have focused on common vocabularies of standards such as MEPG-7 [7], Dublin Core [8] and FOAF [11].

We realised that some of the vocabularies such as Dublin Core are not specifically designed to describe multimedia data. Consequently, using general vocabularies such Dublin Core for metadata description would not be an efficient and sufficient solution. For example there are special spatio-temporal considerations for the media items which are not addressed in the Dublin Core specifications. There are also some other standards like MPEG-7 which are more specifically designed for multimedia data description. MPEG-7 uses XML based descriptors to provide content descriptions for multimedia data. The basic structure of MPEG-7 to describe the multimedia content follows a hierarchical structure. The content descriptors could be represented as a graph in which the description schemas are linked together through relationships between different content specifications [12,13]. The proposed data model benefits from existing meta-data description vocabulary as well as new desired features to provide an efficient annotation structure for our purpose. Utilising the standard vocabularies such as MPEG-7 enables the model to provide more interoperability to the data representation structure. We have defined the attributes for our data annotation model based on standard vocabularies when it has been applicable and we have also added our own detailed properties to define the annotation structure for multimedia objects in the selected discourse domain model. The model implementation utilises semantic web ontology language [14] and description logics to specify the multimedia data.

The content extraction functions and methods are not considered in the model. The content extraction functions are beyond the scope of this research. We assume that the content-specific and technical information are available to the archivist and the system has access to multimedia data through a local repository (or Web links). The following describes each aspect of the proposed annotation model. A simplified structure of the represented features in the model is also shown in Fig. 2. In Fig. 2 "*dc*","*mpeg-7*", and "*mana*" respectively refer to "*Dublin Core*", "*mpeg-7*", and "*mana*" namespaces[1]. The "*mana*" namespaces is developed for the purpose of this research, and "*Dublin Core*", "*mpeg-7*" namespaces are adapted from the existing standards.

---

1 The description of namespaces and OWL documentation of the data model is available at: http://csit.nottingham.edu.my/~bpayam/mana/dm/

4.1 Content Schema

Content-specific information refers to explicit and implicit data obtained from the extracted features and meta-data descriptions. The content of a multimedia object specifies its own intrinsic features. This can be described as explicit data like author, title, keywords, subject, etc. Additionally, the implicit data such as action, scene, etc. could be specified through this schema. The content specific aspects are associated with the ontology concepts. This provides a common vocabulary to annotate the multimedia data, and also to describe meaningful relationships between the data objects in the context of the discourse domain.

4.2 Semantic Schema

The semantics of multimedia data is provided independently of its content-specification structure. Both, content and semantic information, however, are needed for the complete contextual description of a media item. The semantic description and the content description illustrate the perceptual structure of the media items. The semantics are associated to the concepts in the domain ontology, which represents the expert's knowledge of the discourse domain. The semantics represents the contextual subject, and predetermined relationships between the objects (direct associations). The features of this schema are also used for determining the relevancy of objects to the user's contextual perspectives.

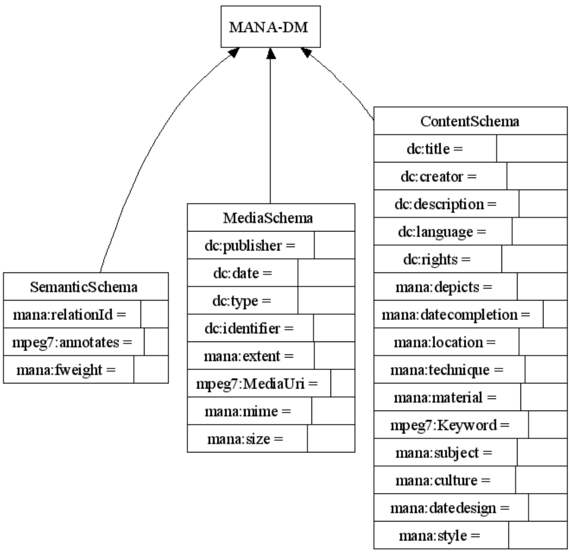


Fig. 2. A simplified structure of the represented features in the model

4.3 Media Schema

The media schema describes technical aspects of multimedia data. Media information for an object specifies the media dependent characteristics of the object in which the specified information is taken from the binary representation of the multimedia object. Each multimedia object has an identifier, a resource link and a set of media features (such as, type, size, duration, etc.) that is dependent on the binary file, or stream representing the media item. An atomic media scheme instance of a specific multimedia object represents spatial and temporal characteristics of the media item. These features lead to define the presentation layout. The media schema in addition to describing the main temporal and spatial features of the object can also describe different versions of a multimedia element such as a part of a video stream, or a part of an image file, etc.

4.4 The Model Representation

The model inherits properties from different description standards and integrates them in a unified structure. The "*mana:depicts*" property is used for associating media objects to the domain concepts. This facilitates the description of the multimedia objects and their relationships according to the domain ontology concepts. Deploying the proposed model and using ontology-based annotation will provide a unified and coherent annotation mechanism and overcome the problem of using vast and inconsistent terminology to describe the multimedia data. Fig. 3 illustrates annotation attributes of an instance multimedia object according to the schemas specified in the data model. The annotations in the system are specified based on OWL-DL [14] serialised in RDF/XML [18] format[2].

5 Presentation Generation Paradigms

In the previous section we discussed multimedia data representation requirements and proposed a multimedia data annotation model. The proposed model indicates the explanatory attributes for multimedia data in order to be used in an automated presentation generation system. The model alone, however, is effective if it is not incorporated with other aspects of the system in order to create an adaptive and automated multimedia presentation. In other words, multimedia data is required to be annotated and described through the data model guidelines, but in addition to such basic requirements, the authoring units need to access and interpret this data to supports other aspects of the presentation generation process, i.e., data collection, organisation of the collected data, narration construction, template design/selection, and transformation. A brief but interesting overview of the current efforts that has been carried out in meaningful multimedia presentation generation is provided by Hardman and Ossenbruggen in [5]. In the current paper, we refer to some of the

---

2 The complete description of the model is available at: http://csit.nottingham.edu.my/~bpayam /mana/MANA.owl

related work carried out to generate knowledge-driven multimedia presentation. We try to highlight different aspects of an automatic presentation generation process and discuss the proposed solutions.

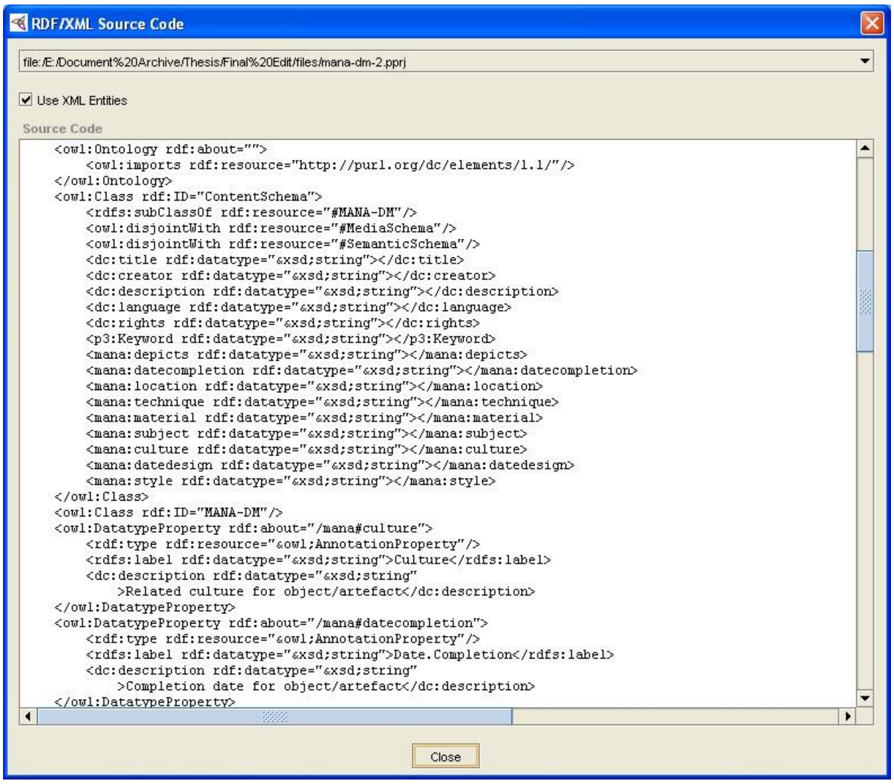


Fig. 3. A fragment of OWL syntax of the model

DISC [15] uses a multimedia data repository and provides a document generator to create automatic multimedia presentations based on a user's query. The system focuses on using higher-level knowledge representation and ontologies to improve the information retrieval as well as the presentation planning and generation processes. The proposed method is domain dependent and is presented based on a biographical knowledge-base. There are discourse, presentation and design ontologies that work together in the presentation generation process. The domain ontology describes the entities and their relationships in the disclosure domain. The multimedia data is annotated based on the domain ontology concepts and relations. The relevant data to a user's query is then collected by referring to these annotations and interpreting the relationships between objects. The object selection process is basically focused on role-based rules which are represented through the design ontology. DISC depicts what type of roles can appear in the presentation and in what type of narration structures. For example a "Personal Life" structure is described with specific domain concepts such as (Person, Artist). This allows the system to decide on the types of data that could be mapped to each specific unit of the predefined presentation templates [15,16].

In the same context, Rutledge et al [17] described a framework to generate hypermedia presentation from annotated multimedia data stored in an RDF repository. Their research focuses on using semantic web technologies and information presentation through the semantics and contextual meta-data. The RDF [18] data is stored in an RDF repository and queried using an RDF Query Language [19]. The user submits the query in text format and then the system encodes the text query to the RDF query language to address the relevant data from the repository. The relevant data items to the query topics are represented as the leaf nodes of a hierarchical structure. The nodes, hierarchy and their sequence are included in a document model which is called structured progression [25]. The significance of the system is forming this enhanced structured progression and generating the final presentation based on the clustering and inferring processes. The clustering process is based on lattice concepts and utilises common features of the leaf nodes to determine the clusters. The semantic relation extraction involves inferring the RDF graph in the repository. The system creates the stories through a combinational approach. The data resources are selected based on the query and clustering mechanism (bottom-up design) and the higher level information (obtained from ontology) are applied to the structure to organise the presentation (top-down design).

Aroyo et al [21] proposed a hypermedia generation system which deals with the external as well as the internal data resources. An information retrieval agent is responsible to retrieve the external data from the Web resources. The system defines three main components to support the adaptive hypermedia presentation generation. The components are represented as, user model, domain model and application model. The domain model represents a semantic structure of the concepts and relationships between the concepts in the system. The concepts and their descriptive attributes are defined in a domain model. The domain model conceptualises the entities and defines how these entities are related to each other in terms of domain concepts. The user model is an overlay model of the domain model which defines the same concepts and associates user-attributes to the represented concepts. For example, the concepts could be represented in different topics and levels in the domain model, and in this case the user model defines the user’s level of knowledge or interest for these concepts. The application model is a set of rules which associate the domain and user model in order to generate the final presentation.

The HERA system [22] provides a model-driven methodology to generate dynamic hypermedia presentations based on ad-hoc queries. The (semi-)automatically generated hypermedia presentation is driven through different sources of intelligence which are embedded into the system. The design knowledge shares the designer’s expertise and expresses guidelines to present the data collections. HERA collects the data from heterogeneous resources and the system utilises high-level model-based abstractions that drive the presentation generation process [22,23].

Little et al [24] described a semi-automatic and intelligent multimedia presentation generation approach through semantic inferring. They proposed a high-level architecture which generates multimedia presentations by using both reasoning and multimedia presentation generation tools. The meta-data schema are represented based on Dublin Core [8] and OAI [25] archives. The system utilises an inference engine which refers to the specifications and logical description to interpret the semantic relations between the objects. As a result, data items and their peer items (inferred based on the logical rules) would be the potential candidates to participate in the individual slides of the presentation structure. In the inferring process, the system refers to a set of predefined rules and utilises a mapping between the extracted semantics and MPG-7 semantic relationships. While the system provides reasoning mechanisms to select the presentation contents, the functionality of this methodology could be enhanced using the domain ontology to conceptualise the semantics and rhetorical relations. The ontology-based knowledge representation could also extend the logical description and reasoning tasks by addressing the concepts and their contextual relationships, while the functionality of the system could also be extended to more flexible queries.

Weitzman and Wittenburg [26] used grammatical rules to provide a mechanism for mapping between content-based features of the media objects and the presentation style. The grammars encapsulate “look and feel” of a presentation and are used to generate the presentation style. The idea of relational grammars focuses on semantic relationships between the multimedia objects. In this context, the grammatical rules are determined by the rhetorical structure of the presentation. In a typical multimedia presentation generation process, the contents are initially addressed using an information search and retrieval agent. The addressed contents represent a set of multimedia objects which have some meaningful relationships between them. The relations are organised based on the presentation scenario which could be defined through the different methodologies. The presentation design process uses the collection of multimedia objects to organise and assign the multimedia elements to spatial and temporal layouts. The main role of relational grammars is to define the constraints in different dimensions based on the semantics. For example, in an authoring process a grammatical rule defines that the “title” of an image should be placed on top of the image (communicates with the spatial layout).

The presentation structure could be specified by predefined templates which are filled-in with media items during the presentation generation process. In this case all the spatial, temporal and navigational structures are fixed and predetermined. The user selects a presentation theme, and then the content specification unit proceeds to collect data items (based on a query or a selected topic). The selected objects are subsequently used to fill-in the predefined template. An enhanced presentation structure includes flexible spatial, temporal specifications. In this context, the design rules are taken into account during the presentation structuring process to define the dynamic template for the presentation. The presentation design rules are translated to the spatial, temporal and navigational constraints during the presentation realisation process. The rhetorical structure specifies how the relationship between the objects could be used in the presentation structure. It is similar to what is called the design ontology in Rutledge et al’s work [17]. The design knowledge is expressed as a set of constraints. Once the objects are selected, the rhetoric structure is used to analyse the collections, and then based on the relationships between the objects, design rules are applied to the presentation structure.

In the discussed works the key idea is in using domain and design knowledge and reasoning in different aspects of automated multimedia presentation generation process. In the next section, we show how the idea of knowledge-driven presentation generation has been adapted to design an integrated automated presentation generation system.

6 Automated Presentation Generation Architecture

We define our architecture based on the proposed data model and reasoning techniques to process machine-interpretable data representations. The design aims to define an adaptive presentation generation framework to answer information request queries. The system architecture is divided into four main layers namely, the resource layer, discourse layer, aggregation layer and presentation layer, as shown in Fig 4. The layers are analogous to the standard reference model's layered architecture [4] (i.e. content, design, presentation and realisation layers in the standard reference model). The ontologies mentioned in Fig. 4 are the common knowledge specifications corresponding to each layer, in that they contain information external to the media items and layers. The detailed specification of ontology-based presentation generation is described in [27]. We focus on describing how the proposed architecture satisfies an adaptive and automatic presentation generation system's requirements.

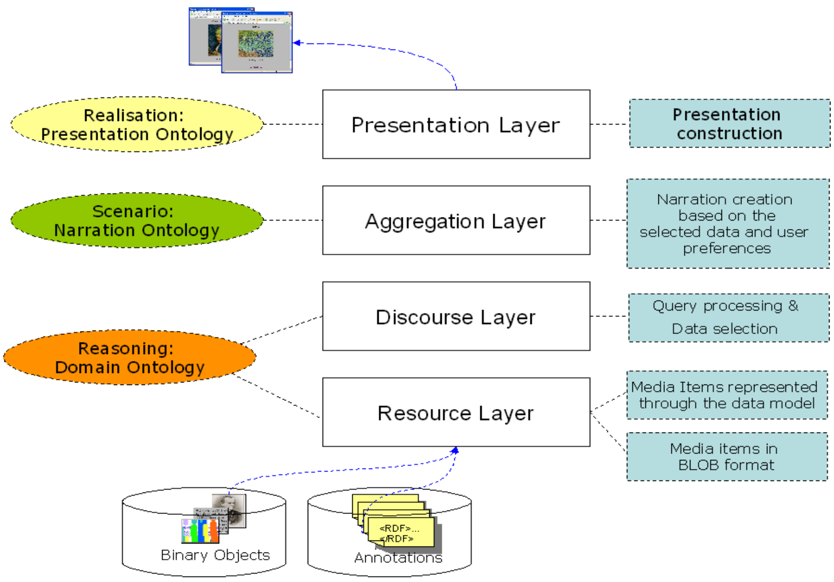


Fig. 4. The layers in the presentation generation architecture

The resource layer contains the multimedia data and the annotations. The multimedia data is stored in a repository as binary objects (local resource) or could be addressed by means of a URI (external resource). The explanatory data represented in terms of the data annotation model specifications and stored in an additional repository. We use Sesame [28] RDF repository to store the ontologies and the annotations. Sesame supports the RDF document manipulation and query. The RDF query in Sesame is provided using a specific RDF query language which is called SeRQL [29].

In order to collect the relevant information of a specific query topic there are two possibilities: either the media items are selected based on the specified metadata and keyword-based search (i.e. direct search), or the selection process addresses the items based on their meaningful relationships to the query topic using an inference process (i.e. indirect search). The inference process results in the extraction of explicit knowledge from implicit information represented through the meaningful relationships between the data objects in the knowledge-base[3]. For example, if the query term is “Rembrandt”, Rembrandt painted a scene that is called “The Mill”. The data items which describe “The Mill” (i.e. image and documents related to this painting in terms of locale, style, etc) are relevant to the query topic to some extent and could be considered as candidate objects to be included in the presentation structure. The data selection process is a recursive progression and the media items would be selected through multiple selection steps based on their semantic association to the query topic. The inferring and content collection processes are both implemented in the discourse layer.

We have implemented a semantic association ranking mechanism to evaluate the complex relationships between the entities in the knowledge-base. In particular, to develop an automatic presentation about a topic, the system attempts to select a main topic from the knowledge-base and then semantic association search mechanism starts spanning over the topic in the knowledge to find the related data. The ranking mechanism measure the robustness of the semantic associations between the main entity and other entities in the knowledge-base. This produces a weighted graph which is used as the fundamental narrative structure to organise the final presentation. We are aware that the discovery query and selection of the main topic is also an essential part of the system. It is necessary that the system is able to find a topic from the knowledge-base and then process the relations to identify the semantic associations. The ranking mechanism is described in detail in [20].

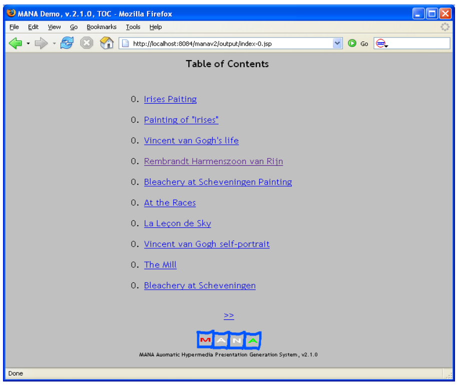
The main components of the presentation narrative are defined in the narration ontology[4]. These components describe the primitive structure of the presentation. The narration construction process develops a progression structure which describes the query topic and also the related information. The progression structure is represented as a graph. This graph is extracted from the main RDF/XML graph (i.e. knowledge-base) based on the narrative components, and selecting the main entity and extending this main topic by analysing the semantic associations. The graph structure describes the related entities to the queried topic and the ranking weights express the degree of relevancy of each related entity to the user’s query. The relevant data is then organised based on the precedence of the events that are defined in a presentation theme. The definition of the presentation aspects and the processing of progression graph are implemented in the presentation layer.

---

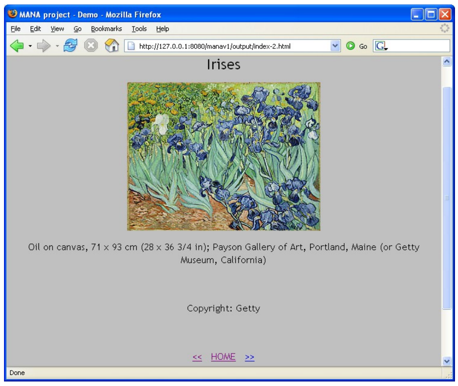
3 The OWL description of the domain ontology is available at: http://csit.nottingham.edu.my /~bpayam/mana/mana-kb.owl

4 The OWL description of the narration ontology is available at: http://csit.nottingham.edu.my /~bpayam/mana/narrationOnto.owl

The main presentation aspects that need to be defined in the construction are defined as: temporal layout, spatial layout, styles and anchor links. The layout definition for the presentations in our architecture is similar to predefined template specifications in HERA [22]. The predefined templates describe the presentation layouts for different themes (i.e. predefined templates for essay, biography, documentary, etc). The system employs a set of XSL style-sheets [30] to specify different presentation layouts. The presentation layer selects the appropriate style-sheet based on the user's selected parameters, such as theme, bandwidth, and size. The selected style-sheet is applied to the results graph in order to generate the final presentation structure. After associating the objects to the presentation template, the system applies XSL transformations [31] to the progression structure in order to generate the final presentation. The generated presentation includes two significant parts: a table of contents, which lists the titles and provides direct links to the slides, and the contents of each particular slide. The navigation structure is provided based on index links from the table of contents, and guided tour which navigates between the slides. Fig. 6 demonstrates instance slides of a presentation which is automatically generated based on a sample query (i.e. "Irises").



a) the table of contents slide



b) a sample content slide

Fig. 5. A sample automatically generated presentation

7 Conclusion

In this paper we discuss using the concepts of existing annotation and multimedia specification standards into a richer document model to represent multimedia data. We use semantic web knowledge representation and reasoning mechanisms to provide enhanced specification and interpretation for the multimedia data. The paper states the requirements and describes the architecture for an automated presentation generation system. We follow a knowledge-driven methodology to generate multimedia presentations. The SRM-IMMPS [4] has formalised the production of multimedia presentations based on different levels of knowledge representation and processing. The crucial role of knowledge-bases and reasoning in automated multimedia presentation is also discussed in different works [5,15,16,17,21,22,24].

The paper describes a knowledge representation method to describe the different aspects of multimedia data. The proposed system employs an inference engine which is responsible to process the attributes and relationships of multimedia data that is represented in the knowledge-base. A ranking mechanism is employed to evaluate the semantic associations between the objects. The system collects the relevant data to a query topic by referring to the knowledge-base and meaningful relationships between the objects. The relevant data is then ranked according to the semantic associations and user selected preferences. The results of the ranking mechanism incorporating with the other components of the system provide an integrated architecture to generate automated multimedia presentations. While we feel that the components designed in the system are functioning adequately, they do not necessarily present the most comprehensible architecture for an automated presentation authoring system. Questions can be asked on how the system will perform for a huge set of data when the result of the reasoning process and semantic associations search may retrieve an enormous number of candidate objects. Another important issue is how the individual phases in the presentation authoring can be improved to deal with external data (with unknown attributes) and how narration structuring can be enhanced to generated more improved presentation constructions.

The work reported in this paper is, however, only a part of a complete environment for the creation, storage, annotation, construction, manipulation, transmission and play-back of a hypermedia presentation. Additional requirements are approaches to manipulating documents compatible with the proposed document model, and methodological approaches to process and interpret the external resources (i.e. the Web data) and analysing them to define the sets of annotated documents. The challenge is to guarantee that these methodologies will cooperate with the other developed parts of the system to create more meaningful and adaptive presentations to answer the user's information request queries.

Reference

1. Brin S., and Page L. (1998). The anatomy of a large-scale hypertextual Web search engine. In Proceedings of WWW1998 Conference, pages 107-117.
2. Kleinberg J. (1999). Authorative sources in a hyperlinked environment. Journal of ACM, no. 48, pages 604-632.
3. Davison B.D., Gerasoulis A., Kleisouris K., Lu Y., Set H., Wang W., and Wu B. (1999). DiscoWeb: Applying Link Analysis to Web Search. In Poster Proceedings of the Eighth International World Wide Web Conference, pages 148-149.
4. Bordegoni M., Faconti G., Feiner S., Maybury M., Rist T., Ruggieri S., Trahanias P., and Wilson M. (1997). A standard reference model for intelligent multimedia presentation systems, Computer Standards and Interfaces, vol. 18, pages 477- 496.
5. Hardman L., and Ossenbruggen J. v. (2006). Creating meaningful multimedia presentations, Technical Report, INS-E0602, CWI, the Netherlands.
6. W3C Semantic Web Activity, available at: <http://www.w3.org/2001/sw/>

7. MPEG-7 (2003). ISO/IEC JTC1/SC29/WG11, Coding of Moving Pictures and Audio, MPEG-7 Overview (version 9), available at:<http://www.chiariglione.org/mpeg/standards/mpeg-7/mpeg-7.htm>
8. Dublin Core Metadata Element Set Version1.1 (1999). Reference Description, Dublin Core Metadata Initiative, July 1999, available at: <http://dublincore.org/documents/1999/07/02/dces>
9. Ossenbruggen J. v., Troncy R., Stamou G., and Pan J. Z. (editors) (2006). Image Annotation on the Semantic Web. W3C Working Draft 22 March 2006, Available at: <http://www.w3.org/TR/swbp-image-annotation/>
10. Getty Museum, The J. Paul Getty Trust, available at: <http://www.getty.edu/>
11. Brickley D., and Miller L. (2005). FOAF Vocabulary Specification, Namespace Document 27 July 2005, Available at: <http://xmlns.com/foaf/0.1/>
12. Manjunath B. S., Salembier P., and Sikora T. (editors) (2002). Introduction to MOEG-7 Multimedia Content Description Interface. John Wiley.
13. ISO/IEC 15938-3:2001 (2001). Multimedia Content Description Interface- Part3: Visual. version1.
14. McGuinness L., and Harmelen F. (editors) (2004), OWL Web Ontology Language Overview W3C Recommendation, available at: <http://www.w3.org/TR/owl-features/>
15. Geurts J., Bocconi S., Ossenbruggen J. v., and Hardman L. (2000). Towards Ontology-driven Discourse: From Semantic Graphs to Multimedia Presentations. In the Proceedings of Second International Semantic Web Conference (ISWC2003), pages 597-612.
16. Ossenbruggen J. v., Geurts J., Cornelissen F., Rutledge L., and Hardman L. (2001). Towards Second and Third Generation Web-Based Multimedia. In Proceeding of the Tenth International World Wide Web Conference, page 479–488.
17. Rutledge L., Ossenbruggen J. v., and Hardman L. (2004). Structuring and Presentation Annotated Media Repositories. Technical Report, INS-E0402, CWI, available at: <http://www.cwi.nl/ftp/CWIreports/INS/INS-E0402.pdf>
18. Resource Description Framework (2004). W3C Recommendation, available at: <http://www.w3.org/RDF/>
19. The RDF Query Language (2003). Institute of Computer Science - Foundation of Research Technology Hellas, Greece, <http://139.91.183.30:9090/RDF/RQL/>
20. Barnaghi P. M., and Sameem A. K. (2006). Relation Robustness Evaluation for the Semantic Associations. Technical Report, the University of Nottingham(Malaysia Campus), available at: <http://csit.nottingham.edu.my/~bpayam/publication/pmb-061030.pdf>
21. Aroyo L., Bra P. D., Houben G., and Vdovjak R. (2004). Embedding information retrieval in adaptive hypermedia: IR meets AHA!. New Review of Hypermedia and Multimedia, vol. 10, no. 1, pages 53-76.
22. Vdovjak R. , Frasincar F., Houben G. J. , and Barna P. . (2003) Engineering Semantic Web Information Systems in HERA. Journal of Web Engineering, vol. 2, no. 1-2, pages 3–26.
23. Rutledge L., Houben G., and Frasincar F. (2004) Combining Generality and Specificity in Generating Hypermedia Interfaces for Semantically Annotated Repositories. Proceedings of First International Workshop on Interaction Design and the Semantic Web, ISWC 2004.
24. Little S., Geurts J., and Hunter J. (2003). Dynamic Generation of Intelligent Multimedia Presentations through Semantic Inferencing, In Proceeding of the 6th European Conference on Research and Advanced Technology for Digital Libraries, pages 158-175.
25. Open Archives Initiative (2004). Available at: <http://www.openarchives.org/>
26. Weitzman L., and Wittenburg K. (1994). Automatic Presentation of Multimedia Documents Using Relational Grammars. In Proceeding of ACM Multimedia 1994. pages 443-451.
27. Barnaghi P. M., and Sameem A. K. (2004). Content-based hypermedia presentation generation: a proposed framework. In Proceedings of the IEEE TENCON 2004 Conference, vol. 2, pages 274- 277.
28. Sesame RDF repository, available at: <http://www.openrdf.org/>

29. Karvounarakis G., Alexaki S., Christophides V., Plexousakis D., and Scholl M. (2002). RQL: a declarative query language for RDF. In Proceedings of the 11th international conference on World Wide Web table of contents, pages 592 - 603.
30. The Extensible Stylesheet Language Family (XSL) (1999). Available at: <http://www.w3.org/Style/XSL/>
31. XSL Transformations (XSLT) (1999). Version 1.0, W3C Recommendation 16 November 1999, < http://www.w3.org/TR/xslt>