*Published in P. Isaias, N. Karmakar, editors, Proceedings of IADIS International Conference WWW/Internet 2004, pages 715--722, Madrid, Spain, October 2004. IADIS Press.*

# A METAMODEL FOR WEB PAGE DESIGN

Olivier Gerbé
*HEC Montréal*
*3000, chemin de la Côte-Sainte-Catherine*
*Montréal (Québec) Canada H3T 2A7*
*Olivier.Gerbe@hec.ca*
Bernard Drolet
*HEC Montréal*
*3000, chemin de la Côte-Sainte-Catherine*
*Montréal (Québec) Canada H3T 2A7*
*Bernard.Drolet@hec.ca*
Brigitte Kerhervé
*Université du Québec à Montréal*
*CP 8888, succursale Centre Ville*
*Montréal (Québec), H3C 3P8*
*Kerherve.brigitte@uqam.ca*

**ABSTRACT**

Since the beginning of the 90's, the phenomenal growth of Internet was supported by the power given to the authors to carry out and distribute documents to a broad audience. These documents, presented in Web pages and grouped into coherent presentations, are usually stored in individual files. The management of their contents then becomes a problem.
In this paper, we propose a metamodel for web pages design allowing to describe the content, structure and layout of Web presentations. This description is done through a Web page model that can be stored in a database and then easily accessed and reused. We then integrate in the same database the Web page models as well as the data composing the Web presentations. We are then able to guarantee the consistency of the information presented and to facilitate the management of contents.

**KEYWORDS**

Metamodel, Web page design, Web modeling

## 1. INTRODUCTION

Embryonic in 1990, but with nearly 110 millions servers and 1.6 billion Web documents in 2001, the extraordinary growth of the World Wide Web was supported by the capacity given to the authors to easily distribute documents to an international audience at a reasonable cost (Bosak 97, Gaedke 99). Specifically designed for that purpose, HTML (Hypertext Markup Language) was proposed by Tim Berners-Lee at the beginning of the nineties. Derived from SGML (Standard Generalized Markup Language), HTML allows to structure texts for visual presentation. The different tags included in the initial model were introduced to identify the text structure (title, heading, paragraph, etc.) and its layout (Ducharme 99, Manola98). HTML allows describing an intra document structure but also an inter document structure through the use of hyperlinks.
A Web document is then composed of several HTML files and its presentation can be updated by adjusting manually the contents of the corresponding files (Hungjun 98). Based on a simple technology, HTML does not allow the abstraction for high level concepts (Gaedke 99). Inside an HTML document, we cannot define modules, functions or data structures which would allow a certain abstraction of the content. Presentations based on HTML documents frequently encounter the limits of HTML infrastructure. Extensibility, structure and validation are not supported by HTML (Bosak97). The main reason is that the most part of HTML tags

have been designed to describe the layout of the document, rather than its structure or content. They do not allow to build a semantic structure of the information units (Manola 98).

To bypass this problem and to extend the possibilities of HTML, several researchers have proposed models which not only focus on the layout of a document, but rather focus on its contents. The principal efforts currently go towards XML (eXtensible Markup Language), a metalanguage also based on SGML, enabling users to define their own customized tags. XML can also be seen as a metamodel since it allows to define the content model through DTD (Document Type Definition) or more recently through XML Schema. In order to use XML, it is necessary to define an ontology in a DTD. An ontology is a representation of knowledge for a specific domain; it makes data understandable not only by humans but also by computers (Liechti 99). Xforms, SVG, MathML, CML and XHTML are examples of ontologies, created to meet specific needs. Similar to HTML, XML has been initially designed to store information in files and the problem of XML files management remains the same (Coda 98, Ceri 2000).

Due to the popularity of Internet, the need to organize great quantities of information in the form of hypertext is increasing. In several cases, the management of Web documents would benefit from data base management techniques (Atzeni 9, Hungjun 98). The classic file systems technology is no more appropriate to manage the organization, storage and access to great amount of information on the Web (Hungjun 98).

The objective of this paper is to propose a metamodel which can be used to model Web pages and to store these models in a database. This metamodel allows the designer to structure the various elements included in a presentation in a logical way and to re-use existing information. The rest of the paper is organized as follows: Section 2 presents the literature review concerning Web presentation design. Section 3 presents the context of our research project and explains in details the metamodel we propose. Finally, Section 4 concludes and points some future work.

## 2. LITERATURE REVIEW

Web presentations originated from hypertext and hypermedia documents. In a paper published in 1988, Akscyn and al. (Akscyn 88) listed the characteristics generally recognized for such a document:

- Information is separated into fragments containing various types of information (text, image, and video).
- Information is presented using several windows.
- Units of information are inter-connected by relationships (links) and the user can navigate between them by selecting the links.

These characteristics are found in Web presentations: Web pages replace fragments of information and hyperlinks the various types of relationships. Using these characteristics, a Web presentation can be described as a coherent collection of Web pages inter-connected by links (Mecca 98). These links cannot only point to elements included in the presentation, but also to any other document integrated into the Web, anywhere in the world (Gloor 97).

Web presentations can be categorized into two families: Web sites and Web applications. Conallen (Conallen 99) differentiates a site from an application by its use: a Web site will be used to disseminate information and its contents is relatively static, whereas an application implies a business process. For example, the site www.hec.ca offers information on the School and the various programs. Www.amazon.com is an example of a Web application: when a user places an order using the site, the accountable system will keep trace of it, a process of delivery is set up and stocks are modified. In this paper we will use the term presentation or Web presentation to describe the concept of Web site.

In a study on tools and approaches for the design of Web presentations, Fraternali (Fraternali 99) describes the characteristics of presentations developed for the Internet in addition to the normal characteristics of an application (safety, extensibility, availability and usability). A Web presentation should support both structured data (i.e. database records) and not-structured data (i.e. multi-media). A presentation should also support user navigation inside the presentation, offer a high quality design and allow content adaptation to the user needs (personalization). Fraternali writes that models and tools suggested for the Web presentation development should be able to describe the following three major dimensions:

- Structure: the organization of the information managed by the semantic presentation and relations between data elements;

- Navigation: the possibilities offered to the user to reach information and to navigate in the presentation content
- Presentation: the way the content and navigation are presented to the user

In Kamps 94, Kamps and Reichenberger define a formal semantics allowing the automatic positioning of objects. The principle is to see the positioning of objects like an organizational process. This process positions the objects according to a formal semantics and to structural properties of the presentation components. Working on the same subject, Weitzman and Wittenburg (Weitzman 96) have developed an approach where a relational grammar is used to automatically organize multimedia presentations. This organization includes graphic constraints (esthetics, font, and size), spatial constraints (relative positioning of the objects) and temporal constraints (sequence of presentation). This grammar permits to separate the constraints describing the presentation layout from its contents. Other rules define the final layout for the presentation.

Hay (Hay 2000) proposes a metamodel to describe an application, using the entity-relationship formalism. The author proposes Display Groups to group the elements which are logically connected. In a Display Group, it is possible to group elements referring to attributes coming from different entities.

Ceri and al. describe in (Ceri 2000) the Web Modeling Language (WebML), a XML model to describe Web presentations at the conceptual level. In WebML, the specification of a presentation is done using four orthogonal dimensions and a DTD has been developed for each of these dimensions:

- Structure model: the structure model is used to describe the information which will be displayed by the presentation.
- Hypertext model: the hypertext model is used to define the pages composing the presentation. This model is integrates two sub-models: one for the composition of the pages and one for the defining the navigation through pages and the elements forming them.
- Layout grid: the layout grid defines the layout of the presentation.
- Personalization model: this model describes users or users groups inside the structure model.

This model makes possible to store information specific to the user, such as preferences or the personalization of the site content.

The WebML specifications, stored into XML files, are translated into a concrete language by a code generator. It generates the composition, navigation and presentation of the pages. Using scripts, the generator organizes the abstract references inside the presentation into concrete instructions to extract information.

Matera and al. introduce in (Matera 2003) a model-driven approach to the design of collaborative Web-based applications. Their approach uses WebML as modeling language.

Koch and Kraus propose in (Koch 2003) a common metamodel for the development of Web applications as an extension of the UML metamodel.

Except the metamodel proposed by Hay that is used to define application data entry screens, no author uses the very well-known entity-relationship formalismto describe a presentation.

In this paper we propose a metamodel to design Web presentations based on the entity-relationship formalism used for both specifications: content specification and layout specification.

## 3. THE WEB DESIGN METAMODEL

The metamodel we propose for Web pages design is the kernel of the Web Content Management System (WCMS) we are currently developing at HEC Montreal and used to support teaching. A simplified architecture of WCMS is presented in Figure 1 and shows the links between the different components. The data model describes the data that will be displayed in a Web presentation. The presentation model is used to describe the presentation layout. The presentation model is directly connected to the data model. Presentation models, data model and data are all stored in a database. A generator interprets the presentation model and the data in order to generate the corresponding HTML presentation.

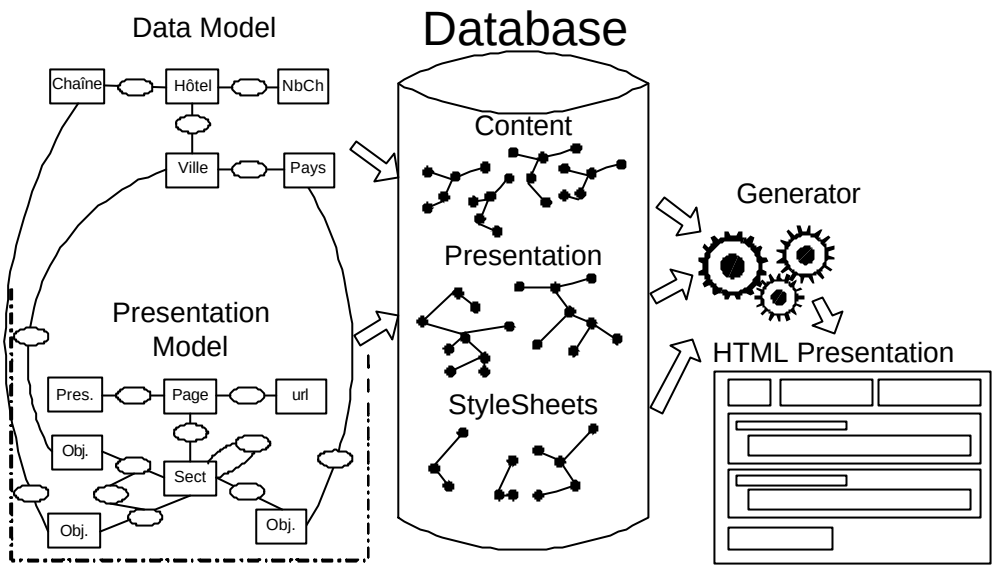


Figure 1. WCMS Architecture

In the current version of WCMS, details related to the visual characteristics of the presentation (dimension, color, character font and other visual characteristics) are managed using style sheets. This solution allows a same presentation to be presented differently.

As we have seen in the previous section, a Web presentation is defined as a coherent collection of Web pages connected by hyperlinks. To describe Web presentations we propose our metamodel which is composed of three parts: Layout, Content and Navigation. In the following sections, we present our metamodel in defining each of these sub-models..

## 3.1 Layout Metamodel

The Layout Metamodel (see Figure 2) introduces the elements the designer uses to structure the Web pages that will be presented to the user. A Presentation contains pages and is characterized by attributes such as URI, title, author and other metadata. The layout of a page is composed of layout objects: section or widget.

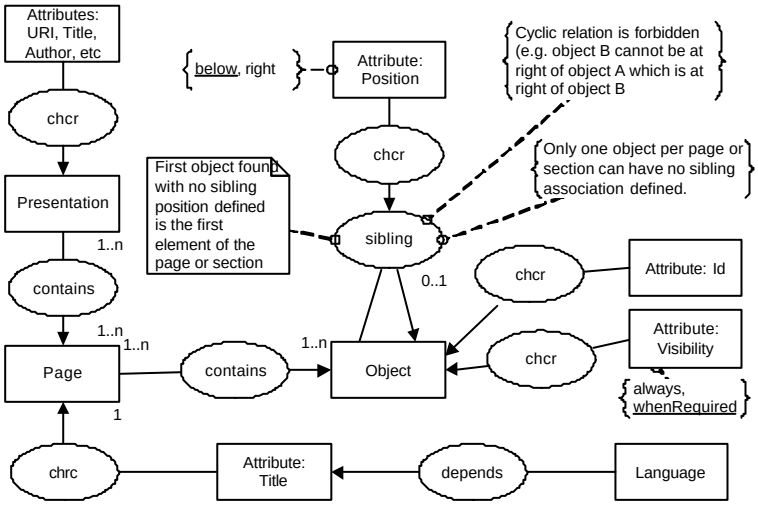


Figure 2. Layout Metamodel

The position of the objects in the page is defined by a sibling relationship between objects. A sibling relationship is characterized by a position attribute that specified if the sibling object is at the right of or below the current object. Constraints have been defined on the sibling relationship as noted in Figure 2.

## 3.2 Content Metamodel

The Content Metamodel (Figure 3) introduces the elements to build a web page and makes the link with the data model. The main element is the object. A layout object is either a section either a widget. A section is a container that semantically groups information units describing a concept. Sections make the link between layout and data. Data are displayed using widgets (text, image, video, sound), each widget may initiate behavior according to a trigger.

A section may be a ForAll section. A ForAll section presents all the occurrences of the concept described by the section. A section may be a ForEach section. A ForEach section presents all the occurrences linked by a relationship to the occurrence described by section. The relationship associated to ForEach section is identified by the limits relationship. In addition, a section may contain header and footer that will be repeated in case of a Repeatable section.

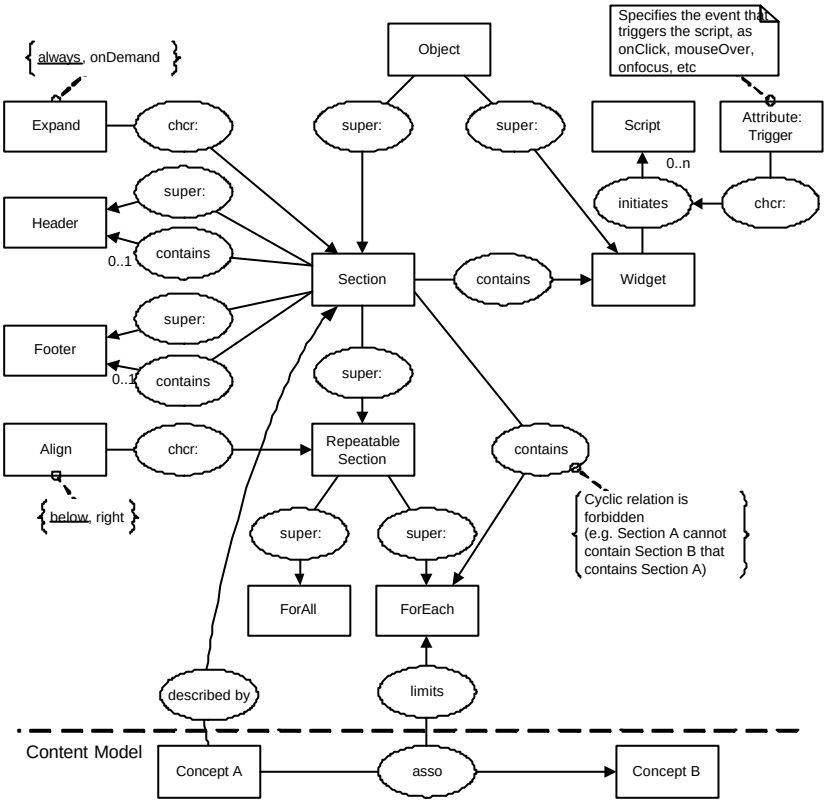


Figure 3. Content Metamodel

## 3.2 Navigation Metamodel

The navigation metamodel (Figure 4) introduces the elements to specify the navigation. Navigation is defined using two types of links: the non-contextual hyperlink and the contextual hyperlink where the destination page is dependent on the content of the source page. For the non-contextual link, we defined three types of hyperlinks: inside the same page, between two pages of the same presentation, or between the presentation and a specific URL. For the contextual link we also defined three types of hyperlinks: between a page and a specific URL stored in the database; between two sections describing the same concept and between two sections describing different concepts semantically connected.

Figure 4 presents the definition of the link element. We can note that the attribute text (value) of an hyperlink depends on the language associated to the text. The text will be different according to the language used during the generation.

One of the objectives of the development of our metamodel is to guarantee the validity of the hyperlinks of the presentation. Non-contextual links which refer to a layout Object or a Page, just as the contextual links referring to a Section" or a ForEach, are guaranteed by the data model; the source and the destination are part of the presentation which is stored in the database. On one hand, deleting an occurrence of Object, Page, Section or ForEach to which a link refers, will be invalidated by the database management system. On the other hand, it is much more complex to validate the integrity of those links where the destination is specified by an URI (a characteristic of the link). Pages specified by an URI can be modified or moved by their owners anywhere on the Web and are usually out of the control of the author of the presentation.

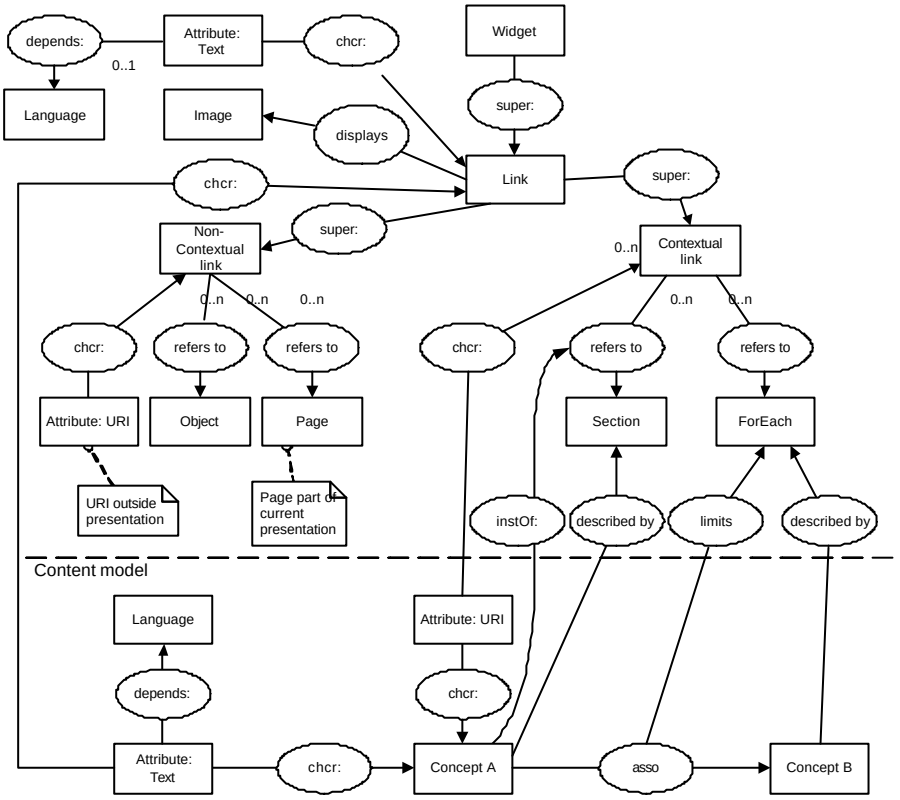


Figure 4. Navigation Metamodel

## 3.2 Widget Table

We detail in this section the special widget Table. Tables are different from other widget because they imply two concepts at the same time (others widgets refer to only one concept described by the section). One of these two concepts describes the occurrences which are found on each line and the other concepts describe the instances that appear in columns. Figure 5 shows an example, lines correspond to provincial parks and columns correspond to sports.

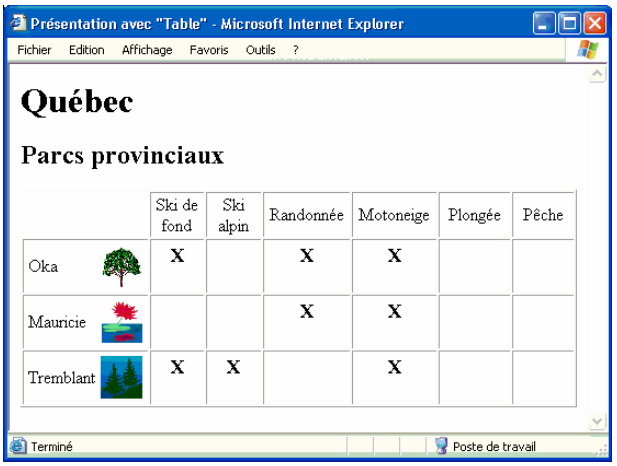


Figure 5. Table

Figure 6 shows the metamodel to describe a table. A table is composed of a RowHeader, ColHeader and cells. RowHeader and ColHeader contain a ForAll or ForEach to explore data. Cell is connected to the relationship that semantically associates concepts of RowHeader and ColHeader.

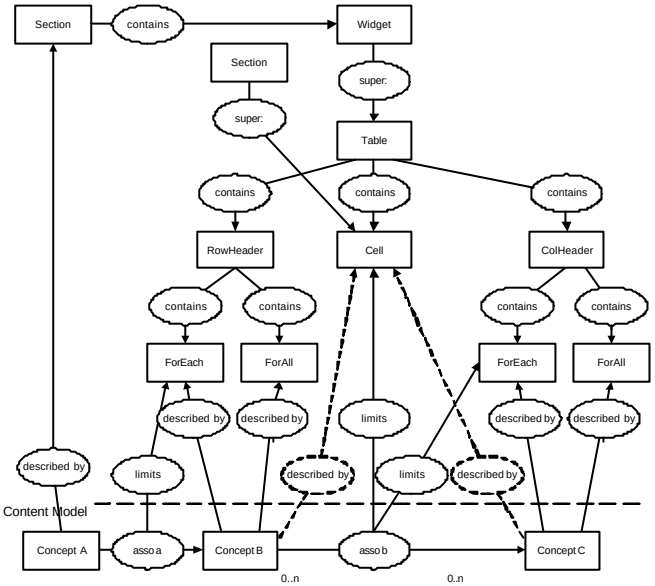


Figure 6. Content Metamodel

## 4. CONCLUSION

We presented a metamodel based on the entity-relationship formalism with some extensions which can be used to model and store Web presentation design. This metamodel allows the specification of a Web presentation as a set of models: layout and content models linked to the data model plus a navigation model.

The work presented in this paper gives interesting inputs on the conceptual structure and semantics existing between the elements displayed in a Web presentation and the underlying data model. We are currently working at developing a graphical interface based on this metamodel to design Web presentations. With such a graphic interface the designer will focus on the tasks of modeling and on the interaction between the presentation and the data model.

Results of this work will be integrated into a web content management system (WCMS) in order to allow a non-specialist to easily develop Web presentations without having to write complex requests to define relationships between the elements of the visual interface and the data model.

In the near future we will enhance the Layout Metamodel to allow more complex specifications of the position of elements in a Web page. We envisage the development of a grammar similar to that proposed by Weitzman and Wittenburg (Weitzman 96), this grammar could increase the power of our metamodel.

## REFERENCES

R. Akscyn and al. KMS: A distributed hypermedia system for managing knowledge in organizations. *Communications of the ACM*, 31(7), pages 820-835, July 1998.

P. Atzeni and al. Design and Maintenance of Data-Intensive Web Sites. In *Proceedings of the 6th International Conference on Extending Database Technology*, pages 436-450, Valencia, Spain, 1998. Springer-Verlag.

J. Bosak. XML, Java, and the future of the Web. *World Wide Web Journal*, volume 2, (4), 1997. http://www.cs.man.ac.uk/ai/ECCS97/srin.html.

S. Cerii and al. Web Modeling Language (WebML) : a modeling language to design Web sites. In *Proceedings of WWW9*, Toronto, Canada, 2000.

F. Coda and al. Towards a Software Engineering Approach to web Site Development. In *Proceedings of the 9th International Workshop on Software Specification and Design*, Ise-shima, Japan, April 16-18 1998.

J. Connallen. Modeling Web Application Architectures in UML. *Communications of the ACM*, 42(10), pages 63-70, October 1999.

B. DuCharme. *XML The Annotated Specifications*. Upper Saddle River, NJ, 1999. Prentice Hall.

P. Fraternali. Tools and Approaches for Developing Data-Intensive Web Applications. *ACM Computing Surveys*, 31(3), pages 227{264, 2000.

M. Gaedke and al. Object-Oriented Web Engineering for Large-scale Web Service Management. In IEEE Computer Society, editor, *Proceedings of the thirty-second Annual Hawaï International Conference on System Sciences*, Hawaï, USA, 1999.

P. Gloor. Elements of Hypermedia Design : Techniques for Navigation and Visualization in Cyberspace. 1997. Birkhauser Boston. 416 pages.

D. Hay. A Repository Model - The Relational Design model. *The Data Administration Newsletter (TDAN.com)*, 2000. Robert S. Seiner. Avalaible at http://www.tdan.com/i014ht01.htm.

L. Hungjun and . Integrating Database and WWW Technologies. *World Wide Web*, 1(2), pages 73{86, 1998.

T. Kamps and T. Reichenberger. Automatic Layout Based on Formal Semantics. In *Proceedings of the workshop on Advanced visual interfaces* , pages 231-233, Bary, Italy, 1994. ACM Press.

N. Koch and A.Kraus. Towards a Common Metamodel for the Development of Web Applications. In *Third International Conference on Web Engineering (ICWE 2003)*. J.M. Cueva Lovelle, B.M. Gonzalez Rodriguez, L. Joyanes Aguilar, J.E. Labra Gayo and M.P. Paule Ruiz, editors, LNCS 2722, Springer Verlag, 497-506, July 2003.

O. Liechti and T. Ichikawa. Metadata, XML and Automation on the WWW. *Journal of the Swiss Information Processing Societies*, pages 30--32, 1999.

F. Manola. Towards a Richer Web Object Model. *SIGMOD Record*, 27(1), 1998.

M. Matera, A. Maurino, S. Ceri, and P. Fraternali. Model-driven design of collaborative Web applications. *Software - Practice and Experience (SPE)*, 33(8) , pages 701--732. July 2003

G. Mecca and al. From Databases to Web-Bases: The Araneus Project. Universita Degli Studi Di Roma, May 1998. Technical report 34-1998.

L. Weitzman and K. Wittenburg. Grammar-based articulation for multimedia documents design. *Multimedia systems*, 4(3), pages 99--111, 1996.