Developing and Deploying an XML-based Learning Content Management System at the FernUniversität Hagen

Gerd Steinkamp, Albert Bullig, Eberhard Heuel, Firoz Kaderali, Heinz Müller, Gunter Schlageter

FernUniversität Hagen
gerd.steinkamp@fernuni-hagen.de

This paper is a report about the project FuXML carried out at the FernUniversität Hagen. FuXML is a Learning Content Management System (LCMS) aimed at providing a practical and efficient solution for the issues attributed to authoring, maintenance, production and distribution of online and offline distance learning material. The paper presents the environment for which the system was conceived and describes the technical realisation. We discuss the reasons for specific implementation decisions and also address the integration of the system within the organisational and technical infrastructure of the university.

Keywords: e-learning, learning content management system (LCMS), cross-media publishing, XML

1 Introduction

The University of Hagen (FernUniversität Hagen) is in a unique position in Germany, since it is the only university for distance education. The university is integral part of the regular public higher education structure and offers German degrees in the fields of economics, law, human science, computer science, mathematics and electrical and information engineering. The students receive the printed course material in a two-week rhythm and have to pass assignments before getting admission to final examinations. The university has built an infrastructure of study centers throughout Germany, Austria, Switzerland and Eastern Europe, where students can meet their tutors and get access to small libraries, computers and video conferencing equipment. Besides the remote aspects of the study, the students have to spend several attendance days at Hagen to take part in seminars, laboratories and oral examinations.

The FernUniversität is determined to make best use of the vast possibilities which are offered by the information technology and new media. Diverse activities have been started in the university to evaluate the route from the traditional university for distance teaching towards a new media university. During the recent years, the university has integrated its faculties and central institutions into a concept called *Education and Knowledge Space: Virtual University.* This concept aims at completely representing the university and its services in the virtual space as provided by the Internet. To achieve this goal several internally funded projects are carried out which target aspects of e-learning content, services and infrastructure.

The learning content management system FuXML [FuXML] described in this paper is one of these projects and was funded during 2002 and 2003. The project was executed by the Department of Computer Science I, the Department of Communication Systems, and the university's media centre, the ZFE (Centre for the Development of Distance Teaching).

The FernUniversität Hagen has a long tradition in employing various kinds of media in addition to the printed course material. Media design is thus of high importance within the concept of the virtual learning space. Teachers who want to make use of hypermedia contents

in their courses are supported by the ZFE. The assistance ranges from specific course design for visually impaired students to the production of TV programmes on specific academic topics.

The content creation process at the FernUniversität has not been standardized so far. Although guidelines and style guides were created by the ZFE in the past, the application of these are not mandatory and teachers are free to use any approach which suits them best. Most courses at the FernUniversität only exist as printed material created with standard word-processing and type-setting software like Microsoft Word and LaTeX. There are also some courses which do not come in the shape of printed paper but are published as CBT courses on CDROM. These are typically results from specially funded multimedia course development projects.

With the advent of the Internet in public life during the last decade, the advantages of the new Internet services for distance teaching and learning were soon becoming undisputed. The possibilities offered by hypertext and hypermedia obviously have great potential with regard to new pedagogical models. The new electronic communication techniques like videoconferencing, chat, email and newsgroups were expected to greatly enhance the contact between students and their teachers.

However, the expectations put into e-learning that it would quickly revolutionise the way students learn have not been fulfilled to the extend predicted a few years ago. While it is possible to develop web-based hypermedia course material within individually financed projects, it is a big challenge to make such developments on the basis of the regular resources available to the university. There are 1600 courses offered by the FernUniversität. It is economically infeasible to spend the amount of money usually required to create high quality hypermedia material for each of these courses. Teaching experience also shows that while students happily accept the provided hypermedia content as additional material, it does not provide a satisfactory replacement for the printed course. Currently there are no cross-media publishing systems in use and the maintenance of an online version and a print version for identical course contents is left to the faculty staff. This situation is not satisfying and an economically reasonable solution has to be found.

2 Make or Buy ?

Besides the public academic sector, e-learning is of major significance for enterprises since it is the key-technology that enables life-long learning, which is regarded to be one of the most important concepts in modern information society. The business interest in e-learning technologies and the associated investment capital has created a considerable market for e-learning solutions. There is a wealth of Learning Management Systems (LMS) and Learning Content Management Systems (LCMS) available on the market, so it should be a natural decision for a university, just like for any other company, to buy one of the existing commercial solutions and refrain from starting its own development. Nevertheless, in practise the choice of a commercial solution is not as good as theory may suggest. The main three reasons are: firstly, academic institutions have different requirements than most enterprises. Secondly, commercial e-learning solution providers do not adept their products to these requirements, they do merely adapt their marketing strategies. Finally, in the long run, the licence costs for commercial solutions are too high. There are several examples of universities that bought commercial LMS at a time when the competition was strong and the licensing

fees were low. After the withdrawal of some of the weaker competitors, licensing prices were suddenly much higher than at the time the products were chosen.

3 Goals and Requirements

Since the FernUniversität is a university for distance teaching, the requirements towards a content creation and management system are different than those of conventional universities. The course material is provided in form of an elaborated textbook which is divided into several course units including self-exercises and assignments. This stands in contrast to the rather ad hoc nature of lecture material used at conventional universities mostly comprised of sketchy Powerpoint slides and lecture notes. Consequently, the content creation process must support substantial more options than assembling course modules out of existing presentation slides and PDF documents.

The high level goals the system should meet are:

- Simplification and standardisation of the course material production process.
- Reduction of the technical overhead in the course creation process so that authors and faculty members can focus on content and didactics.
- Separation of contents and design which allows the independent work of authors and design experts.
- Provision of a cross-media publishing process that gives teachers the option to choose multiple output formats of the course without generating substantial additional workload.
- Utilisation of a standardised document format with a half-life cycle much greater than that of standard word processors.

At the time the FuXML project was started, an LCMS that would meet the university's requirements could not be seen on the market. Besides basic editorial and multimedia features some important requirements are:

- Cross-media publishing that delivers a print version very similar to the existing printed version of the course. The print version of an existing course represents an established standard and a deviation from its original structure and layout is often not accepted.
- High flexibility with regard to design and layout. University professors do not like to be forced to use a specific visual presentation. They want to establish their own style and often make only little compromises. At the same time, the customisation costs must be very low and adaptation must not require special programming skills (e.g. XSLT programming).
- Correct academic implementation of numbering schemes, bibliographies, footnotes and marginal notes. Especially for the faculty of law the correct realisation of these editorial devices is an important requirement.
- Extensive support of scientific mathematical notation in large volumes and adequate presentation on printed paper and screen. This is a requirement that many LCMS do not account for adequately.

4 Media Concept

In the traditional pedagogical concept of the FernUniversität, students use the course material as textbook and working material. The course text is enriched with didactical elements like learning objectives, summaries, prerequisites, examples, comments, etc. The text is based on basic principles regarding structure and format which have been established especially for the distance learning material. Typography and design are chosen to reflect certain pedagogical goals. Hence, although detailed mandatory style guides do not exist, all courses must abide to a minimal common quality standard.

This didactical methodical and design-oriented course profile has proven itself for many years now. The basic principles of traditional course development are also applied for the courses in the Virtual University. Although completely new pedagogical models can be realized in modern e-learning environments, this was not the goal of the FuXML project. The project's emphasis was more on establishing economical work processes than on introducing novel didactical concepts in the distance learning approach of the FernUniversität. Today, this rather conventional approach seems to have been the right choice. Slipping in new methods in teaching through the backdoor of technology without organisational incentive and support might not be the correct procedure. The adoption of new working methods based on structured editing of XML course documents already presents a barrier to untrained academic staff that should not be underestimated. The target user group does not consist of euphoric media and computer experts who love to use fancy e-learning tools for daily enjoyment, but are pragmatic academic employees that are interested in software tools that deliver efficient solutions for their work. Acceptance of the system would have been very low, if extensive training towards new pedagogical methods had been necessary and existing didactically structured material could not have been directly imported into the new system.

Consequently, first applications of FuXML come from existing courses which employ the traditional didactical concepts of the FernUniversität and are imported into the system. Even a simple imported electronic course without any hypermedia features provides added value to students by its online availability and digital versatility. Visually impaired students have always been a target group of the FernUniversität but require specially prepared material. The cross-media production process of FuXML can in future be extended to provide an electronic version of each course that can be processed by braille screen readers. The HTML screen version produced by FuXML conforms to the accessibility guidelines of the W3 consortium and the German directive to ensure accessibility without barriers ("Barrierefreiheit"). Online digital courses also have the advantage that they present an inexpensive way of supplying "demo" material for prospective new students. Parts of regular courses can be made public to offer a first insight into typical university courses for everyone who is interested but undecided whether to start a university study programme or not. An even more radical step would be to completely open the course material for the public, as done by the MIT in their OpenCourseWare project [OCW].

The digital online courses can be "enhanced" in several steps to provide added value to the students. In the first step, the extensive application of hypertext mechanisms by linking the course material is an inexpensive way of introducing an extra quality into the course. This includes linking of contents and special directories like index, glossary, list of abbreviations, etc. Some of the linking can be automated by the system to relieve the author as much as possible from this tedious task. Many students will still use the printed version of a course for the first reading. A hypertext course can be used by a student in a later phase to recapitulate

important concepts and quickly find keywords, definitions, interesting figures, frankly everything the author might consider worthy of indexing. Furthermore, learning objectives and summaries can be easily accessed for every course unit which helps students to get a quick overview of the course's main goals and achievements. The didactical markup provided by the author helps to visually pronounce important concepts and improve the attractiveness of the course. Pleasant visual design may also lead students to preferring the screen version over the printed version, even if no further multimedia contents are available.

The integration of multimedia elements like animations, videos and simulations into the course material is the second step that can be taken to improve the quality of the original course. However, the costs introduced with this step are considerably higher than those attributed to hypertext features. If high quality is demanded, professional support by media experts is required. Care must be taken with regard to the selected media formats. Open standardised formats are better suited than proprietary formats since the latter are often changed with the products they are created with and hold the risk that they cannot be accessed with modern versions of the authoring software even after only a few years. The annoying experience that expensive multimedia content, which is still appealing, cannot be reused due to technical inaccessibility has already been made with CDROM courses that were created in the early years of the last decade.

5 Technical Realisation

While technology should not be the main driving motivation to start an e-learning project, choosing a wrong technological approach will have severe impacts to the practical utility of the project's results.

For FuXML the following technological decisions were made in the inception phase of the project:

- As document format for the textual contents of a course, the Extensible Markup Language (XML) was chosen. A document type definition (DTD) was developed according to the special needs of the course production at the FernUniversität. The supported media types consist only of open standardised formats. Specifically, Microsoft Office documents can only be included as external objects.
- The FuXML system is realized as client server solution and is divided mainly in three parts: authoring (client), data management and cross-media-publishing (server). Clients can connect to the system via the Internet from virtually everywhere.
- The software architecture is based on the Java J2EE enterprise platform, which currently is the leading software technology for developing distributed and scalable enterprise applications.
- For the realisation of server components open-source solutions were selected where appropriate.
- For the client side, commercial solutions were chosen to provide good usability to the users of the system.

The following diagram presents an overview of the system architecture.

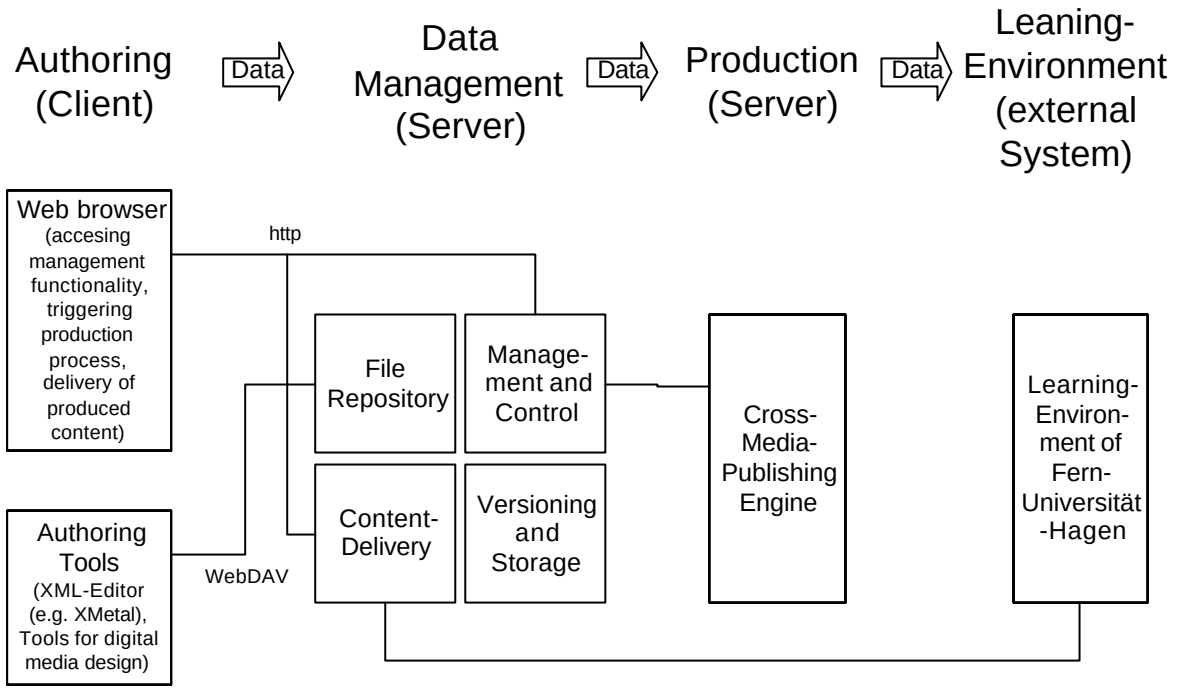

Fig. 1: System overview

The component architecture and the use of standard protocols facilitate development and maintenance. To support the heterogeneity of the IT-infrastructure at the FernUniversität Hagen, there is no commitment to use a specific tool or platform on the client side. This architectural choice improves sustainability. New trends and technologies on the client as well as on the server side can be integrated more easily.

Editing of the XML-documents is the most frequently performed authoring activity during course development. Hence usability of the XML-Editor is a key issue. The project team chose the commercial XMetaL-Editor that was specially developed for XML-content creation by users who do not have previous knowledge of XML. It was extended by style sheets to provide a view that hides the XML markup and that looks similar to the resulting HTML-representation. Also macros individually tailored to facilitate course creation were developed. Users are not constrained to use XMetaL and may use any other XML editor. However, special customisations for the FuXML DTD are currently only provided for XMetaL. It must be noted that the customisations provided for the editor are necessary to achieve good usability for the authors. Hence, before choosing a different or additional XML-Editor, the costs of customisation should be carefully examined.

Authoring of media files, such as graphics, animations and sounds, is done with tools appropriate for the application in question. The media files are linked to the XML documents by a file reference.

The files that constitute a course are managed in a central repository and can be accessed by the standardized WebDAV-Protocol [WebDAV]. On the client side directory based access to the files is provided by a WebDAV-Client. In Windows such a client is part of the standard distribution (Web folder feature). In addition applications that directly support the WebDAV-protocol (e.g. the XMetaL-Editor) make use of its locking facility to avoid concurrent access to documents.

The repository is part of the data management server that manages all information related to the development and maintenance process of the courses. For tasks related to authoring, e.g. triggering the production process, viewing the result or generating consistency information, as well as for system and project management tasks, a Web GUI is realized that can be accessed by a standard web browser.

The data management server implements a role system to facilitate fine grained access control to files and functionality. Versioning of the course files is provided on the level of single resources and complete courses. There are several functions for consistency control as described in section 5.3.

The production server implements the cross-media publishing process. At the moment a PDF and an HTML output can be produced for each course. A course can be published in the learning environment of the FernUniversität Hagen.

5.1 FuXML Documents

Since XML was chosen as a base technology for documents, a document type definition (DTD) that defines the structure of FuXML documents had to be developed. Existing DTDs were found to be not suitable in the context of German distance learning material. The DocBook DTD [DocBook] has too strong an emphasis on computer science and does not provide visual and didactical markup. The Educational Modelling Language (EML) [EML] is targeted towards describing complex pedagogical models but does not consider the editorial and didactical elements required by the course authoring process of the FernUniversität.

So its was a natural decision to develop a FuXML DTD that reflects the university's specific requirements. The FuXML DTD contains

- editorial elements like section, paragraph, quotation, etc.,
- didactical elements like course, course unit, learning objective, prerequisite, exercise, example, etc.,
- faculty-specific elements like for instance formula, theorem, proof, etc. for mathematicians and case, adjudication, etc. for the lawyers,
- multimedia elements like graphic, animation, video, simulation, audio, text object.

The editorial compartment also provides elements for visual markup, like bold, italic, underlined. In theory these elements could be replaced by semantic markup, in practise these elements are urgently required. Similarly, different types of paragraphs are supported that differ in the space they create at the end of the paragraph.

Multimedia elements can also be combined into compound multimedia objects. Such compound objects can be used to account for the different characteristics of the output media formats. For example, a compound object that contains an animation, a graphic and a text object can be transformed to an HTML page containing only the animation and a PDF page containing the graphic and the text object.

As already mentioned, for a LCMS deployed in the academic environment of a university the comprehensive support of mathematical expressions is of utmost importance. There are two problems in this area that have to be solved. The first problem is that of entering and displaying the formula in a user-friendly way in the XML-Editor. The second problem is that of displaying the mathematics in a high quality in the output documents. Although the

international standard MathML [MathML] has been created by the W3C to support mathematical notation in XML and XHTML documents, there are still some shortcomings which currently limit the applicability of MathML in practice. In 2002, when the FuXML project was started, a satisfying solution based on MathML could not be seen. Therefore, the decision to use LaTeX notation instead of MathML was taken. This has the advantage that existing mathematical and technical courses could be readily imported into the system. Recently however, the situation concerning MathML support in browsers and XML editors has improved, with standard browsers supporting MathML notation (current versions of Internet Explorer, Netscape, Mozilla) [WebMathML]. Also on the editor side the availability of commercial MathML plugins (see e.g. [MathFlow]) for the most popular XML editors like BlastRadius' XMetaL and Abortext's Epic is a major step towards the practicability of MathML. However, for documents containing a large number of mathematical expressions, the LaTeX notation will still be superior to MathML, because it does not need a cumbersome equation editor for even the simplest inline mathematical expressions.

FuXML allows the modular composition of courses from several independent XML documents. This is a first step towards the concept of learning objects. For example, this feature can be used to realize a master and a diploma course on the same topic. The common parts of the two courses can be provided by a single XML document. This can considerably reduce the maintenance overhead in comparison to the case that identical changes have to be made in two different documents.

5.2 The process of course development

Course development is divided into several phases: initiation, authoring, publication and maintenance.

Initiation

At first the course to be developed is analysed with regard to its requirements. The design templates and configuration parameters are set to accommodate the wishes of the course provider. This consultation work is done by the ZFE. The course and a project manager are registered in the system. The project manager performs further administrative actions like registering co-workers for specific activities or controlling publication deadlines.

If the course already exists as a Microsoft Word document the word import feature of the XMetaL-Editor can be used to generate an XML document master for the course. The master is particularly of good quality, if the standard word template provided by the ZFE is used.

An important issue to be addressed at the beginning of course development is the specification of the granularity of XML documents that capture the structural and textual contents of the course: Parts that are used in more than one course (e.g. bachelor and diploma course on the same subject) should be isolated in a separate document. To allow several authors to work simultaneously on the course, separation of course parts according to the individual responsibilities of the authors is necessary.

Authoring

After the design and the configuration parameters have been fixed (initiation phase) the course developers can make arbitrary changes within the content and can simply produce an updated version of the course by the press of a button. This procedure is a great improvement in comparison to the complex traditional course development process which required close

interaction of media and design experts with the authors of the course during the whole authoring process.

The main authoring activity is the editing of the XML-documents and the generation of media files (pictures, animations, etc.). If a developer makes a mistake or different co-workers make interfering changes, the system reports on the consistency issue and provides hints for solving the problem as described in section 5.3.

Publishing

In face of a publication deadline, authoring activities change from content creation to proofreading. Also the finishing editorial touch has to be added: if the automatic page layout does not deliver a satisfying result, line and page breaks can be inserted manually where necessary.

Currently, publishing of the print version is done by sending the printed course unit to the technical production for print and postal delivery to the students. Publishing in FernUniversität Hagen's virtual learning environment is much easier: the course can be directly imported.

Maintenance

After the course is delivered to the students, it resides in the system for later revision. If the course contains URLs and the URL-redirect option is activated (see section 5.3) URLs in published material can be replaced by a proper message or an alternative URL if they are no longer valid or the referenced content underwent a major change.

If the course provider decides to offer a new revision, the authoring-publishing-maintenance cycle restarts.

5.3 Centralized data management and consistency control

As mentioned before, the files that make up a course are held in a central repository accessed by the WebDAV-protocol. For each course there is a shared workspace accessed by all co-workers. If more than one course developer works on the course at the same time, managing concurrent access to resources becomes an issue.

Using WebDAV-protocol's locking facility a user can obtain an exclusive writing access to a resource. Neat integration of locking is provided by some authoring tools, e.g. the XMetaL-Editor: If a resource is opened a lock is acquired that prevents other users from opening the same resource for writing access. The lock is released when the resource is closed. This procedure can be regarded as on the fly check-out and check-in coupled with the traditional open and close operations. Locking allows simultaneous work on different course parts provided they are captured in different documents, and consecutive work on the same part without the risk of interference.

If some of the co-workers are temporarily or for a long time disconnected from the system (offline editing) they work on a local workspace. Then locking cannot be used and consolidation of the changes in the local workspace with the changes in the global shared workspace is performed afterwards by file based synchronization. To avoid lost updates in this situation the repository performs checks and prompts conflict resolution.

For consistency control of the global shared workspace consistency reports are available on the level of single XML documents and the complete course. Issues like a broken internal

link, duplicated IDs, an invalid or changed URL, a syntactically incorrect LaTeX fragment representing a mathematical formula or an invalid XML-document are reported and rated by relevance. Hints are given for resolving an issue. For changed resources referenced by URLs, a difference view is provided for change tracking.

To retain control over changes of URLs in courses that are already published a redirect service exists: in the course the original URL is replaced by one to a redirect server. A request to a replaced URL is normally redirected to the original one but in case of changes it is possible to display a proper message and link to another page.

The centralized data management with access control and consistency features described above shows clear advantages over the rigid provident style of interaction that collaborating users are constrained to when using a standard word processor.

5.4 Cross-Media Publishing

The cross-media publishing process produces the desired output documents in various formats. Currently, two major formats are available:

1. A PDF format optimised for printing.
2. An HTML format targeted for viewing the course material on screen.

With these two formats the three major publication channels of the FernUniversität can be utilised (see Fig. 2).

- The traditional offline paper version is realized by sending the printed manuscript or the electronic PDF file to the technical production facility of the FernUniversität. Here the course material is printed, packaged, and sent to the students by conventional mail.
- An offline electronic version can be published on CDROM. As mentioned in section 5 the production process can replace the URLs contained in the course text with URLs linking to a redirect web service.
- An online electronic version can be published via the university's LMS. The production component produces a ZIP file which can be imported by the LMS.

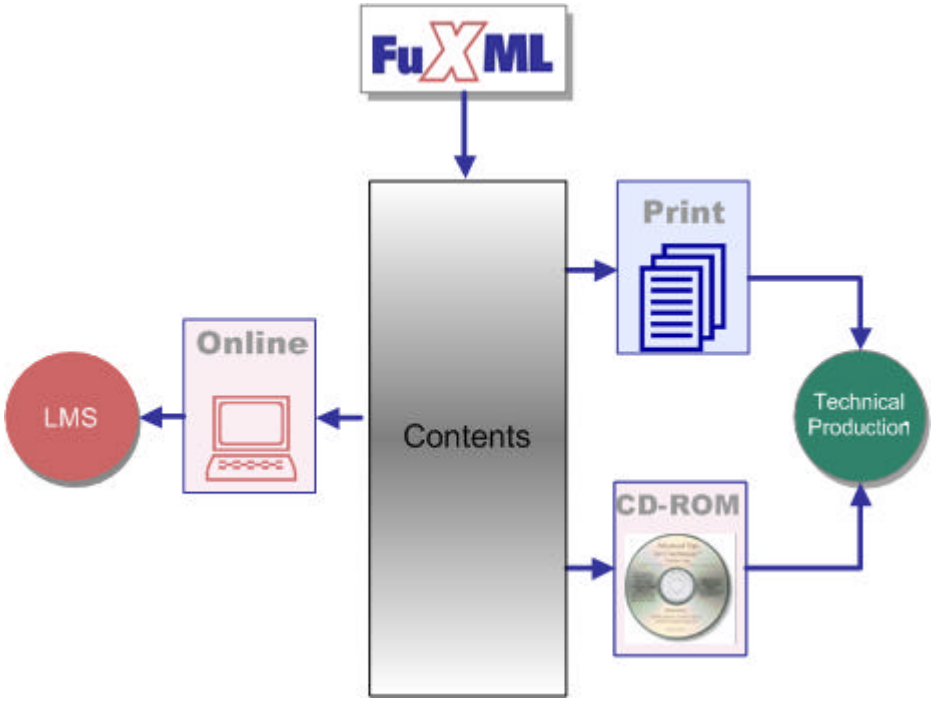

Fig. 2: Primary production channels

An important requirement regarding the production system was, that it should support different design and layout variants for each course. Additionally, the cost of the solution should be moderate. This demand for an efficient solution excluded time-consuming programming of individual transformation for each course. The chosen approach is depicted in Fig. 3. It consists of a rule-based production engine and a design interface which is realised in form of various configuration files.

These configuration files are edited by the configuration managers and designers of a course, not the authors themselves. The design interface provides a way to customise all relevant attributes of a course, e.g.

- the labelling and numbering scheme for sections and other elements like figures, tables, examples,
- the layout and structure of the different XML elements like animation, paragraph, definition, axiom, quotation, exercise,
- appearance and layout of table of contents, index, table of figures, etc.

Furthermore, the output documents to be produced can be configured individually for each XML document. For instance, a single XML document containing the text of a course unit may be transformed into three different output documents: the course unit text for the students, the assignments for the student, the assignments with solutions for the teacher.

The production engine for the PDF format is based on the LaTeX typesetting system, thus the required high quality print version can be achieved. Multimedia elements which cannot be displayed on paper can be represented by textual content, e.g. a note to the students that additional multimedia material is available in the HTML version of the course.

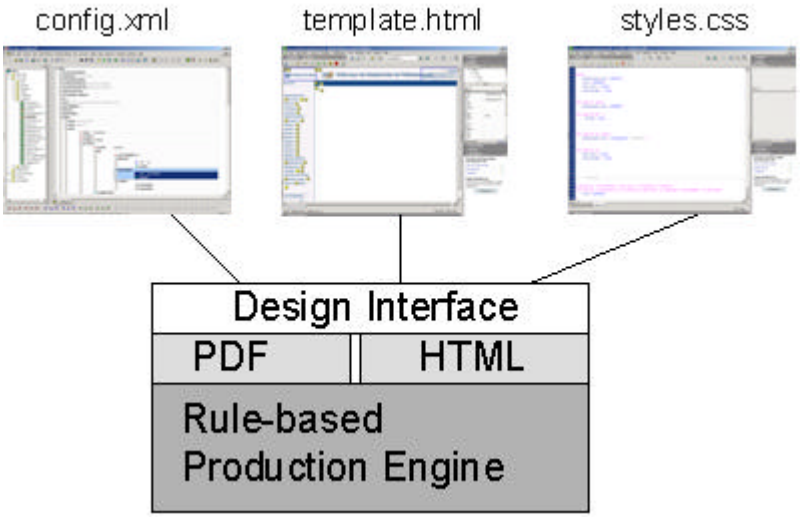

Fig. 3: The production system

6 Integration of the System within the Organisational and Technical Infrastructure of the University

The realisation of sustainability and longevity in e-learning projects at universities has always been a big problem. Many projects do not survive even a couple of months after the official end of the project. The biggest problem is that of finding adequate personal that is able to operate and maintain the system and provide user support over a longer period of time. The members of the project teams are often not available and are assigned to new tasks or leave the university after finishing their academic work. The central institutions of a university like the computer centre or media centre normally do not have the required resources to take over every completed e-learning project.

In the case of the FuXML project the following issues had to be solved:

1. Adequate user support has to be offered. Furthermore, since the employed authoring technology is new, at least rudimentary training and tutorials for university staff have to be provided.
2. Availability and stability of the system has to be guaranteed.
3. Knowledge transfer from the project team to the organisational units has to take place to save the experiences made in the project.
4. Software maintenance and further developments have to be assured after the end of the project.

For the first two issues solutions have been found which required commitments from the media centre and the computer centre. Since the ZFE was part of the project from the very beginning, the third issue is also addressed adequately. Although some knowledge will inevitably be lost when the project team breaks up, the major insights and achievements will be saved for the university.

For the fourth issue a satisfactory solution still has to be found. A solution proposed by the CampusSource [CampusSource] initiative is to make the code open source and let the

software support be provided by small and medium-sized businesses. This approach has already shown to be successful for other e-learning projects.

At the moment, FuXML is introduced to university personal of several departments. Especially the departments with no technical background show great interest for the system.

7 Summary and Conclusions

In this paper we described the approach taken by the FernUniversität Hagen to establish a novel organisational and technical procedure for developing course material in the field of distance learning. The university decided to build its own LCMS within the internally funded FuXML project. The requirements of the university towards a content management system were discussed and essential parts of the developed system were sketched out. Practical considerations and experiences as well as the integration of the system within the organisational and technical infrastructure of the university were presented.

In conclusion it can be stated, that the technical feasibility of the new approach was successfully demonstrated. The future acceptance of the system within the university will largely depend on the political and financial decisions taken by the university's administration with regard to further supporting actions.

References

[FuXML] Home page of the FuXML project, http://www.fuxml.de/

[DocBook]The DocBook DTD, http://www.docbook.org/

[EML] Educational Modelling Language (EML), http://eml.ou.nl/

[MathML] Mathematial Markup Language, http://www.w3.org/Math/

[WebMathML] Putting mathematics on the Web with MathML, http://www.w3.org/Math/XSL/

[OCW] MIT's OpenCourseWare project, http://ocw.mit.edu/

[MathFlow] MathFlow MathML Editor, http://www.dessci.com/en/products/mathflow/

[CampusSource] The CampusSource initiative, http://www.campussource.de/org/

[WebDAV] Web-based Distributed Authoring and Versioning, http://www.webdav.org/