

# Web Site Engineering with Content Provider Participation

Dr. Manfred Bogen, Mrs. Inke Kolb, Mrs. Simone Lahme  
GMD - German National Research Center for Information Technology  
IMK - Institute for Media Communication  
Schloss Birlinghoven, 53754 Sankt Augustin, Germany  
<http://imk.gmd.de/vas/>  
{manfred.bogen, inke.kolb, simone.lahme}@gmd.de

## ABSTRACT

For a successful implementation of a Web site, content providers being non-experts have to cooperate with designers and WWW experts from the onset. Before implementation, a documentation and visualization of a Web site conception is urgently needed. An appropriate *Phase Model* with stress on the conception phase has to be developed. Content Management Systems (CMS) are appropriate tools to develop a Web site in a cooperative effort. In addition, they facilitate maintenance and administration after its completion. At GMD, we have developed a Phase Model and implemented our own IMK Web site and the Web site of the Beethoven Museum in Bonn following the principles outlined. In this paper, we present our approach and the experience gained implementing it.

## KEYWORDS

Web site engineering, Content Management System (CMS), content description table, site map, Phase Model, content centered development.

## 1 INTRODUCTION: THE GMD PHASE MODEL

GMD has been responsible for setting up Internet services including Web servers for different cooperation partners such as *Deutsche Welle*, *Süddeutsche Zeitung*, and the German Government in the past (see [1],[2]). As discussed in [3] we found that it was useful to have three different parties, i.e., content providers, designers, and WWW experts, in the respective task force develop a Web site together.

A Web site development process basically consists of the *conception process* and the *implementation process*. The emphasis of our *Phase Model* is on the *conception process* that is divided again into four phases: *rough planning*, *fine tuning*, *selection of a Content Management System*, and *screen design*. In order to enable a content provider to participate effectively, it is necessary to make graphic representations of all steps of the conception process. The principles of structuring and modularity are taken from general software development [4]. In the fine tuning phase the work is split into modules. In the implementation process, we have to differentiate between HTML template development, program development, and structuring of a content management database. Some basic rules of software engineering [5] have to be applied: the process of implementation is iterative and incremental. All individual model phases require a different methodology in the cooperation between developers (designers, WWW experts) and customers (content providers).

The following paper presents a methodology to create a Web site with participation of content providers in each and every development phase. We elaborated it at our own institute, the Institute for Media Communication of GMD. At present, we apply it to the Web site to be developed for the Beethoven Museum [6] in Bonn. This paper is structured as follows: Section 2 describes the conception process and its different phases, i.e. rough planning, fine tuning, selection of a Content Management System, and screen design. Section 3 explains the implementation process. The application of the GMD Phase Model to construct the IMK and the Beethoven Web Site is subject of section 4. In the conclusion we emphasize the advantages of Content Management Systems not only during the conception and implementation processes, but also in the day-to-day operation of a Web site, and the benefits of our Phase Model.

## 2 THE CONCEPTION PROCESS

In the first phase of the conception, the rough planning phase, people with different areas of competence come together: *WWW experts*, *designers*, and *content providers* (see figure 1). *WWW expertise* includes the conception of Web sites in general, the installation and configuration of hardware, software, and a network infrastructure, HTML knowledge, and experience with appropriate Internet/WWW tools. A *designer* knows how to arrange a Web page and is able to implement a company's CI (corporate identity). A customer is a *content provider* in many cases. In other cases there is a special person, who prepares the Web content on behalf of a customer. We understand both as content providers. Given this scenario, it is very important to respect each other's competence from the beginning. Unfortunately, this natural condition is often neglected.

After defining the target of the Web site and the potential user group (audience), the different conception phases can begin.

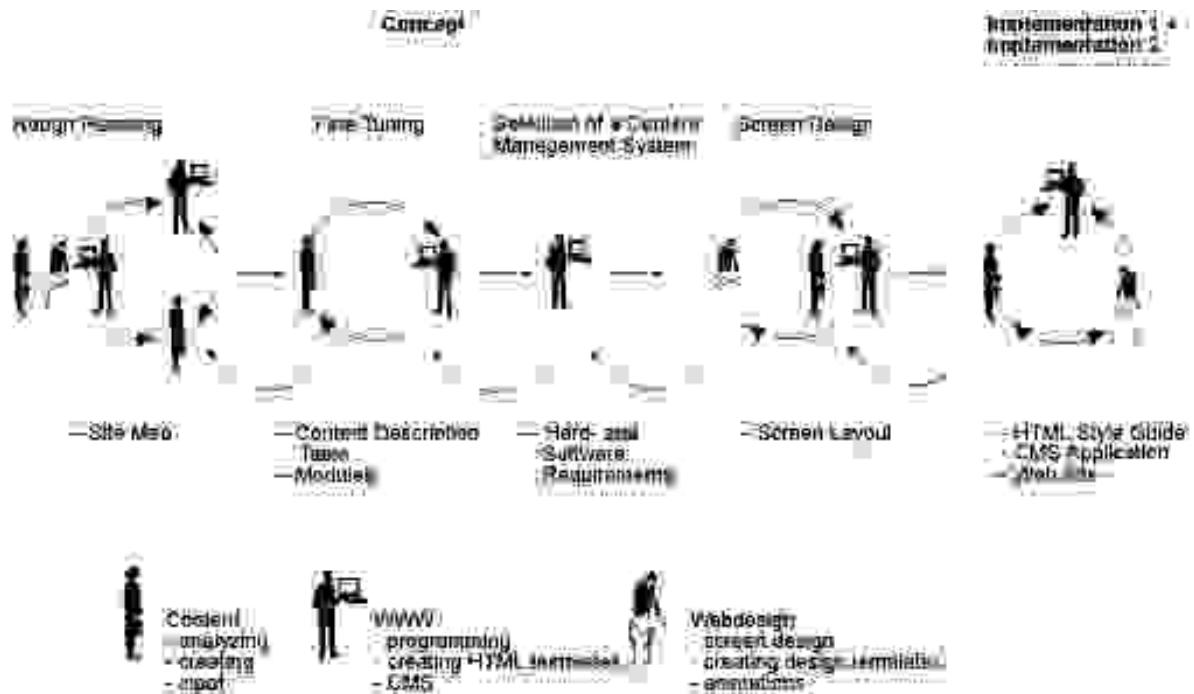


Figure 1: The GMD Phase Model

## 2.1 Conception Phase I: Rough Planning

A good first step in the rough planning phase is to get all areas of competence, i.e. at least three different persons – around a table and let them assemble the desired content of a Web server by means of brainstorming and metaplan techniques [7] (see figure 1). Apart from the usual advantages of this method (participation of all users, etc.) the opportunity to define a common terminology is worth being mentioned. Thus all parties will limit their specialized jargon and technical vocabulary or at least will have to explain it. For example, the term "news" may be understood as an interface to a news group or as some new information about a company. Metaplan cards should not be hung up unless there is a common understanding and acceptance by all participants in this session. In this phase, no restrictions should be made at all and fantasy should run freely. Only in this way, really new and funny ideas will be found. After their collection, the terms are sorted. What belongs together? What is on the first level of a Web site structure, i.e. can be seen by a visiting user first? In some cases generic terms should be introduced.

In order to define the main navigation structure of a site, we put the collected content into a tree diagram (see figure 2) similar to the description in [8]. The leaves are different symbols depending on what kind of content they represent: normal Web pages, download files, e.g., pdf-files, hyperlinks, or pages to be linked to from all other pages, e.g. "search". This gives a good idea of the depth of a site and helps to define the main menu, the submenu, and other elements that should be reached from all pages of a site.

By visualization (see figure 2) we obtain a documentation of the conception, on the one hand. On the other hand, visualization enables a detailed discussion between Web technicians, graphic designers, and content providers. After some re-design, the diagram could be regarded as a *site map*. Each node and each leaf represents a Web page. It is important to confirm that all pages can be filled with content.

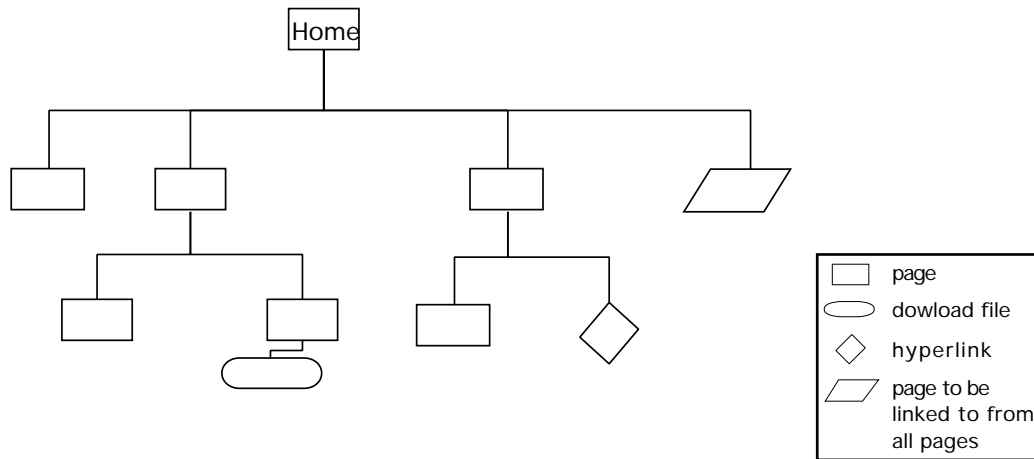


Figure 2: The Site Map: Visualization of a Site Structure

## 2.2 Conception Phase II: Fine Tuning

After the rough conception phase we are now able to do fine tuning. Each small box in our visualization (see figure 2) corresponds to a Web page. Now it is the task of a content provider to specify the content of the pages in more detail. Afterwards, the work will be split into modules (see figure 1).

**2.2.1 Content Description Table.** We have prepared a *content description table* (see table 1) which is similar to the one described in [9]. Here, however, we do not set milestones per page, but, as presented below, per module.

Page	Links	Content	Content available? Source Format	Responsibility	Number of printed pages	Update frequency	Language	Release by	Notes

Table 1: The Content Description Table

In order to fill the tables, the content provider should be clear about each individual page. A short description of *content* should be entered, as well as a list of all *links*. The internal links should result from the site map; otherwise the rough conception must be revised again. Links to other servers should also be specified here. This will give an overview of all links to other sites.

The column *source* is of special importance. Often it is intended to use already *available content*, e.g., texts and pictures. The format in which the data are delivered is important to estimate the implementation expenditure. So the developers can determine, which hardware and software e.g., scanner and conversion tools are needed. If the content is not yet available, it must be created. The table thus gives also an overview of the work to be done for the WWW presentation. The column *responsibility* determines who must supply content and who will later be responsible for maintenance.

The *number of printed pages* column roughly shows how comprehensive the content will be. It helps to avoid inadequately long pages and a Web presentation as a *clicking machine* (only structure pages, no content). A closer investigation of the fields *number of printed pages* and *source* can lead in the end to a decision that a file should be offered for download. In order to show not only the work to be done currently, but also the maintenance expenditure during operation to the content providers, there is the column *update frequency*. Often, this entry leads to a revision of the concept as maintenance is expected to become too cumbersome. Since GMD resides in a non-English-language country, the question about the language arises with the establishment of a Web site. A double maintenance expenditure of a multilingual implementation deters many providers. Basically, a customer or reader determines on the language of the entire site. We found out however that it can be quite useful to offer individual pages and documents in different languages. If that is required, it must be entered into the column *language*.

Sometimes the person who supplies the content is not the person who has the right to publish them. The specification of the person, who is allowed to *release* information, is needed later for the user management in the Content Management System. The work will be finished when the tables are complete for all pages.

**2.2.2 Defining Modules.** The site map is split into small and easy to overview branches, which can be implemented as independent modules. For each module, a time schedule is set up and agreed upon with the designer and with the content providers. How long does the development need? When do the converted and cut pictures have to be supplied by the designer? When do the content providers have to supply the content? When and how long will the test phase be? Are there external requirements, which prescribe a certain publication date? From now on the process is iterative and incremental, as we know it from software engineering [5]. This means that each module passes the entire implementation phase and each module has a Web presence as a result that is online (if necessary access-protected). At last the content provider starts testing and possibly it will be necessary to fall back on an earlier phase.

There are two advantages with this approach: one will get the first feedback from end users while implementing and one is able to look back to the conception and change things before it is too late. Secondly, the content provider will be introduced to the Content Management System. This may save a lot of time to insert and change content in future modules.

### **2.3 Conception Phase III: Selection of a Content Management System**

In this conception phase, the software and tools needed for the implementation process have to be identified, evaluated, if necessary, and selected (see figure 1). *Content Management Systems (CMS)* promise to facilitate Web site management and collaboration among the different contributors (i.e., WWW experts, designers, and content providers) by

- separating content from layout,
- automatically generating navigational links, and
- providing the ability to define a publishing workflow.

Today, there is a vast number of Content Management Systems available, ranging from quite simple systems to high-end enterprise solutions [10]. As with every complex software system, there is no 'best' CMS that fits every possible need. To help in a purchasing decision to be made, a benefit value analysis has proven to be helpful (see also [11]). Even if a complete analysis is not possible, using these concepts will help ensuring transparency.

From our experience, the first step in choosing such a system is to identify the specific needs of the parties involved. Here we have three main user groups: administrators (or Webmasters), developers (implementing the templates), and authors contributing content. Each group will have different requirements for a CMS; these needs have to be identified and afterwards weighted according to their importance. With these lists in hand, systems can be evaluated and scored according to the criteria drawn up. Such assessment can be carried out in one of the following ways, listed in descending order of preference:

- Installing an evaluation version of the system and creating a demonstration project.
- Getting an account on a demo system supplied by the vendor.
- Reviewing the available product documentation and investigating reference customers.

Though being the most time consuming method, installing an evaluation system offers the most valuable unbiased information. Since the introduction of a CMS is not a short-term decision, the time and money required for such an evaluation will pay off later on. With the lists of criteria filled in, it should be possible to identify the CMS that fits one's demands best.

### **2.4 Conception Phase IV: Screen Design**

A prerequisite for this phase is the termination of the conception phases specified above. Only from the complete site map one can see, how many and which branches go off the homepage and the underlying pages. The navigation items such as links, menus, and buttons must be structured accordingly. Length and type (e.g. many pictures) of page content have also an influence on the design. We take this information from the content description table.

It is the task of a designer now to sketch a Web interface that is adapted to the target user group, that fits the corporate identity (CI) of the content provider, that offers a clear navigation structure, and that attains the determined target of the Web site. The diagram will then be discussed with the WWW experts and the content provider (see figure 1). In most cases there will be a design for the homepage and one or two basic designs for all further pages. At this point no HTML code has been produced yet. The sketches of the Web interface are discussed in the development team of WWW experts, content providers, and designers. Finally the designer

adapts them. This process is iterated until a common solution is found. The results of this phase are the pictures of the screen design and the graphic design style guide that documents the design decisions made.

### **3 THE IMPLEMENTATION PROCESS**

As a result of the conception process, the visualized structure of the site and the content description table are available. Modules have been defined, that can now be implemented individually. We deliberately omitted in the conception the creation of HTML pages or thinking in terms of HTML tags, browser windows, and "MouseOvers". The content is in the center of conception. Of course we assume all parties involved are experienced in using the Internet. That is mostly the case. If not, the developers should give an Internet introduction with many examples even before beginning the first conception phase.

#### **3.1 Implementation Phase I: CMS Application**

During the conception one meets content of different type. There are structured data, on the one hand, and text and graphic data on the other hand, whose structure arises only as a result of the Web page structuring. A page structure can be described roughly, for example, in such a way: *heading, short text, figure 200x200 pixels, on the right with caption*. Partially the structure of the pages is defined by the designer, such as the positioning and size of pictures within the text. A close cooperation with the designers is therefore recommended. Content-wise different pages, which have the same structure, are collected and implemented in one *CMS application*, i.e. there is a single table in the underlying database package, a kind of relational database.

Examples of structured data are personal pages of the employees. A company could determine that all employees have to publish their names, telephone numbers, e-mail addresses, photos, and descriptions of their function in a company. If exactly these data are available in a database or an existing database can be adapted to these needs; an interface between the Web and this database will be programmed for reasons of consistency. (If no questions of security or data protection are against it.)

If the data are needed however exclusively for the Web server, they are stored in the CMS. This enables the editors (in our example possibly even the employees themselves) to process all Web content in an identical way using the same software package.

It is important to understand that web site implementation does not replace a fundamental well-structured data administration system. So one should not enlarge a small e-commerce solution to a small commodity management system or to take - returning to our first example - the personal Web pages as a basis for an employee database. Technically the databases may be hosted on the same hard- and software configuration. Conceptionally however a distinction is necessary. Otherwise you may be forced as Web site developer to have to change the complete data management of a content provider.

While developing a CMS application we also operate according to an important rule of software engineering: re-use. An application is being implemented as generally as possible and as specifically as necessary, so that a re-use - at least in other modules of a Web site - will be possible. The results of this implementation phase are the CMS applications and their complete documentation.

#### **3.2 Implementation Phase II: HTML Templates and Programs**

It has to be determined, which WWW standard is to be used (HTML, XHTML, XML, etc.), and whether a script language such as JavaScript or PHP can be applied. These decisions are documented in a *technical style guide*. For each module, the screen design pictures are converted and cut for their integration into the templates e.g., HTML templates. The templates are created (as generally as possible and as specifically as necessary) and integrated into the CMS. For the sake of completeness it has to be mentioned here that in this phase programs and scripts have to be designed, implemented, and documented.

### **4 APPLYING THE GMD PHASE MODEL TO THE IMK AND BEETHOVEN WEB SITE**

Following our methodology, we built the Web site of the Institute for Media Communication (IMK) of GMD - German National Research Center for Information Technology (<http://imk.gmd.de>). In this case, the content providers were the researchers of the institute. The objective of the Web server is the publication of the institute's work for scientific exchange and public relations. Target user groups were mainly scientists from all over the world and potential project partners from industry and science. In addition, we addressed European and German promoters as well as potential job applicants. As language we chose English. In the following we address only the major issues of our Phase Model neglecting the less important ones.

#### **4.1 Rough Planning and Fine Tuning**

All people involved participated in the first phases of the conception: the existing group of Web editors, the department managers, the institute director, the marketing manager, the designers, and the implementors (Webmasters). After several iterations of these phases we obtained the IMK site map.

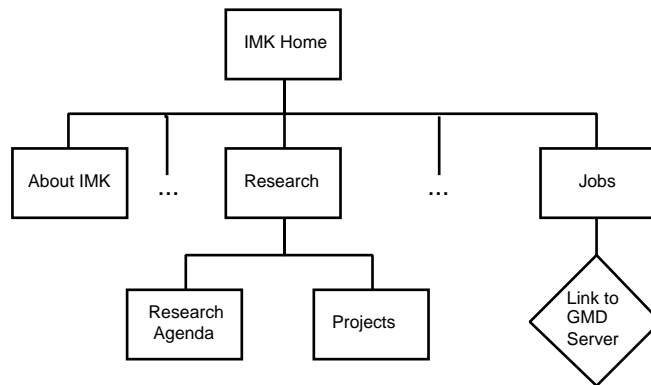


Figure 3: Visualization of the IMK Web Site Structure (part)

Experience gained from the operation of the previous institute Web server has shown that more importance has to be attached to content maintenance. When completing the content description table (see table 1) we observed the following, which is crucial for the structure of a Web site: The amount of content, which has to be produced exclusively for an online presentation has to be minimal. Only few resources are available for the production and maintenance of this content. Content has to be derived from other documents. There is some content in the responsibility of the marketing director. The representation of the research projects is in the responsibility of the project teams themselves. Basically, all employees have to be enabled to act as Web editors. Some content is in German and is exclusively of interest for German-speaking readers. On the Web server, whose official language is English, it may appear in German.

We defined eleven modules to be implemented. Since easy maintenance was seen as one of the most important topics, we decided to also generate the structuring elements from the CMS database. In this way it would be easy to let the marketing department change, add or delete, e.g. the menu text whenever needed. A special module was defined for this purpose.

#### 4.2 Selection of a Content Management System

For the IMK Content Management System, we have found the following items to be decisive:

- User interface: Authors should not need any software apart from a Web browser, providing a simple and intuitive interface to the system.
- Costs: Not only the price for the software - CMS range between a few and some hundred thousand Euros - has to be taken into account but also the expenses for support or additional hard- or software requirements.
- Dynamic page generation: To ensure topicality, we think that a CMS should generate pages on the fly instead of being subject to predefined export cycles or requiring manual static page generation to reflect site updates.

SixCMS (<http://www.six.de/>) scored well with respect to the above mentioned items. We found the pricing model (licenses are sold on a per server basis, regardless of the hardware or number of clients) to be appealing and due to the system being based on open source components, no additional software fees (e.g. for a database) were needed.

We started our work with SixCMS version 3. SixCMS is written in the Web scripting language PHP (<http://www.php.net>). It uses the Apache Web server (<httpd.apache.org>) and MySQL (<http://www.mysql.com>) as data repository. While all of these components are also available on the Windows platform, a Unix server is recommended for stability and performance reasons.

All pages are generated dynamically; nevertheless performance should not be an issue even on low-end hardware. Access to the system is through a Web browser (JavaScript and Cookies enabled). A single SixCMS server can be divided into several separate sites, thus giving us the possibility to run not only the external Web site but also our internal pages on a different 'site'. Each 'site' can be divided into user-definable content sections that can be divided into categories. Access to the content is controlled on group basis for user and group management; a Web interface is available.

#### 4.3 Screen Design

The Web design was developed to vividly portray the dynamic modern nature of our Institute. Since IMK is doing research in several fields such as Virtual Environments, Virtual Studio systems, and Internet Broadcast we

wanted to express the character of the institute by using many pictures and different types of multimedia elements.

The style of the old Web presentation should be retained, but it should be easier to make changes to the pages without having to involve a designer to do the changes whenever necessary.

Since an easy maintenance was seen as one of the most important aspects, the design includes as many menu elements as possible directly from the CMS. Those text elements combined with graphic elements are used for all submenu parts. The few navigation elements that are pictures (main menu) have enough space to add more elements if needed.

The *graphical style guide* defines the partitioning of a page, the space for text and pictures, and the colors. All pages have the main layout in common: the institute name and logo on the top and below that the main menu. On the left there is the submenu, the main content is in the middle, and a picture might be on the right. This general layout (see figure 4) is used for most pages.



Figure 4: Graphic Design Proposal for the IMK Web Server

#### 4.4 Implementation

A generic CMS application that could be re-used for most modules was generated. Structured data are contained in the pages of the projects, the r&d groups, and, the individual employees. Also for this purpose, CMS applications were produced. For the publication lists, an interface to the database of the GMD library was implemented.

We decided to stick to HTML 4.01 as far as possible and to use JavaScript for the animations. In accordance with the general layout, a general HTML template was created. Due to some restrictions of SixCMS version 3, the re-use of templates in different modules was limited so that quite a lot of templates had to be built. As mentioned above we decided to generate the structuring elements from the CMS database. A special CMS application and special HTML templates were created for this purpose.

#### 4.5 The Beethoven Museum Web Site

We are about to apply the GMD Phase Model to the Web site development of the *Digital Beethoven House* project [6]. This differs from the IMK Web project in three main points:

- It is interdisciplinary, as the content providers are musicologists and librarians.
- It is much larger.
- The Web site will not only represent the departments of the Beethoven Museum, but will also be the portal to a new digital archive about Beethoven.

The first conception phase, i.e., the rough planning phase, has been completed successfully. For most of the pages the content description table was filled in the second conception phase, the fine tuning. It was very useful to define terminology together. We defined one module to be implemented first and wrote down a time schedule. The other modules will be defined afterwards so that experience gained from the first can be incorporated.

The following needs made us to choose SixCMS as Content Management System again. Due to the expected complexity, the CMS should run under the operating system UNIX. It should be able to manage large amounts of data and multimedia documents (graphics, audio and video files). The clients are to run on Windows PCs; the WWW browser is the preferred user interface. They must be usable without detailed technical knowledge, since the musicologists and librarians of the Beethoven house will use them. Apart from the costs of a Content Management System also the costs of the pertinent database and for additional software have to be considered. Maintenance that is to be outsourced should be minimal. Because of the long project duration and to minimize maintenance costs, the CMS should support future-proof techniques and current Internet standards (such as HTML, XHTML, XML, and WML).

At present, the screen design proposals are being discussed in the fourth conception phase.

Within the *Digital Beethoven House* project, the Phase Model will be refined and completed. We will find out where more content provider participation is necessary and where the developers have to make more decisions on their own. The technical development will show to what extent the re-use of CMS applications and HTML templates is possible.

## 5 CONCLUSION

As any other software package Content Management Systems have a few non-particular disadvantages: the investment needed for purchasing and ongoing maintenance, a certain amount of dependency on a specific product and database, and the training effort for the content providers, the designers, and even the Web experts. However, the advantages predominate. With a CMS it is much easier for content to change, add, or delete the content of a Web site. Webmasters are relieved from routine and monotonous work. Once a Web site is set up with a Content Management System, neither the Webmaster nor the content providers have to worry about consistency. In particular, Content Management Systems are an ideal tool to complement our Phase Model in establishing a Web site.

The GMD Phase Model has proved to be viable. Based on the experience made in earlier Web developments, we applied it to the IMK Web site and are currently using it to develop the Beethoven Museum Web site. Gone are the days when you developed a product and presented it to your customer or cooperation partner after its completion. Today, a participation of all parties involved is needed from the very beginning. Someone has to structure the workflow and the mode of cooperation among these different parties. Though everybody is talking about the Internet and the World-wide Web, people have a completely different understanding and use a different terminology. Clarification and a certain amount of precision are needed. Our Phase Model helps in these aspects. It emphasizes visualization, structuring, modularization, and iteration issues while developing a Web site. Following the different phases of the model, each party involved is able to bring in know-how, experience, and ideas according to their competence, producing a Web site meeting a common acceptance today and retaining convincing topicality in the future.

## REFERENCES

1. M. Bogen, M. Lenz, S. Zier, "Deutsche Welle: on the air," Computer Networks and ISDN Systems: Proceedings of the Fifth International World Wide Web Conference in Paris, Volume 28, Numbers 7-11, Elsevier, Amsterdam, pp. 1187-1196
2. M. Bogen, M. Lenz, S. Zier, "A broadcasting company goes Internet," Proceedings of WebNet 96 - World Conference of the Web Society, pp. 27-32
3. C. Liu, J. Peek, R. Jones, B. Buus and A. Nye, "Managing Internet Information Services", O'Reilly & Associates, Inc.
4. Helmut Balzert, "Die Entwicklung von Software-Systemen:Prinzipien, Methoden, Sprachen, Werkzeuge", Reihe Informatik/34, Bibliographisches Institut, 1982
5. Philippe Kruchten, "The rational unified process: an introduction", Addison Wesley, 2000
6. M. Bogen, S. Lahme, and U. Lechner, "The Digital Beethoven House: Conception and Demonstrator," INET'99: The 9th Annual Internet Society Conference, San Jose, California, USA, 22-25 June 1999, Proc., ISBN 1-891562-06-1, CD, [http://www.isoc.org/inet99/2b/2b\\_2.htm](http://www.isoc.org/inet99/2b/2b_2.htm)

7. Josef W. Seifert, "Visualisieren. Präsentieren. Moderieren.", Gabal Verlag, 1995
8. Louis Rosenfeld, Peter Morville, "Information Architecture for the World Wide Web", O'REILLIY 1998, pp.154-156
9. David Siegel, "secrets of successful Web Sites", 1997, Hayden Books Hayden, USA, 1997
10. Andreas Karajannis, Torsten Bissel: Webenslauf - Content-Management-Systeme im Vergleich. In: iX - Magazin für professionelle Informationstechnik, Ausgabe 11/November 2000, Heise Verlag, Hannover, 2000
11. T. Bissel, M. Bogen, M. Borowski, A. Karajannis, and J. Löffler. Content Management Systems for Editors - In Search of Quality. *Proceedings of WebNet 99 - World Conference on the WWW and the Internet*, Honolulu, Hawaii, Oct. 25-30, 1999, Volume 1, pp. 106-111

#### **AUTHORS INFORMATION**

*Manfred Bogen* has a M.Sc. in Computer Science from the University of Bonn in Germany and a Ph.D. from the University of Nottingham in the UK. He has been working at GMD since 1980 in different positions as student, scientist, and manager. He heads the competence center VaS (Value-added Solutions) of GMD. At present he leads a 5 years digital library project for the conception and installation of the 'Digital Beethoven House' in Bonn.

*Inke Kolb* is computer scientist at IMK - Institute for Media Communication of GMD. She was involved in the first launch of GMD's Web server in 1994. In 1995 Inke Kolb worked in a team at ESI - European Software Institute, Spain (<http://www.esi.es>) to start the ESI Web server. Currently, she is webmaster at IMK and is responsible for the "online presentation" of the project "The Digital Beethoven House".

*Simone Lahme* is currently working as an educational designer in the field of e-learning for Viviance new education GmbH in Potsdam/Germany. From 1991 to 1993 Simone Lahme studied History of Art at the University of Bonn. In 1993 Simone Lahme started her graphic and communication design studies at the Rhein-Sieg-Kunst-Akademie in Hennef, where she successfully graduated in 1998. From 1997-2000 she was a member of the Institute for Media Communication at GMD - National Research Center for Information Technology.