

„Eine wissenschaftliche Informations-, Kommunikations- und Publikationsplattform für die Forschung (eSciDoc)“

Schlussbericht

Ausführende Stelle: Fachinformationszentrum Karlsruhe,
Gesellschaft für wissenschaftlich-technische Information mbH

Förderkennzeichen: 01C5957

Bewilligungszeitraum: 01.08.2004 – 31.07.2009

Projektleiterin: Dr. Leni Helmes

Das diesem Bericht zugrundeliegende Vorhaben wurde mit Mitteln des Bundesministeriums für Bildung und Forschung unter dem Förderkennzeichen 01 C5957 gefördert. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt beim Autor.

Berichtsblatt

1. ISBN oder ISSN	2. Berichtsart (Schlussbericht oder Veröffentlichung) Schlussbericht
3. Titel „Eine wissenschaftliche Informations-, Kommunikations- und Publikationsplattform für die Forschung (eSciDoc)“	
4. Autor(en) [Name(n), Vorname(n)] Razum, Matthias Helmes, Leni Dr.	5. Abschlussdatum des Vorhabens 31.07.2009
	6. Veröffentlichungsdatum
	7. Form der Publikation
8. Durchführende Institution(en) (Name, Adresse) Fachinformationszentrum Karlsruhe, Gesellschaft für wissenschaftlich-technische Information mbH Hermann-von-Helmholtz-Pl. 1 76344 Eggenstein Leopoldshafen	9. Ber. Nr. Durchführende Institution
	10. Förderkennzeichen 01 C5957
	11. Seitenzahl 32
12. Fördernde Institution (Name, Adresse) Bundesministerium für Bildung und Forschung (BMBF) 53170 Bonn	13. Literaturangaben 30
	14. Tabellen
	15. Abbildungen 9
16. Zusätzliche Angaben	
17. Vorgelegt bei (Titel, Ort, Datum)	
18. Kurzfassung	
19. Schlagwörter	
20. Verlag	21. Preis

I. Übersicht

1. Aufgabenstellung

Das Projekt „eSciDoc“ hatte die Schaffung einer integrierten Informations-, Kommunikations- und Publikationsplattform für die Forschung zum Ziel. Damit lag der Schwerpunkt – im Gegensatz zu den Projekten im Umfeld der D-Grid-Initiative – eher im Umfeld des Wissensmanagements und der virtuellen Forschungsumgebungen. Zwei wichtige Prämissen für das Projekt bildeten die Anforderung, ein produktionsreifes System bereitzustellen (also keinen Prototypen), sowie die Berücksichtigung eines multidisziplinären Umfelds. Beide Prämissen beeinflussten die Auswahl der Projektpartner Max-Planck-Gesellschaft als eine der großen deutschen, multidisziplinär aufgestellten Forschungsgesellschaften sowie von FIZ Karlsruhe als einer etablierten Serviceeinrichtung für Wissenschaft und Forschung.

Im Antrag noch als integrierte Informations-, Kommunikations- und Publikationsplattform bezeichnet, stellt eSciDoc am Ende der Projektlaufzeit eine e-Science-Plattform (oder, internationalen Gepflogenheiten entsprechend) eine e-Research-Plattform dar. Die Ziele, die im Antrag formuliert und in der jetzt vorliegenden Plattform umgesetzt wurden, stimmen dabei überein:

- Beschleunigte Verbreitung von Forschungsergebnissen
- Einfacher und schneller Zugriff zu relevanten Informationen und Daten
- Vereinfachte Wissensgenerierung (durch Vermeidung von Redundanz, Unterstützung verteilter, kollaborativer Arbeit, interdisziplinäre Forschung)
- Sofortige Evaluierung von wissenschaftlicher und Forschungsleistung (statt Evaluierung durch indirekte Indikatoren) durch Verknüpfung von Publikation mit Primärdaten

Vertiefter Prozess der Wissensgenerierung durch

- Effektive Kommunikation durch neue Medien
- Allgemeiner Zugriff auf komplexe Datenquellen
- Kollaborative Arbeit mit Daten
- Höhere Transparenz des Forschungsprozesses

2. Voraussetzungen

Das Projekt eSciDoc war zu Projektbeginn Mitte 2004 visionär. Dies brachte eine hohe Unsicherheit hinsichtlich der konkreten Ausprägung der einzelnen Arbeitspakete mit sich. Tatsächlich stellte sich im Verlauf der ersten Projektphase mit dem Schwerpunkt Anforderungsanalyse heraus, dass viele im Antrag getroffene Annahmen nicht haltbar waren. Dies führte zu substantiellen Abweichungen vom ursprünglichen Projektplan (siehe auch folgenden Abschnitt 3).

Auch national und international begannen kurz zuvor oder zeitgleich mit eSciDoc verschiedene Forschungsgruppen, sich mit für eSciDoc relevanten Themen wie Umgang mit Primärdaten, standardisierten Objektmodellen, virtuelle Forschungsumgebungen und Langzeitarchivierung wissenschaftlicher Daten auseinanderzusetzen. Damit hatten sich zu Projektbeginn in vielen Bereichen noch keine Standards und Best Practices etabliert, auf die das Projekt hätte aufsetzen können. Dies führte zu einer hohen Unsicherheit und oftmals schwierigen Entscheidungsfindung.

Aufbauen konnte das Projekt auf den bereits innerhalb der Max-Planck-Gesellschaft eingeführten eDoc-Server als institutionellem Repository, der wichtige Erkenntnisse für die deutlich erweiterte und robustere Applikation „Publication Management“ lieferte, die inzwischen Schritt für Schritt die noch bestehenden eDoc-Installationen ablöst. Auch hatten einzelne Institute der Max-Planck-

Gesellschaft bereits Vorarbeiten geleistet, auf die aufgebaut werden konnte, etwa ECHO¹ und CDLI² des Max-Planck-Instituts für Wissenschaftsgeschichte oder LAMUS³/LEXUS⁴ des Max-Planck-Instituts für Psycholinguistik.

3. Planung und Ablauf

Eine integrierte Informations-, Kommunikations- und Publikationsplattform erfordert eine den gesamten Forschungsprozess abdeckende Datenhaltung, eine konsistente Rechteverwaltung und konfigurierbare Arbeitsabläufe. Diese grundlegenden Dienste finden sich bei eSciDoc in der Infrastruktur wieder („eSciDoc Infrastructure“, bestehend aus Basis- und höherwertigen Diensten). Auf dieser Infrastruktur setzen dann die fachlichen Applikationen („eSciDoc Solutions“) auf, mit denen die Wissenschaftler, Bibliothekare und Datenarchivare in den Instituten vor Ort oder in zentralen Serviceeinrichtungen arbeiten. Die einzelnen eSciDoc Solutions decken dabei unterschiedliche Schritte in den disziplinspezifischen oder auch übergreifenden Forschungsprozessen ab. Abbildung 1 verdeutlicht diese grundlegende Herangehensweise.

Grundlegender Ablauf

Wie im ursprünglichen Projektantrag beschrieben, übernahm die Max-Planck-Gesellschaft, vertreten durch die Max Planck Digital Library, die Gesamtsteuerung des Projekts. Zur Abstimmung zwischen den beiden Partnern wurde ein paritätisch besetzter und auf Leitungsebene angesiedelter eSciDoc Lenkungsausschuss etabliert, der wichtige strategische Entscheidungen traf und sich regelmäßig über die Ergebnisse der Arbeitsebene informieren ließ und Konflikte ausräumte.

Weiterhin führte die MPDL bzw. die Vorläuferinstitution ZIM (Heinz-Nixdorf-Zentrum für Informationsmanagement in der Max-Planck-Gesellschaft) die Anforderungsanalyse durch, aus denen sich die Requirements für das Gesamtsystem ergaben. Parallel dazu untersuchten beide Partner für ausgewählte Kernkonzepte (Persistente Identifier, Objektmodelle, Repository-Systeme, Open Access, Langzeitarchivierung) den Stand der Wissenschaft und hielten die Ergebnisse in „state-of-the-art“-Reports fest.

Die Erstellung der Anwendungsszenarien teilten sich beide Partner. FIZ Karlsruhe beschrieb, in enger Zusammenarbeit mit dem Max-Planck-Institut für Wissenschaftsgeschichte, die sogenannte „Scholarly Workbench“, eine Arbeitsumgebung für Geisteswissenschaftler.

Aus der Anforderungsanalyse und den Anwendungsszenarien entwickelten beide Partner gemeinsam die Gesamtarchitektur des Systems. Diese beschreibt eine service-orientierte Architektur, die sich grob in drei Bereiche aufteilt: Basisdienste (übergreifend), höherwertige Dienste (aggregieren Basisdienste und/oder stellen spezifische Funktionen bereit), sowie Applikationen („eSciDoc Solutions“). Darauf aufbauend entwarf und implementierte FIZ Karlsruhe die Basisdienste sowie einige der höherwertigen Dienste, MPDL entwarf und implementierte einige höherwertige Dienste sowie die Applikationen.

Seit Dezember 2007 veröffentlichten die Partner die erarbeitete Software unter einer von der Open Software Initiative (OSI) anerkannten Open-Source-Lizenz. Seitdem wird die Software

¹ European Cultural Heritage Online; <http://echo.mpiwg.mpg.de/>

² Cuneiform Digital Library Initiative; <http://cdli.ucla.edu/>

³ Daan Broeder, Andreas Claus, Freddy Offenga, Romuald Skiba, Paul Trilsbeek, Peter Wittenburg (2006). *LAMUS – the Language Archive Management and Upload System*. Abgerufen von <http://www.lat-mpi.eu/papers/papers-2006/lamus-paper-final2.pdf>

⁴ Marc Kemps-Snijders, Mark-Jan Nederhof, Peter Wittenburg (2006). *LEXUS, a web-based tool for manipulating lexical resources*. Abgerufen von <http://www.lat-mpi.eu/papers/papers-2006/lexus-paper-final.pdf>

regelmäßig aktualisiert. Für Nachnutzer stehen sowohl vorkompilierte Pakete als auch der Programmcode zum Download bereit.

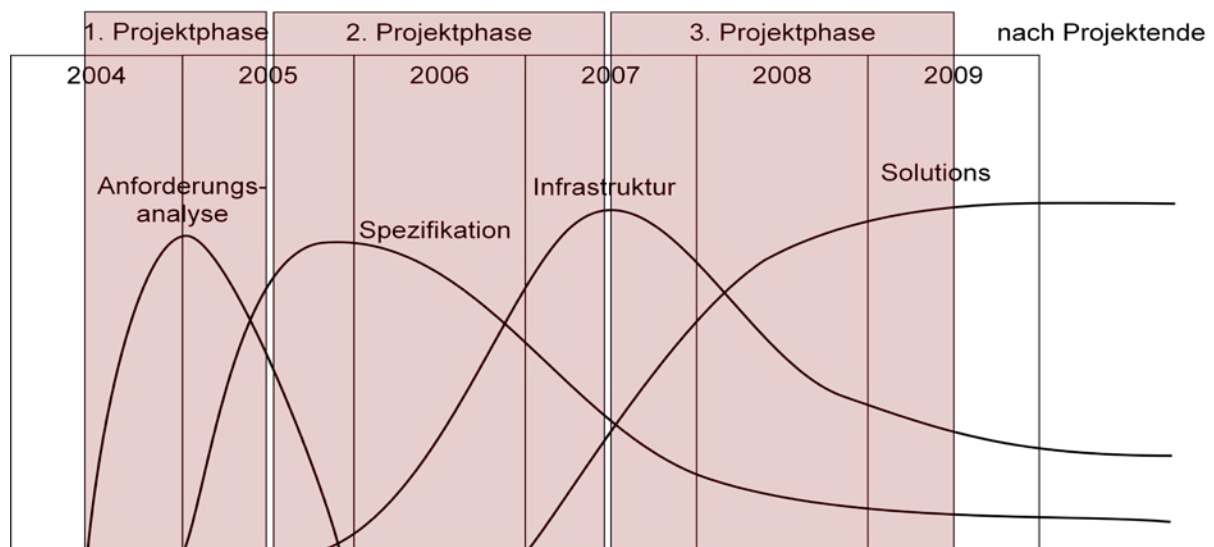


Abbildung 1: Darstellung der Projektphasen mit den Arbeitsschwerpunkten

Abweichungen von der ursprünglichen Planung

Wie im Verlauf eines fünfjährigen Projekts mit solcher Komplexität nicht anders zu erwarten, gab es auch bei eSciDoc einige Abweichungen gegenüber der ursprünglichen Planung, wie sie sich im Projektantrag findet. Die folgenden Arbeitspakete wichen substantziell von der Planung ab:

Arbeitspaket 5: GUI-Design

Die grafische Benutzungsschnittstelle (GUI) als Teil der eSciDoc Solutions bearbeitete die MPDL, da sie Entwurf und Entwicklung der eSciDoc Solutions aufgrund der größeren Nähe zu den Wissenschaftlern in den Instituten komplett übernahm.

Arbeitspaket 6: Tools

Im Sinne der service-orientierten Architektur wurden die Tools als eigenständige und weitgehend unabhängige Dienste realisiert, die überwiegend auch unabhängig von eSciDoc durch Dritte nachnutzbar sind.

Arbeitspaket 8: Portal

Im Rahmen der Anforderungsanalyse zeigte sich, dass die Idee eines zentralen Portals nicht den Ansprüchen der Wissenschaftler entsprach und wurde entsprechend fallen gelassen.

Arbeitspaket 9: Langzeitarchivierung

Die ursprünglich angedachte enge Kooperation mit Kopal kam nicht zustande, da Kopal während der Projektlaufzeit nicht den Produktivbetrieb aufnahm. Das Projekt war jedoch im Projektzeitraum in die Arbeitsgruppe Grid-eScience des Nestor Projekts eingebunden. Die ursprünglich von FIZ Karlsruhe sicherzustellende Bitstream Preservation übernahm die MPDL, da sie den Betrieb des Gesamtsystems entgegen der ursprünglichen Planung selbst übernahm.

Arbeitspaket 10: Grid

Die vorgesehene Integration mit Grid-Technologien entfiel mangels Interesse seitens der Wissenschaftler, aber es erfolgte die Integration mit Shibboleth, so dass zumindest eine einheitliche Authentifizierungstechnologie zum Einsatz kommt.

Insgesamt führten die Abweichungen nicht zu relevanten Beeinträchtigungen der Projektergebnisse gegenüber der im Antrag formulierten Vision einer integrierten e-Science-Plattform.

4. Stand der Wissenschaft und Technik

Im Verlauf des Projekts entwickelte sich der Stand der Wissenschaft und Technik rapide weiter. Diese Fortschritte wurden intensiv beobachtet und fanden Eingang in den Entwurf und die fortlaufende Entwicklung des Systems. Teilweise beeinflusste das eSciDoc-Projekt und die beteiligten Partner auch den Stand der Technik mit den erarbeiteten Ergebnissen, z.B. im Bereich der Repository-Systeme. Abschnitt II.5 beschreibt die Änderungen der relevantesten Technologien bzw. Themen für eSciDoc im Projektverlauf im Detail.

5. Zusammenarbeit mit anderen Stellen

Während der Projektlaufzeit fand eine intensive Zusammenarbeit mit anderen Projekten und Organisationen statt, einerseits um parallel stattfindende Entwicklungen im e-Science-Umfeld mit dem Projekt abzustimmen, andererseits aber auch im Rahmen der Nachnutzung der Ergebnisse von eSciDoc.

Fedora Commons (Ithaka, NY, USA)

Fedora Commons (seit 1. Juli 2009 Duraspace) ist die gemeinnützige Organisation, die unter anderem die Repository-Software Fedora als Open Source bereitstellt. Nachdem die Entscheidung fiel, Fedora als Object Store für eSciDoc zu verwenden, entwickelte sich eine intensive Kooperation zwischen beiden Projekten. Fedora Commons bezeichnet eSciDoc als eines ihrer „flagship projects“ im Bereich e-Research. Auf der anderen Seite stellt FIZ Karlsruhe inzwischen einen Committer, der sich aktiv an der Weiterentwicklung von Fedora beteiligt.

DANS (Data Archival and Networked Services, Den Haag, Niederlande)

DANS war einer der ersten Nachnutzer von eSciDoc und gleichzeitig einer der ersten Partner, die sich aktiv an der Weiterentwicklung von eSciDoc beteiligten. DANS regte nicht nur die Java Client Library an, die den programmatischen Zugriff auf die eSciDoc Infrastructure deutlich vereinfacht, sondern beteiligte sich aktiv an deren Design und Implementierung.

NIMS (National Institute for Material Sciences, Tokyo, Japan)

Ziel der Kooperation mit NIMS war der Aufbau einer web-basierte Plattform, die den Wissenschaftlern einen verbesserten Zugriff und höhere Sichtbarkeit ihrer Forschungsergebnisse ermöglicht. NIMS unterstützte hierzu die Weiterentwicklung der eSciDoc Plattform durch eigene Entwicklungskapazität und fokussierte sich dabei auf die Implementierungen eigener Erweiterungen und Anpassungen.

Ergebnisse dieser Kooperation sind z.B. die japanische Sprachversion der PubMan Solution oder Beiträge zur Nutzung der FACES Solution im Bereich der Materialwissenschaften.

TIB Hannover (Hannover, Deutschland)

Gemeinsam mit der TIB Hannover baut FIZ Karlsruhe zurzeit ein auf dem eSciDoc Publication Management basierendes System für die Leibniz-Gemeinschaft auf, um die Open-Access-Publikationen der Institute der Sektion D und ausgewählter Institute der Sektion E verfügbar zu machen.

Max-Planck-Institut für Wissenschaftsgeschichte (Berlin, Deutschland)

Das Max-Planck-Institut für Wissenschaftsgeschichte war eine der treibenden Kräfte hinter eSciDoc und hat während der gesamten Projektlaufzeit wertvolle Anregungen gegeben. Seit

Frühjahr 2009 kooperieren das Institut und FIZ Karlsruhe mit amerikanischen Partnern bei der Entwicklung eines White Papers, das zu einem großen NSF-Infrastrukturprojekt für die Wissenschaftsgeschichte führen soll und zu einem großen Nachnutzer der eSciDoc Software werden könnte. Darüber hinaus haben das Institut und FIZ Karlsruhe einen gemeinsamen Antrag bei der DFG über die Entwicklung einer virtuellen Forschungsumgebung im Bereich der Wissenschaftsgeschichte auf Basis von eSciDoc gestellt.

DFN-AAI

Schon relativ früh während der Projektlaufzeit fiel die Entscheidung, bei der Authentifizierung auf Shibboleth zu setzen, nicht zuletzt, um ein zur Grid-Middleware kompatibles System zu wählen und damit einerseits perspektivisch Interoperabilität herzustellen und andererseits eine einheitliche föderierte Authentifizierungsinfrastruktur für die Wissenschaft und Forschung in Deutschland zu befördern. Deshalb kam es früh zu einer Abstimmung mit dem ebenfalls vom BMBF geförderten Projekt AAI, aus dem sich inzwischen die deutsche Shibboleth-Föderation DFN-AAI entwickelt hat.

Weitere, auf eSciDoc aufbauende Projekte

Weiterhin bekundeten bereits während des laufenden Projekts verschiedene deutsche und internationale Partner ein Interesse an der Zusammenarbeit und einer Nachnutzung von eSciDoc. Hervorgehoben seien hier zwei zwischenzeitlich bewilligte Projekte, die auf den Ergebnissen von eSciDoc aufbauen: BW-eLabs und BW-eSci(T).

BW-eLabs ist ein gemeinsames Projekt mit den Universitäten Stuttgart und Freiburg sowie der Hochschule der Medien Stuttgart zur Unterstützung von virtuellen und Remote-Laboren in der Nanotechnologie.

BW-eSci(T) ist ein gemeinsames Projekt mit der Universität Tübingen mit dem Ziel, eine e-Science-Plattform an einer Universität zentral über Rechenzentrum und Bibliothek auszurollen, am Beispiel der Computerlinguistik den Nutzen für eine vereinheitlichte Bereitstellung von Daten und Methoden zu evaluieren und später auf weitere Disziplinen auszuweiten.

Weiterhin erproben z. B. das Geoforschungszentrum Potsdam, UNED in Spanien, die Bayrische Staatsbibliothek und die königliche Bibliothek in Dänemark den Einsatz von eSciDoc für den produktiven Einsatz innerhalb der eigenen Organisation.

Darüber hinaus wurden im letzten Call der DFG zu virtuellen Forschungsumgebungen neun Anträge (von ca. 50 insgesamt) eingereicht, die explizit eSciDoc als Grundlage für die Implementierung vorsahen. Dies zeigt einerseits die Vorteile einer solchen Plattform für e-Science-Anwendungen, andererseits aber auch den bereits erreichten Bekanntheitsgrad von eSciDoc innerhalb der deutschen Forschungsszene auf.

II. Eingehende Darstellung

1. Erzielte Ergebnisse

eSciDoc – eine e-Research-Plattform für die Wissenschaft

eSciDoc wurde als eResearch-Plattform speziell für die Benutzung in Forschung und Wissenschaft entwickelt, um die geografisch verteilte und disziplinübergreifende Zusammenarbeit in virtuellen Arbeitsgruppen zu unterstützen. Es umfasst Kernfunktionen inklusive eines zentralen Repositories (“eSciDoc Infrastructure”), diversen ergänzenden Diensten (“eSciDoc Services”) und Applikationen, die auf der Infrastruktur aufsetzen und nach Bedarf ergänzende Dienste einbinden (“eSciDoc Solutions”). Diese Applikationen stellen bereits einige interessante

Lösungen für Forscher bereit, aber der tiefere Nutzen von eSciDoc liegt in der Bereitstellung eines „Baukastens“ im Sinne einer service-orientierten Architektur (SOA), über die sich einfach und schnell weitere e-Research-Szenarien (zum Beispiel sogenannte Virtual Research Environments) realisieren lassen. Kern all dieser Ansätze ist die durchgehende Speicherung aller im Verlauf des Forschungsprozesses anfallenden Artefakte und deren Verknüpfung („continuum of data“), wie in Abbildung 2 dargestellt. Dabei muss eSciDoc nicht notwendigerweise die Primärdaten tatsächlich speichern, sondern kann die Daten auch referenzieren und sie damit auf besser geeigneten Speichersystemen belassen. Wissenschaftler, Bibliothekare und Softwareentwickler können dank eSciDoc einfach mit Forschungsdaten arbeiten, neue und innovative Publikationsformen erproben und neue Methoden in der wissenschaftlichen Kommunikation etablieren.

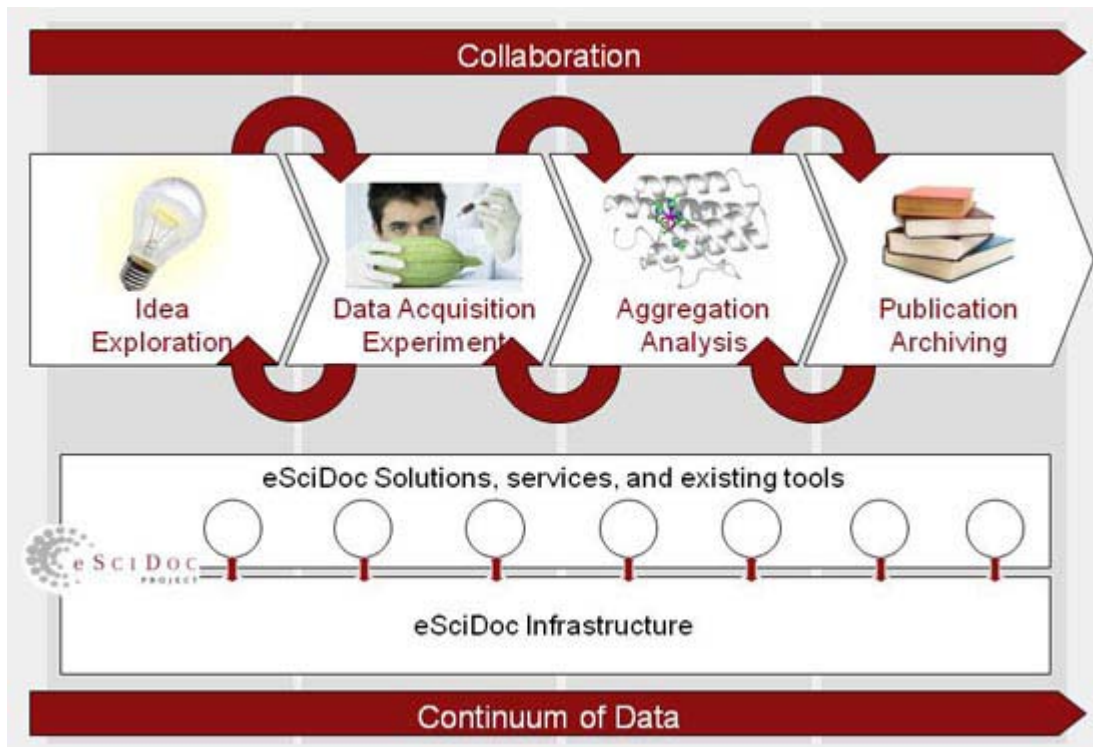


Abbildung 2: eSciDoc speichert alle im Verlauf des Forschungsprozesses anfallenden (digitalen) Artefakte und stellt für die verschiedenen Schritte in diesem Prozesse geeignete Werkzeuge zur Verfügung bzw. bindet bestehende Werkzeuge ein.

Anforderungsanalyse (Needs Analysis)

Dieses Ergebnis wurde von der MPDL (bzw. deren Vorgänger-Organisation ZIM) erarbeitet. Es enthält viele MPG-interne Informationen und wurde deshalb nicht publiziert. Aus der Anforderungsanalyse heraus kristallisierten sich jedoch die beiden primär verfolgten Anwendungsszenarien „PubMan“ und „Scholarly Workbench“ heraus. Das weiterhin dort beschriebene Anwendungsszenarium „eLib“ wurde bei der Auslegung der Systemarchitektur berücksichtigt. Die eigentliche Umsetzung als Dienst für die Max-Planck-Gesellschaft erfolgt außerhalb des Projekts eSciDoc.

State-of-the-Art Reports

In der ersten Projektphase erarbeiteten beide Partner sogenannte „State of the Art“-Reports, die jeweils zu einem ausgewählten Thema den Stand der Forschung darlegten. Diese Reports dienten zum Aufbau von Knowhow innerhalb des Projektteams und wurden nicht veröffentlicht. Insbesondere die Reports zu Repository-Systemen und zu Objektmodellen bildeten die

Grundlage für die spätere Entscheidung, das Repository-System Fedora (Flexible Extensible Digital Object Repository Architecture)⁵ als Kernkomponente für eSciDoc auszuwählen. Der Report zu Persistent Identifier beeinflusste die Entscheidung, Handle in einer ersten Implementierung auszuwählen, gleichzeitig die Architektur des Systems gegenüber konkurrierenden Ansätzen offen zu halten.

Insgesamt wurden zu folgenden Themen State-of-the-Art Reports erstellt:

- Persistente Identifier
- Objektmodelle
- Repository-Systeme
- Open Access
- Langzeitarchivierung

Anwendungsszenarien

Es wurden zwei umfängliche Anwendungsszenarien (Usage Scenarios) gemeinsam mit den Wissenschaftlern und Bibliothekaren aus den beteiligten Max-Planck-Instituten entwickelt und beschrieben: „Publication Management“ und „Scholarly Workbench“. Jedes Anwendungsszenario beschreibt in mehreren Anwendungsfällen die Art und Weise, in der ein Wissenschaftler bzw. Bibliothekar mit dem System interagiert. Diese orientieren sich stark am tatsächlichen Arbeitsfluss, um eine möglichst hohe Akzeptanz zu erreichen.

Insgesamt wurden für Publication Management (PubMan) neunzehn Anwendungsszenarien erstellt⁶. Viele dieser Szenarien haben sich im Verlauf des Projekts fortentwickelt und durch weitere ergänzt. Ein aktueller Stand findet sich stets im Collaboratory der Max Planck Digital Library⁷.

Die in fünfzehn Usage Scenarios beschriebene Scholarly Workbench wurde nicht als eigenständige Solution umgesetzt, sondern in thematisch abgegrenzte und besser handhabbare Komponenten aufgegliedert (Arbeiten mit Bildern mittels der eSciDoc Solution FACES, Arbeiten mit Digitalisaten, Erstellung digitaler Editionen mittels der Solution ViRR). Weitere Aspekte der Scholarly Workbench werden zukünftig durch neue bzw. erweiterte Solutions angegangen.

Entwicklung einer Gesamtarchitektur

Ausgehend von der Anforderungsanalyse und den erarbeiteten Anwendungsszenarien erarbeitete eine gemischte Arbeitsgruppe, bestehend aus Mitgliedern beider Projektpartner und unterstützt durch externe Beratung iterativ eine Gesamtarchitektur für eSciDoc sowie darauf aufbauend die wichtigsten Konzepte. Wichtige Schritte dabei waren

- die Identifizierung von Diensten im Sinne einer service-orientierten Architektur (SOA);
- die Erstellung eines logischen Datenmodells;
- die Auswahl des Objektcontainers und eines Repository-Systems (FOXML, Fedora);
- die grundlegende Konzeption von Content Models (von spezifischen Inhalten abhängige Ausprägung von Objekten, z.B. für Publikationen, Videos, Digitalisate usw);
- die Definition von Metadaten-Profilen.

Die im ersten Schritt identifizierten Dienste untersuchte das Architekturteam auf ihre Generalisierbarkeit. Über mehrere Anwendungen hinweg nutzbare Dienste wurden im Folgenden als Teil der eSciDoc Infrastructure betrachtet. Das Team bestimmte für diese Dienste notwendige Programmierschnittstellen und Datenaustauschformate.

⁵ <http://www.fedora-commons.org>

⁶ <https://www.escidoc.org/JSPWiki/en/Publications>

⁷ <http://colab.mpdl.mpg.de/mediawiki/Portal:PubMan>, siehe „Functionalities/Basic Concepts“

Software-Entwurf und Implementierung

Von diesen Vorarbeiten ausgehend, teilte sich die eigentliche Softwareentwicklung auf die beiden Projektpartner auf. Die Max Planck Digital Library konzentrierte sich auf die Umsetzung der eSciDoc Solutions, also der sich an die Wissenschaftler und Bibliothekare richtenden Anwendungen. FIZ Karlsruhe realisierte derweil die darunterliegenden Kerndienste (eSciDoc Infrastructure). Die im Rahmen der Gesamtarchitektur vereinbarten Schnittstellen, eine enge Abstimmung beider Teams sowie eine iterative Entwicklungsmethodologie mit z.T. wöchentlichen Releases sicherten ein funktionierendes Zusammenspiel der einzelnen Komponenten auch bei einer räumlich verteilt stattfindenden Softwareentwicklung. Trotzdem stellten der hohe Termindruck, die fehlende Vorlaufzeit bei der Implementierung der Kerndienste und sich im Projektverlauf neu ergebende Anforderungen beide Teams vor große Herausforderungen, die insbesondere in der zweiten Projektphase zu zum Teil erheblichen Verzögerungen führten. Diese konnten aber größtenteils in der letzten Projektphase wieder aufgeholt werden.

Ein wichtiger Aspekt war von Anfang an die Devise, nicht das Rad neu zu erfinden und wo immer möglich vorhandene Software in das Gesamtsystem einzubinden. Interessanterweise ist der dadurch eingesparte Ressourceneinsatz nur relativ gering, da einem verminderten Entwicklungsein deutlich erhöhter Einarbeitungsaufwand gegenübersteht, der durch fehlende, unvollständige oder fehlerhafte Dokumentation und Fehler in der Software selbst im ungünstigsten Fall die notwendigen Ressourcen für eine Eigenentwicklung erreichen oder gar übersteigen kann. Allerdings sind üblicherweise Softwarekomponenten von Drittanbietern ausgereifter. Ebenfalls nicht zu unterschätzen ist der verminderte Aufwand zur Pflege dieser Komponenten. Einen Überblick über in eSciDoc eingesetzter Fremdsoftware gibt die folgende Aufstellung:

- JBoss Application Server (Ablaufumgebung für Java Enterprise Software)
- Fedora (Repository für digitale Objekte)
- PostgreSQL (relationales Datenbanksystem)
- Sun XACML Engine (Auswertung von Zugriffsrechten)
- Shibboleth (verteilte Authentifizierung und Autorisierung)
- SRW/U (verteilte Suche im Bibliotheksumfeld)
- Lucene (Volltextindexierung und –suche)
- OAI-PMH (Harvesting von Metadaten für Aggregationszwecke)
- JHOVE (Extraktion technischer Metadaten)
- arXiv Duplicate Detection (Erkennung von Duplikaten aufgrund von Wortmustern im Volltext)
- Digilib Image Library (effiziente Darstellung und Manipulation von Bildern im Browser)
- Handle Server (persistente Identifizierung von digitalen Objekten)
- jBPM Workflow Engine

Im Rahmen der Software-Entwicklung kristallisierte sich eine Klasse von weiteren Diensten heraus, die weder den Kerndiensten noch den Applikationen zuzurechnen sind. Die meisten dieser „eSciDoc Services“⁸ stellen bereits höherwertige Dienste bereit und erleichtern so die Entwicklung von Applikationen. Viele der Dienste können auch unabhängig von eSciDoc durch Dritte eingesetzt werden. Sie wurden sowohl von der Max Planck Digital Library wie auch von FIZ Karlsruhe implementiert.

Kerndienste (eSciDoc Infrastructure)

Die eSciDoc Infrastructure stellt die Kerndienste für die e-Research-Umgebung zur Verfügung. Auf ihr setzen die eSciDoc Solutions auf. Sie stellt basale Dienste wie die Speicherung und Verwaltung von Objekten (Objekt Handling and Persistence), die Authentifizierung, Autorisierung und Verwaltung und Prüfung von Zugriffsrechten (Authentication and Authorization), die

⁸ <https://www.esdoc.org/JSPWiki/en/Services>

Erfassung von Events auf allen Ebenen des Gesamtsystems sowie deren Auswertung und Aufbereitung (Statistics), die Volltext- und Metadatenindexierung und Suche (Search Service), die Verwaltung von dynamischen Sets für OAI-PMH⁹ (Set Service) sowie die Administration der eSciDoc Infrastructure (Admin Service).

Alle Dienste der eSciDoc Infrastructure bieten ausschließlich Programmierschnittstellen (Application Programming Interface, API). Sie besitzen keinerlei Benutzungsoberfläche und richten sich daher nicht an Endanwender, sondern an Entwickler von höherwertigen Diensten und Solutions.

Die Dienste der eSciDoc Infrastructure lassen sich thematisch gliedern: Objektspeicherung und -manipulation, Sicherheit, Statistik, Suche und Administration.

Objektspeicherung und -manipulation

Die Objektspeicherung und -manipulation stellt alle notwendigen Dienste zur Verwaltung digitaler Objekte bereit. Diese unterstützen die Typisierung von Objekten durch Content Models, die Aggregationen durch Container, die Versionierung von Objekten sowie deren Lebenszyklus. Die insgesamt fünf Dienste stellen die Verbindung zum zentralen Repository (Fedora) sicher und bilden so die zentralen Bausteine der eSciDoc Infrastructure.

eSciDoc kennt fünf grundlegende Objekttypen: Contexts, Containers, Items, Organizational Units und Content Models. Diese Objekttypen bilden das Grundgerüst für jedes Datenmodell, d.h. jede Applikation, die auf der eSciDoc Infrastructure aufsetzt, muss ihre Daten derart strukturieren, dass sie sich durch eine Kombination dieser Typen darstellen lassen. Was nach Einschränkung klingt, hat sich in der Praxis als äußerst flexibel herausgestellt, da es sich bei den fünf Typen um generische Muster handelt. Die Objekttypen für Inhaltsobjekte (Items) und Aggregationen (Container) lassen sich mittels Content Models weiter an die eigenen Anforderungen adaptieren (typisieren), um so zu einem passgenauen Datenmodell zu kommen. Für jeden der fünf Objekttypen stellt ein eigener Dienst eine spezifische Programmierschnittstelle bereit:

- **Context Service**
Ein Context repräsentiert die rechtliche und organisatorische Entität, der die zugeordneten Objekte im Repository gehören und für die sie Verantwortung übernimmt. Jeder Container und jedes Item muss genau einem Context zugeordnet sein. In einem Repository können beliebig viele Context-Objekte nebeneinander angelegt sein, z.B. um unterschiedliche Verwaltungseinheiten abzubilden (wichtig für die Max Planck Gesellschaft mit ihren Instituten) oder aber in sich geschlossene Räume für die Daten von Benutzergruppen oder Applikationen vorzusehen.
- **Container Service**
Container setzen das Konzept der Aggregation um, also der Gruppierung von Objekten. Container "enthalten" andere Container und/oder Items. Tatsächlich werden die aggregierten Objekte in einer sogenannten StructMap referenziert, die einen einfachen und schnellen Zugriff auf die Mitglieder einer solchen Aggregation erlauben.
- **Item Service**
Items repräsentieren den eigentlichen Inhalt des Repositories. Items können einfache oder komplexe Objekte darstellen, etwa Artikel, Bilder oder auch komplette Datensätze. Neben beliebigen Inhalten nehmen Items auch einen oder mehrere Metadatensätze auf, die das Objekt oder seine Inhalte beschreiben.
- **Organizational Unit Service**
Organizational Units (OU) bilden die Struktur von Organisationen ab. Damit

⁹ Open Archives Initiative – Protocol for Metadata Harvesting; <http://www.openarchives.org/>

lassen sich zum Beispiel die einzelnen Institute einer Universität und darunter die Abteilungen und Arbeitsgruppen der einzelnen Institute abbilden. OUs sind nicht notwendigerweise hierarchisch angeordnet; so kann eine Arbeitsgruppe durchaus mehrere Institute als übergeordnete Einheiten definieren. Gleichzeitig verwaltet der Organizational Unit Service auch die Entwicklung der Organisationsstruktur über die Zeit hinweg, so dass auch Umstrukturierungen, Umbenennungen und Ähnliches nachvollzogen und z.B. bei der Suche mit berücksichtigt werden.

- **Content Model Service**

Content Models beschreiben die Struktur und das Verhalten von digitalen Objekten im Repository. Bei der Erstellung eines neuen Objekts und bei allen folgenden Änderungen überprüft das System, ob das Objekt seinem Content Model entspricht. Durch Content Models lässt sich das sehr generische Item für eine spezielle Anwendung oder einen spezifischen Datentyp modellieren. Konkret spezifizieren Content Models zum Beispiel die zu verwendenden Metadaten-Profile, die zwingend erforderlichen Inhalte und Formate (z.B. könnte ein Content Model für ein Bild einen Metadaten-Record gemäß MIX sowie eine Repräsentation des Bildes als TIFF und als JPEG vorschreiben).

Sicherheit

In eSciDoc verwaltete Forschungs- und Publikationsdaten sind nicht immer frei zugänglich. Wissenschaftler veröffentlichen aus verschiedenen Gründen nur ungern Zwischenergebnisse vor der abschließenden Publikation. Manche Primärdatensätze können z.B. aus Datenschutzgründen überhaupt nicht publiziert werden. Selbst innerhalb eines Projektteams sind oftmals fein abgestufte Zugriffsrechte auf die verschiedenen Artefakte des Forschungsprozesses notwendig.

eSciDoc bietet daher ein umfangreiches und mächtiges Sicherheitssystem, das Authentifizierung, Autorisierung und “policy enforcement”, also die Durchsetzung der vergebenen Zugriffsrechte, kombiniert. Gleichzeitig unterstützt es auch verteilte Benutzer- und Rechteverwaltung über Shibboleth. eSciDoc erweitert dazu das bereits von Fedora mitgelieferte Sicherheitsmodell um speziell für virtuelle Forschungsumgebungen relevante Funktionalität. Es implementiert dazu einen zentralen Punkt zur Prüfung der Zugriffsrechte, den sogenannten “Policy Decision Point”. Jeder Zugriff auf eine Methode der Kerndienste muss den Policy Decision Point durchlaufen, bevor er die eigentliche Geschäftslogik erreicht. Dieser zentrale Ansatz vereinfacht die Entwicklung von Applikationen auf Basis von eSciDoc, da diese sich nicht um Sicherheitsaspekte kümmern müssen und die Sicherheit der Daten auch bei fehlerhafter Programmierung nicht kompromittieren können.

- **User Account Service**

Jeder Benutzer im System wird durch einen User Account repräsentiert. Dieser enthält beim Einsatz der verteilten Authentifizierung (über Shibboleth) nur wenige Basisdaten und dient damit nur als Proxy-Objekt, an dem Rechte festgemacht werden können. Der User Account Service ermöglicht die programmgesteuerte Verwaltung dieser Benutzerobjekte.

- **User Group Service**

Gruppen fassen mehrere Benutzer (oder auch andere Gruppen) zusammen. An diese können dann, anstatt an einzelne Benutzer, Rechte (definiert über Policies) vergeben werden. Der User Group Server erlaubt die programmgesteuerte Verwaltung von Gruppen.

- **Role Service**

Rollen ermöglichen es, mehrere Policies zusammenzufassen und zu benennen. Dadurch vereinfacht sich die Zuordnung von Policies zu Benutzern erheblich, da statt einzelner Policies nun Rollen vergeben werden. Auch können dadurch Rollen zentral gepflegt und ein konsistentes Verhalten über mehrere Benutzer hinweg

sichergestellt werden. Der Role Service erlaubt die programmgesteuerte Verwaltung von Rollen.

- **Policy Decision Point**

Der Policy Decision Point dient, wie oben schon beschrieben, zur zentralen Überprüfung der Zugriffsrechte und wird von allen Kerndiensten genutzt. Er hat nur eine einzige Methode, die üblicherweise auch nur intern durch die eSciDoc Infrastructure aufgerufen wird und die eigentliche Überprüfung durchführt.

Statistik

Statistiken und Reports können entscheidend dazu beitragen, das Verhalten eines komplexen Systems zu verstehen und zu überwachen, Probleme im laufenden Betrieb zu identifizieren und Schwachstellen in der Systemarchitektur aufzudecken. Typische Schritte in diesem Zusammenhang sind das Erfassen der Benutzeraktivitäten und anderer Systemereignisse zur Laufzeit, deren Aggregation und Analyse sowie die Erstellung geeigneter Reports und Auswertungen. Die Ergebnisse können helfen, einen zuverlässigen und performanten Service anzubieten. Im wissenschaftlichen Umfeld kommt oft noch hinzu, dass z.B. Zugriffszahlen auf Veröffentlichungen sowie Zitationen als Grundlage für Erfolgskontrollen verwendet werden.

Applikationen sammeln typischerweise die Statistikdaten zur Laufzeit pro Methodenaufruf in Form von Ereignissen ('event logging'). Jeder Methodenaufruf oder jede Benutzerinteraktion resultiert in einem oder mehreren Ereignissen, die in Form von Statistik-Rohdaten festgehalten werden. In service-orientierten Umgebungen, wie sie eSciDoc bietet, kann eine einzelne Benutzerinteraktion zum Aufruf mehrerer Dienste führen, die jeweils unabhängig voneinander Ereignisse festhalten. Um diese Ereignisse über alle Dienste hinweg zusammenzuführen und auswerten zu können, stellt eSciDoc hierfür eine Gruppe von zentralen Diensten zur Verfügung, die von der Speicherung der Ereignisse bis zur Generierung von Reports alle notwendigen Bearbeitungsschritte abdecken.

- **Scope Service**

Jeder eSciDoc Service und jede eSciDoc Solution besitzen ihren eigenen Scope in der Statistikdatenbank. Innerhalb eines Scopes kann das Format, mit dem die Ereignisse festgehalten werden, innerhalb gewisser Begrenzungen festgelegt werden. Der Scope Service erlaubt die Einrichtung solcher Scopes sowie die Festlegung des zu verwendenden Formats für Ereignisse.

- **Statistics Data Service**

Der Statistics Data Service sammelt alle Ereignisse der Dienste und speichert sie als Statistik-Rohdaten in einer Datenbank. Dabei verwendet er jeweils die über den Scope Service definierten Datenformate. Im nächsten Arbeitsschritt werden die Rohdaten aggregiert bzw. verdichtet, und zwar gemäß der Aggregationsregeln, die im Aggregation Definition Service (siehe unten) festgelegt wurden. Diese Verdichtung erfolgt asynchron, um die Laufzeitumgebung nicht auszubremsen.

- **Aggregation Definition Service**

Da die resultierenden Rohdaten üblicherweise zu umfangreich sind, um sie in Echtzeit oder zumindest mit vertretbarer Wartezeit auszuwerten, erlaubt der Aggregation Definition Service die Hinterlegung von Aggregationsregeln, die festlegen, wie die Rohdaten für die spätere Auswertung und das Reporting verdichtet werden sollen.

- **Report Definition Service**

Reports sind Abfragen über aggregierte Daten. Die meisten Reports werden mehrfach abgefragt, etwa täglich oder monatlich. Deshalb kann man solche Reports als gespeicherte Anfragen im Report Definition Service hinterlegen und verwalten.

- **Report Service**

Der Report Service schließlich erlaubt den Abruf von Reports, die auf hinterlegten

Report-Definitionen basieren. Typischerweise bezieht sich dabei ein Report auf Daten aus einem Scope, es sind allerdings auch Scope-übergreifende Reports möglich.

Suche

Die Suche wird durch einen Such- und Indexierungsdienst bereitgestellt, der sowohl alle (publizierten) Inhaltsobjekte wie auch Organizational Units indexiert. Er bietet mit dem Search Retrieval via URL/Webservice (SRU/W) Protokoll eine standardisierte Schnittstelle für Suchanfragen, die auch für verteilte Suchen verwendet werden kann.

Der Indexierungsprozess ist konfigurierbar und lässt sich so an die speziellen Bedürfnisse, aber auch Objektformate und -eigenschaften anpassen. Der Suchdienst unterstützt beliebig viele Indexe, um z.B. Recherchemöglichkeiten für spezielle Objekttypen umzusetzen.

Der Suchdienst indexiert in seiner Standardkonfiguration alle Metadaten-Records sowie den Volltext, soweit vorhanden. Momentan extrahiert der Dienst Informationen aus den Formaten Text, HTML, XML, Microsoft Word und PDF. eSciDoc verwendet durchgehend Unicode-Codierung für alle Objekte, so dass auch außereuropäische Sprachen wie z.B. Japanisch korrekt behandelt werden.

Administration

The Admin Handler provides a programmatic interface to some methods implementing administrative tasks like re-indexing of the whole repository (e.g., because of changes to the indexing configuration or because of a migration to a newer software version). The service is mainly targeted at the eSciDoc Infrastructure Administration Tool that is part of the standard installation.

Höherwertige Dienste (eSciDoc Services)

OAI-PMH Provider Service

Um den Austausch und die Interoperabilität von Metadaten zwischen Bibliotheken und Betreibern von Repositories zu vereinfachen, definierte die Open Archives Initiative das "Protocol for Metadata Harvesting" (OAI-PMH). Der Standard spezifiziert neben dem eigentlichen Protokoll auch mögliche Austauschformate für die Metadaten und zwei Softwaresysteme: einerseits den Provider, der die OAI-PMH-Schnittstelle implementiert und entsprechende Daten bereitstellt, und den Harvester, der von einem oder mehreren Providern die bereitgestellten Daten einsammelt. eSciDoc stellt einen Provider zur Verfügung, der die Metadaten der im System verwalteten Objekte über seine Schnittstelle zum Harvesten anbietet. Aufgrund der Vielzahl unterschiedlicher Objekttypen, die potenziell in eSciDoc vorliegen können, bietet der Provider mächtige Filterwerkzeuge an, damit Harvester nur die sie interessierenden Daten erhalten (sogenannte Sets). Diese können zur Laufzeit dynamisch definiert und abgefragt werden.

PID Service

Persistente Identifier (PID) identifizieren eindeutig und dauerhaft Objekte, unabhängig vom verwendeten Softwaresystem zu deren Speicherung und Verwaltung. Sie abstrahieren damit die Adressierung der Objekte von ihrem physischen Speicherort. Persistente Identifier ermöglichen dadurch den zuverlässigen und dauerhaften Zugriff auf Ressourcen auch bei einem Wechsel des Softwaresystems, indem ein sogenannter Resolver die persistenten Identifier auf die momentan gültigen Systemidentifizierer des Speichersystems abbildet.

eSciDoc's PID-Service unterstützt die Erzeugung von PIDs ('minting'), die Vergabe von PIDs an eine Ressource und die Registrierung einer vergebenen PID bei einem Resolver-Service. Dabei ist der Service sehr flexibel und muss nicht alle Schritte unter eigener Kontrolle durchführen. Er akzeptiert zum Beispiel auch vorab erzeugte oder bereits registrierte PIDs zur Vergabe an

Ressourcen. Auch unterstützt er verschiedene PID-Systeme (z.B. Handle, DOI, URN, PURL). Damit lässt er sich in bestehende PID-Infrastrukturen von Organisationen integrieren.

Duplicate Detection Service

In großen eSciDoc-Systemen arbeiten viele Benutzer mit den Objekten. Gerade in verteilten Arbeitsgruppen kann es dazu kommen, dass mehrere Benutzer dasselbe Objekt mehrfach in das System einpflegen wollen. Der Duplicate Detection Service versucht, solche Duplikate frühzeitig zu erkennen und den Benutzern einen Hinweis zu geben. Dazu berechnet er die Wahrscheinlichkeit, dass es sich bei einem neu zu ladenden Objekt um ein Duplikat eines bereits im System vorliegenden Objekts handelt. Die Berechnung der Wahrscheinlichkeit basiert auf dem Vergleich von Prüfsummen, die auf Basis von Wortgruppen errechnet wurden. Er identifiziert also nicht nur exakte Übereinstimmungen, sondern berechnet einen Ähnlichkeitsfaktor, anhand dessen der Benutzer dann intellektuell entscheiden kann, ob es sich tatsächlich um ein Duplikat oder z.B. um eine abgeleitete Version handelt. Momentan ist der Duplicate Detection Service aufgrund des verwendeten Algorithmus auf textuelle Objekte ausgelegt, lässt sich aber zukünftig durch Hinzufügen weiterer Algorithmen auch auf andere Objekttypen ausdehnen.

Technical Metadata Extraction Service

E-Research-Plattformen müssen üblicherweise mit einer Vielzahl von Dateiformaten umgehen. Jedes Dateiformat hat dabei seine spezifischen Eigenschaften, die sowohl für die Anzeige (z.B. der Mime-Type) als auch für die Langzeitarchivierung wichtig sind.

Der Technical Metadata Extractor Service extrahiert automatisch die wichtigsten Eigenschaften aus Dateien und stellt sie als zusätzliche Metadaten zur Verfügung. Dazu bindet er den JHOVE-Dienst ein, der die wichtigsten Dateiformate erkennt und auslesen kann, darunter reine Textdateien, XML, HTML, PDF, Soundformate wie AIFF und WAV, sowie Bildformate wie TIFF, JPEG, JPEG2000 und GIF.

Der Technical Metadata Extractor Service wird in typischen Arbeitsabläufen vor dem Speichern eines Objekts aufgerufen. Er gibt eine XML-Struktur mit den extrahierten Metadaten zurück, die üblicherweise als zusätzlicher Metadaten-Record in das Objekt aufgenommen und mit ihm zusammen gespeichert wird.

Image Service (Digilib)

Der Image Service bindet Digilib, eine vielseitige Bildbetrachtungssoftware für das Internet¹⁰, in eSciDoc ein. Digilib besteht aus Bildmanipulationsmethoden auf der Serverseite sowie einer browserbasierten Anzeige auf der Clientseite. Bilder werden zur Laufzeit vom Server derart prozessiert, dass nur die gerade im Client sichtbaren Teile bearbeitet und übertragen werden. Das erlaubt die schnelle und effiziente Navigation selbst in sehr großen Bildern mit hoher Auflösung, wie sie im wissenschaftlichen Bereich häufig vorkommen, etwa als Ergebnisse hochauflösender Scans oder medizinischer bildgebender Verfahren. Digilib unterstützt eine Vielzahl von Bildformaten und Ansichtsoptionen auf der Serverseite, während auf der Benutzerseite nur ein Browser mit Javascript und einer schmalbandigen Internetanbindung erforderlich ist.

Validation Service

Der Validation Service bietet Schnittstellen zur Validierung von Item- und Container-Ressourcen hinsichtlich definierbarer Regeln. Durch den Einsatz von Schematron als Beschreibungssprache für diese Regeln können in Verbindung mit Content-Models, Validierungsschemata, Metadaten

¹⁰ <http://digilib.berlios.de/>

und Validierungspunkten sehr komplexe Einstellungen z.B. für die Metadatenanreicherung ermöglicht werden.

Transformation Service

Mit Hilfe des Transformation Service können unterschiedliche textliche Formate ineinander überführt werden. Zurzeit unterstützt dieser Dienst Transformationen für Metadatenformate, wie z.B. arXiv, eSciDoc, PubMed Central und Biomed Central, Transformationen für Mikroformate, wie z.B. Coins, und Referenzformaten, wie z. B. BibTeX oder Endnote. Durch den stetigen Ausbau der Formatunterstützungen wird die Interoperabilität zu anderen Systemen und Standards zunehmend erweitert.

Data Acquisition Service

Der Dienst bietet Schnittstellen zum Austausch von Datenströmen mit externen Repositorien und Systemen über eine unAPI-Schnittstelle. Zurzeit unterstützt der Dienst das Laden von Daten von BioMed Central, Spire, eSciDoc, PubMed Central und arXiv Repositorien

Citation Style Service

Die Formatierung von Publikations-Referenzen und -Listen in unterschiedliche Zitierstile wird durch den Citation Style Service in generischer und erweiterbarer Form als Dienst angeboten. Der Dienst ermöglicht die gezielte Selektion von Metadaten zu unterschiedlichen Austauschformaten und Formatierungen. Derzeit werden die Formate RTF, ODT, PDF und HTML unterstützt. Mögliche Zitierstile sind zurzeit APA und AJP.

Deposit Service (SWORD)

SWORD¹¹ ist ein leichtgewichtiges Protokoll zur Ablage von Inhalten in entfernte Systeme und basiert auf einem Atom Publishing Protocol¹² Profil. Der Dienst unterstützt die Übertragung komplexer Pakete vollständiger Publikationsdaten zwischen unterschiedlichen Repositorien und ist ein wichtiger Baustein für die Interoperabilität.

Control of Named Entities (CoNE) Service

Zur Hebung der Datenqualität von Metadateneinträgen bietet der CoNE Service Methoden zur Abfrage von kontrollierten Vokabularen, um eine möglichst hohe Datenqualität zu erzielen. Die Daten können über JQuery, JSON, HTML oder RDF disseminiert werden. Zurzeit sind kontrollierte Entitäten für die Bereiche Mimetypes, Sprachen, DDC¹³, Journals, Creative Commons Lizenzen und Personen eingebunden.

Search & Export Service

Aufsetzend auf dem Search Service der Kern-Infrastruktur bietet dieser Dienst Funktionen zur Suche von Ressourcen und zum Export in unterschiedlichen Ausgabeformaten. Hierdurch werden Szenarien zum unkomplizierten Austausch von Daten der eSciDoc-Infrastruktur mit externen Systemen, wie z.B. Blogs, Webseiten oder Content Management Systemen ermöglicht und Re-Use Szenarien gefördert. Als Formate werden PDF, HTML, RTF, ODT und ein maschinenlesbares Format mit eingebundenen Zitierstilen unterstützt.

¹¹ <http://www.swordapp.org/>

¹² <http://wwwmm.ch.cam.ac.uk/~ojd20/sword-profile-1.3-20081007.html>

¹³ <http://www.oclc.org/dewey/>

Applikationen (eSciDoc Solutions)

Publication Management

Das Publication Management (PubMan) unterstützt Forschungsorganisationen im Management, der Verbreitung und der Nachnutzung von Publikationen und ergänzendem Material („supplementary material“). Die Anwendung PubMan ist eine Komponente der e-Research-Infrastruktur der Max-Planck-Gesellschaft und basiert auf der service-orientierten Architektur von eSciDoc.

Die Anwendung PubMan bietet alle notwendigen Funktionalitäten zur Verwaltung und Nachnutzung von Publikationsdaten und kann als instituts- oder individuelles Archiv genutzt werden. Weiterhin bietet die Anwendung umfangreiche Funktionalitäten zur Nachnutzung der erfassten Daten in unterschiedlichen Kontexten und der Präsentation virtueller Kollektionen nach unterschiedlichen Kriterien in externen Systemen, wie z. B. Blogs, Wikis oder Content Management Systemen.

Die folgenden Abbildungen sollen einen Eindruck der Anwendung vermitteln:

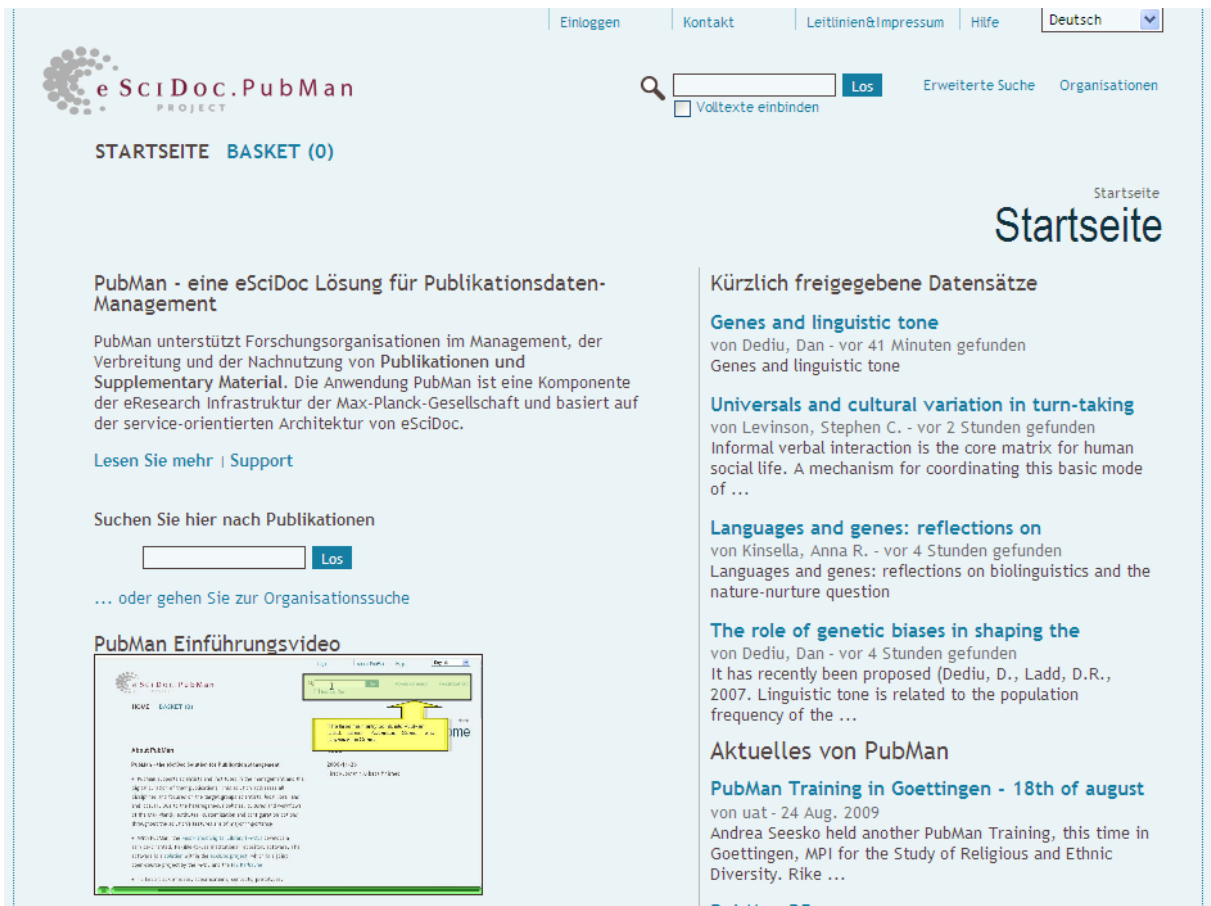


Abbildung 3: PubMan Homepage

demo Ausloggen Kontakt Leitlinien&Impressum Hilfe Deutsch

eSciDoc.PubMan PROJECT

Suche Los Volltexte einbinden Erweiterte Suche Organisationen

STARTSEITE MEINE DATENSÄTZE QA BEREICH DATENEINGABE BASKET (0)

Datensatz bearbeiten > Meine Datensätze > Suchergebnisse

ANSICHT SORTIERUNG EXPORT ZUM BASKET HINZUFÜGEN

Bibliographische Liste | Tabellarische Ansicht

333 Datensätze gefunden (absteigend sortiert nach Änderungsdatum)

25 Treffer je Seite Vorhergehende 1 2 3 4 5 6 7 Nächste Gehe zu Seite von 14

Status	Titel	Urheber	Genre	Datum	Dateien
<input type="checkbox"/> Freigegeben	Radiation effects on the mechanical properties of the material for the BESIII b...	Zheng, Lifang; Wang, Li; Wu, Ping; Ji, Quan, ...	Konferenzbeitrag	2008, Online veröffentlicht	
<input type="checkbox"/> Freigegeben	Cross-species mapping of bidirectional promoters enables prediction of unannota...	Piontkivska, Helen; Yang, Mary; Larkin, Denis; Lewin, Harris ...	Zeitschriftenartikel	2009, Erstellt	
<input type="checkbox"/> Freigegeben	Liver-enriched transcription factors are critical for the expression of hepatoc...	Kheolamai, Pakpoom; Dickson, Alan	Zeitschriftenartikel	2009, Erstellt	
<input type="checkbox"/> Freigegeben	Stilbene glycosides are natural product inhibitors of FGF-2-induced angiogenesis	Hussain, Sajjad; Slevin, Mark; Ahmed, Nessar; West, David ...	Zeitschriftenartikel	2009, Erstellt	
<input type="checkbox"/> Freigegeben	The readout system of the ATLAS liquid argon calorimeters	Biondel, A; La Marra, D; Leger, A; Perrot, G ...	Konferenzbeitrag	2002, Erstellt	

Abbildung 4: Listenansicht

demo Ausloggen Kontakt Leitlinien&Impressum Hilfe Deutsch

eSciDoc.PubMan PROJECT

Suche Los Volltexte einbinden Erweiterte Suche Organisationen

STARTSEITE MEINE DATENSÄTZE QA BEREICH DATENEINGABE BASKET (0)

Einfache Eingabe Detaillierte Eingabe Importieren Massenimport Import Bereich

Startseite > Start der Dateneingabe

Eingabemethode

Einfache Eingabe	Beschreibung Über die einfache Dateneingabe können Sie schrittweise Ihren Datensatz beschreiben. Am Ende der Eingabe haben Sie die Möglichkeit entweder den Datensatz abzuspeichern oder in die ausführliche Eingabemaske zu wechseln um weitere Angaben zu machen.
Detaillierte Eingabe	Beschreibung Die detaillierte Eingabemaske bietet eine dokumentenspezifische Maske zur Eingabe von Publikationsdaten. Des weiteren können ein oder mehrere Volltexte hochgeladen werden, sowie eine URL zu extern abgelegtem Material (z.B. Supplementary Material) gesetzt werden.
Importieren	Beschreibung Diese Funktion bietet zum einen den Import einer Datei, die einen BibTeX Eintrag enthält, zum anderen kann man hier Metadaten und Volltext(e) aus anderen Systemen, wie arXiv oder PubMed Central in die eSciDoc Eingabemaske laden.
Massenimport	Beschreibung Diese Funktion bietet zum einen den Import einer Datei, die mehrere BibTeX Einträge enthält, zum anderen kann man hier Metadaten (und Volltexte) aus anderen Systemen, wie SPIRES oder BioMed Central in die eSciDoc Eingabemaske laden.

Abbildung 5: Überblick der Eingabemethoden

Ein Beispiel der Nachnutzung von Publikationsdaten in virtuellen Kollektionen ist die Kollektion des Lebenswerks von Reinhard von Sengbusch aus dem Bereich der Züchtungsforschung:



About



On this website you find the publications of Prof. Dr. Reinhold von Sengbusch. Reinhold Oskar Kurt von Sengbusch (* 16 February 1898 in Riga; † 13 June 1985 in Hamburg) was a Director at the Max Planck Institut für Kulturpflanzenzüchtung in Hamburg-Volksdorf and one of the most important breeding researchers of the last century. In addition to many of his publications you can find out more about his [favorite topic](#), his [breeding achievements](#) and the [scientific impact](#) he had.

Reinhold von Sengbusch began to publish in 1924 and continued until the nineteen-eighties. Sengbusch alone and together with his colleagues published about 300 scientific articles. His institute produced about 600 publications. Over time the fulltext availability on this site will increase.

This collection is presented by the [Max Planck Digital Library](#). The

Pages

- » [About](#)
- » [Favorite Topic](#)
- » [Legal](#)
- » [Most Successful Topic](#)
- » [Most Yielding Topic](#)
- » [Scientific Impact](#)
- » [Vita](#)

Archives

Select Year

Categories

- » [Breeding](#) (19)
- » [Chemistry](#) (1)
- » [Educational](#) (3)
- » [Fish](#) (14)
- » [Carps](#) (6)
- » [Kidney stones](#) (17)
- » [Methods](#) (34)
- » [MPI](#) (11)
- » [Mushrooms](#) (37)

Abbildung 6: Virtuelle Kollektion Reinhold von Sengbusch

Hierbei werden die Publikationen durch PubMan verwaltet und die Präsentation durch ein Blog mit entsprechendem PubMan-Plugin ermöglicht. Die chronologische und Kategorien basierende Selektion wird hierbei direkt durch das Plugin angeboten.

Die Anwendung PubMan wird derzeit innerhalb der MPG eingeführt und wird das bestehende System ablösen.

Das System unterstützt die folgenden Szenarien:

Eingabe

- Easy Submission (kurze Eingabe) und Full Submission (detaillierte Eingabe)
- Dokumenttyp-spezifische Eingabe
- Metadaten abholen von: arXiv, PubMed Central, BioMed Central und SPIRES
- Dateien über eine externe Referenz importieren
- Mehrfach-Import inkl. Batch-Operationen von EndNote, WoS, RIS, BibTeX (inkl. Link zum Volltext)
- Batch Operationen beim Import
- Copyright Informationen, Vergabe von CC -Lizenzen zu Volltexten, Embargo-Zeiten*
- Vergabe von Zugriffsrechten auf Volltexte public/private/restricted (für definierte Nutzergruppe)
- Eingabe Vorschläge für Autorennamen, Journals und Subjects (basierend auf CoNe Identifier)
- Definierbare Validierungsregeln für einzelnen Kontexte

Suche

- Quick Search - Advanced Search
- Organization Search (inkl. Historie)
- Suche in Volltexten (pdf, doc, txt, xml), inkl. Markierung der Fundstellen

Researcher Portfolio

- Alle Publikationen pro Autor, optional: mit Angaben zur Person, Photo, CV, etc.
- Look-up in WorldCat und google Scholar

Versionierung

- Automatische Working Versionen (bei Änderung an Metadaten oder Anzahl Volltexte)
- Intellektuelle Versionen (Revisionen) für Relationen

Hilfe/ Komfort

- Blog Integration auf der Startseite (News)
- Kontext-sensitive Hilfe
- Screencast / Einführungsvideo
- Mehrsprachigkeit (Deutsch, Englisch, Japanisch)
- Automatische Mail für neuen Nutzer
- Change Password – Funktion

Workflows

- Simple Workflow: Submit und Release durch Depositor, Änderungen sind durch den Moderator nachträglich möglich
- Standard Workflow: Submit durch Depositor, Accept oder Send back for rework durch den Moderator – Release durch Moderator. Änderungen sind durch den Moderator nachträglich möglich (Änderungen auch durch Depositor nachträglich möglich)
- E-Mail Kontakt für bestimmtes Item (konfigurierbare Mail-Adresse)
- Arbeitsplätze: My Items, QA Workspace, Import Workspace
- Collaboration Scenario (Zugriff für bestimmte Nutzer vor dem eigentlichen Release)

Export/ Nachnutzung

- Export einzelner Datensätze oder Listen (sortiert) als EndNote, BibTeX, xml
- Export als download oder per Mailversand
- Unterstützung von Zitierstilen APA oder AJP als RTF, ODT, PDF, HTML oder Snippet (XML)
- Basket Export
- Verschiedene Listen-Ansichten (bibliografisch, tabellarisch)
- RSS/ATOM feeds, OAI-PMH
- Bookmarking Services (delicious, connotea, CiteULike)
- Schnittstelle Search&Export (REST) zur Einbindung in Webseiten
- Schnittstelle zu Open Search und Zotero
- Möglichkeit zur Stylesheet-Änderung im Browser
- Zugriff auf „private“ Volltexte für „privileged viewer“

Statistiken

- Angaben zu Zugriffen auf Metadaten und/oder Volltexte (Eingeloggte/ Anonyme Nutzer)
- Visualisierung

In Zusammenhang mit den flexiblen Möglichkeiten der eSciDoc Architektur findet PubMan international zunehmendes Interesse und wird verstärkt auch für Szenarien eingesetzt, die

ursprünglich nicht direkt vorhergesehen wurden, wie z.B. der Einsatz am Geoforschungszentrum Potsdam im Rahmen der Einstellung von eBooks¹⁴.

<http://pubman.mpdl.mpg.de/>

Virtueller Raum Reichsrecht (ViRR)

Der Virtuelle Raum Reichsrecht bietet eine kooperative Autorenumgebung zur Erstellung und Publikation von digitalen Kollektionen der vielfältigen Artefakte aus der Zeit des Heiligen Römischen Reichs Deutscher Nation. Die einzelnen Kollektionen werden digitalisiert, nach dem METS-Standard strukturiert, transkribiert und mit weiterer wissenschaftlich relevanter Literatur verknüpft.

Die folgende Abbildung soll einen Eindruck der Möglichkeiten der Anwendung vermitteln:

Schauroth, Eberhard Christian Wilhelm von:
Vollständige Sammlung Aller Concluserum, Schreiben Und anderer übrigen Verhandlungen Des Hochpreißlichen Corporis Evangelicorum nnnnn. Vol. 1, 1751

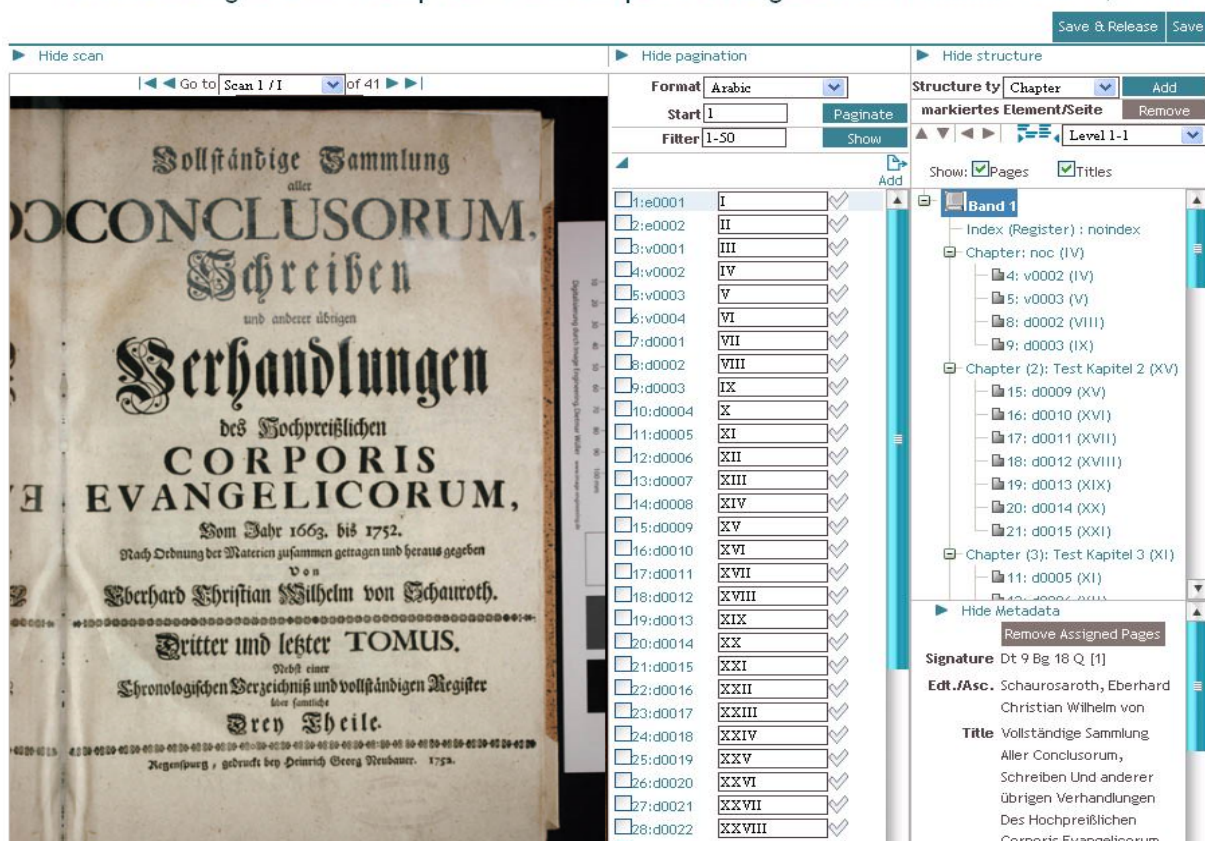


Abbildung 7: ViRR Editor

ViRR wird gemeinsam mit dem Max-Planck-Institut für Europäische Rechtsgeschichte entwickelt.

<http://virr.mpdl.mpg.de/>

FACES


FACES ist eine Fotosammlung von Gesichtern 171 verschiedener männlicher und weiblicher Personen unterschiedlicher Altersgruppen. Jede Person zeigt sechs unterschiedliche Emotionen durch Gesichtsausdrücke: Neutral, Traurigkeit, Ekel, Angst, Ärger und Glück. Für jede Person und jeden Gesichtsausdruck stehen jeweils zwei Aufnahmen bereit.

¹⁴ <http://ebooks.gfz-potsdam.de/>

Gleichzeitig ist FACES aber auch eine generische Bilddatenbank, die inzwischen z.B. auch in den Materialwissenschaften mit Fotos von Diamantschnitten befüllt wird. Über eine sehr flexible Handhabung von Metadaten lassen sich disziplinspezifische Anforderungen einfach umsetzen. Alben erlauben Wissenschaftlern, beliebige Zusammenstellungen von Bildern für ein Projekt oder eine Fragestellung vorzunehmen. Diese Alben können als Datenpublikation wie eine klassische Publikation mit den entsprechenden bibliographischen Angaben publiziert und persistent verfügbar gehalten werden.

Home Create Album My Albums Published Albums Statistics Browse Search

Expression Research on Middle Aged Men

published 

Album Url	http://test-faces.mpdl.mpg.de/album/escidoc:41047	
Dates	Date Created: 2009-03-27 at 11:39:59 CET	Date published: 2009-03-27 at 13:49:11 CET
Author	Kleinfercher , Friederike Büchner , Kristina	Max Planck Digital Library Max Planck Digital Library
Description	This album contains a selection of middle aged men with different facial expressions.	

Export Album

Show of 8 hits Previous Next

Go to page of 1

Sorted by **Emotion** Person-ID Picture Set Sort

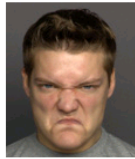

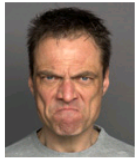









 	 	 	 	 	 
---	---	---	---	--	---

Abbildung 8: FACES Publikation von Alben

FACES wird gemeinsam mit dem Max-Planck-Institut für Bildungsforschung entwickelt.

<http://faces.mpdl.mpg.de/>

Administrationswerkzeug

Das Administrationswerkzeug dient zur Verwaltung der eSciDoc Infrastructure und zur Konfiguration der Solutions. Hierunter fallen Aufgaben wie das Re-Indexieren des Suchindex, das Neubefüllen des Caches, das Laden von Beispielobjekten oder das Löschen von Einträgen wie das Anlegen von Contexts, Benutzern sowie die Verwaltung ihrer Rechte.

Das Administrationswerkzeug wird laufend ausgebaut und stellt einen wichtigen Baustein für den Produktivbetrieb dar. Aus Sicherheitsgründen gibt es keine öffentlich zugängliche Installation.

2. Zahlenmäßiger Nachweis

Einhaltung der Kosten- und Zeitplanung

Das Projekt wurde im vorgegebenen Zeit- und Kostenplan durchgeführt. Die während des Projekts entstandenen Verzögerungen in einzelnen Themenbereichen konnten im weiteren Projektverlauf wieder aufgeholt werden. Eine detaillierte Kostenübersicht ist im Erfolgskontrollbericht und den entsprechenden Formblättern des BMBF dargestellt.

Einhaltung der Bewilligungsbestimmungen

Die allgemeinen Bewilligungsbestimmungen des BMBF wurden eingehalten. Das zuständige Referat im BMBF wurde laufend auf dem Stand der aktuellen Planungen gehalten. Der PTF wurde ebenfalls informiert, insbesondere durch die aktuellen Planungspapiere. Auch geplante größere Beschaffungen wurden jeweils vorab mit dem PTF diskutiert.

3. Notwendigkeit und Angemessenheit der geleisteten Arbeit

Auch nach Projektende gibt es nur wenige vergleichbare, integrierte und tatsächlich im täglichen Einsatz befindliche e-Science-Plattformen, die mehrere Arbeitsumgebungen gleichzeitig unterstützen, eine durchgehende und konsistente Datenhaltung („continuum of data“) über den gesamten wissenschaftlichen Arbeitsprozess hinweg bieten. Damit ist eSciDoc auch fünf Jahre nach Projektbeginn immer noch ein weitgehend singuläres Angebot, das im akademischen Umfeld auf hohe Aufmerksamkeit und großes Interesse stößt.

Vergleichbare Projekte (etwa Bamboo¹⁵ in den USA) sind noch in der Diskussionsphase. Andere Projekte, wie etwa die EU-geförderten Diligent¹⁶ und BRICKS¹⁷, fokussieren sich auf Digitale Bibliotheken, also auf statische Inhalte. eSciDoc jedoch unterstützt daneben auch dynamische Inhalte, wie sie im wissenschaftlichen Arbeitsprozess laufend entstehen.

Gleichzeitig zeigt das hohe Interesse in der nationalen und internationalen Forschungsgemeinde die Notwendigkeit einer solchen e-Science-Plattform. Abbildung 9 zeigt die zum Projektende von eSciDoc bereits die Software einsetzenden oder evaluierenden Partner in Deutschland und Europa, wie sie in der folgenden Liste wiedergegeben sind:

- Academy of Athens, Athen, Griechenland
- Alfred-Wegner-Institut (AWI), Bremerhaven
- Bayerische Staatsbibliothek, München
- CLARIN Project (EU FP7)
- CNRS, Frankreich
- DANS, Den Haag, Niederlande
- Deutsches Archäologisches Institut, Berlin
- DTU, Lyngby, Dänemark
- Fachhochschule Potsdam
- FAO (Food and Agriculture Organization of the United Nations), Rom, Italien
- Verbundzentrale/Gemeinsamer Bibliotheksverbund, Göttingen
- GeoForschungsZentrum Potsdam
- Technische Informationsbibliothek (TIB) Hannover
- Humboldt Universität, Berlin
- IRCHSS, Dublin, Irland
- King's College, London, Großbritannien
- National Institute of Material Science, Tokyo, Japan
- Universität Nihon, Japan
- RWTH Aachen
- SURF Foundation, Niederlande
- UNED, Madrid, Spanien
- University of Amsterdam
- University of Copenhagen, Dänemark

¹⁵ <http://projectbamboo.org/>

¹⁶ <http://diligent.ercim.org/>

¹⁷ <http://www.brickcommunity.org/>

- Universität Freiburg
- Universität Heidelberg
- Universität Münster
- Universität Stuttgart
- Universität Tübingen

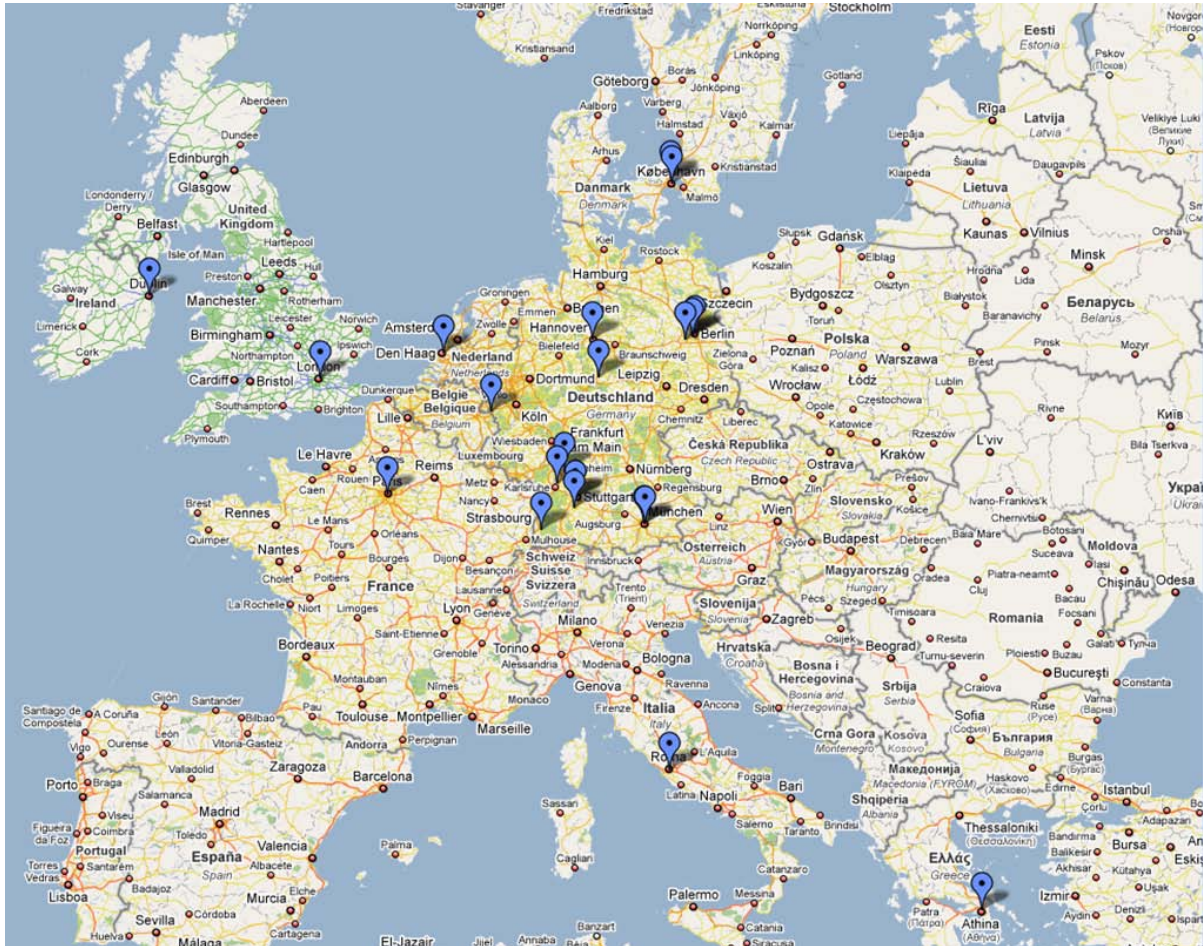


Abbildung 9: Geografische Verteilung der Institutionen in Europa, die bis Mitte 2009 eSciDoc bereits einsetzen oder evaluieren. Weitere Nutzer außerhalb Europas finden sich vorwiegend in den USA und Japan.

Damit scheint die Notwendigkeit der geleisteten Arbeit hinreichend belegt. Hinsichtlich der Angemessenheit wurde während der Projektlaufzeit stets darauf geachtet, geforderte Funktionalität gegen bestehende Ergebnisse anderer Projekte abzugleichen, um Doppelarbeit zu vermeiden und so viele Komponenten wie möglich nachzunutzen. In diesem Sinne wurden folgende Komponenten eingebunden, statt sie selbst zu entwickeln:

- Repository: Fedora Commons
- Datenbank: PostgreSQL
- Semantic Store: MPT und Mulgara
- Application Server: JBoss Application Server
- Duplikatserkennung: Code von arXiv, Cornell University
- Extraktion technischer Metadaten: JHOVE und FITS, Harvard University
- Bildbearbeitung im Browser: Digilib, Universität Bern und MPI für Wissenschaftsgeschichte
- Rechteverwaltung: XACLM Engine, Sun Microsystems
- Verteilte Authentifizierung: Shibboleth, Internet2
- Diverse Programmierbibliotheken und Frameworks

Die breite Verwendung bestehender Komponenten sicherte einerseits den effizienten Einsatz der Projektressourcen, andererseits die Qualität des Gesamtsystems durch Einsatz weitgehend ausgereifter und verlässlicher Software.

4. Verwertbarkeit der Ergebnisse

Wie bereits im vorangegangenen Abschnitt erläutert, besitzen die Ergebnisse des Projekts eSciDoc eine hohe Relevanz für die deutsche und internationale wissenschaftliche Community. Bereits vor Abschluss des Projekts nutzen Institutionen weltweit die Ergebnisse nach. Diesen Erfolg beförderten die Projektpartner FIZ Karlsruhe und Max-Planck-Gesellschaft durch diverse Maßnahmen:

- Diverse Präsentationen und Publikationen;
- Veröffentlichung der gesamten Software als Open Source unter einer vom Open Software Institute (OSI) anerkannten Lizenz (CDDL, Common Development and Distribution License). Seit Ende 2007 sind der Programmcode sowie vorkompilierte Pakete öffentlich unter <https://www.escidoc.org/> verfügbar;
- Veröffentlichung umfangreicher Dokumentation zu den Konzepten, dem Design der Software sowie zu den Programmierschnittstellen;
- Angebot von Mailinglisten sowie Schulungen;
- Veranstaltung der eSciDoc Days mit jeweils 80-100 Teilnehmern, denen jeweils 1 ½ Tage Fachvorträge, Tutorials und Workshops geboten wurden.

Mit diesen Maßnahmen konnte bereits während der Projektlaufzeit eine Benutzercommunity innerhalb und außerhalb der Max-Planck-Gesellschaft aufgebaut werden.

Eine Besonderheit beim Aufbau einer solchen e-Research-Umgebung liegt in der strikten Forderung der Nutzer, also der Wissenschaftler und Bibliothekare, nach Offenheit des Systems. Das bezieht sich nicht nur auf Programmierschnittstellen, um das System mit anderen Programmen und Geräten zu verknüpfen, sondern auch in der Offenheit des Source Codes im Sinne von Open Source. Damit scheidet eine kommerzielle Verwertung der Projektergebnisse im Sinne einer zu lizensierenden Software aus. Diese hätte aber im akademischen Umfeld ohnehin einen schweren Stand gehabt. Die Entscheidung für Open Source befördert die Verbreitung der Software und schafft damit gleichzeitig einen Markt für Dienstleistungen um die Software herum, etwas im Bereich Beratung, Schulung, Betrieb der Software für Dritte („software as a service“, SaaS) und Supportverträgen. Auch hier konnte FIZ Karlsruhe bereits während der Projektlaufzeit erste, wenn auch geringe, Einnahmen erzielen. Dieses Geschäftsfeld wird nach Projektende kontinuierlich ausgebaut und intensiviert.

5. Fortschritte auf dem Gebiet des Vorhabens

Repository-Systeme

Repository-Systeme bilden das Rückgrat der meisten digitalen Bibliotheken, Forschungs-umgebungen und e-Science-Infrastrukturen. Während zu Anfang des Projekts das Open Society Institute¹⁸ noch neun unterschiedliche Softwaresysteme untersuchte, die die Kriterien „Open Source“, „unterstützt OAI-PMH“ und „verfügbar zum Zeitpunkt der Studie“ erfüllten, sind aus heutiger Sicht davon nur noch drei international relevant: DSpace, ePrints und Fedora. In Deutschland ist OPUS nach wie vor weit verbreitet, konnte allerdings nie außerhalb von Deutschland Fuß fassen.

¹⁸ Open Society Institute (2004): *A Guide to Institutional Repository Software*. Abgerufen von http://www.soros.org/openaccess/pdf/OSI_Guide_to_IR_Software_v3.pdf

Die zu Beginn des Projekts noch als hochinteressante Entwicklung eingestufte Software aDORe¹⁹ des Los Alamos National Laboratory konnte sich in den folgenden Jahren nicht weiter durchsetzen und spielt heute keine nennenswerte Rolle.

Im Mai 2009 haben die beiden Organisationen hinter DSpace und Fedora beschlossen, zusammenzugehen, so dass man insgesamt von einer starken Konsolidierung der internationalen Repository-Landschaft ausgehen kann.

Umgang mit Forschungsdaten (Langzeitarchivierung, Publikation, Erschließung)

Forschung – insbesondere in den Naturwissenschaften – ist immer stärker datengetrieben. Es fallen gemessene Primärdaten, kalibrierte und analysierte Daten sowie abgeleitete und aus Simulationen gewonnene digitale Daten an. Im Sinne „guter wissenschaftlicher Praxis“ wird heute national wie international von Wissenschaftlern erwartet, in Projekten verwendete und erzeugte Daten längerfristig vorzuhalten und auch anderen Forscherteams zur Nachnutzung zur Verfügung zu stellen. Daten sollen Teil von Publikationen werden bzw. leicht mit diesen verlinkt werden können oder Datensätze werden selbst in Form einer Publikation verfügbar gemacht – wie z.B. im neuen Earth System Science Data Journal²⁰.

Eindeutige Identifier garantieren die langfristige Identifizierbarkeit von Datensätzen. So bietet die TIB Hannover bereits seit 2005 die Vergabe von DOI (Digital Object Identifier) für Datensätze an – zunächst für Geo- und Umweltwissenschaften, künftig auch für weitere Disziplinen. Diese Datensätze können über die Identifier wiederum in verschiedene Repositories bzw. digitale Forschungsumgebungen eingebunden werden. International sind vor allem die Initiativen des ANDS (Australian National Data Service) sowie der NSF (National Science Foundation (USA): Programm DataNet) zu nennen, in denen systematisch die langfristige Verfügbarkeit und Nachnutzbarkeit von Forschungsdaten gefördert wird.

Die für eine Vielzahl von Objekttypen geeignete eSciDoc-Infrastruktur bietet den Wissenschaftlern die notwendige Unterstützung ihre Daten systematisch zu verwalten, mit Dokumenten und Publikationen zu verlinken, speziellen Anwendungen zur Verfügung zu stellen, über die Rechteverwaltung bestimmten Nutzergruppen verfügbar zu machen oder auch als eigenständige Identität und mit beschreibenden und technischen Metadaten versehen für eine allgemeine Veröffentlichung vorzubereiten. Darüber hinaus sind damit auch Voraussetzungen für das Einbinden von Datensätzen in Langfristarchivierungs-Aktivitäten geschaffen.

Service-orientierte Architekturen in der Wissenschaft, insbesondere Service Registries

eSciDoc als service-orientierte Architektur gliedert die angebotene Funktionalität in von einander weitgehend unabhängige Dienste, die auch einzeln angesprochen werden können. Damit erreicht man mehrere Ziele: statt eines monolithischen Systems erhält man kleine, leichter wartbare und austauschbare Einheiten; die Gesamtfunktionalität kann durch Hinzufügen weiterer Dienste einfach ergänzt und das Gesamtsystem kann räumlich verteilt installiert und angesprochen werden. Dienste müssen also nicht notwendigerweise an einem Ort installiert sein, sondern können über das Internet verbunden miteinander kommunizieren. Auch kann eine geforderte Funktionalität von mehreren konkurrierenden Diensten angeboten werden. Damit stellt sich die Frage der Auffindbarkeit von Diensten („Discovery“). Hierfür dienen Service Registries, wie sie z.B. der Universal Description, Discovery and Integration (UDDI) Standard beschreibt. Während der Antragstellung erschien dieser Standard noch zukunftssträftig, doch konnte er sich

¹⁹ <http://african.lanl.gov/aDORe/projects/adoreArchive/>

²⁰ <http://www.earth-system-science-data.net/>

zumindest außerhalb von geschlossenen Systemen (wie z.B. innerhalb einer Firma) nicht durchsetzen. Insbesondere im wissenschaftlichen Umfeld mit hochgradig verteilten und weitgehend unkoordinierten Arbeiten in vielen, voneinander vollkommen unabhängigen Projekten entstehen zwar eine Vielzahl wertvoller Dienste, aber eine einheitliche Beschreibung und eine zentrale oder föderierte Suche nach solchen Diensten gibt es bis heute nicht.

Das leider nur als Vorprojekt geförderte Vorhaben links4science²¹ hatte die Integration der resultierenden technologischen Lösungen und Dienste aus einer Vielzahl von im Rahmen der BMBF-Initiative "e-Science-Wissensvernetzung" geförderten Projekte zu "e-Science und vernetztes Wissensmanagement" zum Ziel. Letztlich wäre ein Ergebnis des Projekts eine nationale Service Registry gewesen, die aber aufgrund des nicht mehr geförderten Hauptprojekts nie angegangen wurde.

Gleichzeitig fand in den letzten Jahren aber auch ein Paradigmenwechsel im Bereich der Service-Schnittstellen statt. Getrieben durch das Web 2.0 mit seinen vielfältigen Diensten und der Idee von sogenannten „Mashups“ (also der einfachen Verknüpfung mehrerer Dienste zu einem aggregierten Dienst) reduzierte sich die Komplexität von Schnittstellen und Protokollen. Heute bieten Dienste meist eine REST-Schnittstelle²², die auf dem HTTP-Protokoll basiert. So hat auch eSciDoc neben einer SOAP für alle Dienste auch eine REST-Schnittstelle implementiert. Einige der neueren Dienste implementieren nur noch die REST-Schnittstelle, die sich über herkömmliche Service Registries nur schwer abbilden lassen.

Virtuelle Forschungsumgebungen

Zum Zeitpunkt der Antragsstellung wurde unter e-Science ganz überwiegend Grid-Technologie für rechen- und speicherintensive Aufgaben im Bereich der Naturwissenschaften verstanden. Aber nicht alle Fragestellungen lassen sich mit Mitteln der Grid-Technologie gleichermaßen gut bearbeiten. Auch die tiefergehende Integration der dort erzielten Ergebnisse mit Publikationen spielte damals noch keine Rolle.

eSciDoc setzte sich zum Ziel, diese Integration voranzutreiben und dabei auch für andere Disziplinen bzw. Fragestellungen geeignete Werkzeuge zu entwickeln, weniger in Konkurrenz zu Grid-Technologien, sondern vielmehr in Ergänzung zu diesen.

Im Projektverlauf entwickelte sich der Begriff der „virtuellen Forschungsumgebungen“ für solche Ansätze. Die Office of Science and Technology e-Infrastructure Working Group definiert eine virtuelle Forschungsumgebung als *“set of online tools, systems and processes interoperating to facilitate or enhance the research process within and without institutional boundaries. The purpose of a VRE is to provide researchers with the tools and services they need to do research of any type as efficiently and effectively as possible. This means VREs will help individual researchers manage the increasingly complex range of tasks involved in doing research. In addition [VREs] will facilitate collaboration among communities of researchers, often across disciplinary and national boundaries.”*²³ Diese Definition charakterisiert gleichzeitig wesentliche Eigenschaften von eSciDoc.

Authentifizierung und Autorisierung

Einer der Ziele von eSciDoc war, einen Kristallisationspunkt für eine nationale e-Science-Infrastruktur zu bilden. In diesem Sinn war von vorneherein klar, dass für möglichst viele Dienste auf (sich teilweise noch entwickelnde) Standards gesetzt werden sollte. Weiterhin ist eines der

²¹ <http://www.lgmmia.fernuni-hagen.de/projects/links4science/index.html>

²² Roy Fielding (2000). *Architectural Styles and the Design of Network-based Software Architectures*. Abgerufen von: http://www.ics.uci.edu/~fielding/pubs/dissertation/fielding_dissertation.pdf

²³ VRC Final Report (2006). *Report of the Working Group on Virtual Research Communities for the OSI e-Infrastructure Steering Group*. Abgerufen von <http://www.nesc.ac.uk/documents/OSI/vrc.pdf>

herausragenden Merkmale von e-Science (bzw. e-Research) die Zusammenarbeit über Instituts- und nationale Grenzen hinweg. Aus beiden Anforderungen ergab sich die Notwendigkeit, auf eine föderierte Authentifizierung und Autorisierung zu setzen. Die Wahl fiel dabei auf Shibboleth²⁴, das auch die D-Grid-Initiative einsetzt. Allerdings hat sich Shibboleth auf breiter Basis auch in den fünf Jahren des Projekts nicht durchsetzen können. Das liegt einerseits an der Komplexität des Protokolls und der darunterliegenden Software, aber auch an den hohen Anforderungen an die Partner in Shibboleth-Föderationen. Gleichzeitig gibt es einen Trend zu breiter genutzten Technologien wie z.B. OpenID²⁵ oder Information Card²⁶. Aus diesem Grund setzt eSciDoc nicht ausschließlich auf Shibboleth, sondern setzt ein erweiterbares Framework ein, über das momentan auch eine lokale Benutzerdatenbank und über LDAP herkömmliche Directory-Server ansprechbar sind. Weitere Authentifizierungsmethoden lassen sich recht einfach integrieren.

Auch im Bereich der Zugriffsrechte hat sich in den letzten Jahren mit XACML (eXtensible Access Control Markup Language)²⁷ ein Standard etabliert, der zwar sehr mächtig, aber gleichzeitig auch sehr komplex ist. Viele Projekte limitieren die Mächtigkeit (z.B. Muradora²⁸), um auch Endbenutzern die Möglichkeit zu geben, eigene Policies zu formulieren und zu prüfen. Bei Zugriff auf große Objektmengen (z.B. bei Suchen oder Datenanalysen) kann die Auswertung der komplexen Regeln für jedes einzelne Objekt zu Geschwindigkeitsproblemen führen, weshalb zum Teil auch andere Technologien wie z.B. RBAC (Role-based Access Control)²⁹ zum Einsatz kommen. Auch für eSciDoc sind diese Überlegungen relevant.

6. Veröffentlichungen

Publikationen

Leni Helmes. 2006. E-Science: Web 2.0 für die Wissenschaft. Zu den konzeptionellen Grundlagen eines künftigen Arbeitsschwerpunktes von FIZ Karlsruhe. Password 11/2006

Malte Dreyer, Ulla Tschida, Natasa Bulatovic, Matthias Razum. 2007. eSciDoc – a Scholarly Information and Communication Platform for the Max Planck Society. In: German e-Science Conference, Baden-Baden.

Matthias Razum, Frank Schwichtenberg, Rozita Fridman . 2007. Versioning of Digital Objects in a Fedora-based Repository. In: German e-Science Conference, Baden-Baden.

Leni Helmes, Karl-Heinz Weber. 2007. E-Science: erweiterte Formen netzbasierten wissenschaftlichen Arbeitens. openource Wissenschaftsmagazin, Heft Okt/Nov 2007, S. 60-62.

Sabine Brünger-Weilandt. 2007. e-Science – Advancing new ways of scientific communication. Information Services & Use, Special Issue: APE 2007, Volume 27, Number 4, p. 161-166.

Ute Rusnak. 2008. MPG und FIZ Karlsruhe öffnen e-Science-Plattform für die wissenschaftliche Gemeinschaft – Bericht über die eSciDoc Days in Berlin. Information Wissenschaft & Praxis IWP, Vol. 59 August 2008, Nr.5.

2008. Projekt eSciDoc: MPG und FIZ Karlsruhe öffnen e-Science-Plattform für die Forschungsgemeinde. IDW online [about eSciDoc Days](#)

²⁴ <http://shibboleth.internet2.edu/>

²⁵ <http://openid.net/>

²⁶ <http://informationcard.net/>

²⁷ <http://www.oasis-open.org/committees/xacml/>

²⁸ <http://www.muradora.org/>

²⁹ <http://csrc.nist.gov/groups/SNS/rbac/>

³⁰ <http://idw-online.de/pages/de/news?print=1&id=260351.>

Natasa Bulatovic, Ulla Tschida, Andreas Gros. 2008. eSciDoc – A Service Infrastructure for Cultural Heritage Content. 14th International Conference on Virtual Systems and MultiMedia (VSMM), Limassol, Cyprus.

Malte Dreyer, Andreas Aschenbrenner. 2008. Digital Library Architecture: Service Patterns for Large-scale Digital Libraries. First Workshop on Very Large Digital Libraries In conjunction with the 12th European Conference on Research and Advanced Technologies on Digital Libraries (ECDL).

Malte Dreyer, Ulla Tschida. 2009. eSciDoc – Das Repository-Konzept der Max Planck Digital Library. CMS-Journal Computer- und Medienservice.

Ute Rusnak. 2009. Wissenschaft digital: Von elektronischen Publikationen zu virtuellen Forschungsumgebungen. ABI Technik, 3. Vj.

Matthias Razum, Frank Schwichtenberg, Steffen Wagner, Michael Hoppe. 2009. eSciDoc Infrastructure: A Fedora-Based e-Research Framework. M. Agosti et al. (Eds.): ECDL 2009, LNCS 5714, pp. 227–238.

Präsentationen

2005

Gerhard Beier. 2005. eSciDoc - eine integrierte Informations-, Kommunikations- und Publikationsplattform für netzbasiertes wissenschaftliches Arbeiten. 22. DV-Treffen der Max Planck Gesellschaft, Göttingen, Deutschland.

Gerhard Beier. 2005. eSciDoc - eine integrierte Informations-, Kommunikations- und Publikationsplattform für netzbasiertes wissenschaftliches Arbeiten. ASpB-Tagung 2005, München, Deutschland.

Andreas Barth. 2005. E-Science: Concepts and Challenges. 12th German-Japanese Workshop on Chemical Information, Dresden, Deutschland.

Sabine Brünger-Weilandt. 2005. New Ways of Working in Science. ICI Conference, Nîmes, Frankreich.

Sabine Brünger-Weilandt. 2005. Innovative Konzepte als Chancen für eine nationale e-Science-Infrastruktur. E-Science-Forum, Berlin, Deutschland.

2006

Andreas Barth, Leni Helmes, Matthias Razum. 2006. eSciDoc. An Integrated Information, Communication, and Publication Platform for Net-based Scientific Collaboration. E-Science Workshop, Tokyo, Japan.

Andreas Barth, Leni Helmes, Matthias Razum. 2006. eSciDoc. An Integrated Information, Communication, and Publication Platform for Net-based Scientific Collaboration. I & D Panel, Tokyo, Japan.

Matthias Razum. 2006. eSciDoc – Aufbau einer integrierten Informations-, Kommunikations- und Publikationsplattform für die Forschung. Wissenschaftsvernetzung – Konzepte, Erfahrungen, Perspektiven, Netzwerk Technologiefolgenabschätzung, Berlin, Deutschland.

Matthias Razum. 2006. eSciDoc Projektübersicht. Kick Off e-Science Wissensvernetzung, Bonn, Deutschland.

Matthias Razum. 2006. eSciDoc - Fedora for e-Scholarship and Scholarly Publication. The Fedora Service Framework - Advanced Applications, Joint Conference on Digital Libraries, Chapel Hill, USA.

Andreas Barth. 2006. E-Science. Allianz-Treffen der Bibliothekssprecher, Hannover, Deutschland.

Leni Helmes. 2006. eSciDoc – Stand und Ausblick. Deutsche Physikalische Gesellschaft (Arbeitskreis Information – AKI), Berlin, 12.10.2006

Matthias Razum. 2006. The eSciDoc Project- an Overview. Fedora User's Meeting, Charlottesville, USA.

Matthias Razum. 2006. Versioning of Digital Objects in a Fedora-based Repository. Workshop on Digital Object Repository Systems in Digital Libraries, European Conference on Digital Libraries, Alicante, Spanien.

Matthias Razum. 2006. eSciDoc – a Scholarly Information and Communication Platform in the Age of e-Science. Workshop on Digital Libraries and eScience, Alicante, Spanien.

Leni Helmes, Matthias Razum. 2006. eSciDoc – Stand und Ausblick. 70. Jahrestagung der DPG (Arbeitskreis Information – AKI), München, Deutschland.

2007

Natasa Bulatovic. 2007. eSciDoc – Content Model Requirements. Fedora Architecture Summit, Ithaca, USA.

Malte Dreyer. 2007. eSciDoc – A Service Oriented Architecture. eScience-Seminar "Service and Information Registries", Göttingen, Deutschland.

Ulla Tschida. 2007. Solution Publication Management – Cooperation with Early Adopters. Kickoff Workshop "Early Adopter – PubMan", Nijmegen, Niederlande.

Malte Dreyer. 2007. eSciDoc Status. XXX. Bibliothekstagung der Max-Planck-Gesellschaft, Marburg, Deutschland.

Malte Dreyer. 2007. eSciDoc. Allianz-Meeting der Bibliotheken der Max-Planck-Gesellschaft, München, Deutschland.

Malte Dreyer. 2007. eSciDoc. D-Grid All Hands Meeting, Göttingen, Deutschland.

Malte Dreyer. 2007. eSciDoc Cooperative Open Source Software Development. 24. DV-Treffen der Max-Planck-Gesellschaft, Jena, Deutschland.

Malte Dreyer. 2007. eSciDoc – eScience-Infrastruktur für digitale Assets. Arbeitstagung Digital Asset Management, Wien, Österreich.

Matthias Razum. 2007. The eSciDoc Project – an Overview. Open Repositories, San Antonio, USA.

Matthias Razum, Frank Schwichtenberg, Rozita Fridman. 2007. Versioning of Digital Objects in a Fedora-based Repository. Fedora User Group Meeting, Open Repositories, San Antonio, USA.

Matthias Razum. 2007. Fedora als Repository für E-Science-Szenarien. Sun Bibliotheken Summit, Frankfurt, Deutschland.

Matthias Razum. 2007. Eine disziplinübergreifende Arbeitsumgebung für die MPG. II. Göttinger Grid-Seminar, Göttingen, Deutschland.

Sabine Brünger-Weilandt, Leni Helmes. 2007. e-Science – Advancing New Ways of Scholarly Communication. Academic Publishing in Europe / Innovation & Publishing, Berlin, Deutschland.

Matthias Razum. 2007. eSciDoc: una plataforma de nueva generación para la información y la comunicación científica. VII Workshop REBIUN sobre proyectos digitales, Madrid, Spanien.

Matthias Razum. 2007. eSciDoc: a Next-generation e-Research Infrastructure. E-Infrastructures in the Arts and Humanities and Social Sciences: Grid-enabling Data Sets, Edinburgh, Großbritannien.

Ute Rusnak. 2007. FIZ Karlsruhe und seine e-Science-Dienstleistungen. 3. Leipziger Kongress für Information und Bibliothek, Leipzig, Deutschland.

Ute Rusnak. 2007. e-Science-Dienstleistungen von FIZ Karlsruhe. IuK 2007 & ISI 2007, Köln, Deutschland.

Frank Schwichtenberg. 2007. eSciDoc and Fedora Digital Repository. Source Talk Tage 2007, Göttingen, Deutschland.

2008

Malte Dreyer. 2008. Das Projekt eSciDoc. Deutsche Forschungsgemeinschaft (DFG), Workshop "Förderung der wissenschaftlichen Informationslandschaft in Deutschland – Chancen und Strategien zum Aufbau vernetzter Repositorien, Berlin, Deutschland.

Malte Dreyer. 2008. eSciDoc – Überblick und Solutions. XXXI. Bibliothekstagung der Max-Planck-Gesellschaft, Jena, Deutschland.

Ulla Tschida. 2008. Stand eSciDoc Solutions – Migration eDoc-PubMan. XXXI. Bibliothekstagung der Max-Planck-Gesellschaft, Jena, Deutschland.

Natasa Bulatovic. 2008. eSciDoc – Service Infrastructure and Solutions. eSciDoc Days, Berlin, Deutschland.

Ulla Tschida. 2008. Current eSciDoc Solutions – an Overview. eSciDoc Days, Berlin, Deutschland.

Malte Dreyer. 2008. eSciDoc Community Model-Draft. eSciDoc Days, Berlin, Deutschland.

Frank Schwichtenberg. 2008. eSciDoc Infrastructure. eSciDoc Days, Berlin, Deutschland.

Malte Dreyer, Matthias Razum. 2008. eSciDoc Introduction. eSciDoc Days, Berlin, Deutschland.

Matthias Razum. 2008. eSciDoc Functional Roadmap. eSciDoc Days, Berlin, Deutschland.

Natasa Bulatovic. 2008. The Role of dLTP in the eSciDoc Project. eScience-Seminar "Aspects of Long-term Archiving", Göttingen, Deutschland.

Malte Dreyer. 2008. Persistent Identifier for Long-term Archived Data. eScience-Seminar "Aspects of Long-term Archiving", Göttingen, Deutschland.

Wolfgang Voges, Jürgen Renn. 2008. eSciDoc – Anwendungen in der wissenschaftlichen Praxis in der MPG. Workshop des Ministeriums für Wissenschaft, Forschung und Kunst Baden-Württemberg: Informationsvernetzung im wissenschaftlichen Forschungsprozess, Stuttgart, Deutschland.

Malte Dreyer. 2008. Max Planck Digital Library (MPDL) – Supporting the Scientific Information Workflow within the Max Planck Society. International Colloquium "Autoren, Bibliotheken und Verlage – Regionale und nationale Herausforderungen der Digitalisierung in Europa", Lissabon, Portugal.

Malte Dreyer. 2008. Langzeitarchivierung in der MPG – eSciDoc. Workshop "Archivierung von wissenschaftlichen Daten", Helmholtz-Zentrum Berlin für Materialien und Energie, Berlin, Deutschland.

Malte Dreyer. 2008. eSciDoc Overview. National Institute for Materials Science, Tsukuba, Japan.

Ulla Tschida. 2008. eSciDoc in der MPG: Wohin geht die Reise? Jahrestagung der Arbeitsgemeinschaft für medizinisches Bibliothekswesen AGMB e.V., Magdeburg, Deutschland.

Malte Dreyer. 2008. eSciDoc in der MPG: Wohin geht die Reise? Herbstworkshop der IuK-Initiative, Berlin, Deutschland.

Malte Dreyer. 2008. Stand und Ausblick eSciDoc. MPG PubMan Days, Berlin, Deutschland.

Nicole Kondic. 2008. Aktuelle Release PubMan – Demonstration. MPG PubMan Days, Berlin, Deutschland.

Tom Endres. 2008. eSciDoc Build and Development Environment. 25. DV-Treffen der Max-Planck-Gesellschaft "Diversität der Ideen 2.0", Göttingen, Deutschland.

Malte Dreyer. 2008. eSciDoc: The Open Source eResearch Environment. Library Fair & Forum 2008, Yokohama, Japan.

Malte Dreyer. 2008. eSciDoc Cooperation Meeting with the Institute for Archaeology of the Ruhr-Universität Bochum, Dortmund, Deutschland.

Ute Rusnak. 2008. KnowEsis, e-Science–Dienstleistungen für die Wissenschaft. 72. Jahrestagung der DPG (Arbeitskreis Information - AKI), Berlin, Deutschland.

Ute Rusnak. 2008. KnowEsis, die e-Science-Solutions von FIZ Karlsruhe. LEARNTEC, Karlsruhe, Deutschland.

Ute Rusnak. 2008. KnowEsis, die e-Science-Solutions von FIZ Karlsruhe, 10. InetBib-Tagung, Würzburg, Deutschland.

Michael Hoppe, Matthias Razum. 2008. An SRW/U-compliant search service for Fedora. Fedora User Group Meeting, Open Repositories, Southampton, Großbritannien.

Matthias Razum. 2008. eSciDoc – an e-Research Infrastructure. DRESNET Workshop, Newcastle, Großbritannien.

Ute Rusnak. 2008. e-Science – der wissenschaftliche Arbeitsplatz der Zukunft. Buchmesse Forum Wissenschaft, Frankfurt/Main, Deutschland.

Matthias Razum. 2008. eSciDoc – an e-Research Infrastructure. Adding Value to Data – Digital Repositories in the e-Science World, IEEE e-Science Conference, Minneapolis, USA.

Rainer Stuike-Prill, Ute Rusnak. 2009: eSciDoc – Aufbau von digitalen Forschungsumgebungen, Achema – Discussion Corner, Frankfurt, Deutschland.

2009

Malte Dreyer. 2009. eSciDoc Eine wissenschaftliche Informations-, Kommunikations- und Publikationsplattform für die Forschung. HyperImage-Konferenz, Berlin, Deutschland.

Malte Dreyer. 2009. eSciDoc A scientific collaboration, communication and publication platform for research. Asia-Pacific Conference on Library & Information Education and Practice (A-LIEP 2009), Tsukuba, Japan).

Ulla Tschida. 2009. Creating an eResearch environment - Lessons learned. A-LIEP Conference 2009, Tsukuba/Japan.

Natasa Bulatovic. 2009. eSciDoc - Object and content modelling experiences. DARIAH Project Meeting WP7/WP8, Deutschland.

Nicole Kondic, Kristina Büchner. 2009. eSciDoc Solutions. eSciDoc Days, Karlsruhe, Deutschland.

Natasa Bulatovic. 2009. eSciDoc Services. eSciDoc Days, Karlsruhe, Deutschland.

Malte Dreyer. 2009. eSciDoc Overview. eSciDoc Days, Karlsruhe, Deutschland.

Matthias Razum. 2009. eSciDoc 2009 and beyond. eSciDoc Days, Karlsruhe, Deutschland.

Frank Schwichtenberg. 2009. eSciDoc Infrastructure: Current Status of Core Services. eSciDoc Days, Karlsruhe, Deutschland.

Malte Dreyer. 2009. eSciDoc, Infrastructure and Solutions for eResearch and eScience. Kyoto, Japan.

Malte Dreyer. 2009. eSciDoc Data Repository Infrastructure in Science Trends and Practice in Europe. Kyoto, Japan.

Malte Dreyer. 2009. eSciDoc als Virtuelle Forschungsinfrastruktur Szenarien und Anwendungen. 10. Jahrestreffen des Arbeitskreises Bibliotheken und Informationseinrichtungen der Leibniz-Gemeinschaft, Leipzig, Deutschland.

Malte Dreyer. 2008. eSciDoc als Plattform für die Wissenschaft Anwendungen und Szenarien. DINI Jahrestagung, Kassel, Deutschland.

Ute Rusnak. 2009. 2009: eSciDoc – Lösungen zum Aufbau einer digitalen Forschungsumgebung, 73. Jahrestagung der DPG (Arbeitskreis Information - AKI), Hamburg Deutschland.

Ute Rusnak. 2009. eSciDoc –Lösungen für eine digitale Forschungsumgebung. Bielefeld Conference, Bielefeld, Deutschland.

Frank Schwichtenberg, Matthias Razum. 2009. eSciDoc Infrastructure: a Fedora-based e-Research Framework. Fedora User Group Meeting, Open Repositories, Atlanta, USA.

Matthias Razum. 2009. eSciDoc: Von der digitalen Bibliothek zu e-Research. Universität Tübingen, Deutschland.