



Centrum voor Wiskunde en Informatica

**REPORT***RAPPORT*

*INS*

Information Systems



*Information Systems*

Context aware guidance for multimedia authoring:  
harmonizing domain and discourse knowledge

K.I. Falkovych, F.-M. Nack

**REPORT INS-E0502 FEBRUARY 2005**

CWI is the National Research Institute for Mathematics and Computer Science. It is sponsored by the Netherlands Organization for Scientific Research (NWO).

CWI is a founding member of ERCIM, the European Research Consortium for Informatics and Mathematics.

CWI's research has a theme-oriented structure and is grouped into four clusters. Listed below are the names of the clusters and in parentheses their acronyms.

Probability, Networks and Algorithms (PNA)

Software Engineering (SEN)

Modelling, Analysis and Simulation (MAS)

**Information Systems (INS)**

Copyright © 2005, Stichting Centrum voor Wiskunde en Informatica

P.O. Box 94079, 1090 GB Amsterdam (NL)

Kruislaan 413, 1098 SJ Amsterdam (NL)

Telephone +31 20 592 9333

Telefax +31 20 592 4199

**ISSN 1386-3681**

# Context aware guidance for multimedia authoring: harmonizing domain and discourse knowledge

## ABSTRACT

This paper presents an approach to assist authors during the authoring of multimedia presentations. We extend existing authoring support by integrating processes of topic identification, content collection and discourse structure building in a single environment. This integration allows identification of the context of the authoring process. Our approach combines this process context awareness with explicit domain and discourse knowledge to steer system suggestions. We evaluate our approach with an experimental system prototype.

*2000 Mathematics Subject Classification:* none

*1998 ACM Computing Classification System:* H.5.1 Multimedia Information Systems H.5.4 Hypertext/Hypermedia (Architectures, Navigation, Theory, User issues), I.7.2 Document Preparation (Hypertext/hypermedia, Markup languages, Multi/mixed media)

*Keywords and Phrases:* Context aware multimedia authoring; Discourse structures for multimedia authoring; Semantic support; User assistance



# Context Aware Guidance for Multimedia Authoring: Harmonizing domain and discourse knowledge

Kateryna Falkovych and Frank Nack

Centrum voor Wiskunde en Informatica  
P.O.Box 94079  
NL-1090 GB Amsterdam  
The Netherlands  
[Firstname.Lastname@cw.nl](mailto:Firstname.Lastname@cw.nl)

**Abstract.** This paper presents an approach to assist authors during the authoring of multimedia presentations. We extend existing authoring support by integrating processes of topic identification, content collection and discourse structure building in a single environment. This integration allows identification of the context of the authoring process. Our approach combines this process context awareness with explicit domain and discourse knowledge to steer system suggestions. We evaluate our approach with an experimental system prototype.

## 1 INTRODUCTION

Authors construct multimedia presentations by collecting, organizing and manipulating various sorts of media items with the goal of assembling them into a coherent audience experience [41]. The outcome of this process is a document that is meaningful to the author as well as to readers. Authoring is an active and complex process that one only masters through ongoing learning: selecting topic and content enriches authors' knowledge about a domain; manipulating content helps authors investigate different presentation structures, and thus different ways of articulation [30].

Existing authoring tools for multimedia presentations provide functional support for the authoring process in the form of tools or languages, enable automatic presentation building on request or aid a specific phase of the authoring process. To make use of these systems, authors must have knowledge about what content they want to author and how they want to author this content. Thus, the learning aspects of knowledge acquisition and expertise gaining are left solely to the authors.

The field of distributed cognition suggests that many types of conscious human activity are structured by the use of external tools [23]. This viewpoint is supported by the developments in HCI [24, 25, 26]. Using intelligent tools and guidance during learning allows us to go beyond current levels of skills and knowledge [43]. Given this, our goal is to aid the authoring process. We aim to provide tools that support intellectual development, thus supporting the mastering of how to author as well as what to author.

In this paper, we extend existing approaches for multimedia authoring with the processes of topic identification and discourse structure building. For multimedia, precise discourse rules do not yet exist. We explore the application of the well-defined rules from traditional text-based authoring to multimedia authoring. Enabling authors to carry out their tasks in a single environment allows building a support strategy that uses outputs of earlier stages of the process as inputs for the following stages. The context of the author's work in this flow becomes known to the system, allowing more focused support.

Our approach is explored in the design of a **Semi-Automatic Multimedia Presentation**

authoring Environment (SampLe) [10]. SampLe exploits large media-aware semantic spaces through semantic- and discourse-aware authoring methods. The domain knowledge space used to illustrate SampLe covers the fine arts.

We first outline four top-level authoring phases common for multimedia authoring systems and text-based authoring approaches. We then describe existing multimedia authoring approaches and give an analysis of authoring facets they support. Next, we present our approach to aid all top-level authoring phases on the cognitive level. Finally, we discuss the implementation of our approach in a semi-automatic authoring system SampLe. The paper concludes with an evaluation of the implementation and future work.

## 2 TOP-LEVEL AUTHORING PHASES

We define a presentation as a multimedia-based structured discourse that allows an author to communicate the intended message to the target audience. The precise strategies for building discourse for multimedia presentations have not yet been defined. Thus, we refer to existing guidelines for constructing a coherent text-based discourse [15] and discuss these in terms of constructing a textual presentation:

**Topic.** There needs to be a topic – the theme for the final presentation. There are various ways of determining a topic of the presentation. A topic can be assigned, for example ‘De Stijl art movement’, or the author can be free to select a topic of interest. In any case, the author must be familiar, or become familiar, with the topic domain to be able to proceed with the authoring process.

**Structure.** The author must build a discourse structure for the presentation to ensure a coherent text rather than a collection of unrelated sentences. The structure typically follows a genre. For example, a biography is a genre usually chosen for describing a person’s life. Other examples of commonly used literature genres include essay, article, memoir and monograph. A genre defines an overall structure of the presentation.

**Content.** Next, the author has to write the actual text to express the argumentation of the story. Each genre communicates information with a particular discourse structure by providing a level of detail and different strategies for building up arguments. In general three basic units can be distinguished [15]: introduction, body and conclusion. The detailed discourse structure emerges by filling in basic units with argumentation appropriate for a particular selected genre. For example, an article usually starts with a statement in the introduction unit. The elaboration and arguments are placed in the body. The author then summarizes the argumentation in the conclusion unit. The author must carefully structure arguments within the discourse structure to achieve the coherence of the story.

**Style.** If the author is also the publisher of the story she has to think about the appropriate layout and typography [47]. Well-selected style of the print is important because not many people will read badly formatted material.

These top-level phases are common to all design processes [27, 47] but vary in detail once they are applied to different media presentation forms, such as interaction design [32], web design [24], experience design [38], film and audio production [8], or multimedia production [3, 4]. In the domain of automatic presentation generation these phases also play a key role [7, 19, 29, 35]. In the multimedia presentation authoring process the interplay of these four phases becomes even more complex due to richer means for representing the same concept, different expressiveness levels of media types and more intricate design issues. Design of multimedia presentations should be not only visually appealing but also consistent with the genre and multimedia content.

Before introducing our approaches to support all four stages of the multimedia authoring process, we discuss existing authoring systems and analyse the types of support they provide.

### 3 EXISTING SYSTEMS FOR MULTIMEDIA AUTHORIZING

We differentiate between three general classes of support systems: systems that support strict *manual presentation authoring*, systems that support mainly *automatic presentation generation* with limited user influence and *semi-automatic authoring environments* that provide semantic-based support during at least one of the phases described in Section 2.

#### 3.1 Manually crafted authoring

Manual presentation authoring is currently supported by a number of commercial environments. Microsoft's PowerPoint [22] is used for building linear slide shows. Macromedia's Director/Shockwave, Flash [20] and Oratrix' GriNS [14] enable animated and interactive presentations mainly on the web. Macromedia's Dreamweaver [20] and Microsoft's FrontPage [21] are designed for supporting the authoring of web pages.<sup>1</sup> Each of these applications tries to ease the authoring process by introducing authoring metaphors. Macromedia Director, for example, models the authoring process on a stage production where media items form the cast that can perform certain actions. It also provides a set of stylistic means for designing a final look and feel of the authoring process. It is left to the author, though, to figure out how these metaphors help in the creation of her presentation structure. None of these tools assist in clarifying what the provided structures mean in a larger presentation context, nor do they support any relations between established structures and potential content. It is assumed that the author not only knows about the discourse structure to be established but also how the provided presentation functionalities can be integrated into this discourse. With respect to content collection and meaningful integration of content into a discourse structure these tools do not offer any support whatsoever.

#### 3.2 Automatic authoring

The automatic generation of multimedia presentations has been a focus of multimedia research for over a decade. Research into the automated generation of multimedia presentations has resulted in a number of knowledge-based systems that are able to generate multimedia presentations with minimal or no human intervention. Some of these systems focus on innovative presentation techniques [2, 6, 7, 18, 45] that facilitate the synthesis of multimedia documents and plan how to present this material to various users. These approaches are insightful as they model the authoring process from a planning point of view but they are not sufficient for the authoring processes described in Section 2 as their top-down planning approach is too limited for scenarios where neither the individual user requirements nor the requested material can be predicted in advance.

More recently, there are other approaches that apply semantic-web technology to overcome these restrictive planning problems, such as Cuypers [29], Artequakt [19], DISC[11] and Topia [35]. These approaches are restrictive with respect to our aims because they allow author involvement only at the beginning of an otherwise fully automatic process. An author has limited control over the structure of a presentation, the style and the exact content. The search engine of an automatic system is fully responsible for content supply, while mechanisms for discourse generation try to arrange the collected content into a coherent presentation structure. The complexity and richness of the presentation structure can vary from a simple document structure in Topia to more genre-oriented template presentation structures as supported by DISC. DISC offers flexible rules that evolve an initial template into an iteratively created presentation structure.

---

<sup>1</sup> Note: we address here only those tools that are mainly designed for the creation of multimedia presentations. Tools that provide media for such authoring tools, such as non-linear video editing systems (FAST 601, Softimage DS or the MAD system [3]), image editing tools (Photoshop, Illustrator, GIMP, or Maya), audio systems such as Cubase VST are not covered.

### 3.3 Semi-automatic authoring systems

Established work closest to our work has been carried out by Bailey [4], Gibbins [12] and Hunter [16].

The work by Bailey et al. supports the early stages in the design process, when the user struggles to explore and compare ideas. The DEMAIS approach is based on the notion of an informal interactive storyboard that addresses the needs of skilled designers to develop interactive multimedia applications first in a sketchy manner. Though this work covers a range of user types and task approaches it provides little help for the range of users that require a more formal work environment where the authoring process leads eventually to the final presentation. The major shortcoming of DEMAIS with respect to our work is that the system has no notion of the author's activities.

Related to this approach is the work by Barry [5]. She suggests a set of tools to support documentary videography, such as a script network populated by video clips, where the network displays the position of a clip in a collection of events related to the subject of the documentary; a common sense annotation for each video clip, providing an expanded context for a clip; a suggestion prompt that delivers a direct suggestion for the next shot taken; a display of story structures related to the documentary subject that can be used for video organization into story threads. It is in particular the suggestion prompt that strikes us as potentially useful for the work described in this paper.

The work of Gibbins et al. is of interest because mSpace addresses the mapping of user-determined interaction onto a high dimensional space represented by an ontology. The aim of mSpace is to explore a multidimensional space by extracting meaningful views on domain aspects interesting for a user. For a task of building a final presentation an exploration level is insufficient. An author should be able to collect and structure the material. Moreover, a greater variety of media types need to be handled than those supported by mSpace (text and images).

There are also approaches to collaborative authoring of media materials in the eScience domain. The FUSION system [16] investigates various approaches to facilitating data exploration by allowing user involvement in a directed browsing process. The system also supports the process of multimedia presentation building by allowing a user to define a sequencing of a particular parameter in numeric data and different presentation formats and modes. Leaving a large part of computational, scaling and representational work to the system facilitates analysis of the scientific data. FUSION also aids knowledge discovery by presenting semantically associated information visually. This discovery comes only after the presentation is built, so that during the presentation building process a user has to rely on own expertise.

### 3.4 Evaluation of authoring systems

The majority of the systems for manual production either provide a structural approach, a methodology or a language to put already assimilated knowledge in to a certain format. None of the discussed systems actually supports a process of knowledge assimilation. The burden of getting an idea about a presentation topic, collecting the material, structuring and presenting the final result lies solely on the author.

In automatic authoring environments, an idea about topic and genre has to come from the author while material is selected and structured automatically. Only Topia allows the importance of one or another domain concept to be emphasized enabling the adaptation of the presentation structure according to a user's preferences.

Most existing semi-automatic authoring systems focus on one aspect of the authoring process. The DEMAIS system supports an early stage of the design process by providing a means of efficiently capturing designers' ideas. mSpace mainly supports exploration and provides views on the domain which are tailored to a user needs. FUSION addresses the complexity of analysis of heterogeneous data and mixed-media objects. It facilitates exploration,

presentation and analysis of the *final* presentations. None of these systems support the complete authoring process from the initial exploration of the domain to the final presentation production. The issue about building extensive discourse structures for multimedia presentations is also not investigated.

#### 4 AUTHORIZING PHASES SUPPORT

We develop approaches for supporting authors during all four phases of the presentation building process. In this section we present the type of support we propose for each of the top-level authoring phases presented in Section 2.

**Topic.** In order to build a presentation about a particular topic the author has to be familiar with the topic domain. To aid authors at this stage we propose an approach that facilitates the exploration and knowledge acquisition processes. A choice of the starting point for exploration is simplified by presenting the structured categories of available materials. This allows placing a topic of interest in the knowledge space. At the beginning of an exploration path on acquiring information about a topic, the author is provided with concise information about this topic. Continuing exploration on the same topic should lead to providing more detailed information. This allows gradually increasing the author's domain knowledge and facilitating its acquisition by providing the appropriate level of domain information complexity at each exploration step.

**Structure.** The process of building a discourse structure for the presentation requires support for going from a more abstract to a more specific level. At the abstract level, the author can choose a genre for the presentation that matches an author's ideas about the message that needs to be conveyed. For example, an essay can suit the author if she wants to describe the topic from a personal viewpoint. Then a specific discourse structure should be built that corresponds to the selected genre. The author should be able to choose from a number of discourse structures for the genre. This provides an opportunity to get an idea about possible discourse flows and gives a starting point for developing more complex and rich discourse structures.

**Content.** In multimedia authoring, media items, rather than only text, form the content of the presentation. During this phase the choice of media items and their placement within the envisioned discourse structure needs to be supported. The relevance of the media material should be determined not only by matching the search request with the topic but also by relating it to the specific context of the authoring process. The authoring *process context* represents the current position of the author in the authoring space. The previous two phases determine the current context. Knowing the *topic* and the *specific discourse structure* the author is working on, we can identify what media items can be used in what section of the presentation with regard to their *content* and their role *within the given discourse*.

**Style.** The author has to choose the appropriate design for the final presentation that matches the overall theme. The design elements include a choice of the final document format (e.g. XHTML [49], SMIL [39]) for representing media material, a colour schema, layout and typography issues. Style aspects of multimedia presentation are quite complex due to material diversity. Section 3 shows that solutions can be found in the automatic presentation generation area. Our system's output contains a discourse structure filled with an ordered sequence of media items. This output can be fed into an automatic presentation generation engine, such as Cuypers [29], to produce a final presentation.

#### 5 SAMPLE ARCHITECTURE AND FUNCTIONALITY

This section demonstrates the applicability of the proposed approaches by introducing the architecture and implementation of an experimental system SampLe (Semi-Automatic Multimedia Presentation generation Environment)<sup>2</sup>. In this paper we focus on processes and

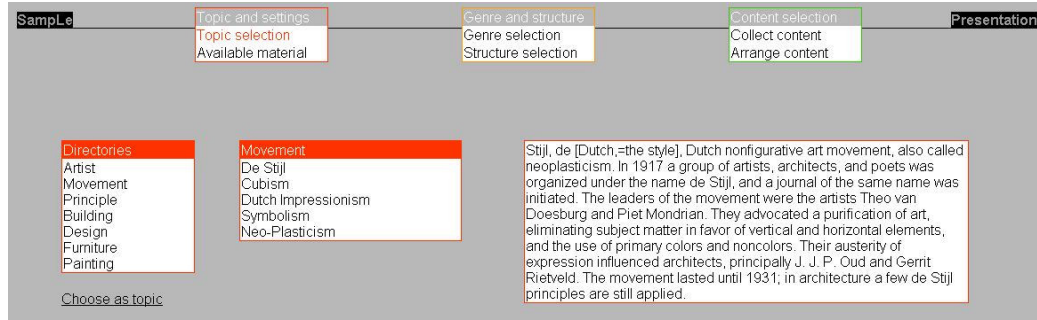
---

<sup>2</sup>See also our test page at <http://www.cwi.nl/~media/projects/CHIME/demos.html>

resulting internal data structures of the system. The interface examples in this section serve as plausible interaction design interfaces for presenting system functionality. The initial interaction design work for the SampLe system is described in [46].

The three main components of the SampLe system are the Exploration, Discourse structure and Material collection support mechanisms. The functionality of each of these three components depends on the context of the authoring process, thus, all actions of the author are maintained by the system’s session management component. For example, actions such as topic selection or setting genre preferences influence the entire authoring process and are thus maintained during the entire session.

An author communicates with the system via a web-based user interface. The user interface reflects the top-level phases of the authoring process. Each phase is represented by a colour-coded menu, as shown at the top in Figure 1. These phases correspond to the three support mechanisms.



**Figure 1.** SampLe topic selection phase and the overall view on the interface

The domain-dependent elements of the web-interface, for example the two left menus ‘Directories’ and ‘Movement’ in Figure 1, are generated from domain-specific meta-data that is attached during system configuration. This enables switching to a different domain without re-implementation.

In the following sections we discuss the implementation of our approach in three support mechanisms. We begin with the description of meta-data structures, as they form the foundation for the system support strategies.

### 5.1 SampLe meta-data structures

The approaches described in Section 4 assume that the system is aware about content of media items and their potential role within a discourse. In order to enable this awareness, media items have to be described and these descriptions have to be made accessible to the authoring environment. Following the approaches taken by the Semantic Web community we use RDF/S-based meta-data structures [33] to achieve this accessibility. The general meta-data framework consists of domain meta-data, discourse meta-data and media type meta-data together with annotation templates:

**Domain meta-data.** Domain meta-data specifies domain concepts (e.g. Artist, Artefact) and relationships between them (e.g. Artist painted Artefact). One or more of these concepts are assigned to a media item to describe its content.

**Discourse meta-data.** Discourse meta-data structures are introduced with a number of purposes. First, they allow content based browsing during the exploration phase based on increasing content density and complexity. Second, they allow deriving relevance of the material in terms of the current discourse structure. The discourse annotations describe what type of discourse function a media item represents with regard to the type of information it expresses about its content. We distinguish the following *discourse functions*: introduction, summary, note, quote, elaboration, description, definition, conclusion, example and comparison. We do not claim that this list is complete or sufficient for describing media

content with regard to its discourse function. We are using this set of concepts to verify our approach.

We share a number of the discourse concepts with Rhetorical Structure Theory [34]. RST was built to provide an explanation for the coherence of text. The authors of RST pursued the goal of describing a narrative (analysis) rather than building it (synthesis). Consequently, the descriptions refer to particular rhetoric figures functioning at the lower level of the narrative structure than the one we are interested in. The authors of multimedia presentations in our system manipulate existing media items. Thus, our system needs descriptions of the items to facilitate their selection and their structuring within higher-level structures.

**Media-type meta-data.** This type of meta-data includes media-specific descriptions of the material such as production properties of a painting or physical characteristics of audio.

**Annotation template.** An annotation template relates different concepts from various meta-data structures to a media item and assigns additional characteristics to it, such as complexity or level of details, the source of material etc.

Our repository contains meta-data structures and media items annotated with concepts from these structures using the annotation template. The next subsections present how the various forms of annotations are exploited during the authoring process.

## 5.2 Exploration and topic selection

**Conceptual exploration.** At the beginning of the exploration process the author has to define a starting point of exploration. The system provides the structured overview of material directories available in the media repository. The directories are inferred from the major classes of domain meta-data. This allows the author to choose a direction of exploration by selecting one of the available directories. Each directory name is a hyperlink that opens a corresponding menu. In Figure 1 the choice of ‘Movement’ in ‘Directories’ causes a menu with different ‘Movements’ to appear. It is the result of a specification traverse of the domain meta-data structure. The choice of a particular movement ‘De Stijl’ opens a content area with information about this movement.

**Content exploration.** The next stage of the exploration process is content exploration, which allows the author to become familiar with the current topic of interest. The exploration support mechanism aims to facilitate this step. For that we apply the strategy of content based browsing with increasing content complexity.

The author establishes the process context by selecting the current topic of interest. During the exploration the first time the author requests information about a particular topic (De Stijl), the system provides her with some introductory materials. Each time the author requests more information about the topic the support mechanism changes the status of current process context of the author and offers more expanded content until the author decides to select the current topic or switches to a different direction of exploration. The key element of our exploration mechanism is that the system combines knowledge about the context with domain and discourse knowledge to solve the support request.

To enable transformation of the author’s request for more information about the topic into the appropriate query to the repository, the system uses domain and discourse knowledge it incorporates about the media material. Discourse knowledge is distributed between discourse annotations of media items and exploration support mechanism<sup>3</sup>. The exploration support mechanism contains knowledge about applicability of various discourse functions in various process contexts. In the context when the topic has to be introduced for the first time to the author, the condensed information about this topic is required. The support mechanism identifies that the discourse function ‘summary’ corresponds to the required type of information. The mechanism constructs a query to the repository, which contains the request for media items annotated with the identified discourse function (summary) and domain

---

<sup>3</sup> During the content selection phase the discourse knowledge of material collection support mechanism is used.

concept (De Stijl). In cases where a large number of media items match the required annotations, the distinction is made based on the level of detail and complexity of the particular media item. An item with appropriate levels is selected and presented to the author. This process repeats for every new step within the exploration of the current topic. The changes of the current process context affect system's decisions about the appropriate discourse functions.

Gradually increasing the complexity levels of media material allows the author to acquire knowledge during exploration in a structured way. At the same time, if the current topic represented does not interest the author, the short and condensed introduction allows the immediate elimination of the topic from consideration. After the topic is selected the author can proceed to the next phase of establishing a genre and a presentation discourse structure.

### 5.3 Discourse structure building

**Genre selection.** In the process of choosing an appropriate genre the author is presented with different genres and their descriptions. At this stage it is the responsibility of the author to make the selection. This step is shown in Figure 2.

**Discourse structure building.** After the genre is selected, for example the Essay, the system presents various discourse structures that correspond to this genre (Figure 3). The author can select one of the proposed structures as the discourse structure for her presentation.

Conventional agreements on how a particular genre should be structured are incorporated into templates within SampLe. Deciding on particular structures for different genres is beyond the focus of our research. Thus, the presented discourse structures for genres should be considered as examples.

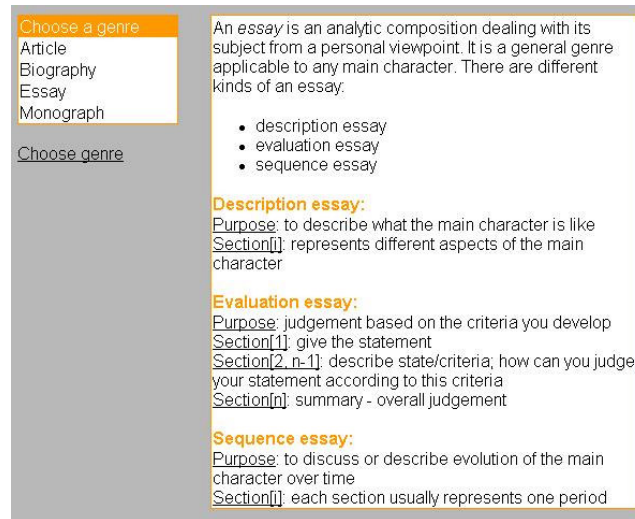


Figure 2. Genre selection

Genre templates are distinguished according to the class of the topic selected. In our example De Stijl belongs to the class Movement. This allows constructing templates with more expanded structure and more suitable for a particular case, since an essay about a movement would differ from an essay about an artist. Knowing the current process context (topic=De Stijl, genre=essay) the discourse structure support mechanism extends a general essay structure for an essay about a movement (description, evaluation and sequence essays) to a discourse structures appropriate for representing De Stijl. For this the mechanism uses domain knowledge. It determines sections in the discourse structure that can be extended based on the structure of the domain. It parses the domain meta-data structure and identifies concepts matching the identified section. The mechanism queries the repository for the instances of the concepts and includes the corresponding sections into a discourse structure. For example, in

the description essay on the left in Figure 3, the section of the general description essay template about the movement “3. Movement members and their works” will be extended into “3. De Stijl members and their works” plus sections 3.1, 3.2 and 3.3 about three major artists of the movement.

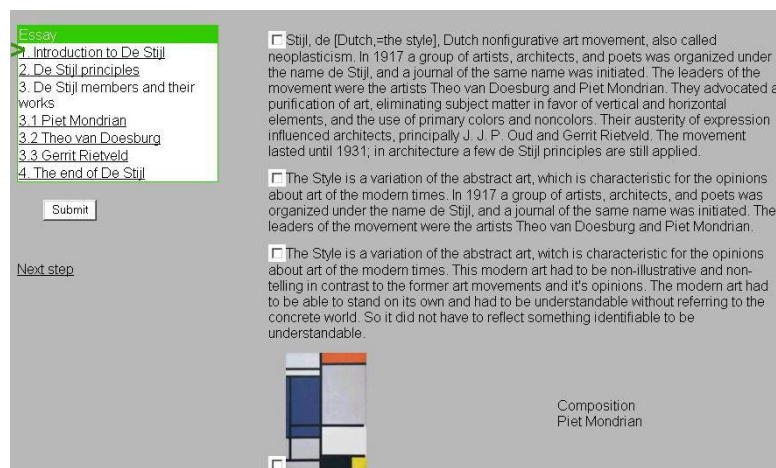
Choose a structure		
--> Essay 1: Description	--> Essay 2: Evaluation	--> Essay 3: Sequence
1. Introduction to De Stijl	1. Statement: De Stijl reflected the needs of its time	1. Introduction to De Stijl
2. De Stijl principles	2. Times when De Stijl appeared: what was needed?	2. First years
3. De Stijl members and their works	3. De Stijl principles	3. De Stijl in blossom: major activities
3.1 Piet Mondrian	4. De Stijl activities	4. The end of De Stijl: final transition
3.2 Theo van Doesburg	5. Summary: What was achieved?	
3.3 Gerrit Rietveld		
4 The end of De Stijl		

**Figure 3.** Discourse structures for essay variations

To enable this process, SampLe makes use of two levels of discourse structure representation. At one level the discourse structure is represented in a human-readable format to be able to present a meaningful text to an author. At the other level it is encoded for the internal representation. This internal representation contains domain and discourse meta-data concepts that specify what kind of content is appropriate for a particular section in the discourse structure, described in the following subsection.

## 5.4 Material collection

**Material collection process.** Having established the discourse structure for her presentation, the author can start selecting media material. The author is able to see the complete structure (on the left of Figure 4) and to work with one section of the structure at a time. The current process context at this phase consists of the presentation topic (de Stijl), genre (Essay), particular discourse structure selected (Description Essay) and a section the author is working on (Introduction to De Stijl). Based on the context the material collection support mechanism takes a decision about content (De Stijl) and discourse function requirements (introduction) of the content and requests appropriate media items from the repository. Figure 4 presents a potential set of material available and suitable for the introduction part of the presentation. Selecting as many of the choice boxes as required does the selection. The selected set is incorporated into the structure, by clicking on the submit button, presented under the structure box.

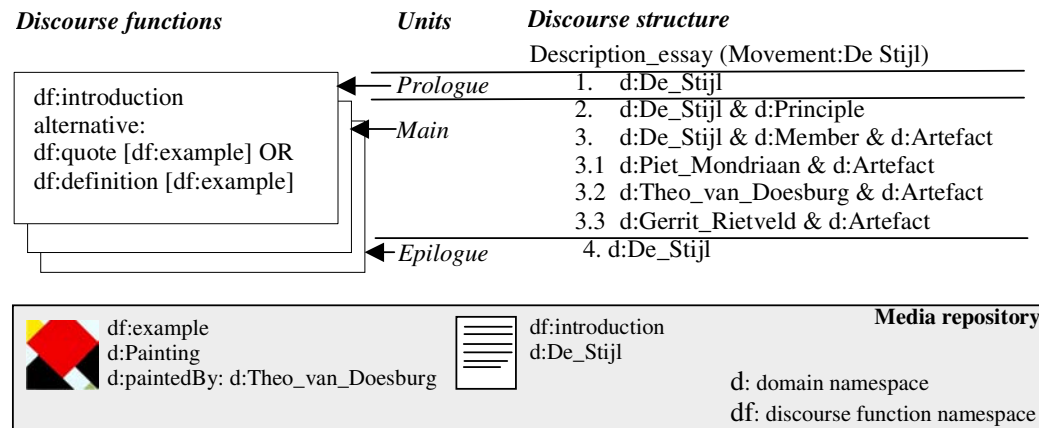


**Figure 4.** Content selection phase

**Internal representation of a discourse structure.** The internal representation of this

discourse structure consists of domain concepts relevant for each section of the structure and specification of the unit the section belongs to. Each discourse structure is conventionally divided into three units: Prologue, Main part and Epilogue. The units serve as an aid for inferring appropriate discourse functions of the material for each section.

The material collection support mechanism contains knowledge about structural aggregation of discourse functions into a discourse structure for a particular genre. Each genre has a different argumentation flow. For example, an article could require a representation of a statement before providing a complete story and arguments. An essay, in contrast, could require an introduction to the topic, followed by elaboration on the major aspects of the topic and conclusion. At the same time the argumentation flow for a particular genre will be similar for all discourse structures of this genre. The units allow expressing the rules for relating discourse functions to a particular flow of the genre. Internal schematic representation of the discourse structure of a description essay about De Stijl from our example is shown in Figure 5.



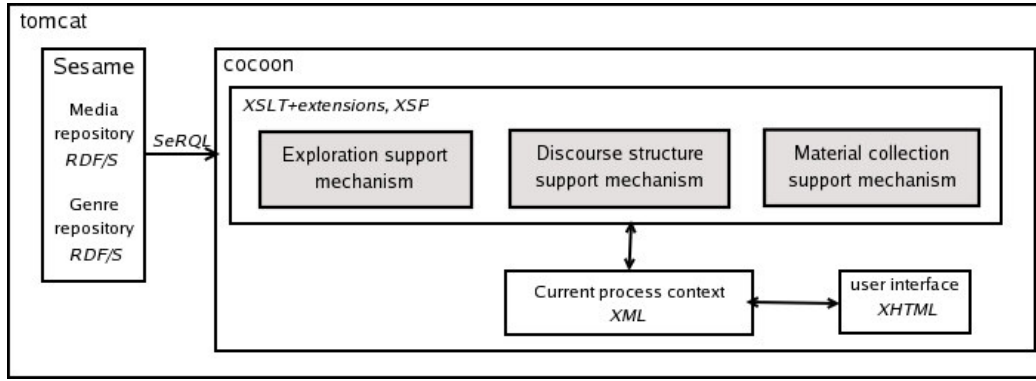
**Figure 5.** Relation between discourse functions, units and a discourse structure

The combination of domain descriptions of the sections and discourse function concepts allows selection of the media material from the repository for that section. All the media items retrieved could be used in the specified section. The author has to make a selection of items that suit best her ideas. In case no media items in the repository fit the specified domain and discourse descriptions, material selection support mechanism will use alternatives for discourse functions.

After the collection process for the whole presentation is completed, the author can make final ordering arrangements of media items within each section to achieve the coherence of the presentation. The output of this stage is an input for the final presentation generation module that takes decisions about style and an output format.

## 5.5 Implementation

Authors communicate with SampLe using off-the-shelf web browser. All application-specific functionality is realized on the server side, using standard components. Figure 6 illustrates the system's basic architecture. The three main components discussed in the previous subsections are currently implemented using XSLT [48] and XSP [50]. The RDF repositories are accessed by using an XSLT extension [28] that allows an RDF query language (in this case SeRQL [36]) to be used in XSLT context. These three components are executed by the Apache Cocoon XML framework [9]. All RDF(S) is modelled using Protégé [31] and is stored using the Sesame open source RDF database [37]. Both Cocoon and Sesame are Java Servlets run by Tomcat [40].



**Figure 6.** SampLe architecture

## 6 EVALUATION AND CONCLUSION

The described mechanisms, which aim to provide tools to assist authors during the four top-level phases of the authoring process, demonstrate the feasibility of our approach towards context aware intelligent guidance of authors during the multimedia authoring process. The experimental SampLe engine is in the position to

- support the exploration of a large media-based repository for the acquisition of knowledge in a structural way by allowing content-based browsing with increasing complexity of the material;
- provide assistance in the understanding of genre structures to facilitate the process of building the presentation discourse;
- suggest a set of relevant media items for a particular section of the discourse structure and thus ease the time-consuming process of investigating large amounts of material.

The main direction of our current work aims to establish more flexible mechanisms for manipulating discourse structures. Currently an author can choose one of the discourse structures provided by the system. We investigate ways to give an author more control over this phase by allowing the alteration of discourse structures. This includes the development of an evaluation mechanism that will ensure genre consistency during alteration process. In addition we work on mechanisms that integrate created presentations into the repository the material was collected from. In this way we increase the ability of the authoring system to provide a larger selection of genre structure types.

The current implementation suggests a sequential order of authoring phases (see section 5.2. - 5.4). Our aim, however, is to allow authors to follow their own authoring strategies resulting in a great variety of workflows. The overall structure of SampLe that modularises the top-level phases as well as the tasks allocated to them already facilitates flexible workflows. The system knows at any time at which authoring stage the author is, what type of task she is performing and what type of content the author wishes to investigate or manipulate. In the ongoing development of the SampLe environment the provision of a less restricted means of moving between authoring stages is a key issue. An important aspect is a better integration between the discourse structure building and material collection phases.

As our environment applies authoring processes on various media we also have to improve the description of the media in the context of their authoring ability. At the moment SampLe has a restricted means to identify, for example, the potential role a media item can play in a discourse.

All solutions to the described problems require that the engine has access to high quality, though not necessarily excessive, annotations of the media units. We are planning to convert the existing meta-data into OWL [44] format, which is more expressive than RDF. We want

to connect existing meta-data to vocabularies for the cultural domain, such as the AAT [1], ULAN [42] or Iconclass [17], to be able to reuse material within different applications.

The approach described in this article is a step towards a context aware semantic-based authoring support. We believe that our view on semantic mediated multimedia authoring provides an essential foundation for environments in which complex domain and discourse information can be studied, discussed, commented, published and demonstrated – thus preserving and developing information artefacts in an intelligent communal way.

## ACKNOWLEDGMENTS

This research is funded by the Dutch National NWO ToKeN2000 CHIME project. The authors wish to thank in particular their colleagues Lynda Hardman, Lloyd Rutledge, Jacco van Ossenbruggen, Stefano Bocconi, Joost Geurts and Katharina Schwarz for useful discussions during the development of this work. The authors would also like to thank Mark van Assen and Guus Schreiber for many useful insights into ontology issues. Finally, we wish to thank Jana Werner who designed the first prototype of the SampLe interface and provided most valuable insights into interaction design.

## REFERENCES

- [1] AAT [http://www.getty.edu/research/conducting\\_research/vocabularies/aat/](http://www.getty.edu/research/conducting_research/vocabularies/aat/)
- [2] Andre, E., Muller, J., and Rist, T. Presenting Through Performing: On the Use of Multiple Lifelike Characters in Knowledge-Based Presentation Systems. In: *Proc. of the Second International Conference on Intelligent User Interfaces (IUI 2000)*, pages 1-8, New Orleans, LA USA, 2000.
- [3] Baecker, R., Rosenthal A., Friedlander N., Smith, E., and Cohen. A. A Multimedia System for Authoring Motion Pictures. In *Proceedings of the 4<sup>th</sup> ACM International Conference on Multimedia*, pp. 31-42, Boston, MA, November 18-22, 1996.
- [4] Bailey, B. P., Konstan, J. A., and Carlis, J.V. DEMAIS: Designing Multimedia Authoring Applications with Interactive Storyboards. In *Proceedings of the 9<sup>th</sup> ACM International Conference on Multimedia*, pp. 241-250, Ottawa, Canada, Sept.30 – Oct. 5, 2001.
- [5] Barry, B. The Mindful Camera: Common Sense for Documentary Videography. In *Proceedings of the 11<sup>th</sup> ACM International Conference on Multimedia*, pp. 648-649, Berkeley, California, USA, November 4-6, 2003.
- [6] Boll, S., Klass, W., and Wandel, J. A Cross-Media Adaptation Strategy for Multimedia Presentations. In *ACM Multimedia '99 Proceedings*, pages 37-46, Orlando, Florida, October 30 - November 5, 1999. ACM, Addison Wesley Longman.
- [7] Bordegoni, M., Faconti, G., Maybury, M., Rist, T., Ruggieri, S., Trahanias, P., and Wilson, M. A Standard Reference Model for Intelligent Multimedia Presentation Systems. *Computer Standards & Interfaces* 18, 6-7 (December 1997), 477–496.
- [8] Bordwell, D., Thompson, K. *Film Art*. McGraw-Hill College; 6th edition, 2000.
- [9] Cocoon <http://cocoon.apache.org/>
- [10] Falkovych, K., Nack, F., van Ossenbruggen, J., Rutledge, L. SampLe: Towards a Framework for System-supported Multimedia Authoring. In *Proceedings of the 10<sup>th</sup> International Multimedia Modelling Conference*, p.362, January 5-7, 2004.
- [11] Geurts, J., Bocconi, S., van Ossenbruggen, J., Hardman, L. Towards Ontology-driven Discourse: From Semantic Graph to Multimedia Presentations. In *Proceedings of the 2<sup>nd</sup> International Semantic Web Conference*, pp. 597-612, October 20-23, Sanibel Island, Florida, USA, 2003.
- [12] Gibbins, N., Harris, S., Dix, A., and schraefel, m.c. Applying mSpace Interfaces to the Semantic Web. <http://eprints.ecs.soton.ac.uk/10027/01/mspace-model-workingpaper.pdf>
- [13] Google <http://www.google.com>
- [14] GRiNS <http://www.oratrix.com/GRiNS/>

- [15] Hacker, D. "A Writer's Reference", 5<sup>th</sup> edition, Bedford/St. Martin's, Boston, MA, 2003. ISBN 0-312-39767-4.
- [16] Hunter, J., Falkovych, K., Little, S. Next Generation Search Interfaces – Interactive Data Exploration and Hypothesis Testing. In *Proceedings of the 8<sup>th</sup> European Digital Library Conference*, pp.86-98, September 2004.
- [17] Iconclass <http://www.iconclass.nl/>
- [18] Kamps, T. Diagram Design : A Constructive Theory. Springer Verlag, 1999.
- [19] Kim, S., et. al. Artequakt: Generating Tailored Biographies with Automatically Annotated Fragments from the Web. In *Proceedings of the Workshop on the Semantic Authoring, Annotation & Knowledge Markup* in conjunction with the Fifteen European Conference on Artificial Intelligence, France, 2002.
- [20] Macromedia <http://www.macromedia.com/>
- [21] Microsoft FrontPage <http://www.microsoft.com/frontpage>
- [22] Microsoft PowerPoint <http://office.microsoft.com/en-us/FX010857971033.aspx>
- [23] Nardi, B. A. (Ed.) 1997. Context and Consciousness, 2nd printing, MIT Press,
- [24] Nielsen, J., and Tahir, M. (2001). Homepage Usability: 50 Websites Deconstructed. New Riders Publishing, Indianapolis.
- [25] Norman, D. A. (1989). The psychology of everyday things. New York: Basic Book.
- [26] Norman, D. A. (1993). Things that make us smart. New York: Addison-Wesley.
- [27] Oldah M. Creativity for Graphic Designers: A Real-World Guide to Idea Generation. North Light Books, 1995.
- [28] van Ossenbruggen, J., Hardman, L. and Rutledge, L.. Combining RDF Semantics with XML Document Transformations. *International Journal of Web Engineering and Technology*, 1(4), 2005, to be published, based on <http://ftp.cwi.nl/CWIReports/INS/INS-E0303.pdf>.
- [29] van Ossenbruggen, J., Geurts, J., Cornelissen, F., Rutledge, L., Hardman, L. Towards Second and Third Generation Web-Based Multimedia. In *Proceedings of the 10<sup>th</sup> International World Wide Web Conference*, pp. 479-488, ACM Press, Hong Kong, May 1-5, 2001.
- [30] Papert, S. (1991). Situating Constructionism. In Edit Harel (Ed.), Constructionism: research reports and essays, 1985-1990. Norwood, NJ: Alex Publishing Corp.
- [31] Protégé <http://protege.stanford.edu/>
- [32] Raskin, J.. The Humane Interface: New directions for designing interactive systems. Addison Wesley, 2000.
- [33] RDFS <http://www.w3.org/TR/rdf-schema/>
- [34] Rhetorical Structure Theory. <http://www.wagsoft.com/RSTTool/RSTDefs.htm>
- [35] Rutledge, L., Alberink, M., Brussee, R., Pokraev, S., van Dieten, W., and Veenstra, M. Finding the Story – Broader Applicability of Semantics and Discourse for Hypermedia Generation. In *Proceedings of the 14th ACM Conference on Hypertext and Hypermedia*, pp.67-76, August 23-27, Nottingham, UK, 2003.
- [36] SeRQL <http://www.openrdf.org/doc/users/ch06.html>
- [37] Sesame <http://sesame.aduna.biz/sesame/>
- [38] Shedroff, N. Experience design I, New Riders Publishing, 2001.
- [39] SMIL 2.0 <http://w3c.org/AudioVideo/>
- [40] Tomcat <http://jakarta.apache.org/tomcat/>
- [41] Tufte, E. R. Envisioning information. Graphic Press, Connecticut, 1990
- [42] Ulan [http://www.getty.edu/research/conducting\\_research/vocabularies/ulan/](http://www.getty.edu/research/conducting_research/vocabularies/ulan/)

- [43] Vygotsky and the Social Formation of Mind by James Wertsch, Harvard University Press 1985.
- [44] Web Ontology Language OWL <http://www.w3.org/TR/owl-features/>
- [45] Weitzman, L. and Wittenburg, K. Automatic presentation of multimedia documents using relational grammars. In: *Proceedings of the second ACM international conference on Multimedia '94*, pp. 443-451, San Francisco, October 15 - 20, 1994.
- [46] Werner, J., Investigation of Methods for User Adapted Visualization of Information in a Hypermedia Generation System. Master thesis, University of Applied Science, Darmstadt, Germany, 2004.
- [47] Williams, R. Non-designer's Design Book. Peachpit Press, Berkeley, California, 1994. ISBN 1-56609-159-4.
- [48] XSLT <http://www.w3.org/TR/xslt>
- [49] XHTML <http://w3c.org/MarkUp/>
- [50] XSP <http://cocoon.apache.org/2.1/userdocs/xsp/>