

# „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	
12. Fördernde Institution (Name, Adresse)  Bundesministerium für Bildung und Forschung (BMBF) 53170 Bonn	9. Ber. Nr. Durchführende Institution	
	10. Förderkennzeichen 01 C5957	
	11. Seitenzahl 32	
16. Zusätzliche Angaben	13. Literaturangaben 30	
	14. Tabellen	
	15. Abbildungen 9	
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 ECHO<sup>1</sup> und CDLI<sup>2</sup> des Max-Planck-Instituts für Wissenschaftsgeschichte oder LAMUS<sup>3</sup>/LEXUS<sup>4</sup> 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

---

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

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

<sup>3</sup> 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>

<sup>4</sup> 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.

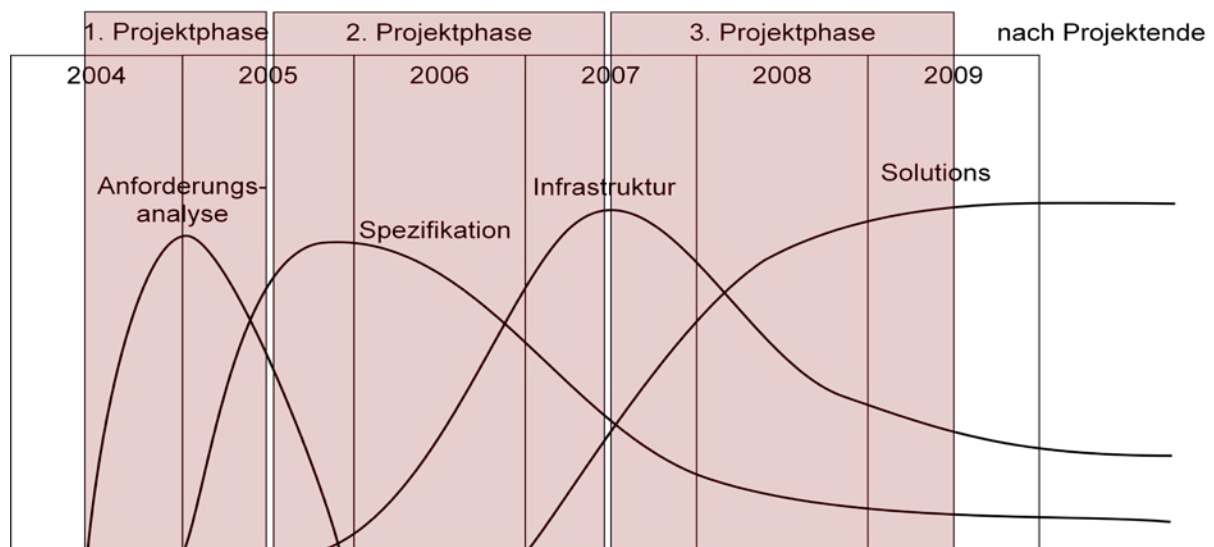


Abbildung 1: Darstellung der Projektphasen mit den Arbeitsschwerpunkten

## 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.