

Eine Serviceplattform für das wissenschaftliche elektronische Publizieren

Uwe Müller
u.mueller@cms.hu-berlin.de

Der Weg zur elektronischen Publikation

Der Publikationsprozess im Allgemeinen und damit auch die technologische und organisatorische Unterstützung, die für Veröffentlichungen auf dem Dokumenten- und Publikationsserver der Humboldt-Universität – kurz: edoc-Server – entwickelt wurde, beginnen bereits mit der Erstellung einer wissenschaftlichen Arbeit. Auch wenn es in dieser Phase scheinbar nur um die inhaltlichen Belange einer Veröffentlichung geht, wird schon beim Schreiben eines Textes der Grundstein für die technische Qualität der späteren elektronischen Publikation gelegt. Für den edoc-Server wurde der Anspruch formuliert, die dort publizierten Arbeiten in einer hohen technischen Qualität zur Verfügung zu stellen [1]. Diese drückt sich unter anderem darin aus, dass die Dokumente über längere Zeiträume nicht nur physisch bereitgestellt werden können, sondern auch les-

bar bleiben. Davon wird später noch die Rede sein (siehe Abschnitt *Stabilität*).

Für die Konzeption des edoc-Servers und des Publikationsprozesses ergab sich aus der Anforderung der Langzeitverfügbarkeit der Dokumente die Schlussfolgerung, XML als zentrales Archivformat einzusetzen [2]. Als offener und firmenunabhängiger Standard [3] bietet der Einsatz dieses für Menschen lesbaren Formats gute Voraussetzungen dafür, dass die derart kodierten Informationen wahrscheinlich auch in 50 oder 100 Jahren noch problemlos interpretiert und dargestellt werden können. Ein weiterer Vorteil der Nutzung von XML ergibt sich aus der Fähigkeit des Formats, Strukturinformationen zu speichern. Damit können der logische Dokumentaufbau (Kapitel, Aufzählungslisten, Tabellen usw.) und sogar semantische Einheiten (z. B. Personennamen, geografische Bezeichnungen, Definitionen, Fachbegriffe) maschinenlesbar vorgehalten und entsprechend weiterverarbeitet werden

Wie wird eine Veröffentlichung zur elektronischen Publikation? Und auf welchem Wege kommt ein mit einem Textverarbeitungssystem erstelltes Dokument auf den Dokumenten- und Publikationsserver? Auch wenn diese Fragen zunächst trivial erscheinen und danach aussehen, über ebenso einfache Antworten zu verfügen, lohnt es sich, genauer hinzuschauen. Denn weder liegen die Lösungen auf der Hand noch sind sie zwangsläufig und allein gültig. Dieser Artikel beschreibt den derzeit eingesetzten Publikationsprozess und gibt einen Ausblick auf zukünftige Entwicklungen.

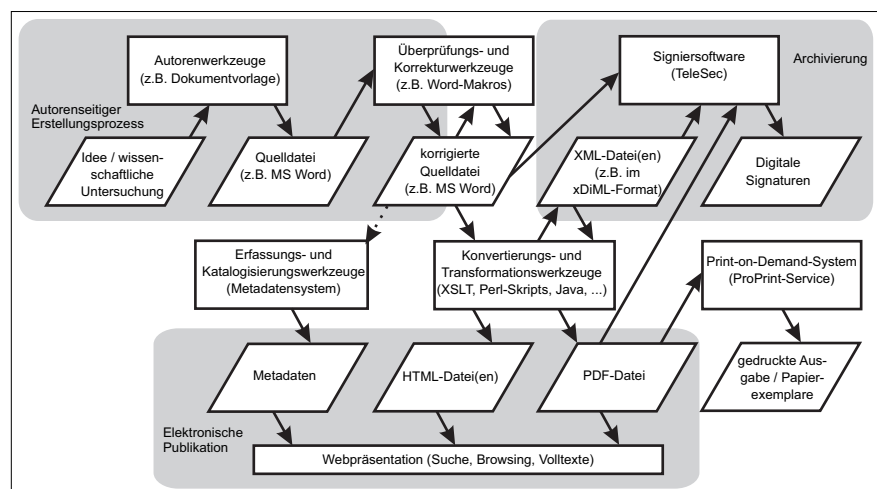


Abb. 1: Schematische Darstellung des Publikationsprozesses für den edoc-Server.

(siehe Abschnitt *Verfügbarkeit und Sichtbarkeit*).

In Abb. 1 wird der Publikationsprozess, wie er für den edoc-Server realisiert ist, dargestellt. Er beginnt beim Autor und führt über diverse Einzelschritte, die in den folgenden Teilabschnitten näher beleuchtet werden sollen, nicht nur zur fertigen elektronischen Publikation, sondern bei Bedarf auch zu einer gedruckten Ausgabe der Veröffentlichung. Die einzelnen Schritte sind dabei als modellhafte Beispiele zu verstehen, die in dieser Form für den edoc-Server realisiert worden sind, jedoch keineswegs die einzig möglichen Lösungen sind.

Erstellen

Wie bereits im vorherigen Abschnitt erwähnt, schafft der Autor bereits während der Erstellungsphase die entscheidenden Voraussetzungen für eine erfolgreiche elektronische Publikation. Dabei kommt es unter anderem darauf an, die der Arbeit innewohnenden und für den menschlichen Betrachter ohne weiteres erkennbaren strukturellen Merkmale so zu kennzeichnen, dass sie auch maschinell ausgewertet werden können. In den gängigen Textverarbeitungssystemen kann dies mit so genannten Formatvorlagen realisiert werden, die Absätze oder einzelnen Zeichen zugeordnet werden können. Abgesehen von diesem für den Autor nicht unmittelbar erkennbaren Nutzen, der sich erst aus der weiteren Verarbeitung des Dokuments ergibt, wird das Abfassen größerer wissenschaftlicher Arbeiten durch die Verwendung von Formatvorlagen erheblich vereinfacht. Als Beispiele seien hier die komfortable und einheitliche Gestaltung des Layouts der Arbeit und die automatische Nummerierung von Überschriften, Tabellen und Abbildungen sowie die Erstellung entsprechender Verzeichnisse genannt. Insofern ist die Nutzung von Dokumentvorlagen nicht erst für die Vorbereitung elektronischer Publikationen, sondern auch für das herkömmliche Publizieren sinnvoll.

Formatvorlagen lassen sich in OpenOffice oder Microsoft Word problemlos selbst erstellen und mit entsprechenden Eigenschaften (z. B. Schriftgröße,

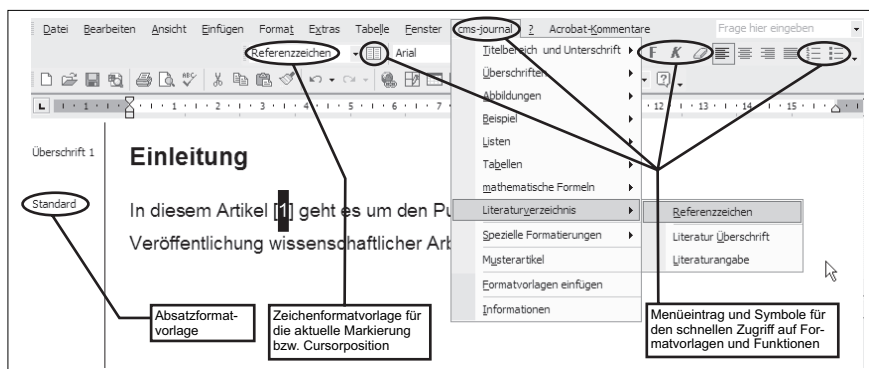


Abb. 2: Erstellen eines Zeitschriftenartikels mit einer Dokumentvorlage.

-schnitt und -art, Zeilen- und Absatzabstand, Einzüge) versehen. Für die sinnvolle Weiterverarbeitung – vor allem für die anschließende Konvertierung nach XML – müssen Name und Bedeutung der Formatvorlagen jedoch eindeutig festgelegt sein. Daher werden für Autoren unterschiedliche Dokumentvorlagen bereitgestellt¹, die neben einem umfangreichen Satz an Formatvorlagen weitere Hilfsmittel enthalten: Eigene Menüeinträge, Makros und eine einfache Literaturverwaltung. Außerdem stehen eine umfangreiche Beschreibung, eine FAQ-Seite und zahlreiche Hinweise für das Erstellen wissenschaftlicher Publikationen – z. B. für den Umgang mit Abbildungen – zur Verfügung. Auch für das Arbeiten mit dem Satzsystem LaTeX [4] und für die Erstellung von PDF-Dateien wurden Vorgaben und Empfehlungen entwickelt.

Im Idealfall fasst der Autor bereits zu Beginn der Erstellungsphase seiner Publikation die elektronische Veröffentlichung ins Auge und arbeitet entsprechend von Anfang an mit der für den jeweiligen Publikationstyp vorgesehenen Dokumentvorlage. In der Realität ist dies allerdings eher die Ausnahme. Die Nachbearbeitung eines bereits teilweise oder vollständig beendeten Dokuments ist etwas aufwendiger, jedoch ohne weiteres möglich.

Außer für Dissertationen und Habilitationsschriften konnten bislang Dokumentvorlagen für Konferenzbeiträge, Öffentliche Vorlesungen und Zeitschriftenartikel bereitgestellt werden. Abb. 2 zeigt den Einsatz der Dokumentvorlage für MS Word, die bei der Erstellung von Artikeln für das cms-journal zum Ein-

satz kommt [5]. Zu sehen sind neben den Anzeigen der verwendeten Formatvorlagen ein zusätzliches Menü sowie Symbole, deren Funktionen von der standardmäßigen Word-Einstellung zum Teil abweichen.

Überprüfen und Einreichen

Die bloße Verwendung der korrekten Dokumentvorlage bietet selbstverständlich keine Gewähr dafür, dass eine Datei auch tatsächlich die technischen Voraussetzungen für eine adäquate elektronische Publikation erfüllt. Die falsche oder unzureichende Verwendung der vorgegebenen Formatvorlagen, die Möglichkeit, eigene Formatvorlagen zu definieren und das unsachgemäße Einbinden von Grafiken stellen nur eine Auswahl möglicher Fehler dar.

Um Dokumente auf ihre Korrektheit hin zu überprüfen und entdeckte Probleme aufzuzeigen und unter Umständen automatisch zu beheben, wurde eine Reihe von Werkzeugen entwickelt, die diese Aufgabe erfüllen. Dazu zählen neben einer umfangreichen Checkliste, die für unterschiedliche Quellformate zur Verfügung steht, vor allem Makros, die als Teil einer speziellen Dokumentvorlage in das Textverarbeitungssystem geladen werden können. Zu diesem Bereich zählt auch die Überprüfung von PDF-Dateien, die – je nach Vereinbarung – bereits durch den Autor erstellt wurden. Vor allem die korrekte Einbettung aller Schriftarten und Zeichensätze ist hierbei von besonderer Bedeutung, da-

¹ siehe z. B. unter http://edoc.hu-berlin.de/e_autoren/

mit die Dokumente langfristig und weitgehend unabhängig von der Systemumgebung angezeigt werden können.

Wie aus der Darstellung in Abb. 1 ersichtlich ist, liegt die Überprüfung und Korrektur der erstellten Dokumente teilweise auf der Autorensseite. Nach der Einreichung der Arbeit zur Publikation auf dem edoc-Server wird die Arbeit nochmals kontrolliert. Die Korrekturen werden dann je nach Vereinbarung durch den Autor oder durch die Mitarbeiter des edoc-Servers vorgenommen. Das Einreichen der Quelldateien, die für die elektronische Veröffentlichung vorgesehen sind, erfolgt über eine Webschnittstelle, über die der Autor außerdem einige Angaben zu seiner Publikation machen muss – z. B. den Titel der Arbeit, den oder die Autorennamen, Schlagwörter und Abstracts.

Konvertieren und Signieren

Wie bereits erwähnt, werden die auf dem edoc-Server zu veröffentlichenden Werke in der Regel zunächst nach XML überführt. Damit werden nicht nur die Voraussetzungen für eine langfristige Verfügbarkeit, eine qualifizierte Recherche und die konsistente Speicherung von Strukturinformationen geschaffen. XML zeichnet sich vor allem dadurch aus, dass es flexibel und mit transparenten, standardisierten Methoden in unterschiedliche Präsentationsformate gewandelt werden kann. Es bildet somit den idealen Ausgangspunkt für verschiedene Anzeigeformate eines Dokuments – z. B. HTML und PDF. Die Konvertierung nach XML und in die unterschiedlichen Präsentationsformate wird durch eine Reihe von Skripten und Programmen realisiert, die von der Arbeitsgruppe *Elektronisches Publizieren* entwickelt wurden. Mit diesem Thema beschäftigt sich der Artikel auf Seite 29 in diesem Heft ausführlich [2].

Einen weiteren Arbeitsschritt im Zusammenhang mit der elektronischen Veröffentlichung bildet die Erstellung von digitalen Signaturen zur Sicherung der Authentizität und Integrität der Dokumente. Diese werden unmittelbar nach der Einreichung der Originaldateien bzw. nach der Erstellung der Archivie-

rungsformate erzeugt. Mithilfe dieser Daten kann nachgewiesen werden, dass es sich bei den im Internet verfügbaren Dateien um die echte und unveränderte Publikation handelt. Im Hinblick auf die sehr leichte Manipulierbarkeit in digitaler Form vorliegender Ressourcen ist dieser Nachweis für die Akzeptanz der elektronischen Publikation von besonderer Bedeutung. Für die technische Umsetzung des Signaturprozesses, der so genannte qualifizierte digitale Signaturen im Sinne des bundesdeutschen Signaturgesetzes erzeugt, wird ein entsprechender Dienst der Firma TeleSec genutzt.

Katalogisieren

Relativ unabhängig von der technischen Verarbeitung einer (elektronischen) Veröffentlichung bildet die Erschließung von Publikationen einen eigenen Themenbereich. Der gestrichelte Pfeil in Abb. 1 deutet darauf hin, dass das Dokument zwar als Vorlage für die Katalogisierung dient, die Metadaten jedoch nicht allein mit technologischen Hilfsmitteln erzeugt werden können.

Analog zu einer Bibliothek werden die auf dem edoc-Server gespeicherten Veröffentlichungen erst über die Metadaten zugänglich und auffindbar. Anschaulich wird dies, wenn man bedenkt, dass die Listen der Veröffentlichungen, die auf den Webseiten des edoc-Servers nach unterschiedlichen Kriterien zur Verfügung stehen, dynamisch aus der mit den Metadaten gefüllten Datenbank generiert werden. Die Verwaltung der Metadaten auf dem edoc-Server wird ausführlich im Artikel ab Seite 43 besprochen [6].

Drucken

Auch im digitalen Zeitalter ist das Papier bekanntlich nicht überflüssig geworden. Es gibt unterschiedliche Gründe, aus denen eine in elektronischer Form erschienene Veröffentlichung auch als Druckexemplar hergestellt werden soll. Viele Publikationen erscheinen beispielsweise regelmäßig parallel als Onlinefassung und in gedruckter Form. Dazu zählen zum Beispiel das cms-journal, aber auch die Öffentlichen Vorlesungen der Humboldt-Universität. Zum anderen neigen

nicht wenige Menschen dazu, wissenschaftliche Arbeiten lieber auf dem Papier als am Bildschirm zu studieren. Wem der Ausdruck loser Blattsammlungen auf dem Arbeitsplatzdrucker zu aufwendig oder unkomfortabel ist, dem steht mit ProPrint ein ausgereifter Print-on-Demand-Service zur Verfügung, über den alle auf dem edoc-Server veröffentlichten Werke angeboten werden können [7]. Somit bringt der Nutzer gewissermaßen dynamisch den Publikationsprozess selbst zum Abschluss.

Daneben erfordern auch rechtliche Regelungen zuweilen die Herstellung von Papierexemplaren: Beispielsweise wird gemäß der gültigen Promotions- und Habilitationsordnungen der einzelnen Fakultäten im Falle der elektronischen Veröffentlichung der Prüfungsarbeiten die zusätzliche Bereitstellung von vier Pflichtexemplaren in gedruckter und gebundener Form verlangt. Sie dienen als Belegexemplare in der Universitätsbibliothek sowie in der Deutschen Bibliothek.

Das Serviceangebot

Der im vorangegangenen Abschnitt skizzierte Publikationsprozess mit den jeweils angedeuteten Variationsmöglichkeiten veranschaulicht die Verarbeitung wissenschaftlicher Arbeiten, die auf dem edoc-Server veröffentlicht werden. Auch wenn dies der Name *edoc-Server* nicht unbedingt nahe legt, handelt es sich bei der durch die Arbeitsgruppe *Elektronisches Publizieren* betreuten Publikationsplattform keineswegs lediglich um die Ansammlung technischer Komponenten. Vielmehr wird ein umfangreicher *Service* rund um das elektronische Publizieren zur Verfügung gestellt, der neben der Beratung von Autoren und Herausgebern die Bereitstellung von Autorenwerkzeugen und den entsprechenden Support enthält und außerdem die umfassende Verfügbarkeit und Sichtbarkeit der Dokumente sowie die Robustheit des Gesamtsystems gewährleistet. Diesen *Service* zu organisieren und für einen größeren Kreis von Autoren und Herausgebern zu etablieren, ist unter anderem das Ziel des DFG-geförderten Projektes

SCOPE – Service Core for Open Publishing Environments. Innerhalb dieses Projektes werden modulare Technologien und entsprechende Werkzeuge entwickelt, die die unterschiedlichen Prozesse realisieren, die beim elektronischen Publizieren eine Rolle spielen. Das individuell anpassbare Serviceangebot steht neben Wissenschaftlern und Arbeitsgruppen an der Humboldt-Universität auch anderen Forschungseinrichtungen – beispielsweise am Wissenschaftsstandort Berlin-Adlershof – zur Verfügung.

Beratung

Während beispielsweise für Dissertationen oder auch für das *cms-journal* der technologische und organisatorische Publikationsprozess weitgehend feststeht und Autoren demzufolge ohne weiteres mit den notwendigen Informationen und Werkzeugen versorgt werden können, müssen bei noch nicht auf dem edoc-Server etablierten Publikationsreihen zunächst die Rahmenbedingungen und Anforderungen abgestimmt werden. Dabei sind unter anderen folgende Fragen zu klären:

- Welche Textverarbeitungssysteme kommen zum Einsatz?
- Welche grundsätzliche Struktur haben die einzelnen Publikationen und wie sind sie in größere Einheiten (z. B. Konferenzbände, Zeitschriftenausgaben) eingeordnet?
- Wer sind die Autoren bzw. wie werden die Beiträge ausgewählt?
- Welche Publikationsformen sollen angeboten werden (HTML, PDF, Druckausgabe)?
- Welche rechtlichen Regelungen gelten für die elektronischen Publikationen?
- Wie soll der Publikationsprozess organisiert sein und wer ist jeweils für die einzelnen Arbeitsschritte zuständig?

Hinzu kommen organisatorische Überlegungen wie die Begutachtung von Arbeiten. Die Mitarbeiter der Arbeitsgruppe *Elektronisches Publizieren* helfen bei der Beantwortung der Fragen und schlagen für den Publikationsprozess entsprechende Komponenten und Verfahren vor.

Autorenwerkzeuge und Support

Die Unterstützung von Autoren und Herausgebern zählt zu den wesentlichen Aufgaben, denen sich das Projekt SCOPE verschrieben hat. Die prinzipielle Funktionsweise dieser Werkzeuge wurde bereits weiter vorn beschrieben. Neben Dokumentvorlagen werden auch PlugIns und Konfigurationsdateien (Styles) für die Integration von Literaturdatenbanken wie etwa *Reference Manager* und *Endnote* in Textverarbeitungssoftware zur Verfügung gestellt. Im Vordergrund steht analog zur Verwendung von Formatvorlagen in Microsoft Word die Übersetzung der einzelnen Strukturelemente von Literatureinträgen (Namen der Autoren, Titel, Erscheinungsjahr, Seitenangabe usw.) in die entsprechenden XML-Elemente des Dokumentenmodells. Gleichzeitig behält der Autor dabei seine gewohnte und komfortable Software zur Literaturverwaltung.

Abgesehen von technischen Hilfsmitteln zur Erstellung elektronischer Publikationen unterstützt die Arbeitsgruppe *Elektronisches Publizieren* die Autoren auch mit Beschreibungen, Hilfesystemen, einer Support-Hotline (per E-Mail und telefonisch) sowie mit regelmäßig stattfindenden Kursen zur Autorenschulung.

Verfügbarkeit und Sichtbarkeit

Wesentliches Merkmal einer Veröffentlichung ist natürlich deren Verfügbarkeit. Im wissenschaftlichen Bereich dienen Publikationen vor allem der effizienten Kommunikation mit Fachkollegen, sie besitzen jedoch noch weitere Funktionen [8]. Im Vergleich zu traditionellen Veröffentlichungsverfahren zeichnet sich das elektronische Publizieren dadurch aus, sehr schnell eine weltweite Verfügbarkeit zu ermöglichen. Das bloße Erstellen einer Webseite mit dem entsprechenden Inhalt ist dafür allerdings meist nicht hinreichend.

Neben der reinen Webpräsentation spielen für die Auffindbarkeit der Publikationen weitere Aspekte eine Rolle. Dazu zählt, dass alle Dokumente unmittelbar verlinkt sind und nicht ausschließlich über Suchmasken erreicht werden können. Nur so können die Websuch-

maschinen – und damit ein nicht unmittelbar auf den Seiten des edoc-Servers recherchierender Nutzer – die Publikationen überhaupt finden. Im konkreten Fall des edoc-Servers heißt dies, dass – beginnend von der Startseite <http://edoc.hu-berlin.de/> – alle Publikationen durch das alleinige Verfolgen von Hyperlinks erreichbar sind.

Abgesehen von dieser »Suchmaschinenfreundlichkeit« bereitet der edoc-Server die Metadaten aber auch gezielt für nationale und weltweite Portale und *Service Provider* auf und stellt sie über entsprechende Schnittstellen zur Verfügung. Dabei kommt dem OAI-Protokoll eine besondere Bedeutung zu, mithilfe dessen sämtliche verfügbaren Metadaten abgeholt werden können (siehe Seite 55 [9]). Durch eine Publikation auf dem Dokumenten- und Publikationsserver werden Veröffentlichungen somit nicht nur über gängige Suchmaschinen, sondern auch über zahlreiche Nachweisinstrumente gefunden, die sich speziell mit wissenschaftlichen Inhalten oder mit einer Fachrichtung befassen.

In den Zusammenhang des Datenexports gehören auch die bereits erwähnte Möglichkeit, edoc-Dokumente über einen Print-on-Demand-Dienst anzubieten, sowie die gerade fertig gestellte Schnittstelle zum Lehr- und Lernmanagementsystem *moodle* [10]. Aus diesem an der Humboldt-Universität inzwischen gern genutzten Teil des Medienportals können edoc-Veröffentlichungen direkt gesucht und als Kursmaterialien für Lehrveranstaltungen zur Verfügung gestellt werden.

Von besonderer Bedeutung ist natürlich auch die Navigierbarkeit auf der Weboberfläche des edoc-Servers selbst. Dank der konsistenten Erfassung von Metadaten und der Speicherung der Publikationen in XML konnten umfangreiche Möglichkeiten des Browsings realisiert werden: Die einzelnen Veröffentlichungen lassen sich in Listen nach Publikationstyp, Erscheinungsdatum, Autorennamen, Einrichtungen und nach fachlichen Kriterien anzeigen und durchblättern. Die eigentlichen Volltexte sind jeweils von einer Startseite erreichbar, die die wichtigsten bibliografischen Daten enthält. Innerhalb einer Publikation

bestehen wiederum weit reichende Navigationsmöglichkeiten, die ein schnelles Wechseln zwischen Abschnitten und zu Abbildungen, Tabellen, Fußnoten, Seiten und Literaturverweisen erlauben.

Stabilität

Eines der wichtigsten Merkmale des durch die Arbeitsgruppe *Elektronisches Publizieren* angebotenen Publikationsservice stellt dessen Stabilität dar. Diese Eigenschaft, die den edoc-Server von den allermeisten dezentral administrierten Systemen (Webseiten, Institutsdatenbanken und dergleichen) abhebt, wird durch die organisatorische Einbettung der Arbeitsgruppe sowie durch die technologische Fundierung des Publikationsprozesses gewährleistet. Sie ist ein wesentliches Kriterium für den Betrieb von Dokumenten- und Publikationsservern und begründet den bibliothekarischen Anspruch der Veröffentlichung elektronischer Dokumente.

Die durch den edoc-Server garantierte Stabilität bezieht sich auf die folgenden Bereiche:

- **Organisatorische Nachhaltigkeit.** Der Betrieb des edoc-Servers ist langfristig gesichert, da er in den Computer- und Medienservice und die Universitätsbibliothek eingegliedert ist und entsprechende personelle Ressourcen für die Wartung und Pflege sowie für den laufenden Betrieb zur Verfügung stehen. Der Service kann unmittelbar auf das an diesen Zentraleinrichtungen vorhandene bibliothekarische und technische Know-how zurückgreifen und ist nicht von Einzelpersonen abhängig.
- **Quality of Service.** Dank der Einbettung des edoc-Servers in den Computer- und Medienservice kann eine hohe Ausfallsicherheit und damit eine hohe Verfügbarkeit gewährleistet werden. Die technischen Grundsysteme befinden sich jeweils auf dem aktuellen Stand. Der edoc-Server ist in die zentrale Datensicherung eingebunden.
- **Langzeitverfügbarkeit.** Durch die bevorzugte Speicherung der elektronischen Dokumente im offenen und firmenunabhängigen Standardformat XML werden die Voraussetzungen für eine spä-

tere Lesbarkeit und Interpretierbarkeit der Daten geschaffen. Auch die nicht als XML gespeicherten Publikationen unterliegen vor der Veröffentlichung auf dem edoc-Server einer strengen Qualitätskontrolle. Dies gilt beispielsweise für die Einbettung der Schriftarten in PDF-Dateien.

- **Authentizität und Integrität.** Wie beschrieben, werden alle eingehenden Veröffentlichungen sowie die archivierten Formatversionen (in der Regel XML und PDF) mit einer digitalen Signatur versehen, mit denen Echtheit und Unversehrtheit der Publikationen nachgewiesen werden können. Außerdem werden auf dem edoc-Server für alle veröffentlichten Formatversionen URNs² vergeben und gemeinsam mit den URLs an *Die Deutsche Bibliothek* gemeldet. Damit besteht in Form eines weltweit eindeutigen Schlüssels ein vom aktuellen technischen System unabhängiger Zugang zu den einzelnen Dokumenten. Im Gegensatz zu Internetadressen (URLs) bleibt dieser beispielsweise auch bei Umstrukturierungen des Servers erhalten. URNs können über einen Resolver³ oder nach der Installation eines entsprechenden Plugins unmittelbar durch Eingabe in die Adresszeile des Browsers aufgelöst werden und führen zur aktuell gültigen URL.

Weiterentwicklungen und Ausblick

Der für Hochschulschriften und exemplarische Publikationsreihen bereits etablierte Service soll im Rahmen des Projekts SCOPE ausgeweitet und allen inneruniversitären Einrichtungen sowie wissenschaftlichen Institutionen außerhalb der Humboldt-Universität angeboten werden. Im besonderen Blickfeld steht dabei der Wissenschaftsstandort Adlershof. Allerdings gibt es auch darüber hinausgehende Kooperationsvereinbarungen. Die momentanen Aktivitäten innerhalb des Projekts konzentrieren sich auf drei Schwerpunkte, die im Folgenden kurz umrissen werden: Die Entwicklung einer Komponentenverwaltung, die Realisierung und Integration erweiterter Workflows und die Unterstützung von Open-Access-Aktivitäten.

Abstraktes Komponentenmodell

SCOPE hat es sich unter anderem zur Aufgabe gemacht, die bereits vorhandenen Publikationsprozesse zu verallgemeinern und den Service rund um das elektronische Publizieren einem größeren Kreis von Autoren und Herausgebern anzubieten. Dabei stellte es sich als notwendig heraus, die inzwischen reichhaltig vorhandenen Autorenwerkzeuge, Konvertierungsskripte, Dokumentmodelle, Styles und dergleichen mit einer formalen Beschreibung zu versehen, um deren Verwaltung, Wartung und Pflege zu erleichtern.

Als kleinste logische Einheit in einem Publikationsprozess wurden dazu so genannte *Publikationskomponenten* definiert, die bestimmte Arbeitsschritte realisieren und entsprechend hauptsächlich durch ihre Funktion charakterisiert sind. Alle in der Abb. 1 gezeigten rechteckigen Kästchen werden als Publikationskomponenten interpretiert. Sie können durch unterschiedliche Softwarelösungen – oder beispielsweise auch durch die Schilderung einer manuellen Tätigkeit – implementiert werden.

Die formalen Beschreibungen der Publikationskomponenten sowie deren Versionen und Abhängigkeiten untereinander werden in einer Datenbank verwaltet und bilden somit die Grundlage für ein noch in der Entwicklung befindliches Managementsystem. Dieses dient in erster Linie der Unterstützung des Service- und Beratungsangebots, das durch SCOPE entwickelt wird. Denn ausgehend von gegebenen Anforderungen und Rahmenbedingungen werden sich mit der Hilfe dieses Systems ganze Publikationsprozesse einschließlich aller erforderlichen Komponenten ableiten lassen. Eine vereinfachte Darstellung des Modells findet sich in Abb. 3.

Das Schema in Abb. 3 geht von einem abstrakten Dokumentbegriff aus. Demnach ist ein Dokument ein abgeschlossenes ideelles Werk, das jedoch nicht an eine bestimmte Erscheinungsform (z. B. ein Buch oder eine Datei) gebunden ist. Vielmehr kann es in unterschiedlichen

² Uniform Resource Names

³ z. B. <http://nbn-resolving.de/>

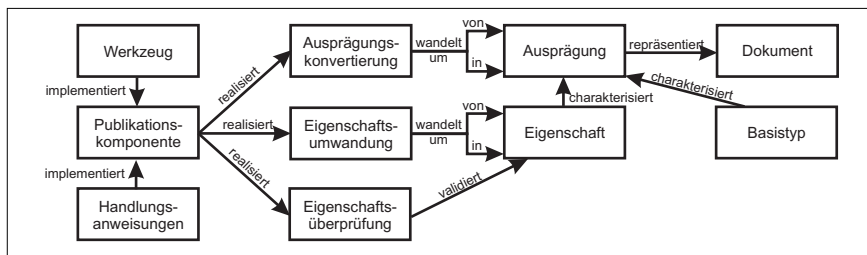


Abb. 3: Formale Beschreibung von Publikationskomponenten (vereinfacht).

Ausprägungen auftreten. Diese sind ihrerseits jeweils durch einen Basistyp – im Falle von Dateien ist dies das Dateiformat – sowie durch weitere Eigenschaften charakterisiert. Zu solchen Eigenschaften zählen beispielsweise die Verwendung von Dokumentvorlagen, das Vorhandensein von Abbildungen in einer Datei und das Papierformat. Das Wesentliche einer Publikationskomponente besteht in deren Fähigkeit, einzelne Eigenschaften oder die gesamte Ausprägung von Dokumenten umwandeln zu können oder eine Eigenschaft einer Ausprägung zu überprüfen.

Erweiterter Workflow

Veröffentlichungsverfahren zeichnen sich unter anderem dadurch aus, dass sie durch eine Vielzahl wiederkehrender Arbeitsschritte, die von unterschiedlichen Personengruppen durchgeführt bzw. veranlasst werden, relativ stark strukturiert sind. Dazu zählen neben dem rein technologischen Publikationsprozess, wie er in Abb. 1 skizziert wurde, auch organisatorische Workflows wie Begutachtungsverfahren bei Zeitschriften, Tagungen und Konferenzen.

Dieser Aspekt wird durch den Einsatz von Workflowsystemen wie das an der Humboldt-Universität entwickelte Konferenzsystem [11] und das Begutachtungssystem GAPWorks, das im Rahmen des DFG-Projekts GAP – German Academic Publishers⁴ – entwickelt wurde, berücksichtigt. Sie werden in nächster Zeit in das Publikationssystem integriert und Herausgebern als zusätzliche Optionen zur Verfügung gestellt. Einen ersten Schritt bildet eine neue Webschnittstelle, die es Herausgebern erlaubt, ihre eigenen Publikationsreihen auf dem edoc-Server autonom zu verwalten, die Metadaten zu erfassen und die

Dokumente einzuspielen. Diese wird bis zum Ende des Jahres 2005 fertig gestellt sein.

Aufbauend auf das im vorhergehenden Abschnitt erwähnte Managementsystem befindet sich außerdem eine Workflowkomponente in der Entwicklung, mit der der technische Publikationsprozess modelliert und gesteuert wird. Damit wird es zukünftig möglich sein, die bislang noch größtenteils manuell angestoßenen Publikationskomponenten weitgehend zu automatisieren und die Ausführung miteinander zu kombinieren. Das beinhaltet auch die Möglichkeit der Einbettung von Konvertierungstechnologien in das webbasierte Publikationssystem.

Unterstützung von Open Access

Wie im ersten Artikel in diesem Heft dargelegt, wird an der Humboldt-Universität eine Open-Access-Initiative vorbereitet, deren Ziel unter anderem darin besteht, durch Mitglieder der Universität bereits auf anderem Wege veröffentlichte Publikationen zusätzlich auf dem edoc-Server zur Verfügung zu stellen [12]. Aus technologischer Sicht erfordert dies vor allem eine Webschnittstelle für Autoren, die das Hochladen eigener Publikationen sowie die Eingabe entsprechender Metadaten erlaubt. Diese Anwendung befindet sich zum Erscheinungstermin dieses Heftes in der Testphase. Im Rahmen von Open Access eingereichte Dokumente erfahren seitens der Arbeitsgruppe *Elektronisches Publizieren* keine Begutachtung oder redaktionelle Bearbeitung mehr. Um die missbräuchliche Nutzung dieser frei zugänglichen Schnittstelle zu verhindern, werden hochgeladene Dokumente allerdings nicht unmittelbar auf dem edoc-Server sichtbar, sondern müssen erst manuell frei geschaltet werden.

Literatur

- [1] Dokumenten- und Publikations-server der Humboldt-Universität zu Berlin, Leitlinien.
URL: http://edoc.hu-berlin.de/e_info/leitlinien.php
- [2] HENNEBERGER, S.: Textverarbeitung und XDiML – Elektronisches Publizieren von Dissertationen und Habilitationsschriften. *cms-journal* 27, August 2005, S. 29–32.
- [3] Extensible Markup Language (XML), URL: <http://www.w3.org/XML/>.
- [4] SALECKER, E.: Von LaTeX zu PDF. *cms-journal* 27, August 2005, S. 33–34.
- [5] SCHÖBEL, C., SCHULZ, M., IVANOVA, N.: In eigener Sache – Die elektronische Publikation des CMS-Journals. *cms-journal* 24, April 2003, S. 71–74.
URL: <http://edoc.hu-berlin.de/docviews/abstract.php?id=20226>.
- [6] MÜLLER, U.: Ein generisches Metadaten-system für den edoc-Server. *cms-journal* 27, August 2005, S. 43–47.
- [7] SCHULZ, M.: ProPrint – Der Print-on-Demand-Service für Dokumenten- und Publikationsserver. *cms-journal* 27, August 2005, S. 65–68.
- [8] BÜTTNER, S.: Inventions are no revolutions. Medienwandel und fachwissenschaftliche Publikationskultur. Vortrag auf der IuK-Jahrestagung 2005, 9.-11. Mai 2005, Bonn.
- [9] MÜLLER, U.: OAI-Protokoll: Metadaten für alle. *cms-journal* 27, August 2005, S. 55–58.
- [10] KRISELEIT, O.: Moodle – Lehre präsent im Web. *cms-journal* 26, März 2005, S. 31–34.
URL: <http://edoc.hu-berlin.de/docviews/abstract.php?id=25281>
- [11] MÜLLER, U.: Konferenzorganisation leicht gemacht. Eine webbasierte Software zur Verwaltung wissenschaftlicher Tagungen. *cms-journal* 27, August 2005, S. 69–72.
- [12] SCHIRMBACHER, P.: Open Access – die Zukunft des wissenschaftlichen Publizierens. *cms-journal* 27, August 2005, S. 3–7.

⁴ siehe <http://www.gap-c.de/>