

# E-Science

 Wissenschaft unter neuen Rahmenbedingungen

Fachkonzept zur Weiterentwicklung der wissenschaftlichen  
Infrastruktur in Baden-Württemberg



Baden-Württemberg

MINISTERIUM FÜR WISSENSCHAFT, FORSCHUNG UND KUNST



E-Science  
Wissenschaft unter neuen  
Rahmenbedingungen



## INHALT

### Teil A: Einführung und Zusammenfassungen

|   |    |
|---|----|
| Vorwort .....                                       | 4  |
| 1. Was ist „E-Science“? .....                       | 6  |
| 2. Zusammenfassung der strategischen Konzepte ..... | 12 |
| 2.1 Lizenzierung .....                              | 12 |
| 2.2 Digitalisierung .....                           | 13 |
| 2.3 Open Access .....                               | 14 |
| 2.4 Forschungsdatenmanagement .....                 | 16 |
| 2.5 Virtuelle Forschungsumgebungen .....            | 17 |
| 3. Glossar .....                                    | 18 |
| 4. Literatur .....                                  | 22 |

### Teil B: Strategische Konzepte für die Handlungsfelder

|   |    |
|---|----|
| AG 1 Lizenzierung .....                   | 24 |
| AG 2 Digitalisierung .....                | 46 |
| AG 3 Open Access .....                    | 58 |
| AG 4 Forschungsdatenmanagement .....      | 78 |
| AG 5 Virtuelle Forschungsumgebungen ..... | 96 |

### Teil C: Mitglieder der E-Science-AG



## VORWORT

Innovation beginnt mit Wissenschaft. Hervorragende Forschung und Lehre an unseren Hochschulen sind wesentlich für die nachhaltige Entwicklung und den Wohlstand unseres Landes und für die Innovationsstärke unserer Unternehmen. Dafür gute Rahmenbedingungen zu schaffen und zu sichern, ist vornehmliche Aufgabe des Ministeriums für Wissenschaft, Forschung und Kunst.

Wissenschaftlerinnen und Wissenschaftler müssen mit den neuesten Methoden und der dafür nötigen Infrastruktur arbeiten können. Digitale Werkzeuge haben die Informationsgewinnung und die Kommunikationsformen revolutioniert und damit die Rahmenbedingungen in Wirtschaft, Alltag und Wissenschaft verändert. Austausch, Vernetzung und Produktivität sind erleichtert und beschleunigt worden. Exponentiell gestiegene Rechenleistung und die Nutzung großer Datenmengen ermöglichen einen qualitativ neuen Umgang mit Komplexität. Durch E-Science erweitern die Wissenschaften ihre Potenziale, stärken interdisziplinäre und interinstitutionelle Kooperationsformen und entwickeln völlig neue methodische Ansätze, Fragestellungen und Lösungsstrategien. Der Zugang zu Datenbanken, E-Journals und weiteren elektronischen Medien sowie die Nachhaltigkeit und Nachnutzung wissenschaftlicher Daten wird daher neben der Geräteausstattung ein immer wichtigerer Faktor für Forschungs- und Innovationsprozesse.

Die von der Gemeinsamen Wissenschaftskonferenz des Bundes und der Länder beauftragte Kommission Zukunft der Informationsinfrastruktur (KII), der Wissenschaftsrat, die Deutsche Forschungsgemeinschaft, die Hochschulrektorenkonferenz und andere Wissenschaftsorganisationen haben seit 2011 über diese Entwicklung beraten, sie strukturiert und eine Reihe von Empfehlungen ausgesprochen.

Das Ministerium für Wissenschaft, Forschung und Kunst Baden-Württemberg hat diese nationale Diskussion aufgenommen – mit dem Ziel, Hochschulen sowie außeruniversitäre Forschungs- und Informationsinfrastruktureinrichtungen zu unterstützen, eine gemeinsame E-Science-Infrastruktur aufzubauen und weiterzuentwickeln, die dem Bedarf der Wissenschaft heute und in Zukunft entspricht.

Gemeinsam sind strategische Konzepte für die konkreten Bedarfe und Entwicklungschancen in den folgenden fünf Handlungsfeldern entstanden:

1. Lizenzierung elektronischer Informationsmedien
2. Digitalisierung
3. Open Access
4. Forschungsdatenmanagement
5. Virtuelle Forschungsumgebungen

Die koordinierte Entwicklung der für diese Handlungsfelder nötigen Infrastruktur bringt in einem Flächenstaat wie Baden-Württemberg mit seiner breiten und differenzierten Hochschul- und Forschungslandschaft die größten Vorteile und bekräftigt das Selbstverständnis Baden-Württembergs als moderner Wissenschaftsstandort.

Es ist mir ein wichtiges Anliegen, für die wissenschaftlichen Einrichtungen des Landes möglichst gute Voraussetzungen zu schaffen, damit sie sich im nationalen und internationalen Wettbewerb behaupten und ihre Absolventen voll ausgestattet mit den „21st century skills“ in ein erfolgreiches Berufsleben entlassen können. In diesen Fähigkeiten ausgebildete Fachkräfte sind sowohl ein Gewinn für Forschung und Lehre als auch für innovative Unternehmen und den High-Tech-Standort Baden-Württemberg.

Dieses Grundsatzpapier zum Thema E-Science soll Wege hierfür aufzeigen. Ich danke allen an der Erstellung Beteiligten für ihr Engagement und ihre Ideen und möchte zur Diskussion hierüber einladen.



Theresia Bauer MdL

Ministerin für Wissenschaft, Forschung  
und Kunst Baden-Württemberg

## 1 · WAS IST „E-SCIENCE“?

Wissenschaftlicher Fortschritt wird daran gemessen, welche neuen Ergebnisse erzielt werden, seien dies Theorien, Erklärungsmodelle oder neuartige Verfahren, Techniken, Medikamente o.ä. Es gibt aber noch eine andere Art von Fortschritt, der die Art betrifft, wie Wissenschaft betrieben wird, d.h. mit welchen Methoden und Werkzeugen Wissenschaftlerinnen und Wissenschaftler arbeiten, um zu ihren Ergebnissen zu gelangen. Hier geht es in der Regel um inkrementelle Verbesserungen des jeweils vorhandenen Instrumentariums, es kommt aber auch vor, dass qualitativ neue Ansätze einen grundlegenden Wandel in der Art, Wissenschaft zu betreiben und zu reflektieren, auslösen. Historische Beispiele dafür sind die Ablösung der verbalen „Mathematik“ durch Leonardo da Pisa im 13. Jahrhundert, der aus dem arabischen Raum die symbolische Mathematik (Algebra) und das heute noch gebräuchliche Zahlensystem einführte, das den römischen Zahlen beim Rechnen weitaus überlegen war, z. B. beim Bau der gotischen Kathedralen. Eine ähnlich grundlegende Wirkung hatte die Entwicklung der Analysis durch Leibniz und Newton im 17. Jahrhundert, die die Basis für die moderne Physik von der Mechanik bis zur Relativitätstheorie legte. Weniger Breitenwirkung aber in den jeweiligen Wissenschaftsbereichen ähnlich nachhaltige Effekte hatten die Entwicklung des Teleskops, des Elektronenmikroskops, des Teilchenbeschleunigers oder der Genomsequenzanalyse. Alle diese neuen Techniken hatten die Folge, dass wissenschaftliches Arbeiten mit ihnen fundamental anders vorstättenging als vorher.

Heute vollzieht sich durch den Einsatz der Informationstechnik ein weiterer grundlegender Wandel in der Art, wie Wissenschaft betrieben wird. Diese Entwicklung hat in den Natur- und Ingenieurwissenschaften begonnen, erfasst aber sehr rasch auch die Lebens-, Geistes- und Sozialwissenschaften – und wirkt bis in unsere alltäglichen Gewohnheiten hinein.

Computer wurden zunächst zur Arbeitserleichterung eingesetzt, wo schon immer gerechnet werden musste, z. B. in der Physik, in den verschiedenen Ingenieur-Disziplinen<sup>1</sup> und später in der

Buchhaltung. Im Laufe der Jahre ist jedoch durch die rasante Weiterentwicklung der Hardware und der Informatik ein Prozess in Gang gekommen, der in mehreren Entwicklungssprüngen qualitativ neue Methoden ermöglicht hat.

- Während anfangs nur stark idealisierende Gleichungen gelöst wurden, können jetzt komplexe Zusammenhänge und Abläufe detailgenau modelliert und auf leistungsfähigen Rechnern („Computern“) simuliert werden. Der Höchstleistungsrechner an der Universität Stuttgart gehört zu den leistungsfähigsten der Welt.
- Seit Jahrzehnten hat sich die bei unverändertem Preis verfügbare Speicherkapazität innerhalb von zwei Jahren verdoppelt. Dadurch können extrem umfangreiche Datensammlungen digital verfügbar gemacht werden ohne die Notwendigkeit, alte (aber nicht veraltete) Daten in periodischen Abständen zu löschen. Man denke etwa an weiträumig ausgebrachte Sensornetze zur Überwachung von Umweltbedingungen, an Klimadatenzentren und Satellitendaten, an die täglich in Printmedien publizierte Sprache.
- Außerdem sind Datennetze eingerichtet worden, die Kommunikationsdienste und die Übertragung relativ großer Datenmengen zu einem geringen Preis verfügbar machen. Das Internet ist zu einer bedeutenden Informationsquelle geworden. Der Zugriff kann von unterwegs mit einem handlichen Gerät erfolgen, in dem die Leistung früherer Großrechner konzentriert ist. In wissenschaftlichen Netzen wie „BelWue“ in Baden-Württemberg können Rechen-, Speicher- und Informationsdienste zentral angeboten und landesweit genutzt werden.
- Die Methodenentwicklung in der Mathematik und der Informatik legt die Grundlagen für

die Entwicklung hochkomplexer Modelle von Wirklichkeitsausschnitten auf der Basis großer Mengen an Messdaten. Dazu gehören Data-Driven Science, Artificial Intelligence, maschinelles Lernen, Sprachanalyse und vieles mehr. Heute erhobene und archivierte Daten können in Zukunft in veränderten Kontexten unter neuen Fragestellungen und mit heute noch nicht bekannten Methoden analysiert werden.

In den letzten Jahrzehnten wurden Forschung und Entwicklung mit großem wirtschaftlichem Erfolg von der Untersuchung und der Erprobung physikalischer Modelle in die digitale Welt verlagert. Mit Hilfe von Simulationen mathematischer Modelle ist die Flugzeugentwicklung so weit, dass bereits der erste Prototyp zuverlässig und sparsam fliegt. Automobilentwickler untersuchen Luftwiderstände, Fahr- und Crashverhalten von Fahrzeugen in allen Varianten, die nur auf dem Bildschirm existieren. Effiziente Autoproduktion wird vollständig digital simuliert, um Arbeitsschritte zu optimieren und die globalisierte Lieferkette bis ans Band zu steuern. Über die inverse Modellierung indirekter Messungen wie der Tomographie werden das Innere von Organen visualisiert oder neue Ölfelder exploriert. Zunehmend komplexere Systeme können auf diese Weise untersucht werden.

So erlauben Klimasimulationen die Bewertung des Einflusses atmosphärischer Rahmenbedingungen wie CO<sub>2</sub>- oder Partikel-Konzentrationen auf regionale Klimaregime. Diese Entwicklungen haben durch ihren rasanten Verlauf den Punkt überschritten, wo Quantität in eine neue Qualität umschlägt. Konnte man früher mit Computern lediglich herkömmliche Berechnungen schneller

lösen, sind jetzt gänzlich andere, komplexere und im wissenschaftlichen Sinne neuartige Berechnungen möglich, die neue Herangehensweisen eröffnen. Ein prominentes Beispiel ist das „Human Brain Project“, eines der zwei von der EU geförderten „Future Emerging Technologies Flagships“. Das Forscherteam an der Universität Heidelberg bildet neurobiologische Strukturen des Nervensystems digital nach. Es entwickelt neuromorphe Rechensysteme, die die Lernfähigkeit und Fehlertoleranz des menschlichen Gehirns adaptieren. Ein anderes Beispiel ist die Bioinformatik, die mit bildgebenden Verfahren die Skalensprünge von den chemischen Prozessen innerhalb einer Zelle über Wachstums- und Zellteilungsprozesse bis hin zur Genese von Organismen und die Funktionsweise von Organen untersucht, um die Aufgabe von Genom-Abschnitten zu identifizieren. Dazu werden große Mengen an Bilddaten in 3D-Modelle umgerechnet und Veränderungen automatisch erkannt.<sup>2</sup> Für die personalisierte Krebstherapie am DKFZ in Heidelberg werden die Genome der Patienten vollständig sequenziert, um die genetischen Ursachen der Erkrankung zu erkennen und Therapien mit passenden Wirkungsmechanismen anzuwenden. Auch in der Pharmazieforschung werden Methoden entwickelt, so dass neue Medikamente nicht mehr an Labortieren oder gar am Menschen erprobt werden müssen, sondern an digitalen Modellen auf ihre Wirksamkeit getestet werden können.

In allen Disziplinen werden sämtliche Daten, Fakten, Hypothesen, Veröffentlichungen usw. in digitaler Form verfügbar und mit anderen Daten – im Idealfall aus anderen Bereichen – verknüpfbar sein. Neue Theorien und Arbeitshypothesen werden dann in der Mehrzahl der Fälle nicht mehr im Labor am realen Objekt überprüft, sondern an digitalen Modellen. Durch die Verknüpfung von Daten, Methoden und Wissen über Disziplinengrenzen hinweg werden neue Zusammenhänge der komplexen Wirklichkeit sichtbar. Die zeitliche und räumliche Entwicklung der Sprache und aktueller Themen lässt sich nachzeichnen, indem Printmedien, Streams von Hörfunksendern und Social-Media-Feeds automatisch auf Muster durchsucht werden. Und neue Theorien können durch die Suche nach statistischen Auffäl-

lichkeiten in Daten unterschiedlicher Wissenschaftsdomänen entstehen, z. B. durch die Verknüpfung des archäologischen Befunds zahlreicher Menschenopfer in der letzten Phase der Maya-Hochkultur mit neuen Klimamodellen, die, weil sie das El-Niño-Phänomen beherrschen, in dieser Zeit eine anhaltende Dürreperiode erklären.

Bei der Entwicklung effektiver Arbeitsumgebungen schafft sich die Wissenschaft eine integrierte digitale Plattform, in der alle Daten, alle Literatur, alle Berechnungsmethoden und weitere Ressourcen verfügbar sind. Sie können von den Wissenschaftlerinnen und Wissenschaftlern mit größtmöglicher Flexibilität genutzt werden. Solche Plattformen sind nicht als Monolithen zu verstehen. Sie können weder isoliert noch universell aufgebaut werden, sondern werden an verschiedenen Orten und in verschiedenen Wissenschaftsbereichen wachsen. Wissenschaftler arbeiten auf diesen Plattformen, sie verbinden Plattformen und entwickeln diese weiter. Dafür müssen sie interoperabel angelegt sein und selbst aus Modulen mit definierten Schnittstellen bestehen. Darum geht es bei „E-Science“.

Abbildung 1 zeigt einen vereinfachten Blick auf diese Welt.

In einigen Teilgebieten der Wissenschaft sind Umgebungen dieser Art bereits in Ansätzen vorhanden (z. B. in den Geowissenschaften, der Teilchenphysik, der Astronomie), doch ihrer Erweiterung über die engen Grenzen des Entstehungsgebietes hinaus und der Nutzung in interdisziplinären Kontexten stehen große Schwierigkeiten entgegen: „The harder problem for the future is heterogeneity, of platforms, data and applications, rather than simply the scale of the deployed resources. The goal should be to allow scientists to ‘look at’ the data easily, wherever it may be, with sufficient processing power for any desired algorithm to process it. Current platforms require the scientists to overcome computing barriers between them and the data.”<sup>3</sup> An dieser Diagnose aus dem Jahr 2005,

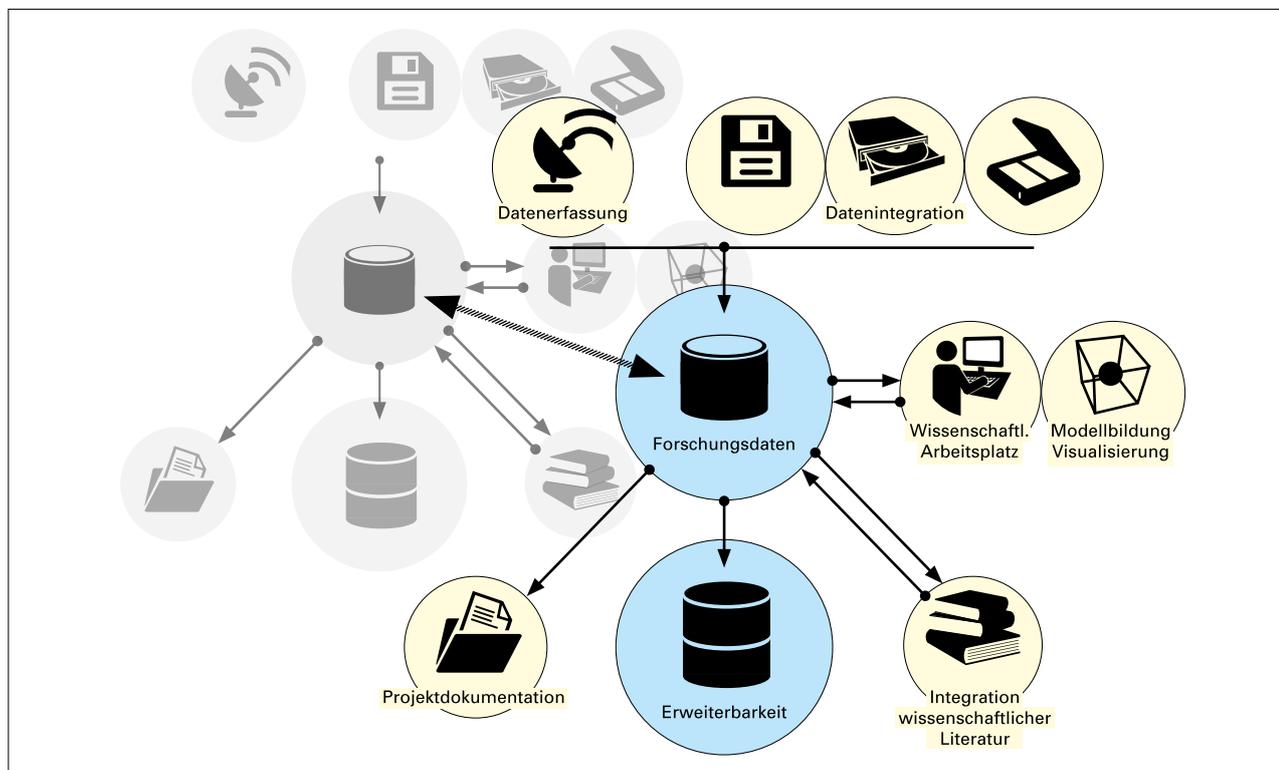


Abb. 1: Schematische Darstellung einer E-Science-Umgebung

die eine mangelnde Interoperabilität beschreibt, hat sich bis heute nichts Wesentliches geändert.

Die Forschung an den Hochschulen in Baden-Württemberg wird nur dann weiterhin leistungsstark und international wettbewerbsfähig bleiben, wenn die Hochschulen auf diese Veränderungen vorbereitet sind und über die notwendige Infrastruktur verfügen, um den Wissenschaftlern die benötigten Ressourcen bedarfsorientiert bereitzustellen. Das MWK konzentriert sich zunächst auf folgende fünf Handlungsfelder:

Wissenschaftliche Publikationen, Datenbanken und Software stehen als grundlegende Ressourcen der Forschung online zur Verfügung, weil nur so der Zugriff flexibel, die Recherche effizient und die Verknüpfungen übergreifend realisiert werden können. Wenn digitale Medien nur lizenzpflichtig zur Verfügung gestellt werden können, sollen die Beschaffungsmittel zum größten Nutzen möglichst aller Wissenschaftlerinnen und Wissenschaftler eingesetzt werden, indem die **Lizenzen** konsortial für alle Hochschulen im Land verhandelt und beschafft und auf dieser Basis Portale bereitgestellt werden.

Analoge Materialien, die die heutige Wissenschaft benötigt, müssen in die digitale Welt transformiert werden. Schriftgut, Bilder, Sammlungsobjekte, Filme werden **digitalisiert** und in die Portale integriert, damit sie ohne Medienbruch den Wissenschaftlerinnen und Wissenschaftlern im Land und weltweit zur Verfügung stehen.

Wissenschaftliche Publikationen und Daten, die Ergebnisse öffentlich geförderter Forschung sind, sollen frei zugänglich sein. Durch **Open Access**

ändern sich Publikationsformen und die Geschäftsmodelle der Verlage. Die Sichtbarkeit der Forschung in Baden-Württemberg wird erhöht, die Verbreitung und weitere Nutzung des Wissens werden erleichtert.

Strategisches Management der Daten, die einer wissenschaftlichen Arbeit zugrunde liegen, sichert die Nachvollziehbarkeit und Nachnutzbarkeit der Forschungsergebnisse. Das **Forschungsdatenmanagement** erstreckt sich über den gesamten Life Cycle der Datengewinnung, Analyse, Publikation und (Langzeit-)Archivierung. Dabei muss der Schutz des geistigen Eigentums des Urhebers gewährleistet sein.

**Virtuelle Forschungsumgebungen** unterstützen den Forschungsprozess durch die Bereitstellung wesentlicher Methoden. Daten werden nahtlos von Laborgeräten in Datenbanken übernommen und stehen über Kollaborationsplattformen, die Schnittstellen zu beliebigen Analysewerkzeugen anbieten, der gesamten – auch über Institutionen und Länder verteilten – Forschungsgruppe zur Verfügung.

Eine solche E-Science-Umgebung verlangt eine neue Infrastruktur. Sie muss aus Sicht der Wissenschaftlerinnen und Wissenschaftler so einfach und zuverlässig verfügbar sein wie die Strom- oder Datensteckdose in der Wand. Unter der wissenschaftlichen Anwendungsebene liegen viele organisatorische Ebenen – von technischen Voraussetzungen und neuen Aufgaben der Informationsinfrastruktureinrichtungen im Sinne von E-Science-Service über die Etablierung geeigneter Geschäftsmodelle bis hin zur Integration der E-Science-Methoden in die Lehre – mit jeweils eigenen **Herausforderungen**, die gelöst werden müssen, wenn die Vision einer durchgängigen E-Science-Infrastruktur bis 2020 Realität werden soll:

**Datenintensives Computing:** Durch immer leistungsfähigere Geräte werden riesige Datenmengen erzeugt. Der Umgang mit Datenmengen im mehrstelligen Petabyte-Bereich – und diese Größenordnung wird weiter steigen – erfordert neue Arbeits-

weisen und technische Konzepte, weil die Datenübertragungsraten der Netze nicht mithalten können. Es ist zu erwarten, dass in Zukunft Datenzentren neben dem Speicher auch Computing-Ressourcen zur Vorfilterung der Daten anbieten. Für das Land werden Betriebsmittelhierarchien analog zur aktuellen Leistungspyramide des High Performance Computings (HPC) entstehen.

**Datenmanagement:** Die abgelegten Datenmengen müssen vorgehalten und so aufbereitet werden, dass sie wiedergefunden und erneut genutzt werden können. Nicht nur in der Raumfahrt ist die Nachnutzung von Daten günstiger als die Neugewinnung. Daten, die einen nicht-wiederholbaren Zustand der Vergangenheit beschreiben, wären ohnehin unwiderruflich verloren. Durch Nachnutzung entstehen Verknüpfungen mit anderen Studien, Kontexten, Wissenschaftsbereichen. Auch ganz praktisch erspart das Wissen über die Existenz einer digital zugänglichen handschriftlichen Urkunde nicht nur das erneute Digitalisieren oder die Reise in die ggf. entfernt gelegene Bibliothek, sondern auch Verzögerungen bei der eigenen Arbeit.

**Workflows für Daten und Publikationen:** Daten entstehen nicht notwendig an dem Ort, an dem sie interpretiert oder weiterverarbeitet werden. Dies gilt für Laborgeräte ebenso wie für die Entstehung einer Monografie in den Geisteswissenschaften. Dies stellt neue Anforderungen an das systematische institutionenübergreifende Qualitäts- und Datenmanagement. Die Methodik der E-Science wird sich auf die Publikationsformen auswirken. Zu den Publikationen in Textform gehören Verweise auf die (unveränderten) Basisdaten und ggf. die eingesetzte Software. Sie müssen zumindest für die Begutachter zugänglich sein, können aber auch selbst – unter Beachtung der geistigen Eigentumsrechte – veröffentlicht werden.

**Archivierung:** Mit dem Datenmanagement kommen neue Fragen zur Archivierung auf, die gelöst werden müssen: Zugriffsrechte müssen bei der Datenarchivierung zuverlässig erkannt und über die Laufzeit der Daten durchgesetzt werden: medizi-

nische und sozialwissenschaftliche Daten bedürfen wegen ihres möglichen Personenbezugs eines besonderen Schutzes, der aber die Nachnutzung der Daten – auch aufgrund neuer wissenschaftlicher Fragestellungen – nicht unnötig erschweren darf. Die Visualisierung oder weitere Nutzung der Daten setzt oft eine bestimmte Software voraus, die möglicherweise selbst archiviert werden muss, weil ihre längerfristige Nutzbarkeit nicht gewährleistet ist.

**Risikomanagement:** Die mit der wachsenden Größe von Speichereinheiten einhergehende Zentralisierung – auch über nationale Grenzen hinweg – macht schon das reine Ausüben von Zugriffsrechten unkalkulierbar: Die Rechte können aufgrund politischer und rechtlicher Vorgaben eingeschränkt werden, die Lizenzgebühren für den Zugriff können sich unkontrollierbar entwickeln, und die Existenz eines Data Providers (z. B. eines Datenzentrums, eines Wissenschaftsverlags) könnte beendet werden. Wissenschaft, die sich auf ein spezielles Szenario verlassen hat, wäre massiv beeinträchtigt. Daher müssen Analysen über die Gefährdungssituation permanent fortgeschrieben werden und Maßnahmen zur Reduzierung der Gefährdung – wie der Aufbau eines eigenen Datenzentrums, das Kopien vorhält – ergriffen werden.

**Finanzierung:** Eine dauerhafte E-Science-Infrastruktur setzt geeignete Geschäftsmodelle voraus. Es muss geklärt werden, wie die Nutzer an den Kosten für Betrieb, Entwicklung, Re-Investition und Beratung beteiligt werden und welche Forderungen und Ansprüche sie gegenüber den Anbietern haben. Gleichzeitig muss sichergestellt sein, dass grundsätzlich alle Wissenschaftlerinnen

und Wissenschaftler an den Hochschulen in Baden-Württemberg auf die Infrastruktur zugreifen können.

**Lehre:** Der wissenschaftliche Nachwuchs in Lehre und Forschung verfügt größtenteils durchaus über die zum Leben in einer E-Science-Welt notwendigen 21st century skills. Die gewinnbringende Nutzung konkreter E-Science-Techniken und -Methoden bedarf aber auch der gezielten Integration in die Lehre und zwingt teilweise zum Umbau existierender Forschungsstrukturen, um Anachronismen zu vermeiden. Beispielsweise erlaubt ein elektronisches Laborbuch als Bestandteil eines konsequent IT-gestützten Labor-Workflows die schnellere Einarbeitung von Doktoranden, da sie auf vorhandenes Wissen aufsetzen und bestehende Arbeitsabläufe nutzen können, die das Forschen auf einem hohen Niveau der Wissenschaftlichkeit unterstützen. Die Ausbildung in den klassischen Disziplinen wird angereichert durch die Kompetenzen des interdisziplinär vernetzt denkenden E-Scientists. Daneben entsteht ein Bedarf für Spezialisten der E-Science, die Big Data Management u.a. für die Wissenschaft beherrschen. Da E-Science auch die außeruniversitäre Forschung erreicht hat, sind E-Science-Kompetenzen auch in Unternehmen gefragt. Sie haben bereits jetzt einen Bedarf entsprechender wissenschaftlicher Weiterbildungsangebote.

Die **Entwicklung vernetzter Informationsstrukturen** wurde durch die Forschung entscheidend geprägt. Heute alltäglich genutzte Techniken wie E-Mail, Hypertext (http), Internet und Suchmaschinen entstanden an Forschungsinstituten, um die Kommunikation unter den Wissenschaftlern und die gleichmäßige Auslastung teurer

Computing-Ressourcen zu erleichtern. Später wurden daraus Plattformen auch für die betriebliche Kooperation und Präsentation, für Marketing und E-Commerce. An dieser Stelle geht es nicht darum, die sozialen und kulturellen Folgen zu bewerten, sondern ausgehend von der Pionierarbeit in Wissenschaft und Bildung an den Hochschulen die Medien- und Gestaltungskompetenz der Bürgerinnen und Bürger weiterzuentwickeln, da neue Mediennutzungen alle Lebensbereiche und Arbeitsweisen erfassen.

Die Verfügbarmachung digitaler Information ist eine **infrastrukturelle Grundversorgungsaufgabe** geworden. Digitale Medienformate ergänzen papiergebundene Wissens-, Kultur- und Unterhaltungsmedien. Dadurch werden wissenschaftlich hochwertige Informationen (Artikel, Bilder, Daten usw.) als Open Access oder über Bezahlmodelle überall und für jeden Bürger zugänglich, sei es zur berufsorientierten Weiterbildung oder aus persönlichem Interesse. Digitale Medien bieten darüber hinaus substanzielle dynamische Mehrwerte wie animierte Darstellungen, Recherchemöglichkeiten, schriftliche oder audiovisuelle Interaktion (Web 2.0 bzw. Web-Conferencing, „skypen“). Derzeit entstehen durch die Einbindung komplexer Simulationen und Visualisierungen neue Anwendungen: Medizinische Diagnosen und Therapien können fernbegutachtet oder die Auswirkungen geplanter Speicherseen oder Stromtrassen auf das Landschaftsbild dargestellt werden. Auf dieser Basis können Handlungsalternativen anschaulicher diskutiert und abgewogen werden. Hier ergeben sich Anknüpfungspunkte für offene Prozesse der Bürgerbeteiligung.

## 2 · ZUSAMMENFASSUNG DER STRATEGISCHEN KONZEPTE

Die strategischen Konzepte zu den fünf Handlungsfeldern finden sich in Teil B. Im Folgenden werden die Handlungsfelder und die

spezifischen Empfehlungen für Baden-Württemberg zusammenfassend dargestellt.

### 2.1 LIZENZIERUNG

In Baden-Württemberg wird 2020 eine Informationsinfrastruktur für Forschung und Lehre und über weitere Wissenschaftseinrichtungen wie die Landesbibliotheken auch für weitere Bevölkerungsschichten zur Verfügung stehen, die bundesweit als vorbildlich anerkannt ist. Sie wird sich durch ebenso hervorragende wie einfach zu bedienende Zugangs- und Recherchemöglichkeiten auszeichnen, die kommerziellen Suchmaschinen in nichts nachstehen, sowie durch eine ausgezeichnete Versorgung mit Volltexten. Zum Studieren, Lehren und Forschen in Baden-Württemberg werden die benötigte Literatur bzw. andere Informationsmedien unmittelbar zur Verfügung stehen, sei es Open Access, mit vorhandener Lizenz oder als elektronischer Sofortlieferdienst. Software für die wissenschaftliche Arbeit kann standortübergreifend genutzt werden („private cloud“), der technisch komfortable und sichere Zugang zu elektronischen Informationsquellen ist für alle Mitglieder der Hochschulen und an den Landesbibliotheken Baden-Württembergs gewährleistet. Um dieses Ziel einer umfassenden Grundversorgung in Baden-Württemberg zu erreichen, bedarf es sowohl von Seiten der Landesregierung als auch der Hochschulen und Landesbibliotheken erheblicher Anstrengungen.

Die Informationsversorgung mit elektronischen Fachinformationsangeboten für Forschung und Lehre wird auf drei Ebenen organisiert: Über lokal beschaffte und vorgehaltene Lizenzen, konsortial erworbene Lizenzen und durch die DFG geförderte National- bzw. Allianz-Lizenzen. Ziel der konsortialen Erwerbung ist es, über eine gesteigerte Marktmacht günstigere Konditionen in Form von Kostenersparnissen über Rabatte oder in Form einer größeren Anzahl zur Verfügung stehender Inhalte zu erzielen sowie den Arbeits- und Verwaltungsaufwand zu reduzieren und das Know-how im Umgang mit elektronischen Informationsressourcen zu bündeln.

Seit 1999 organisiert und koordiniert das Konsortium Baden-Württemberg, ein Zusammenschluss von derzeit 51 wissenschaftlichen Bibliotheken im Geschäftsbereich des MWK, den gemeinschaftlichen Erwerb von Lizenzen für elektronische Informationsmedien im Land. Das Konsortium wird gegenüber den Verlagen und Anbietern von der Universitätsbibliothek Freiburg vertreten. Sie übernimmt die Geschäftsführung, koordiniert die Arbeit des Konsortiums und bewirtschaftet die zur Verfügung gestellten Mittel. Die Ressourcenausstattung ist allerdings auf Grund des sich gravierend verändernden Marktes trotz der bisherigen Förderung durch das MWK und erheblicher Eigenleistungen der Hochschulen und Landesbibliotheken unzureichend.

Auf Basis der turnusmäßig vom Konsortium durchgeführten Bedarfsanalyse muss das Angebotsportfolio künftig deutlich ausgeweitet und um neue Gattungen (z. B. E-Book-Pakete, Open Access Komponenten, Literaturverwaltungstools, Resource Discovery Systems (Indexdienste)) ergänzt werden. Hierfür bedarf es eines deutlichen und nachhaltigen Mittelaufwuchses. Die Finanzierung erfolgt dabei hälftig aus zentralen Mitteln des MWK und lokalen Mitteln der Einrichtungen. Durch die zentralen Mittel wird eine Verbesserung der Informationsversorgung in der Breite gewährleistet. Durch die Beteiligung der Einrichtungen ist die unumgängliche Rückkoppelung der Angebote mit den jeweiligen lokalen Bedürfnissen sichergestellt.

Das Konsortium Baden-Württemberg soll alle benötigten und auf konsortialer Ebene verfügbaren Angebote koordinieren und vertraglich absichern. Hierzu gehört u.a. die logistische und juristische Unterstützung der Verhandler, z. B. bei Fragen der Vertragsausgestaltung oder hinsichtlich der (Nach-)Nutzungsmöglichkeiten im Kontext von Virtuellen Forschungsumgebungen. Zudem sollte es künftig auch verstärkt als Koordinierungs- bzw. Beratungsinstanz für Open Access-Geschäftsmodelle fungieren.

In der laufenden Diskussion um die Weiterentwicklung der Informationsinfrastruktur in Deutschland haben sowohl der Wissenschaftsrat als auch die Gemeinsame Wissenschaftskonferenz des Bundes und der Länder die Ansätze der DFG bzw. der Allianz der deutschen Wissenschaftsorganisationen für eine nationale Lizenzierungsstrategie aufgegriffen und weiterentwickelt. In infrastrukturpolitischer Hinsicht gehen von den genannten Organisationen starke Impulse in Richtung einer stärkeren Koordi-

## 2.2 DIGITALISIERUNG

Digitalisate als Teil einer E-Science-Umgebung ermöglichen es Wissenschaft und Forschung, zeit- und ortsunabhängig mit dem kulturellen historischen Erbe zu arbeiten und dank der internationalen Vernetzung dieses Erbe immer wieder neu zu kontextualisieren. Sie sind ein wesentlicher Beitrag zur Beschleunigung der Forschung, zur Setzung neuer Forschungsimpulse und zur Erschließung des Materials. Bislang unbekanntes oder nur schwer zugängliches Material kann der Forschung sichtbar und zugänglich gemacht werden.

Angesichts der Fülle forschungsrelevanter, nachgefragter Dokumente und Objekte in den baden-württembergischen Archiven, Bibliotheken, Museen, universitären Sammlungen und Filmeinrichtungen im

nierung des Konsortialgeschehens aus. Dabei können und sollen sich auch regionale Konsortialstrukturen zukünftig verstärkt einbringen.

Vor diesem Hintergrund soll sich das Konsortium Baden-Württemberg zu einer noch leistungsfähigeren Organisation entwickeln, die Interessen des Wissenschaftsstandorts Baden-Württemberg auf nationaler Ebene durch ein verstärktes Engagement in den sich neu herausbildenden kooperativen Erwerbungsstrukturen wahren und einen der Bedeutung des Landes angemessenen Beitrag leisten.

Um die Ausweitung des Konsortialportfolios professionell unterstützen und sich im nationalen Versorgungskontext strategisch positionieren zu können, sollte für das Konsortium Baden-Württemberg eine zentrale Geschäftsstelle mit ausreichender personeller Ausstattung eingerichtet werden.

Die konsortiale Beschaffung von Softwarelizenzen soll ebenfalls ausgebaut und organisatorisch gegliedert werden. Eine engere Verzahnung mit dem Konsortium für E-Medien wird dabei angestrebt, um mögliche Synergieeffekte im Beschaffungsprozess zu erreichen und zu nutzen.

Geschäftsbereich des MWK sowie aus urheberrechtlichen Gründen soll sich die Digitalisierungsstrategie des Landes vorerst auf das kulturelle historische Erbe im urheberrechtsfreien Raum beschränken.

Baden-Württemberg verfügt im Bereich der Digitalisierung bereits über beträchtliches Know-how und kann bei der Entwicklung eines Digitalisierungsprogramms auf ein großes Wissenspotenzial zurückgreifen. Langjährige praktische Erfahrungen haben insbesondere das Landesarchiv Baden-Württemberg und die Universitätsbibliothek Heidelberg.

Aus der Ist-Analyse des Status quo bei der Retrodigitalisierung und aus den veränderten Informationsanforderungen der Forschung resultiert das Konzept einer spartenübergreifenden Digitalisierungsstrategie für die genannten Infrastruktureinrichtungen:

Durch Digitalisierungskompetenzzentren kann das Land von den bislang kleinteiligen Digitalisierungsprojekten zu einer großflächigen Struktur mit koordinierten Digitalisierungsvorhaben kommen. Die Digitalisierungskompetenzzentren sollen spartenweise, in Teilen auch spartenübergreifend, Scan- und Beratungsdienstleistungen anbieten sowie neue forschungsrelevante Digitalisierungsprojekte sowohl innerhalb des Landes als auch mit überregionalen Aktivitäten abstimmen.

Zur Digitalisierungsstrategie gehören auch einheitliche Kriterien für die Mittelvergabe, Aussagen zur Nutzung und Verwertung der Digitalisate sowie die regelmäßige Evaluierung der Digitalisierungsstrategie.

Da die im Land vorhandenen zentralen Digitalisierungsmittel weder ausreichen, um den vertraglichen Verpflichtungen gegenüber der Deutschen Digitalen Bibliothek nachzukommen, noch, um gezielt und nachfrageorientiert die forschungsrelevanten Objekte der Forschungsinfrastruktureinrichtungen zu digitalisieren, sollen zusätzliche zentrale Digitalisierungsmittel zur Verfügung gestellt werden. Den derzeit im Netz verfügbaren rund 8 Mio. Images steht ein derzeit bekannter Bedarf von rund 206 Mio. Images gegenüber.

Die Relevanz des digitalisierten kulturellen Erbes für die Forschung lässt sich an den steigenden Zugriffszahlen ablesen: Das Landesarchiv Baden-Württemberg verzeichnete zuletzt jährliche

Zugriffszahlen auf Images in Höhe von 3,75 Mio., die Universitäts- und Landesbibliotheken in Höhe von 7,6 Mio.

Digitalisierung im größeren Stil ist erst möglich, wenn die richtigen finanziellen, organisatorischen und technischen Rahmenbedingungen geschaffen werden. Der größte Teil der Kosten entsteht nicht bei der eigentlichen Image-Digitalisierung, sondern bei der Aufbereitung der Digitalisate für die Präsentation und bei der Qualitätssicherung.

Erst die notwendige Erschließung und Strukturierung, die Aufbereitung und Verwaltung sowie die Langzeitarchivierung der Digitalisate machen diese zu einem zentralen Bestandteil der Informationsinfrastruktur.

Der finanzielle Bedarf für ein mehrjähriges Förderprogramm wurde spartenweise errechnet. Sie umfassen die Investition in die notwendige zusätzliche technische Ausstattung, Personal für die Beratungs- und Scandienste sowie volumenabhängige Aufwendungen für die Produktion des E-Contents und den Speicherplatz.

### 2.3 OPEN ACCESS

Informationstechnik und Internet ermöglichen es, Forschungsergebnisse (Forschungsdaten und wissenschaftliche Publikationen) weltweit zugänglich und nachnutzbar zu machen. Die traditionellen Geschäftsmodelle stehen der Nutzung dieser neuen Möglichkeiten zum Nachteil von Wissenschaft, Wirtschaft und Gesellschaft zum Teil entgegen. Wissenschaftspolitisch erscheint es geboten, das wissenschaftliche Publikationsregime im Wege des Open Access so zu gestalten, dass

- die für Bildung und Wissenschaft aufgewendeten Steuermittel größtmöglichen Nutzen für Wissenschaft, Wirtschaft und Ge-

sellschaft erzielen,

- die Sichtbarkeit der wissenschaftlichen Forschung in Baden-Württemberg weltweit erhöht und damit
- die Wettbewerbsfähigkeit des Landes nachhaltig gesichert und gesteigert wird.

Baden-Württemberg ist damit das erste deutsche Bundesland, das solche Zielvorgaben entwickelt, und übernimmt damit eine bundesweite Vorreiterrolle.

Die Open Access Politik des Landes Baden-Württemberg ist darauf angelegt, im Land die Voraussetzungen zu schaffen, den notwendigen internationalen Umgestaltungsprozess als herausgehobener Akteur zu ermöglichen, zu fördern und aktiv mitzugestalten, um so auch die Wahrnehmung der Forschungsergebnisse aus Baden-Württemberg international zu erhöhen. Die unter den Stichworten „grüner/goldener Weg“ definierten Ansätze für die Implementierung von Open Access an Hochschulen und Forschungseinrichtungen werden im Sinne weitgehender Komplementarität verstanden. Welche Wege die Hochschulen präferieren, ist auch eine Frage der hochschulautonomen Profilbildung und Schwerpunktsetzung.

Mit dem BW-Open Access-Förderprogramm soll ein markanter Beitrag zur nachhaltigen Sicherung

der Informationsinfrastrukturen an den Hochschulen des Landes und der entsprechenden Umgestaltung des internationalen wissenschaftlichen Publikationswesens geleistet werden. Im Fokus steht das Ziel, dass Wissen und Innovation, die in gemeinnützig organisierten Wissenschaftsorganisationen erarbeitet werden, auch „barrierefrei“ zur Nachnutzung durch Wissenschaft, Wirtschaft und Gesellschaft zur Verfügung stehen.

Für die beschleunigte und nachhaltige Implementierung des Open Access-Paradigmas an den Hochschulen in Baden-Württemberg kommt es schließlich auf eine intelligente Verknüpfung „weicher“ und „harter“ Förderfaktoren an. In der Summe wird ein kooperativ mit den verschiedenen Akteuren ausgerichtetes Anreizsystem vorgeschlagen, das zielführend und in mittelfristiger Perspektive den Anteil Open Access-publizierter Forschungsergebnisse in allen Wissenschaftsdisziplinen signifikant zu erhöhen verspricht.

## 2.4 FORSCHUNGSDATENMANAGEMENT

Forschungsdaten sind digitale Daten, die durch wissenschaftliche Quellenforschungen, Experimente, Messungen, Simulationen, Erhebungen oder Befragungen entstehen. Sie umfassen strukturierte Massendaten sowie unstrukturierte Daten unterschiedlichsten Formats und Inhalts, so auch Prozessdaten wie Bearbeitungsschritte, in einem Forschungsvorhaben entwickelte Algorithmen und Programme oder digitalisierte Schriften, Bilder oder Filme.

Häufig sind Forschungsdaten einmalig und unwiederbringlich in dem Sinne, dass sie zeitbedingte Zustände der Umwelt, der Gesellschaft oder eines Individuums repräsentieren, die im Nachhinein nicht mehr neu erhoben werden können, oder dass ihre erneute Gewinnung unangemessen aufwendig ist. Wissenschaft, die Phänomene beschreibt, ihre Bedingungen analysiert und kausale Zusammenhänge erforscht, ist auf solche Daten angewiesen, nicht nur im Rahmen der Fragestellungen, für die sie erhoben wurden, sondern auch im Sinne der Nachnutzung: Analyse mit neuen Methoden, Deutung in neuen, auch transdisziplinären Kontexten.

Forschungsdaten unterliegen einem Lebenszyklus, der viele Schritte des Anreicherns und Aussortierens enthält: von Rohdaten über geprüfte Mess- und Erhebungsergebnisse, verschiedene Stufen der Aufbereitung und Analyse bis hin zu Datensätzen in archivierungsfähiger und nachnutzbarer Form. Die Herausforderung einer Forschungsdaten-Strategie besteht in der ausreichenden Berücksichtigung disziplinspezifischer Anforderungen einerseits und dem Setzen allgemeingültiger Regeln andererseits, gerade auch im Hinblick auf die gewünschte und zunehmende Interdisziplinarität zwischen Natur-, Geistes- und Lebenswissenschaften. Neben bibliographischen und technischen müssen auch disziplinspezifische Attribute von zuverlässig hoher Qualität für die Beschreibung von Inhalt, Kontext und Provenienz der Daten abgelegt werden. Daher muss der Übergang von Forschungsdaten aus der persönlichen in die institutionelle oder die öffentliche Domäne durch einen Workflow unterstützt werden. Der Zugriff der Wissenschaftler auf die Forschungsdaten muss technisch und organisatorisch so konzipiert sein, dass die Nutzenden über benutzerfreundliche Portale eine Meta-Suche in mehreren Fach- und

Daten-Repositoryn durchführen können, die entsprechend standardisiert und vernetzt sein müssen.

Eine neue Publikationsform entsteht durch die Verknüpfung von herkömmlichen textorientierten Publikationen mit den für ihre Erstellung verwendeten primären Forschungsdaten. Die Publikation von Forschungsdaten bedeutet einen Kulturwandel in der Wissenschaft, der sich in den Geowissenschaften und der Astronomie bereits abzeichnet. Dies gilt insbesondere für die Publikation von Text und Daten als Open Access.

Das MWK beabsichtigt, gemeinsam mit den Universitäten, weiteren Forschungseinrichtungen und Forschungsverbänden eine Forschungsdaten-Strategie zu entwickeln, so dass alle Wissenschaftlerinnen und Wissenschaftler im Geschäftsbereich des MWK Zugriff auf die notwendigen Prozesse und die darunter liegende technische Infrastruktur haben. Dadurch soll die strategische Wettbewerbsfähigkeit der Hochschulen und Forschungseinrichtungen bei der Einwerbung von Forschungsmitteln nachhaltig weiterentwickelt und der Wissenschaftsstandort Baden-Württemberg gestärkt werden.

Als Voraussetzung für die Einführung eines Forschungsdatenmanagements muss die bestehende technische Infrastruktur bzgl. Datenspeicher, Datennetze und HPC-Ressourcen weiterhin nach dem Bedarf der Wissenschaft ausgebaut werden. Zu dieser Infrastruktur gehören auch die Authentifizierung und die Berechtigungssteuerung auf Basis föderativer Strukturen der Universitäten. Die Hochschulen, insbesondere die Universitäten, haben die Aufgabe, Kompetenzen des professionellen und nachhaltigen Umgangs mit Forschungsdaten in der Lehre zu ver-

mitteln. Dabei sollten sie sowohl eine breite Zielgruppe in allen Disziplinen ansprechen als auch spezielle Kurse in der Weiterbildung und im Bereich Data Scientist / Data Librarian anbieten.

Zum Aufbau der notwendigen Infrastruktur und ihrer Governance-Strukturen wird ein Förderprogramm vorgeschlagen, um ein landesweites Forschungsdaten-Repository aufzubauen. Hierfür müssen Infrastruktureinrichtungen und wissenschaftliche Communities zusammenarbeiten und das Repository in übergreifende Strukturen integrieren. Folgende Aspekte sind dabei zu berücksichtigen:

1. Metadatenschema und Erschließung
2. Persistente Adressierung
3. Qualitätssicherung und Datenschutz

## 2.5 VIRTUELLE FORSCHUNGsumGEBUNGEN

Die Wissenschaft entwickelt sich in kooperativen, interdisziplinären, internationalen und ortsunabhängigen Strukturen. Die Nutzung digitaler Ressourcen und Werkzeuge und die Verschmelzung digitaler und physischer Ressourcen werden zunehmen und die Methodik der Wissenschaften verändern. Daher werden die Communities verstärkt das Potenzial digitaler Methoden und Virtueller Forschungsumgebungen nutzen.

Virtuelle Forschungsumgebungen sind eine integrative Plattform, in der digital(isiert)e Literatur (insbesondere Open Access), Forschungsdaten, Analysemethoden und andere Forschungsressourcen zusammengeführt werden. Sie haben das Ziel, alle Schritte des Forschungsprozesses von der Projektidee über die Datensammlung und Auswertung bis zur Veröffentlichung und Nachnutzung zu unterstützen.

4. Open Access für Forschungsdaten
5. Archivierung
6. Entwicklung von Geschäftsmodellen

Dabei können Data Life Cycle Labs entstehen, die für spezifische Wissenschafts-Communities im Land Methoden und Workflows für Forschungsdatenmanagement vorantreiben und in internationale Initiativen eingebunden sind. In Kooperation mit geeigneten Vereinigungen auf der nationalen Ebene wird ein Zertifikat für die Nachvollziehbarkeit und Reproduzierbarkeit von Forschungsdaten angeregt.

Die Aktivitäten zum Forschungsdatenmanagement müssen landesweit koordiniert werden, um eine Wissensplattform aufzubauen, Methoden und Best Practices zu verbreiten, rechtliche Fragen zentral zu klären sowie Policies, Methoden und Verrechnungsmodelle für die Langzeitarchivierung landesweit abzustimmen und umzusetzen.

Eine systematische Erschließung von Quellen und eine einheitliche umfassende Dokumentation der Ergebnisse leisten darüber hinaus einen wesentlichen Beitrag zur Qualitätssicherung.

Da die Wissenschaftsbereiche unterschiedliche Datenarten erzeugen und unterschiedliche Methoden anwenden, sind die Bedarfe an unterstützenden Werkzeugen – abgesehen von einigen grundlegenden Diensten – disziplinspezifisch. Dies bedeutet, dass Virtuelle Forschungsumgebungen von den Fach-Communities aufgebaut und von Infrastruktureinrichtungen unterstützt werden müssen. Für die zuverlässig planbare dauerhafte Bereitstellung müssen Governance-Strukturen entwickelt und Betriebs- und Geschäftsmodelle ausgehandelt werden, und zwar auf Landes-, Bundes- und EU-Ebene. Zugriffe und Zugänge weltweiter Nutzer sind zu regeln, aber auch darüber hinausgehende rechtliche Rahmenbedingungen.

Die neuen Methoden und Arbeitsweisen müssen über die an den Hochschulen gelebte Verzahnung von Forschung und Lehre in die

Ausbildung der Studierenden einfließen, um eine Verstetigung in der Praxis und eine kontinuierliche Weiterentwicklung zu erreichen. Virtuelle Forschungsumgebungen nehmen einige Aspekte der Lernmanagementsysteme und anderer verbreiteter Kommunikationsplattformen auf, werden aber darüber hinaus Forschungsmethoden und -infrastrukturen integrieren und daher weit über die E-Learning-Werkzeuge hinausgehen.

Baden-Württemberg will – zur Umsetzung von Maßnahmen, wie sie auch im Gesamtkonzept der Kommission Zukunft der Informationsinfrastruktur (KII) und den folgenden Empfehlungen des Wissenschaftsrates empfohlen wurden – eine führende Rolle bei der Entwicklung und Nutzung Virtueller Forschungsumgebungen einnehmen. Daraus folgt die Notwendigkeit, sich verstärkt mit anderen Ländern, dem Bund sowie der DFG abzustimmen.

Außerdem wird ein Förderprogramm zur Unterstützung der Etablierung Virtueller Forschungsumgebungen an den Hochschulen vorgeschlagen. Das primäre Ziel des Programms ist die Entwicklung, Anpassung, Verbreitung, Nutzung und der Betrieb Virtueller Forschungsumgebungen sowie ggf. die Einrichtung eines Kompetenzzentrums. Weitere Ziele sind die Reflexion der methodischen

Veränderung der Wissenschaften und die Integration in die Lehre am Beispiel erfolgreicher Anwendungen Virtueller Forschungsumgebungen. Drei Säulen der Förderung sollen in die Ausschreibung aufgenommen werden:

1. Identifizierung bestehender Infrastruktur und Bedarfe der Communities mit dem Ziel, dass Infrastruktureinrichtungen die Communities proaktiv beraten.
2. Anschub- und Kofinanzierung von Initiativen, die in nationale oder europäische Projekte münden. Ziel ist die Beteiligung an nationalen und internationalen Initiativen und die Profilbildung der Hochschulen.
3. Schaffung von Voraussetzungen für die Weiterführung von Elementen Virtueller Forschungsumgebungen über Förderperioden hinaus: Entwicklung von Governance-Strukturen und Geschäftsmodellen, Überführung in den Dauerbetrieb, Strukturwandel als Voraussetzung für Nachhaltigkeit.

### 3 · GLOSSAR

Die **Berliner Erklärung über offenen Zugang zu wissenschaftlichem Wissen** (engl.: Berlin Declaration on → Open Access to Knowledge in the Sciences and Humanities) wurde am 22. Oktober 2003 von den 19 Initiativmitgliedern von deutschen und internationalen Forschungsorganisationen beschlossen und bis zum Februar 2012 von 363 Institutionen unterzeichnet. Sie gilt als wichtiger Meilenstein der Open Access-Bewegung. Von früheren Open Access-Erklärungen, die sich auf die Forderung nach freier Zugänglichkeit der wissenschaftlichen Zeitschriftenliteratur im Internet beschränkten, unterscheidet sich die Berliner Erklärung durch die Einbeziehung des kulturellen Erbes, also des in Archiven, Bibliotheken und Museen verwahrten Kulturguts. <http://oa.mpg.de/lang/de/berlin-prozess/berliner-erklarung/>

**Big Data:** Große Datenmengen bzw. Algorithmen u. a. Methoden zur Analyse großer Datenmengen

**bwDATA, bwLSDF:** Aktuelle Konzepte und Projekte des MWK für Methoden der Datenspeicherung wissenschaftlicher Einrichtungen (→ LSDF)

**bwFLA:** Projekt des MWK zur funktionalen → Langzeitarchivierung komplexer digitaler Objekte

**bwHPC, bwForCluster, bwUniCluster:** Aktuelle Projekte des MWK zum Aufbau einer Infrastruktur für High Performance Computing (HPC) in Baden-Württemberg. Aufgebaut wird eine dreistufige Architektur mit dem Höchstleistungsrechenzentrum an der Universität Stuttgart (HLRS), dem Forschungshochleistungsrechner am KIT (ForHLR) und auf der Ebene 3 den bwUniCluster als Grundversorgungssystem und bwForCluster für ausgewiesene Fachwissenschaften.

**bwIDM:** Projekt des MWK zum föderativen Identity Management aller Universitäten in Baden-Württemberg. Ziel ist die vereinfachte Nutzung IT-Ressourcen, die den Universitäten landesweit zur Verfügung gestellt werden.

**cc-by:** cc-... sind verschiedene Creative Commons-Lizenzen zur freien Nutzung von Objekten geistigen Eigentums. Die Lizenz cc-by erlaubt die freie Nutzung, wenn der Urheber namentlich genannt wird, und zwar auch zur kommerziellen Nutzung und zur Überarbeitung.

**Data Life Cycle Labs:** Organisierte Gruppen von Wissenschaftlern und Entwicklern, die Methoden für den Data Life Cycle einer Wissenschaftsdisziplin bereitstellen. Der wissenschaftliche Data Life Cycle umfasst die Generierung, Analyse, Verknüpfung, Publikation und Archivierung der → Forschungsdaten

**Datenformat:** die Definition der Art, wie Daten gespeichert werden und wie ggf. ihre innere Logik interpretiert werden muss. Man unterscheidet proprietäre Formate, die von Software-Herstellern definiert werden und die vollständigen Regeln zur Interpretation der inneren Logik nicht offenlegen, und offene Formate, die vollständig dokumentiert sind.

**Deutsche Digitale Bibliothek (DDB):** Nationales Online-Portal für digitalisierte Bücher, Archivalien, Bilder, Skulpturen, Musikstücke und andere Tondokumente, Filme und Noten

**Digitalisierung:** Umwandlung von Informationen wie Ton, Bild oder Text in Zahlenwerte zum Zwecke ihrer elektronischen Bearbeitung, Speicherung oder Übertragung

**E-Content:** Digitale Inhalte

**ERIC:** European Research Infrastructure Consortium

**ESFRI:** European Strategy Forum on Research Infrastructures

**Europeana:** Europäisches Online-Portal für Bild-, Text-, Ton- und Video-Dateien.

**Forschungsdaten:** (digitale) Daten, die in einem wissenschaftlichen Vorhabens z. B. durch Quellenforschungen, Experimente, Messungen, Erhebungen, Befragungen o. ä. entstehen oder dem Vorhaben zugrunde liegen

**Goldener/Grüner Weg von Open Access:** Unter dem goldenen Weg wird → Open Access vor allem über Erstveröffentlichungen durch Open Access-Journals verstanden, die von Verlagen oder Wissenschaftseinrichtungen verlegt werden. Der grüne Weg bezieht sich dagegen in erster Linie auf Zweitveröffentlichungen vornehmlich in institutionellen Repositorien von Hochschulen und Forschungseinrichtungen.

**HPC:** High Performance Computing

**Identity Management:** Verwaltung der Identitäten und Zugriffsrechte von Mitgliedern, Angehörigen und Gästen einer Organisation (→ bwIDM)

**Image:** Digitales Abbild

**Interoperabilität:** Die Möglichkeit, verschiedene Systeme über (offene) Schnittstellen so zu verknüpfen, dass sie Elemente eines umfassenderen Systems sind

**KII-Papier:** Gesamtkonzept für die Informationsinfrastruktur in Deutschland – Empfehlungen der „Kommission Zukunft der Informationsinfrastruktur“ (KII) im Auftrag der Gemeinsamen Wissenschaftskonferenz des Bundes und der Länder (GWK), <http://www.leibniz-gemeinschaft.de/infrastrukturen/kii/>, April 2011

**Kollaboratives Arbeiten (oder E-Collaboration):** Arbeitsweisen (z. B. schreiben, dokumentieren, Wissensmanagement), bei denen mehrere Autoren als gemeinschaftliche Urheber zusammenarbeiten, z. B. in einem Wiki

**Konsortium Baden-Württemberg:** Zusammenschluss von derzeit 51 wissenschaftlichen Bibliotheken zum gemeinschaftlichen Erwerb der Lizenzen für elektronische Medien. Das Konsortium Baden-Württemberg besteht seit 1999. Die Geschäftsführung wird von der Universitätsbibliothek Freiburg geleistet. Die technische Umsetzung des Zugangs zu lizenzierten Medien wird über das Portalsystem ReDI (Regionale Datenbank-Information) bereitgestellt.

**Langzeitarchivierung:** Erfassung, langfristige Aufbewahrung und Erhaltung der dauerhaften Verfügbarkeit und Interpretierbarkeit von digital vorliegenden Informationen (→ bwFLA)

**LEO-BW:** Vernetztes landeskundliches Online-Informationssystem für Baden-Württemberg

**Lizenzierung:** hier: Erwerb der Nutzungsrechte für elektronische Medien (digitale Publikationen, E-Journals, Datenbanken) oder Software

**LSDF:** Large-scale data facility: Großer Datenspeicher für die wissenschaftlichen Einrichtungen in Baden-Württemberg (→ bwLSDF)

**Metadaten:** Strukturierte, maschinell lesbare Daten, die Informationen zu den Datenobjekten (z. B. Publikationen, Digitalisate, Forschungsdaten) enthalten. Sie sind zur Beschreibung, Identifizierung, Klassifizierung und Interpretation der Datenobjekte notwendig, z. B. Autoren, Titel, Stichwort, Lizenzinformation, Datenformat.

Als **Open Access** (engl.: offener Zugang) wird der freie Zugang zu wissenschaftlicher Literatur und anderen Materialien im Internet bezeichnet. Ein wissenschaftliches Dokument unter Open Access-Bedingungen zu publizieren, gibt jedermann die Erlaubnis, dieses Dokument zu lesen, herunterzuladen, zu speichern, es zu verlinken, zu drucken und damit entgeltfrei zu nutzen. Darüber hinaus können über Freie Lizenzen den Nutzern weitere Nutzungsrechte eingeräumt werden, welche die freie Nach- und Weiternutzung, Vervielfältigung, Verbreitung oder auch Veränderung der Dokumente ermöglichen können.

**Opt-in/Opt-out:** Lizenzverhandlungen werden grundsätzlich für das ganze Konsortium oder für ein Teilkonsortium (z. B. Universitätsbibliotheken) geführt. Nach Möglichkeit werden sie so offen gehalten, dass sich einzelne Bibliotheken von der Konsortiallizenz ausschließen (Opt-out) oder sich einem Teilkonsortium anschließen (Opt-in) können. (S. 33)

**Persistent Identifier (PID):** Eindeutige, dauerhafte Adresse digitaler Inhalte, z. B. URN<sup>®</sup>, DOI<sup>®</sup>)

**Repository:** Speicher zur Ablage digitaler Inhalte und ihrer → Metadaten

**Resource Discovery System (RDS):** Portal zur Recherche in elektronischen Medien, insbesondere in digitalen Publikationen. Durch die Nutzung von Indexdiensten, die eigens lizenziert werden, können z. B. Autoren, Titel, Stichwörter gesucht oder auch Abstracts oder Volltexte durchsucht werden.

**Retrodigitalisierung (oder retrospektive Digitalisierung):** Digitalisierung analoger Publikationen (z. B. Handschriften, Printmedien, Filme, Tonbänder).

**Scanner:** Gerät zur → Digitalisierung von physischen Objekten mit optischen Methoden. Für die unterschiedlichen historischen Dokumentenarten werden unterschiedliche Scanner benötigt, um qualitativ hochwertige Digitalisate bei schonendster Behandlung der Objekte erstellen zu können. Unterschieden werden Buchscanner bzw. Aufsichtsscanner, Mikrofilm-Scanner, Dokumenten-Scanner und Flachbett-Scanner.

**Virtuelle Forschungsumgebung (VFU):** Arbeitsplattform, die eine kooperative Forschungstätigkeit durch mehrere Wissenschaftler an unterschiedlichen Orten zu gleicher Zeit ohne Einschränkungen ermöglicht

**ZVDD:** Portal für in Deutschland erstellte Digitalisate von Druckwerken vom 15. Jahrhundert bis heute

**Zweitveröffentlichungsrecht:** Darunter ist das Recht des Autors (persönliches Mandat) zu verstehen, bei öffentlich geförderten Forschungsergebnissen nach Ablauf einer Frist von in der Regel 6 bis 12 Monaten seine Publikation unbeschadet von Rechten Dritter anderweitig nichtkommerziell öffentlich zugänglich zu machen. Für eine entsprechende urheberrechtsvertragliche Regelung im Urheberrechtsgesetz setzen sich die Allianz der Wissenschaftsorganisationen und die Kultusministerkonferenz, aber auch die EU-Kommission im Rahmen von Horizont 2020 ein. Mitunter wird hierfür synonym der Begriff „Zweitverwertungsrecht“ verwendet; dies ist jedoch missverständlich, da darunter das Recht der Institution, bei der der Autor als Wissenschaftler beschäftigt ist, verstanden wird, dessen Publikation anderweitig nichtkommerziell zu veröffentlichen (sog. institutionelles Mandat).

## 4 · LITERATUR

### Deutsche Forschungsgemeinschaft (DFG):

- Positionspapier „Die digitale Transformation weiter gestalten – der Beitrag der Deutschen Forschungsgemeinschaft zu einer innovativen Informationsinfrastruktur für die Forschung“, [http://www.dfg.de/download/pdf/foerderung/programme/lis/positionspapier\\_digitale\\_transformation.pdf](http://www.dfg.de/download/pdf/foerderung/programme/lis/positionspapier_digitale_transformation.pdf), Juni 2012
- Übergreifende Empfehlungen zu Informationsinfrastrukturen, <http://www.wissenschaftsrat.de/download/archiv/10466-11.pdf>, Januar 2011.

### Enquete-Kommission Internet und digitale Gesellschaft:

- Handlungsempfehlungen der Projektgruppe „Bildung und Forschung“, Ausschussdrucksache 17(24)052, [http://www.bundestag.de/internetenquete/dokumentation/Sitzungen/20120625/A-Drs\\_17\\_24\\_052\\_-\\_PG\\_Bildung\\_und\\_Forschung\\_Handlungsempfehlungen.pdf](http://www.bundestag.de/internetenquete/dokumentation/Sitzungen/20120625/A-Drs_17_24_052_-_PG_Bildung_und_Forschung_Handlungsempfehlungen.pdf), Juni 2012.

### EU-Kommission:

- Empfehlung der Kommission vom 17. Juli 2012 über den Zugang zu wissenschaftlichen Informationen und deren Bewahrung (2012/417/EU), <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2012:194:0039:0043:DE:PDF>, Juli 2012.
- Riding the wave – How Europe can gain from the rising tide of scientific data. Final report of the High Level Expert Group on Scientific Data. A submission to the European Commission, <http://cordis.europa.eu/fp7/ict/e-infrastructure/docs/hlg-sdi-report.pdf>, Oktober 2010.

### Gemeinsame Wissenschaftskonferenz des Bundes und der Länder (GWK):

- Gesamtkonzept für die Informationsinfrastruktur in Deutschland – Empfehlungen der Kommission Zukunft der Informationsinfrastruktur im Auftrag der GWK, <http://www.leibniz-gemeinschaft.de/infrastrukturen/kii/>, April 2011.

### Hochschulrektorenkonferenz (HRK):

- Hochschule im digitalen Zeitalter: Informationskompetenz neu begreifen – Prozesse anders steuern. Entschließung der 13. Mitgliederversammlung vom 20.11.2012, [http://www.hrk.de/uploads/tx\\_szconvention/Entschl.\\_Informationskompetenz\\_final\\_20\\_11.pdf](http://www.hrk.de/uploads/tx_szconvention/Entschl._Informationskompetenz_final_20_11.pdf), November 2012.

**Wissenschaftsrat (WR):**

- Empfehlungen zu einem Kerndatensatz Forschung, <http://www.wissenschaftsrat.de/download/archiv/2855-13.pdf>, Januar 2013.
- Empfehlungen zur Weiterentwicklung der wissenschaftlichen Informationsinfrastrukturen in Deutschland bis 2020, <http://www.wissenschaftsrat.de/download/archiv/2359-12.pdf>, Juli 2012.
- Übergreifende Stellungnahme zu Informationsinfrastrukturen, <http://www.wissenschaftsrat.de/download/archiv/10466-11.pdf>, Januar 2011.

<sup>1</sup> Der Computer-Pionier Konrad Zuse, ein Bauingenieur, begann die Entwicklung der Computer-Technik, um sich selbst von den ermüdenden baustatischen Routineberechnungen zu befreien.

<sup>2</sup> Eugene Myers, Direktor des Max-Planck-Instituts für molekulare Zellbiologie und Genetik, Dresden, bei einem Vortrag am 16.10.2012 am HITS, Heidelberg: „Combined with new modalities of light microscopy, this allows us to observe molecular mechanisms within the cell, observe the developmental trajectory of growing organs, and to map the cellular anatomy of organisms and organs such as the brain, the heart, or the stem of a plant. All this increasingly requires computation to either extract information or to quantitatively measure an effect in the vast sea of images produced by such explorations. This is creating the growing sub-field of bioimage informatics.“

<sup>3</sup> „Towards 2020 Science“, <http://research.microsoft.com/en-us/um/cambridge/projects/towards2020science/downloads.htm>, Microsoft Corporation, 2006.

## AG 1 · Lizenzierung

# Innovationspotenzial und Perspektiven für die Lizenzierung von E-Medien und Software in Baden-Württemberg

## INHALT

|   |    |
|---|----|
| 1. Einleitung .....   | 27 |
| 2. System Lizenzen .....  | 27 |
| 3. Nutzeranforderungen .....  | 27 |
| 3.1 Erwartungen der Wissenschaftlerinnen und Wissenschaftler .....  | 27 |
| 3.2 Erwartungen der Studierenden .....  | 28 |
| 3.3 Nachfrageorientierte Bedarfserhebung .....  | 29 |
| 4. Status Quo in Baden-Württemberg .....  | 29 |
| 4.1 E-Literatur und Datenbanken .....   | 29 |
| 4.2 Software .....  | 31 |
| 5. Nationaler und internationaler Kontext .....   | 32 |
| 6. Innovationspotenzial und Perspektiven für die Lizenzierung von E-Medien<br>und Software in Baden-Württemberg ..... | 32 |
| 6.1 Konsortium Baden-Württemberg zur Versorgung mit E-Medien .....  | 32 |
| 6.2 Neues Konsortialmodell Softwarelizenzen .....   | 34 |
| 7. Konkreter Handlungsbedarf (2013-2017) und Empfehlungen an ... .....  | 35 |
| 7.1 ... die Universitäten, Hochschulen (LRKs) und Landesbibliotheken .....  | 35 |
| 7.2 ... die Landesregierung .....   | 36 |
| Anhang 1: Die Versorgungsstruktur für elektronische Medien .....  | 37 |
| Anhang 2: Am Konsortium BW teilnehmende Einrichtungen .....   | 38 |
| Anhang 3: Alternative Konsortialstrukturen .....  | 38 |
| Anhang 4: Technische Infrastruktur des Konsortiums .....  | 39 |
| Anhang 5: Etatentwicklung des Konsortiums Baden-Württemberg .....   | 40 |
| Anhang 6: Ausgaben der wissenschaftlichen Bibliotheken für digitale Inhalte .....                                     | 41 |
| Anhang 7: Übersicht Landes-Softwarelizenzen .....   | 42 |
| Anhang 8: Qualitätssicherungsmaßnahmen bei Vergabe zentraler Mittel .....   | 43 |
| Anhang 9: Schnittstelle zu Virtuellen Forschungsumgebungen .....  | 43 |
| Anhang 10: Transformationsprozess von Subskriptions- zu Open Access-Modellen .....                                    | 44 |

## MITGLIEDER DER ARBEITSGRUPPE

**PROF. DR. OLAF GEFELLER**

Institut für Medizininformatik, Biometrie und Epidemiologie,  
Universität Erlangen-Nürnberg

**KARL-WILHELM HORSTMANN**

Kommunikations-, Informations- und Medienzentrum,  
Universität Hohenheim

**DR. JOCHEN JOHANNSEN**

Badische Landesbibliothek Karlsruhe

**DR. ANTJE KELLERSOHN**

Bibliothek, Universität Freiburg

**PER KNUDSEN**

Bibliothek, Universität Mannheim

**ATO RUPPERT**

Bibliothek, Universität Freiburg

**ASTRID SCHÖNSTEIN**

Ministerium für Wissenschaft, Forschung und Kunst  
Baden-Württemberg (Vorsitz)

**DR. CHRISTIANE SPARY**

Bibliothek, Pädagogische Hochschule Ludwigsburg

**PROF. DR. THOMAS WALTER**

Zentrum für Datenverarbeitung, Universität Tübingen

## 1 · EINLEITUNG

*„Der Zugang zu Datenbanken und E-Journals sowie die Nachhaltigkeit und Nachnutzung wissenschaftlicher Daten wird neben der Geräteausstattung ein immer wichtigerer Faktor für Forschungs- und Innovationsprozesse. Wir werden deshalb verstärkt in die Informationsversorgung investieren.“*

*Koalitionsvertrag „Der Wechsel beginnt“, S. 14, 2011*

Als Teil der Landesstrategie Informationsinfrastruktur in Baden-Württemberg beschreibt das vorliegende Papier den Versorgungsstand der Universitäten und Hochschulen des Landes sowie der Landesbibliotheken mit elektronischen Informationsmedien. Es analysiert die sich wandelnden Nutzererwartungen an die Bibliotheken als Anbieter bzw. Vermittler der digital verfügbaren Inhalte und formuliert unter Berücksichtigung

der nationalen und internationalen Entwicklungen in diesem Bereich Strategien zur Professionalisierung der Konsortialtätigkeit. Außerdem werden Bedarfe genannt, die die Grundversorgung der baden-württembergischen Universitäts- und Hochschulangehörigen sowie der Nutzerinnen und Nutzer der Landesbibliotheken mit elektronischen Informationsquellen nachhaltig sicherstellen. Das Strategiepapier skizziert zudem ein innovatives Konsortialmodell für Softwarelizenzen und beschreibt die hierfür notwendigen Strukturen.

## 2 · SYSTEM LIZENZEN

Bei der Erwerbung von Lizenzen für elektronische Informationsmedien bzw. Software gibt es ausgeprägte gemeinschaftliche Erwerbungsstrukturen sowohl auf regionaler als auch auf nationaler Ebene. Kompetenzen und Ressourcen verschiedener Institutionen werden gebündelt. Ziel der konsortialen Erwerbung ist es, über eine gesteigerte Marktmacht günstigere Konditionen in

Form von Kostenersparnissen über Rabatte oder in Form einer größeren Anzahl zur Verfügung stehender Inhalte zu erzielen. Zwischen nationalen und regionalen Konsortien gibt es in dieser Hinsicht keine Unterschiede; diese bestehen letztlich nur im Hinblick auf eine mögliche Förderung durch die DFG und in dem größeren Kreis an potenziellen Teilnehmern, inklusive beispielsweise der außeruniversitären Forschungseinrichtungen (zur Versorgungsstruktur vgl. **Anhang 1**).

## 3 · NUTZERANFORDERUNGEN

### 3.1 ERWARTUNGEN DER WISSENSCHAFTLERINNEN UND WISSENSCHAFTLER

Nutzungsintensität und Bedeutung von elektronischen Medien für die wissenschaftliche Forschung unterscheiden sich zwischen den einzelnen Fächern gravierend. Außerhalb des geisteswissenschaftlichen Bereichs ist jedoch zunehmend die nahezu ausschließliche Verwendung elektronisch verfügbarer Inhalte festzustellen. Diese Entwicklung, die sich durch den anstehenden Generatio-

nenwechsel innerhalb der aktiv tätigen Wissenschaftlerinnen und Wissenschaftler noch verstärken und sich vermutlich auch auf den geisteswissenschaftlichen Bereich ausdehnen wird, hat direkte Auswirkungen auf die Erwartung der Wissenschaftlerinnen und Wissenschaftler an die Bibliotheken als Anbieter bzw. Vermittler der digital verfügbaren Inhalte. Klassische Bibliotheksaufgaben wie Bestandsaufbau und -archivierung von Buch- und Zeitschriftenbeständen sowie die Pflege von bibliothekarischen

Dokumentations- und Klassifikationssystemen zur Erschließung dieser Bestände treten in den Hintergrund und machen dem Verständnis einer Bibliothek als breit aufgestellter Informationsinfrastrukturanbieter, der auch die Welt digital verfügbarer Information zugänglich macht, in fortschreitendem Maße Platz.

Einheitliche Nutzeranforderungen, wie dies konkret in optimaler Weise zu geschehen hat, lassen sich nicht ohne einen Bezug zu einzelnen Fächern formulieren. Allen Fächern gemeinsam ist allerdings der Nutzerwunsch nach einem technisch komfortablen elektronischen Zugang zu digital verfügbarer Information, der ohne komplizierte Authentifizierungsverfahren möglich sein muss (umfassende Nutzbarkeit, zeitlich und mengenmäßig unbegrenztes Abspeichern, unbeschränkte Ausdruck- und Kommentierungsfunktionen, Verfügbarkeit in Standardformaten wie PDF oder EPUB). Auch eine direkte Einbindung in Virtuelle Forschungsumgebungen und andere Systeme der Forschungskollaboration muss möglich sein.

Zudem ist die Integration der elektronischen Bibliotheksangebote in bereits bestehende fachspezifische elektronische

Nachweis- und Portalsysteme von besonderer Bedeutung für die Steigerung der Akzeptanz und praktischen Nutzung innerhalb der Wissenschafts-Community.

Alle kontinuierlich voranschreitenden Entwicklungen auf diesem Gebiet, durch das Angebot von Informationsinfrastrukturanbietern einen komfortablen Zugang zu digital verfügbarer Information mittels einfach zu bedienender elektronischer Werkzeuge zu ermöglichen, sind stets daraufhin zu überprüfen, ob sie von der Zielgruppe, den forschenden Wissenschaftlerinnen und Wissenschaftlern unterschiedlicher Fächer, breit genutzt werden. Eine parallel zur technischen Weiterentwicklung stattfindende Information und Fortbildung der Wissenschaftler über die sich bietenden Möglichkeiten muss daher unverzichtbarer Bestandteil der Aktivitäten sein.

### 3.2 ERWARTUNGEN DER STUDIERENDEN

Studentische Nutzeranforderungen decken sich in weiten Teilen mit den unter 3.1 beschriebenen Erwartungen der Wissenschaftler, insbesondere hinsichtlich eines technisch komfortablen elektronischen Zugriffs auf digital verfügbare Information im Sinne einer Plug & Play-Lösung. Die Akzeptanzfrage stellt sich bei Studierenden in noch geringerem Maße als im Wissenschaftlersegment, da hier keine tradierten Nutzungsgewohnheiten verändert werden müssen. Studentische Nutzung digitaler Angebote unterscheidet sich nicht in der Frage, wie die Angebote genutzt werden, sondern nur – insbesondere bei Studierenden in einem frühen Studienabschnitt – in der Frage, was inhaltlich im Fokus der Nutzung steht. Insbesondere das E-Book-Segment, in dem auch einführende Lehrbücher und Standardwerke eines Fachgebiets für

eine digitale Nutzung zur Verfügung stehen, erfreut sich bei Studierenden einer immer stärkeren Nachfrage und bedarf daher des besonderen Augenmerks beim Ausbau digitaler Angebote für Studierende. Dabei ist zu beachten, dass es sich hierbei um durchaus hochpreisige Angebote handelt, da sich die Anbieter den Wegfall von Mehrfachexemplaren mit einem deutlich höheren E-Book-Preis refinanzieren lassen. Mit fortschreitendem Stadium des Studiums und fachlicher Spezialisierung gleichen sich die studentischen Anforderungen auch in inhaltlichen Aspekten immer stärker an wissenschaftliche Nutzeranforderungen an.

### 3.3 NACHFRAGEORIENTIERTE BEDARFSERHEBUNG

Bibliotheken und Informationszentren in den Hochschulen des Landes steuern Lizenzierung und Bereitstellung gedruckter und digitaler Medien nachfrageorientiert in enger Abstimmung mit den Wissenschaftlern und Studierenden. Die Bedarfsermittlung erfolgt z. B. anhand von Bedarfslisten aus Instituten, Fakultäten und/oder Lehrstühlen, durch Beschaffungsvorschläge von Studierenden, durch Gespräche von Fachreferenten mit den jeweiligen Universitätseinrichtungen, durch Senatsausschüsse für Bibliotheken bzw. Informationseinrichtungen oder weitere Gremien zur Abstimmung des Medienbedarfs in Forschung, Lehre und Studium. Eine weitere Quelle zur Identifizierung des Medienbedarfs sind zudem Berufungsverhandlungen. Auch die Landesbibliotheken gehen in ihrer Erwerbspolitik nutzerorientiert vor und stimmen sich

dabei auch mit den Universitäts- und Hochschulbibliotheken vor Ort ab.

Die Nutzung bereitgestellter digitaler Medien wird routinemäßig über Nutzungsstatistiken (Suchanfragen, Downloads) erfasst und dient ebenfalls als wesentliche Grundlage bei den Beschaffungs- bzw. Lizenzierungsentscheidungen. Der auf diese Weise lokal erhobene Bedarf an digitalen Medien wird über die Bibliotheken und Informationszentren in das baden-württembergische Erwerbungsconsortium oder in bundesweite Konsortien zur gemeinsamen Lizenzierung überführt.

Der Bedarf der Wissenschaftler/innen und Studierenden wird somit systematisch und lokal erhoben und soweit bei beschränkten Ressourcen möglich, lokal oder über konsortiale Beschaffung und Bereitstellung befriedigt. Die Auswahl der zu lizenzierenden Inhalte erfolgt also *bottom up* und nicht *top down*.

## 4 · STATUS QUO IN BADEN-WÜRTTEMBERG

### 4.1 E-LITERATUR UND DATENBANKEN

Das Konsortium Baden-Württemberg, ein Zusammenschluss der wissenschaftlichen Bibliotheken im Geschäftsbereich des Ministeriums für Wissenschaft, Forschung und Kunst Baden-Württemberg, besteht seit 1999. Derzeit nehmen 51 Bibliotheken am Konsortium teil (vgl. **Anhang 2**). Es organisiert und koordiniert den gemeinschaftlichen Erwerb von Lizenzen für Datenbanken, E-Books und elektronischen Zeitschriften und erzielt dabei z. T. beträchtliche Rabatte bzw. einen Zugewinn an zur Verfügung stehenden Inhalten. Darüber hinaus konnten durch die Konsortialbeschaffung der Arbeits- und Verwaltungsaufwand reduziert und das Know-how im Umgang mit elektronischen Informationsressourcen gebündelt werden.

Das Konsortium wird gegenüber den Verlagen und Anbietern von der UB Freiburg vertreten. Sie übernimmt die Geschäftsführung, koordiniert die Arbeit des Konsortiums und bewirtschaftet die zur Verfügung gestellten Mittel. Die Entscheidung über die konsortiale Beschaffung eines Produktes fällt stets das Konsortium. Die Verhandlungen mit den Anbietern werden von den Konsortialteilnehmern arbeitsteilig geführt<sup>4</sup> (zu alternativen Konsortialstrukturen vgl. **Anhang 3**).

Seit 1999 stellt das Wissenschaftsministerium jährlich eine feste Summe zur Beschaffung von elektronischen Fachinformationsangeboten und für den technischen Betrieb von ReDI (Regionale Datenbankinformation) zur Verfügung (zu ReDI vgl. **Anhang 4**). Die Mittelzuteilung ist aus historischen Gründen in die folgenden Titel aufgeteilt:

Zuweisungen für das Konsortium Baden-Württemberg (2011)

| KAPITEL / TITEL | ZUWEISUNGSBETRAG<br>IN EURO | BEMERKUNGEN  |
|-----------------|-----------------------------|--|
| 1407 428 01     |                             | 2 Personalstellen E13TVL<br>(Stuttgart, Freiburg) für ReDI |
| 1407 534 02     | 119.100                     |  |
| 1407 74*        | 1.056.500                   | Summe 1407 abzüglich 70.000 Euro GMA                       |

\* Die Ausgabeermächtigung erhöht oder vermindert sich um die Mehr- oder Wenigereinnahmen bei Titelgruppe 74 i.H.v. 103.500 Euro

Die zur Beschaffung von Lizenzen verfügbaren Mittel sind seit 1999 nicht erhöht worden. Dem gegenüber stehen aber Preissteigerungen im Bereich der elektronischen Angebote, die deutlich über dem allgemeinen Preissteigerungsniveau liegen (z. T. mehr als 8 % p.a.<sup>5</sup>). Die Zahl der teilnehmenden Einrichtungen steigt, so dass sich auch die anfallenden Lizenzkosten für die lizenzierten Angebote kontinuierlich erhöhen. Das Angebot an elektronischen Fachinformationen muss parallel mit dem Aufkommen neuer Informationsprodukte und dem entsprechenden Bedarf der Hochschulen stetig ausgeweitet werden. Die hieraus entstehende Finanzierungslücke muss das Konsortium seit 2006 durch eine Umlage auf die nutzenden Einrichtungen decken. Die Umlagebeträge sind von anfangs 14 % auf inzwischen 39 % dramatisch angestiegen (vgl. **Anhang 5, Tabelle 1**). Das MWK hat seit 2006 immer wieder Sondermittel bereitgestellt, um einzelne Bibliothekssparten bei der Beschaffung von elektronischen Zeitschriften zu unterstützen (vgl. **Anhang 5, Tabelle 2**). Dennoch wird die Diskrepanz zwischen den zentral durch das MWK finanzierten Ausgaben und den lokal von den baden-württembergischen wissenschaftlichen Bibliotheken für digitale Inhalte insgesamt aufgewendeten Mitteln immer größer. So sind allein zwischen 2005 und 2010 die Ausgaben für digitale Angebote von knapp 4 Mio. auf knapp

16 Mio. Euro angestiegen und werden bald die 20 Mio.-Marke überschreiten (vgl. **Anhang 6**). Der Transformationsprozess bei den Zeitschriften – weg von der gedruckten Ausgabe, hin zu den E-Journals – ist bereits weit gediehen und insbesondere durch den Kauf großer Pakete in den STM-Bereichen<sup>6</sup> stark befördert worden. Die drastisch steigenden Ausgaben für E-Journals gehen aber immer mehr zu Lasten der Etats für Printmedien, insbesondere für den Monografienkauf (vgl. **Anhang 6**). Dieser Trend wird in besonderem Maße die Geistes- und Kulturwissenschaften treffen, die im Vergleich zu den STM-Fächern wenig Bedarf an Zeitschriften haben, aber besonders abhängig von einer guten Ausstattung an – in der Regel noch gedruckten – Monografien sind. Auch im Bereich der Zeitschriften und Datenbanken mussten aufgrund fehlender Ressourcen bereits substanzielle Abbestellungen durchgeführt werden. Als besonders drastisches Beispiel ist hier die Datenbank REAXYS zu nennen, die auch an Hochschulen mit forschungsstarken Chemie-Fakultäten nicht mehr flächendeckend zur Verfügung steht.

## 4.2 SOFTWARE

Die Rechen- und Informationszentren der Universitäten des Landes haben die Vorteile des konsortialen Lizenz- und Software-Managements beim Einsatz kommerzieller Softwareprodukte früh erkannt und seit mehreren Jahren konsequent ausgebaut<sup>7</sup>. Die Vorteile sind primär wirtschaftlicher Art, darüber hinaus ergeben sich durch ein abgestimmtes Vorgehen eine wesentlich vereinfachte Verwaltung, eine Vereinheitlichung des lokalen Angebotes sowie eine wesentlich stärkere Verhandlungsposition gegenüber dem Lizenzgeber. Als sinnvoll hat sich dieses Vorgehen in allen Fällen des Volumengeschäfts erwiesen, wenn also eine große Anzahl an Lizenzen abgenommen wird. Je nach Lizenzgeber kommen verschiedene Modelle zum Einsatz: Beitrittsmodelle (nochmal unterschieden zwischen ganzen Einrichtungen und Teileinrichtungen), Abgabe von Lizenzbündeln oder nur Angebot einheitlicher Vergabebedingungen. Der Kreis der Teilnehmer geht z. T. über den der staatlich getragenen Hochschulen hinaus (private Hochschulen).

Im Jahr 2012 bestehen konsortiale Lizenzverträge für alle Microsoft Desktop-Produkte, für das integrierte Bibliothekssystem aDIS/BMS, für GIS-Software (ESRI arcGIS), für Statistik (IBM SPSS), für Viren-Software (McAfee) und Backup-Software (IBM Tivoli TSM) (vgl. **Anhang 7**). Weitere werden verhandelt (Matlab, VMWare). Das Gesamtvolumen dieser Lizenzverträge liegt jährlich im Bereich von 3 Mio. Euro. Softwarelizenzen erwerben die Hochschulen mit eigenen Mitteln. In einzelnen Fällen gibt es für die Anfangsphase (maximal 3 Jahre) einen Zuschuss des Landes, um leichter genügend Teilnehmer für ein wirtschaftlich tragfähiges Modell zu finden.

Die Umsetzung von Landeslizenzen im Bereich Software ist grundsätzlich vergleichbar mit der der E-Medien und erfolgt in einem mehrstufigen Verfahren:

(1) Bedarfsermittlung:

Im Kreise der Leiter der Informations- oder Rechenzentren werden regelmäßig die in Frage kommenden Produktbereiche bewertet und wird ggf. ein Teilnehmer mit der genaueren Analyse betraut.

(2) Entwicklung einer Umrechnungspauschale:

Komplex gestaltet sich die Ermittlung einer Kostenverrechnung, so dass möglichst viele Teilnehmer eine potenzielle Konsortiallizenz im Bereich Software nutzen und andererseits geringe Kosten für die Teilnehmer anfallen. Hier liegt ein umfassender, hochschultypübergreifender Prozess zu Grunde, der auch Vorverhandlungen mit dem potenziellen Lizenzgeber beinhaltet.

(3) Im Falle der temporären Mitfinanzierung durch das Land nach Art. 143c GG ist eine Begutachtung im Programm "Großgeräte der Länder" der DFG durchzuführen.

(4) Wenn ein akzeptiertes Verteilmodell gefunden ist und die belastbaren Beitrittserklärungen der Teilnehmer vorliegen, sind die eigentlichen Vertragsverhandlungen mit dem Lizenzgeber durch die federführende Einrichtung zu führen und ist der Vertrag abzuschließen.

(5) Häufig verkauft der Lizenzgeber nicht direkt an die federführende Einrichtung, sondern bietet seine Dienste nur über Reseller an; dann ist eine Ausschreibung, bei den Kostenvolumina meistens EU-weit, notwendig.

(6) Nach Erwerb der Lizenzen sind diese abzurechnen.

(7) Parallel zur Abrechnung sind die Lizenzen technisch zu verteilen. Hierzu kann auch der Betrieb von Portalen zur Softwareverteilung, wichtiger aber noch der Betrieb von Servern für die Verwaltung von Lizenzkeys (wie KMS von Microsoft) gehören. Hier und bei Schritt (5) wird in manchen Fällen auch auf bestehende Strukturen industrieller Partner für den Softwarevertrieb zurückgegriffen, wobei dann die Umsatzsteuerrechnung besonders zu beachten ist.

## 5 · NATIONALER UND INTERNATIONALER KONTEXT

Sowohl der Wissenschaftsrat als auch die Gemeinsame Wissenschaftskonferenz des Bundes und der Länder haben in der laufenden Diskussion um die Weiterentwicklung der Informationsinfrastruktur in Deutschland die Ansätze der DFG bzw. der Allianz der deutschen Wissenschaftsorganisationen für eine nationale Lizenzierungsstrategie aufgegriffen und in Richtung des Aufbaus eines „stabile[n] strukturelle[n] Rahmen[s]“ für die gemeinschaftliche Bedarfserhebung und zentralisierte Verhandlungsführung für wissenschaftliche Literatur und Informationen in digitaler Form weiterentwickelt.<sup>8</sup> Bis auf weiteres sollen die AG Lizenzen der Allianzinitiative und die DFG hier eine koordinierende Rolle auf der nationalen Versorgungsebene übernehmen. Die Lizenzkosten wie auch die Overheadkosten sollen von den lizenznehmenden Einrichtungen (Bibliotheken) selbst bzw. als zentrale Mittel von deren Unterhaltsträgern aufgebracht werden. In infrastrukturpolitischer Hinsicht gehen von den genannten Organisationen starke Impulse in Richtung einer stärkeren Koordinierung des Konsortialge-

schehens aus. Dabei können und sollen sich auch regionale Konsortialstrukturen zukünftig verstärkt einbringen.

Im Rahmen des neuen Förderprogramms der DFG „Fachinformationsdienste für die Wissenschaft“ sollen mittelfristig Kompetenzzentren zur Lizenzierung elektronischer Medien aufgebaut werden. Diese Einrichtungen sollen für alle Bibliotheken, die Verantwortung für ein oder mehrere Fachinformationsdienste tragen, einen Vollservice zur Beschaffung kostenpflichtiger digitaler Ressourcen und bei der Abwicklung technischer Prozesse aus einer Hand anbieten<sup>9</sup>. Es wird derzeit geprüft, inwieweit das Konsortium Baden-Württemberg sich als ein solches Kompetenzzentrum in die Versorgungsstruktur der Fachinformationsdienste einbringen kann.

## 6 · INNOVATIONSPOTENZIAL UND PERSPEKTIVEN FÜR DIE LIZENZIERUNG VON E-MEDIEN UND SOFTWARE IN BADEN-WÜRTTEMBERG

### 6.1 KONSORTIUM BADEN-WÜRTTEMBERG ZUR VERSORGUNG MIT E-MEDIEN

#### **Vision 2020 – Vollversorgung für Forschung und Lehre**

In Baden-Württemberg steht 2020 eine Informationsinfrastruktur für Wissenschaft, Forschung und Lehre und über weitere Wissenschaftseinrichtungen wie die Landesbibliotheken auch für weitere Bevölkerungsschichten zur Verfügung, die bundesweit als vorbildlich anerkannt ist. Sie zeichnet sich durch ebenso hervorragende wie einfach zu bedienende Zugangs- und Recherchemöglichkeiten aus, die kommerziellen Suchmaschinen in nichts nachstehen, sowie durch eine ausgezeichnete Versorgung mit Volltexten. Studieren, Lehren und Forschen in Baden-Württemberg bedeutet, dass die benötigte Literatur bzw. andere Informationsmedien unmittelbar zur Verfügung stehen, sei es Open Access, mit vorhandener Lizenz oder als elektronischer Sofortlieferdienst. Diese Informationsquellen sind bruchlos integriert in Virtuelle Forschungs-, Arbeits- und E-Lear-

ning-Umgebungen, die bei einem Hochschulwechsel weiterverwendet werden können. Der Zugang zu elektronischen Informationsquellen ist innerhalb Baden-Württembergs auf Grund der existierenden Vollversorgung gewährleistet, so dass Universitäten und Hochschulen in diesem Bereich innerhalb des Landes nicht konkurrieren, während exzellente Forschende, Lehrende und Studierende deshalb in Baden-Württemberg bleiben bzw. nach Baden-Württemberg kommen wollen.

#### **Das Finanzkonzept als Beteiligungsmodell**

Um eine derartige ausreichende Versorgung mit elektronischen Fachinformationsangeboten durch

das Konsortium Baden-Württemberg sicherstellen zu können, ist ein schrittweiser Mittelaufwuchs für die Grundversorgung auf insgesamt 10 Mio. Euro notwendig.

Die Lizenzierung von Produkten über das Konsortium Baden-Württemberg setzt grundsätzlich eine 50% Eigenbeteiligung der lizenznehmenden Einrichtungen voraus. Die Lizenzierung erfolgt – analog zur nationalen Versorgungsebene – über Opt-in/Opt-out-Verfahren. Im Rahmen des eigenen aber auch des Konsortialbudgets können die Einrichtungen einzelne Produkte bedarfsgerecht lizenzieren.

Nach Verfügbarkeit eines Rahmenangebotes eines Anbieters/Verlages, in dem die Lizenzbedingungen abschließend beschrieben sind, werden die tatsächlich interessierten Einrichtungen informiert und aufgefordert, ihr Votum abzugeben. Der endgültige Lizenzvertrag wird dann vom Konsortium, vertreten durch die Geschäftsstelle bzw. ihre Trägerorganisation, im Namen und Auftrag der beteiligten Einrichtungen abgeschlossen. Der Nutzungs- und Lizenzzeitraum beträgt jeweils mindestens ein Jahr. Die Kostenabrechnung gegenüber den Anbietern/Verlagen erfolgt über die Geschäftsstelle.

Durch dieses Konzept entsteht eine Angebotsplattform, die zentral organisiert, aber dezentral entschieden wird. Mit diesem Geschäftsmodell bleibt die Entscheidung zur Lizenzierung eines Produktes grundsätzlich in der einzelnen Einrichtung.

Sofern die zentral zur Verfügung gestellten Mittel nicht ausreichen, um den gesamten Bedarf an digitalen Angeboten zu decken, wird derzeit im Konsortium schrittweise ein umlagebasiertes Finanzierungsmodell eingeführt, damit eine möglichst faire und

gerechte Verteilung der zentralen Mittel erfolgt (zu Qualitätssicherungsmaßnahmen bei Vergabe zentraler Mittel vgl. **Anhang 8**).

### **Professionalisierung der Geschäftsstelle**

Um die Ausweitung des Konsortialportfolios unterstützen zu können, muss eine zentrale Geschäftsstelle personell neu aufgebaut werden. Sie koordiniert die Verhandlungen und pflegt Kontakte zu den Anbietern. Sie unterstützt die Verhandler logistisch und juristisch, z. B. bei Fragen der Vertragsausgestaltung oder hinsichtlich der Nutzungsmöglichkeiten im Kontext von Virtuellen Forschungsumgebungen. Die Geschäftsstelle sollte künftig auch verstärkt als Koordinierungs- bzw. Beratungsinstanz für Open Access-Geschäftsmodelle fungieren. Sie ist für die ordnungsgemäße Finanzabwicklung aller Vorgänge verantwortlich.

Hierzu ist zum einen eine Stelle des höheren Dienstes erforderlich, vorteilhaft ist eine Ausbildung zum wissenschaftlichen Bibliothekar mit juristischem Hintergrund. Zum anderen sollte zur Abwicklung administrativer Aufgaben (Abrechnung, Rechnungsbearbeitung, u.a.) die Geschäftsführung durch eine/n Verwaltungsmitarbeiter/in (gehobener Dienst) unterstützt werden.

Nach wie vor sind die Verhandler in den einzelnen Einrichtungen angesiedelt. Nur dort ist die notwendige, fachspezifische Expertise zu finden, die zur Steuerung des tatsächlichen Bedarfs erforderlich ist. Die unter Kap. 3.3 beschriebenen Verfahren zur nachfrageorientierten Bedarfserhebung haben sich bewährt und sollen weiter genutzt werden.

### **Das Konsortium als Entscheidungsgremium**

Das Konsortium ist die Gruppe aller teilnehmenden Einrichtungen und seiner Verhandler. Im Konsortium soll jeder Hochschultyp durch mindestens eine Vertretung repräsentiert sein – auch wenn diese Person nicht als Verhandler auftritt. Diese Organisationsform soll zweierlei sicherstellen: Zum einen müssen alle Hochschultypen angemessen vertreten werden und zum anderen brauchen die engagierten Personen und Einrichtungen ein Forum zum Erfahrungsaustausch.

Das Konsortium entscheidet über das Produktportfolio. Dieses wird von den Verhandlern im Konsortium jährlich neu zusammengestellt und mit den Anbietern/Verlagen ausgehandelt. Welche Produkte dabei in die zentrale Förderung aufgenommen werden, entscheidet das Konsortium auch unter Zugrundelegung der Grundsätze der DFG, beschrieben in „Grundsätze für den Erwerb DFG-geförderter überregionaler Lizenzen“<sup>10</sup>.

Die Verhandlungsführer sind die erfahrenen Einkäufer in den Bibliotheken. Ein Rahmenangebot schreibt die Konditionen fest. Die Konditionen werden den interessierten Einrichtungen transparent gemacht.

### **Verstärkte Positionierung im nationalen Versorgungskontext**

Das Konsortium Baden-Württemberg soll sich künftig zu einer noch leistungsfähigeren Organisation zur Unterstützung der wissenschaftlichen Bibliotheken des Landes Baden-Württemberg entwickeln und möglichst alle benötigten und auf konsortialer Ebene verfügbaren Angebote koordinieren und vertraglich absichern. Die Vertragsverhandlungen werden künftig verstärkt auch (Nach-)Nutzungsmöglichkeiten in Virtuellen Forschungsumgebungen und Open Access-Komponenten beinhalten (vgl. **Anhänge 9 und 10**).

Die Stärkung und Professionalisierung der baden-württembergischen Konsortialstrukturen dient nicht nur der Verbesserung der wissenschaftlichen Informationsversorgung im Bundesland,

## **6.2 NEUES KONSORTIALMODELL SOFTWARELIZENZEN**

Die wirtschaftlichen und organisatorischen Vorteile der konsortialen Softwarebeschaffung wurden in Kap. Software erläutert. Es ist aus den dort beschriebenen Gründen sinnvoll und angestrebt, die konsortiale Erwerbung von Softwarelizenzen weiter auszubauen und organisatorisch zu gliedern. Hierbei ist eine Bündelung mit dem bestehenden Konsortium für E-Medien naheliegend und sinnvoll.

sondern hat auch zum Ziel, die Rolle des Konsortiums im Hinblick auf die nationale Versorgungsebene weiter zu verstärken. Baden-Württemberg ist eine der bedeutendsten Hochschulregionen in Deutschland und verfügt über zahlreiche innovative Bibliotheken und über ein leistungsfähiges Konsortium, das im Zuge der bundesweiten Arbeitsteilung der GASCO-Konsortien auch überregionale Aufgaben übernimmt.

Dennoch ist das Land in den DFG-geförderten Beschaffungsstrukturen bisher weitgehend unterrepräsentiert: Zwar entsendet z. B. die Universitätsbibliothek Heidelberg seit 2012 eine Vertreterin in die AG Nationale Lizenzierung der Schwerpunktinitiative „Digitale Information“ der Allianz der deutschen Wissenschaftsorganisationen, aber hinsichtlich der Verhandlungsführung bei den Allianzlizenzen sollte das Konsortium Baden-Württemberg zukünftig eine deutlich aktivere Rolle spielen. Ein schlagkräftiges Konsortium mit einer professionellen Geschäftsstelle sollte das Ziel haben, die Interessen der Wissenschaftsregion Baden-Württemberg in Deutschland durch ein verstärktes Engagement in den sich neu herausbildenden kooperativen Erwerbungsstrukturen zu wahren und einen der Bedeutung des Landes angemessenen Beitrag zu leisten.

Die unter Software beschriebenen Schritte (1) und (2) sind individuell und unabhängig vom vergleichbaren Prozess bei den E-Medien zu sehen, welche ja ebenfalls die Verhandlungsphase dezentral strukturieren; dies gilt noch stärker für eine Begutachtung durch die DFG in Schritt (3). Zu (4) gehört als wichtiger Bestandteil die juristische Prüfung der teilwei-

se sehr umfangreichen und international verfassten Vertragsentwürfe; hier soll eine Koordination mit der bestehenden Infrastruktur für E-Medien angestrebt werden, primär unter Nutzung der dort vorhandenen Kompetenz, unter besonderen Bedingungen (erhöhte Auslastung, spezielle Fragestellungen) ist eine externe juristische Beratung zu erwägen. Die Abrechnung der Lizenzen in Schritt (6) erfolgt durch die federführende Einrichtung, eine weitere Konzentration ist zu prüfen. Der Schritt (7) erfolgt sehr produktspezifisch und ist unabhängig vom Vorgehen bei den E-Medien. Eine verstärkte Kooperati-

on ist in Schritt (5) gefordert: der – häufig EU-weiten – Ausschreibung. Seitens des bestehenden Konsortiums für E-Medien besteht hierfür keine Notwendigkeit, weshalb auch keine Infrastruktur vorgehalten wird. Bisher erfolgte diese Arbeit durch diejenige Universitätsverwaltung, deren Rechen- oder Informationszentrum die Verhandlungen geführt hat. Künftig sollte die Aufgabe dauerhaft an einer Stelle – an einer Beschaffungsabteilung einer Landesuniversität – gebündelt werden, um dort spezielle Kompetenz zu vertiefen und nicht einzelne Universitäten zu stark zu belasten. Die Zahl der Ausschreibungen ist nicht sehr hoch, so dass hierfür eine anteilmäßige Entlastung im Umfang von 25 % einer Stelle E11 veranschlagt wird.

## 7 · KONKRETER HANDLUNGSBEDARF (2013-2017) UND EMPFEHLUNGEN AN...

### 7.1 ... DIE UNIVERSITÄTEN, HOCHSCHULEN (LRKs) UND LANDESBIBLIOTHEKEN

Die vorgenannten Ziele sind nur erreichbar, wenn sie auch von den Universitäts- und Hochschulleitungen konsequent mitgetragen werden. Die Einrichtungen müssen im Rahmen ihrer Autonomie einen entsprechenden Schwerpunkt setzen und die erforderlichen Mittel für die E-Literatur bereitstellen. So muss vor allem die 50 % Kofinanzierung sichergestellt sein, wobei eine direkte Belastung der Etats der Hochschulbibliotheken zu vermeiden ist. Hierzu kann und sollte auch eine konsequente Konsolidierung der bisher noch zweischichtigen Bibliothekssysteme der Universitäten hin zu einem zentral an der zentralen Universitätsbibliothek angesiedelten Budget für digitale Inhalte beitragen.

Die Universitäten, Hochschulen und Landesbibliotheken sollten darüber hinaus in ihrer Beschaffungspolitik den konsequenten Umstieg auf „e-only“, wo immer entsprechende Produkte ohne qualitative Einschränkungen nachhaltig angeboten werden, vorantreiben und Doppelbeschaffungen in gedruckter und elektronischer Form vermeiden.

Die neue Konstruktion des Konsortiums basiert auf einer Arbeitsteilung in den teilnehmenden Einrichtungen und erfordert das Engagement der dort beschäftigten Mitarbeiterinnen und Mitarbeiter bei der Verhandlungsführung mit den Verlagen und Anbietern. Diese Verhandlungsaufgabe erfolgt im Rahmen des jeweiligen lokalen Dienstvertrages und wird gegenüber dem Konsortium nicht abgerechnet.

Um eine nachhaltige Zusammenarbeit auf dem Sektor der Beschaffung von elektronischen Publikationen sicherzustellen, wird empfohlen, diese verstärkte Kooperation und auch die damit zusammenhängende Kofinanzierung in den Struktur- und Entwicklungsplänen sowie den Medienentwicklungsplänen der Einrichtungen festzuschreiben.

Eine exzellente Informations- und IT-Infrastruktur für Forschung, Lehre und Studium muss den Forschenden, Lehrenden und Studierenden auf einfachem Wege und kostengünstig die erforderliche Software für die wissenschaftliche Arbeit zur Verfügung stellen. Dabei ist darauf zu achten, dass Software standortübergreifend genutzt werden kann („private cloud“), wie z. B. auf den Rechenclustern der HPC-Landesstrategie (z. B. bwUniCluster, bwForCluster).

Grundlegende Software, die von sehr vielen Anwendern im Land genutzt wird, sollte über Landeslizenzen bereitgestellt werden, da sich hier sowohl wirtschaftliche als auch organisatorische Vorteile ergeben. In diesen Fällen ist eine Finanzierung aus zentralen Mitteln der Hochschulen notwendig. Software, die nur von einem kleineren Teil der Wissenschaftlerinnen, Wissenschaftlern und Studierenden genutzt wird, sollte weiterhin fallweise aus den Mitteln der Hochschulen beschafft werden.

## 7.2 ... DIE LANDESREGIERUNG

Eine ausreichende Versorgung mit elektronischen Fachinformationsangeboten ist durch das Konsortium Baden-Württemberg auf der Basis der gegebenen Ressourcenausstattung trotz der bisherigen Förderung zuzüglich der bereitgestellten Sondermittel des Ministeriums für Wissenschaft, Forschung und Kunst nicht möglich. Allein um die Preissteigerungen auf dem Fachinformationsmarkt und die steigenden Betriebskosten auffangen zu können, ist eine dynamische Erhöhung des Konsortialbudgets von 10 % pro Jahr unbedingt notwendig. In dieser Steigerungsrate ist der Bedarf der Wissenschaft an weiteren, bisher noch nicht lizenzierten Informationen noch nicht enthalten.

Auf Basis der turnusmäßig vom Konsortium durchgeführten Bedarfsanalyse muss das Angebotsportfolio künftig in erheblichem Maße ausgeweitet und um neue Gattungen (z.B. E-Book-Pakete, Open Access Komponenten, Literaturverwaltungstools, Resource Discovery Systems (Indexdienste)) ergänzt werden. Daher sollte ein schrittweiser Mittelaufwuchs für die Grundversorgung auf insgesamt 10 Mio. Euro angestrebt und nachhaltig sichergestellt werden (davon 50 % zentrale und 50 % lokale Mittel). Um eine solche Finanzierungsform künftig haushaltsrechtlich korrekt und organisatorisch effektiv verwalten zu können, muss die Arbeit des Konsortiums professionalisiert werden. Hierzu ist der Aufbau einer Geschäftsstelle mit geeigneter Personalausstattung notwendig (s. Kap. 6.1 Konsortium Baden-Württemberg zur Versorgung mit E-Medien). Als Standort für die Geschäftsstelle bietet sich die Universitätsbibliothek Freiburg an, da sie bisher den Vorsitz im Ent-

Solch klare Rahmenbedingungen verbessern die Möglichkeiten konsortialer Verhandlungen zur Erzielung kostengünstiger Lizenzverträge und damit zur Bereitstellung der erforderlichen Software für Forschung, Lehre und Studium.

scheidergremium des Konsortiums innehatte, sie das Konsortium im Außenverhältnis vertreten hat und so die bisher gewonnenen Erfahrungen wie auch die Nähe zur technischen Betriebseinrichtung ReDI optimal genutzt werden können. Darüber hinaus ist die Ansiedlung der Geschäftsstelle an einer Universitätsbibliothek wegen ihrer Nähe zur Wissenschaft und zu den Fakultäten, Rektoraten und über sie zur LRK, die in einem hohen Maße die wissenschaftsstrategischen und finanziellen Entscheidungen verantworten, geboten (vgl. auch **Anhang 3**).

Schließlich ist eine Konsolidierung der derzeitigen Finanzkapitel wünschenswert. Das Zustandekommen neuer Software-Landeslizenzen bedarf in der Regel einer Anschubfinanzierung durch das MWK für maximal 3 Jahre. Hierfür sind entsprechende jährliche Fördermittel vorzusehen.

## ANHANG 1: DIE VERSORGUNGSSTRUKTUR FÜR ELEKTRONISCHE MEDIEN

### 1. NATIONAL-/ALLIANZLIZENZEN

Die Deutsche Forschungsgemeinschaft (DFG) finanziert seit 2004 den Erwerb von nationalen Lizenzen für elektronische Medien, um die Versorgung an deutschen Hochschulen, Forschungseinrichtungen und wissenschaftlichen Bibliotheken mit elektronischer Fachinformation nachhaltig zu verbessern. Wissenschaftler, Studierende und wissenschaftlich interessierte Privatpersonen können durch diese Lizenzen auf Datenbanken, digitalen Textsammlungen, elektronischen Zeitschriften und E-Books zugreifen. Die DFG hat bereits mehr als 100 Mio. Euro aufgewendet, um abgeschlossene digitale Publikationen, d.h. digitale Text- und Werkausgaben, von wissenschaftlichen Verlagen angebotene Digitalisierungen zurückliegender Zeitschriftenjahrgänge (sog. Backfile-Archive) sowie spezielle Fachdatenbanken dauerhaft zu lizenzieren.

Im Jahr 2008 wurde die Schwerpunktinitiative „Digitale Information“ der Allianz der deutschen Wissenschaftsorganisationen gegründet. Aus dieser Initiative heraus wurden ab 2011 die Nationallizenzen zu „Allianz-Lizenzen“ weiterentwickelt. Der Fokus dieser Allianz-Lizenzen liegt auf dem Erwerb dynamischer Produkte (laufende Zeitschriften, Datenbanken und E-Books) für den wissenschaftlichen Grundbedarf.

Der Beitritt zu den Allianz-Lizenzen setzt eine finanzielle Eigenbeteiligung der Bibliotheken von ca. 75 % voraus. Die Lizenzen werden im Rahmen einer mehrjährigen Vertragslaufzeit als nationale Opt-in-Konsortien angeboten. Während der Zugang zu den aktuellen Jahrgängen nur den jeweiligen Konsortialteilnehmern vorbehalten ist, werden die lizenzierten Inhalte mit einem zeitlichen Verzug (Moving Wall) von in der Regel einem Jahr deutschlandweit und dauerhaft im Sinne einer klassischen Nationallizenz bereitgestellt. Die verhandlungsführenden Einrichtungen reichen Anträge für Allianzlizenzen bei der DFG ein. Die Begutachtung der eingereichten Anträge erfolgt durch Fachwissenschaftlerinnen und Bibliothekare. Die Begutachtung orientiert sich an schriftlich festgelegten Kriterien (AL-Grundsätze, 2010)<sup>11</sup>.

Für National- und Allianzlizenzen gelten Open Access-Regelungen. Autoren aus autorisierten Einrichtungen sind ohne Mehrkosten berechtigt, ihre in den lizenzierten Zeitschriften erschienenen Artikel in der Regel in der durch den Verlag publizierten Form (z. B. PDF) zeitnah in ein Open Access-Repositoryum ihrer Wahl einzupflegen.

(Quellen: DFG-Publikationen und Webseite [www.nationallizenzen.de](http://www.nationallizenzen.de))

### 2. REGIONALE KONSORTIALLIZENZEN

#### a) E-Literatur und Datenbanken

Seit Ende der 1990er Jahre haben sich in fast allen deutschen Bundesländern Konsortialstrukturen für den Erwerb von E-Medien herausgebildet. Diese Konsortien haben sich mit den institutionellen Konsortien aus den außeruniversitären Forschungsverbänden (Max Planck, Fraunhofer, Helmholtz, Leibniz) und mit den österreichischen, Schweizer

und seit 2010 auch Luxemburger Konsortien seit 2000 in der Arbeitsgemeinschaft GASCO – German, Austrian and Swiss Concoctia Organisation zusammengeschlossen. In diesem Kreis gehört das Konsortium Baden Württemberg zu den besonders aktiven und – ebenso wie die Konsortien in Bayern, Berlin-Brandenburg, Hessen und Nordrhein-Westfalen – zu den auch überregional tätigen Konsortien. Über die GASCO werden mit den anderen Konsortien gemeinsame Strategien für die Verhandlungen mit

den Verlagen koordiniert oder – soweit dies praktikabel ist – auch gemeinsame Verhandlungen geführt.

teilweise nehmen nur die Landesuniversitäten teil, teilweise alle staatlich anerkannten Hochschulen.

### b) Software

Die Rechen- und Informationszentren des Landes haben die Vorteile konsortialer Softwarebeschaffungen seit längerem erkannt. Bedarfsgerecht wird jeweils geprüft, ob für die relevanten Produkte eine Konsortiallizenz sinnvoll ist. Je nach Produkt sind die Finanzierungs- und Beteiligungsmodelle dabei stark unterschiedlich;

Im Jahr 2012 bestehen konsortiale Lizenzverträge für alle Microsoft Desktop-Produkte, Anti-Viren-Software, die gängigen Pakete für Statistik und Geoinformationsdaten, Literaturverwaltung, das Bibliotheks-System aStec|aDIS BMS und Backup-Software, weitere sind in Verhandlung.

## ANHANG 2: AM KONSORTIUM BW TEILNEHMENDE EINRICHTUNGEN (STAND: 09/2012)

### Universitäten:

Freiburg, Heidelberg, Hohenheim, Karlsruher Institut für Technologie (KIT), Konstanz, Mannheim, Stuttgart, Tübingen, Ulm

Weingarten), Reutlingen, Rottenburg, Schwäbisch-Gmünd, Stuttgart (HfT), Stuttgart (HdM), Ulm, Villingen-Schwenningen (HfPol),

### Landesbibliotheken:

Badische Landesbibliothek Karlsruhe, Württembergische Landesbibliothek Stuttgart

Duale Hochschule Baden-Württemberg mit den Standorten:

Heidenheim, Karlsruhe, Lörrach, Mannheim, Mosbach, Ravensburg, Stuttgart, Villingen-Schwenningen

### Pädagogische Hochschulen:

Freiburg, Heidelberg, Karlsruhe, Ludwigsburg, Weingarten (mit HS Ravensburg), Schwäbisch Gmünd

### Kunst- und Musikhochschulen:

Hochschule für Musik Freiburg, Staatliche Hochschule für Gestaltung Karlsruhe, Hochschule für Musik Karlsruhe, Hochschule für Musik und Darstellende Kunst Mannheim, Staatliche Hochschule für Musik Trossingen

### Hochschulen für Angewandte Wissenschaften:

Aalen, Albstadt-Sigmaringen, Biberach, Esslingen, Furtwangen, Heilbronn, Karlsruhe, Kehl, Konstanz, Ludwigsburg, Mannheim, Nürtingen-Geislingen, Offenburg, Pforzheim, Ravensburg (mit PH

## ANHANG 3: ALTERNATIVE KONSORTIALSTRUKTUREN

Das Konsortium Baden Württemberg sticht im Vergleich mit den weiteren führenden deutschen Konsortien durch seine besonders kooperativ angelegte Arbeitsweise mit verteilter Verhandlungsführung hervor, die zu einem in der Region besonders weit verbreiteten konsortialen Know-how führt. Damit unterscheidet es sich von Konsortien mit zentralisierter Verhandlungsführung, wie es etwa

bei den an der Bayerischen Staatsbibliothek, der hessischen Verbundzentrale oder dem Hochschulbibliothekszentrum NRW organisierten Konsortialstellen der Fall ist. Eher mit Baden-Württemberg vergleichbar ist unter den großen Konsortien noch das als eingetragener Verein organisierte Friedrich-

Althoff-Konsortium für Berlin-Brandenburg, das zwar über ein am Konrad-Zuse-Zentrum angesiedeltes hauptamtliches Back Office verfügt, aber ebenfalls eine unter den beteiligten Bibliotheken verteilte Verhandlungsführung praktiziert.

Geschäftsstellen mit eigenem Personal sind für eine schlagkräftige Organisation des Konsortiums unverzichtbar: Sie ermöglichen schnelles Reagieren und den Aufbau von Spezialwissen. Wichtig ist jedoch die feste Verankerung im Gefüge der Konsortialteilnehmer und ein Selbstverständnis als Dienstleister: Dies garantiert, dass die Geschäftsstelle sich weder zu stark von dem alltäglichen Informationsbedarf vor Ort entfernt, noch ein zu starkes Eigenleben mit eigenen Interessen entwickelt. Diese Gefahren bestehen insbesondere, wenn die Geschäftsstelle selbst für den Zuschnitt und die Verhandlung der einzelnen Angebote zuständig ist. Erfahrungen aus anderen Bundesländern zeigen, dass Geschäftsstellen, die a) institutionell weitgehend entkoppelt sind

(d.h. nicht an einer Hochschule oder einer Bibliothek mit eigenem Bestand und eigenen Nutzern angesiedelt) oder die b) an einer übermächtigen Zentralinstanz angesiedelt sind, regionale Gremien und kooperative Arbeitsstrukturen zur Seite gestellt werden sollten, um eine strikte Ausrichtung an den Nutzererwartungen, eine konsequente Qualitätssicherung und die Integration in das regionale Gesamtsystem zu gewährleisten. Dies gilt besonders für Konsortien, die über Zentralmittel für den Erwerb von Lizenzen verfügen. Auch wird dadurch eine Überlastung der jeweiligen Konsortialstelle vermieden, die ohnehin nicht in der Lage ist, die informationelle Bedürfnislage an einzelnen Institutionen detailliert zu analysieren – dies ist und bleibt Aufgabe der dafür vorgesehenen lokalen Informationseinheiten (Bibliotheken, Medien- und Rechenzentren) als Dienstleister für die Endnutzer. Hauptaufgabe der Konsortialstelle ist deshalb in den genannten Beispielen stets die transparente, möglichst softwaregestützte Erhebung des regionalen Bedarfs, das Management des gesamten konsortialen life cycles, sowie die Schnittstellenfunktion zu anderen überregionalen oder nationalen Strukturen. Beispiel für letzteres wäre etwa die Aggregatorenfunktion der regionalen Geschäftsstellen für die Bedarfserhebung im Hinblick auf Allianzlizenzen.

#### ANHANG 4: TECHNISCHE INFRASTRUKTUR DES KONSORTIUMS

Die technische Infrastruktur des Konsortiums wird über das Portalsystem ReDI (Regionale Datenbankinformation) bereitgestellt, das an den beiden Standorten UB Freiburg und UB Stuttgart angesiedelt ist. An diesen Standorten ist das Know-how zum Betrieb komplexer Angebotsformen von elektronischer Information gebündelt. ReDI ist der Dienst, der im Auftrag des Konsortiums Baden-Württemberg die Lösung der technischen Fragen zum Angebot von heute fast 1000 Informationsquellen übernimmt. Die Arbeitsgruppe ReDI setzt sich aus Mitarbeitern der UB Freiburg und UB Stuttgart zusammen. Die Leitung der Arbeitsgruppe liegt bei der UB Freiburg.

Die Aufgaben von ReDI umfassen:

- Vermittlung externer Angebote von Verlagsservern
- Bereitstellung CD-basierter Angebote auf eigenen Serversystemen
- Verwaltung aller DB-Lizenzen in Baden-Württemberg
- Projektarbeit mit anderen bibliothekarischen Einrichtungen in und außerhalb Baden-Württembergs

Die Zusammenfassung der technischen Koordination zum Angebot der lizenzierten Fachinformationen in einer zentralen Stelle bei ReDI bedeutet erhebliche Synergien und eine höhere Transparenz bei den Lizenzen und der Nutzung. Bis 1999 hat jede einzelne Einrichtung mit eigenen Ressourcen (Personal, Serversysteme) die Angebote organisiert. Heute wird dies zentral von ReDI für über 50 Einrichtungen des Landes Baden-Württemberg mit zwei Stel-

len durchgeführt. Das seit 1999 aufgebaute zentrale Managementsystem für elektronische Lizenzen (ERM) stellt allen Bibliotheken des Landes Informationen zur Lizenzierung und ggf. Kosten zur Verfügung. Ergänzt wird diese zentrale Dienstleistung von ReDI noch durch die zentral zur Verfügung gestellten Statistikdaten, die es den teilnehmenden Einrichtungen ermöglichen, Nutzungsanalysen und damit bedarfsgesteuerte Erwerbung durchzuführen.

Die ReDI-Gruppe betreut das System zur Verwaltung der Lizenzen<sup>12</sup> (ERM) und organisiert die statistische Auswertung<sup>13</sup> der Nutzung. Die Zusammenfassung der technischen Kompetenz für das Angebot inkl. Lizenzkontrolle, Authentifizierung und Link-Resolving hat sich in der Vergangenheit bewährt und soll fortgeführt werden.

Die Arbeitsgruppe ReDI bleibt unverändert erhalten und garantiert die lizenzgerechte technische Bereitstellung der Produkte.

## ANHANG 5: ETATENTWICKLUNG DES KONSORTIUMS BADEN-WÜRTTEMBERG

Tabelle 1: Etatentwicklung des Konsortiums von 2006 - 2012 (nur Kap. 1407)

| JAHR | JÄHRLICHE ZUWEISUNG (NETTO*): EURO | V&B-KOSTEN <sup>14</sup> EURO | EINNAHMEN AUS CD-SERVICE-VERTRÄGEN EURO | VERFÜGBARE MITTEL EURO | DECKUNGS-LÜCKE EURO | EIGENANTEIL IN % |
|------|------------------------------------|-------------------------------|---|------------------------|---------------------|------------------|
| 2012 | 1.072.100                          | 121.550                       | 37.520                                  | 988.070                | 623.448             | 39               |
| 2011 | 1.002.100                          | 111.929                       | 40.180                                  | 930.351                | 267.728             | 22               |
| 2010 | 1.002.100                          | 125.031                       | 32.040                                  | 909.109                | 410.093             | 31               |
| 2009 | 1.002.100                          | 93.854                        | 21.700                                  | 929.946                | 342.035             | 27               |
| 2008 | 1.002.100                          | 73.865                        | 18.900                                  | 947.135                | 220.857             | 19               |
| 2007 | 1.002.100                          | 94.818                        | 3.360                                   | 910.642                | 237.411             | 21               |
| 2006 | 1.002.100                          | 66.738                        | 0                                       | 935.362                | 148.170             | 14               |

\* ab 2009 unter Berücksichtigung des Einnahmesolls in Höhe von 103.500 Euro

Ergänzend dazu hat das MWK seit 2006 immer wieder Sondermittel bereitgestellt, die aber nur einzelne Bibliothekssparten bei der Beschaffung von elektronischen Zeitschriften unterstützt haben.

Tabelle 2 zeigt die Entwicklung dieser Förderung seit 2006:

Tabelle 2: Sondermittel des Konsortiums von 2006 - 2012 (nur Kap. 1499)

| JAHR | VERWENDBAR FÜR EURO |     |      |    |    | MWK-ZUWEISUNGEN EURO | EIGEN-BETEILIGUNG EURO | VERFÜGBARE SUMME EURO |
|------|---------------------|-----|------|----|----|----------------------|------------------------|-----------------------|
|      | UB                  | HAW | DHBW | PH | LB |                      |                        |                       |
| 2012 | X                   |     |      |    | X  | 440.000              | 440.000                | 880.000               |
| 2011 | X                   |     |      |    | X  | 440.000              | 649.000                | 1.089.000             |
| 2010 | X                   |     |      |    | X  | 440.000              | 649.000                | 1.089.000             |
| 2009 |                     | X   |      |    |    | 0                    | 0                      | 0                     |
| 2008 | X                   | X   |      | X  |    | 592.000              | 0                      | 592.000               |
| 2007 | X                   | X   |      | X  |    | 588.500              | 0                      | 588.500               |
| 2006 | X                   |     |      | X  |    | 536.000              | 0                      | 536.000               |

## ANHANG 6: AUSGABEN DER WISSENSCHAFTLICHEN BIBLIOTHEKEN FÜR DIGITALE INHALTE

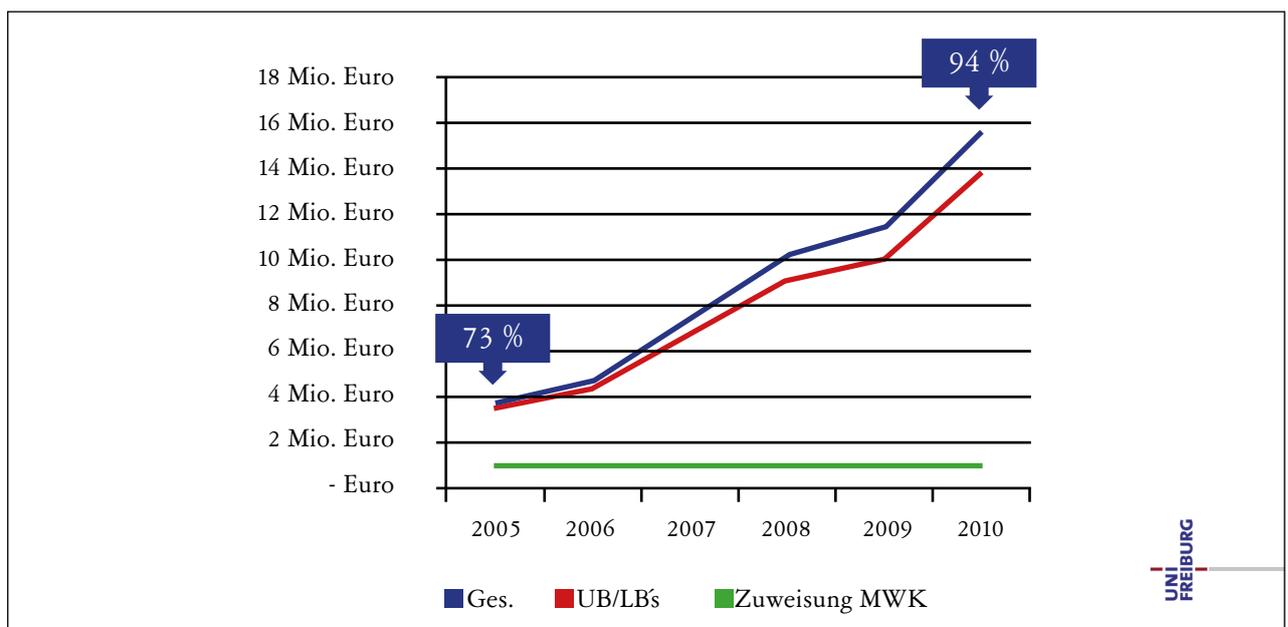


Abb. 2: Verhältnis von zentraler Förderung und dezentraler Finanzierung elektronischer Lizenzen in Baden-Württemberg: Die Zuweisung aus zentralen Haushaltsmitteln des MWK an das Konsortium Baden-Württemberg ist in Relation zu den gesamten Erwerbungsmitgliedern des Konsortiums von 27 % im Jahr 2005 auf 6 % im Jahr 2010 gesunken. (Diagramm: Antje Kellersohn)

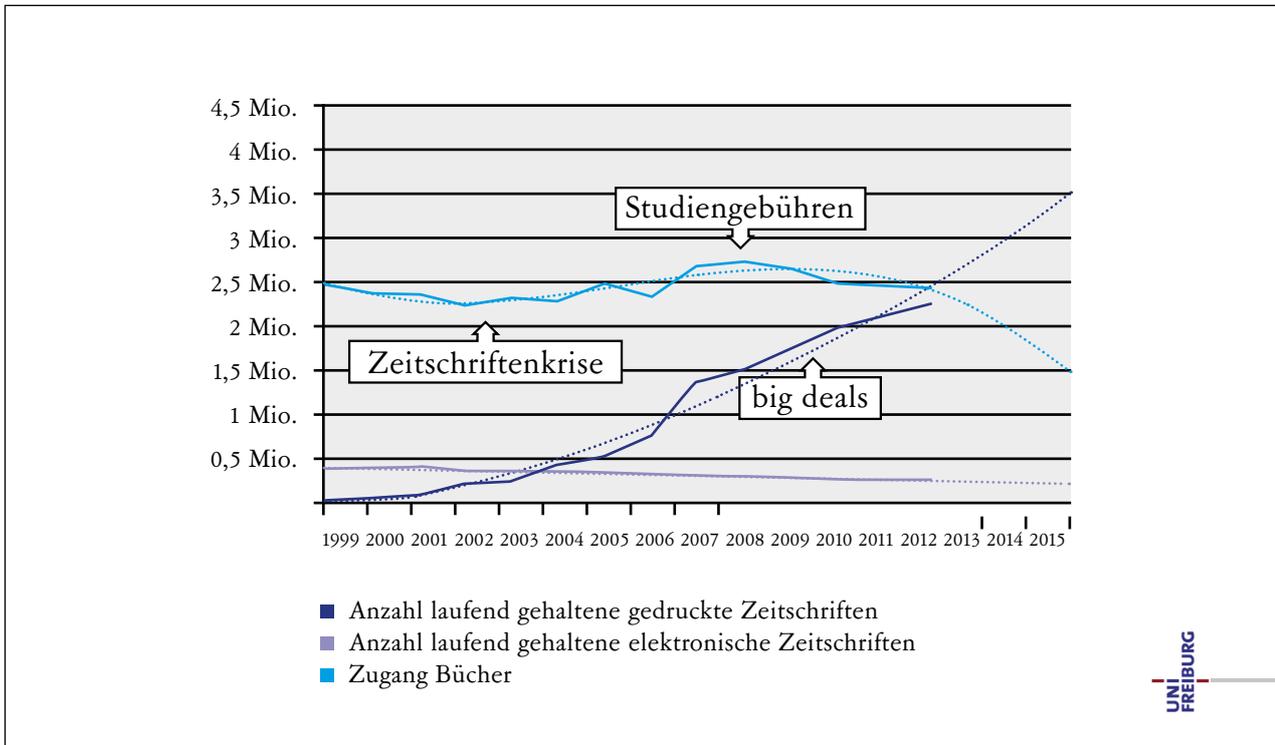


Abb. 3: Zeitschriften und Monografien-Zugang der Universitätsbibliotheken in Deutschland. (Quelle: Deutsche Bibliotheksstatistik. Diagramm: Antje Kellersohn)

**ANHANG 7: ÜBERSICHT LANDES-SOFTWARELIZENZEN**

| PRODUKTE                      | TEILNEHMER  |
|-------------------------------|---|
| Microsoft Desktop core        | alle staatl. anerkannten Hochschulen  |
| Microsoft additional products | alle staatl. anerkannten Hochschulen  |
| ESRI                          | einzelne Hochschulen  |
| aStec aDIS/BMS                | 5 Universitäten, alle Pädagogischen Hochschulen, alle HAWen, Duale Hochschule Baden-Württemberg, beide Landesbibliotheken |
| IBM Tivoli                    | Universitäten   |
| IBM SPSS                      | ausgewählte Hochschulen   |
| McAfee Antivirus              | ausgewählte Hochschulen   |

Eine Mitfinanzierung des Landes besteht für ESRI und IBM Tivoli im Umfang von 33 % für die ersten drei Jahre der Vertragslaufzeit.

## ANHANG 8: QUALITÄTSSICHERUNGSMASSNAHMEN BEI VERGABE ZENTRALER MITTEL

Nutzungsstatistiken sind ein unverzichtbares Instrument zur Kontrolle, ob die Lizenzierung eines Produktes ziel- und bedarfsgerecht war. Die Auswertung von Nutzungszahlen trägt daher sehr stark zum bedarfsgerechten Einsatz von Haushaltsmitteln bei. Daher hat ReDI in den Jahren 2011/2012 gemeinsam mit dem Hessischen Verbund (HeBIS) einen Statistikserver entwickelt, der automatisch die Nutzungsdaten elektronischer Angebote in einem standardisierten Verfahren von den Verlagsservern abrufen, aufbereiten und den Einrichtungen zur Verfügung stellt. Diese können dann produktgenau ihre lokalen Nutzungszahlen analysieren und *cost per download* Berechnungen anstellen. Innerhalb des Konsortiums kann dann eine Auswertung vorgenommen werden, die gegebenenfalls Auswirkungen auf die künftige Kostenverteilung haben oder dazu führen kann, dass ein bestimmtes Produkt stärker, weniger stark oder gar nicht mehr zentral gefördert wird. Durch die Orientierung an internationalen Standards wie COUNTER<sup>15</sup> besteht eine annäher-

nde Vergleichbarkeit von digitalen Angeboten, die dem Konsortium auch eine vertiefte Marktanalyse ermöglicht.

Während die infrastrukturellen Grundlagen für eine verstärkte Integration von statistischen Analysen in die Qualitätssicherung bei der Konsortialarbeit durch die Entwicklung des REDI-Statistikportals gemeinsam mit HeBIS aktuell bereits gelegt sind, so ist noch sicherzustellen, dass dem Konsortium ein Beirat als beratendes Gremium zur Verfügung steht, der die strategische Ausrichtung des Konsortiums überprüft. Dabei soll es nicht um einzelne Lizenzierungsentscheidungen gehen, sondern um eine Überprüfung der Aktivitäten des Konsortiums darauf hin, inwieweit diese den in der E-Science-Initiative des Landes formulierten übergeordneten Zielen entsprechen und ob sie geeignet sind, das Standing des Konsortiums innerhalb des größeren Zusammenhanges der neu entstehenden bundesweiten Informationsinfrastruktur zu stärken. Ein entsprechender Beirat müsste neben dem MWK sowohl Personen umfassen, die Expertise im Bereich der praktischen Konsortialarbeit im deutschsprachigen Raum mitbringen, als auch solche, die einen guten Überblick über die wissenschaftliche Informationsversorgung in Baden-Württemberg und in Deutschland haben.

## ANHANG 9: SCHNITTSTELLE ZU VIRTUELLEN FORSCHUNGSUMGEBUNGEN

Digital vorliegende Publikationen sind prädestiniert, im Rahmen Virtueller Forschungsumgebungen von einzelnen Wissenschaftlerinnen und Wissenschaftlern, Projektteams oder der Fach-Community über die reine Lese- und Ausdruckfunktion hinaus für die wissenschaftliche (Zusammen-)Arbeit genutzt zu werden. Solche Nutzungsszenarien umfassen die Setzung von Lesezeichen und Farbmarkierung, Annotierung und Kommentierung, die privat, projektintern oder Community-weit wiederverwendet

bzw. bekannt gemacht werden und als Basis des wissenschaftlichen Austauschs im Sinne des Web 2.0 dienen. Maßnahmen, die zum Aufbau der dafür notwendigen Infrastruktur führen können, sind Gegenstand der AG Virtuelle Forschungsumgebungen.

Das Konsortium BW wird in der Zukunft verstärkt in seinen Lizenzverträgen mit den Anbietern für E-Medien darauf hinarbeiten, dass möglichst weitreichende (Nach-)Nutzungsmöglichkeiten für die lizenzierten Inhalte in Virtuellen Forschungsumgebungen – und auch Lehr- und Lernplattformen – vereinbart werden.

## ANHANG 10: TRANSFORMATIONSPROZESS VON SUBSKRIPTIONS- ZU OPEN ACCESS-MODELLEN

Der Transformationsprozess zu OA-Modellen ist vor einigen Jahren angelaufen. Mit speziellen Förderprogrammen fördert z. B. die DFG (z. B. „Open Access Publizieren“) gezielt diese Entwicklung. Dennoch liegt der Anteil von OA-Publikation heute noch im einstelligen %-Bereich aller Publikationen. Einzelne Fachrichtungen – insbesondere die Geisteswissenschaften – agieren noch sehr zurückhaltend, andere – vor allem die naturwissenschaftlichen Fächer und die Medizin – setzen in zunehmendem Maße auf OA-Publikationen. Langfristig ist die Wirkung auf die Kostenentwicklung zur Beschaffung elektronischer Informationen nach jetziger Einschätzung durchaus positiv zu bewerten.

Allerdings ist das Einsparpotenzial, wenn es denn tatsächlich so ausfällt, nur schrittweise und mit Verzögerung zu erreichen. Vielmehr führt der Umstieg zum Open Access zunächst über Mehrkosten, da die OA-Publikationen in den ersten Jahren nicht direkt

zu Abbestellungen von Zeitschriften führen können und so anfangs Doppelfinanzierungen für die wissenschaftlichen Bibliotheken unumgänglich sind.

Daher sieht der beantragte Mittelaufwuchs für E-Medien auch einen Betrag für die Finanzierung von Kosten für Open Access im Grünen und Goldenen Weg (Publikationsfonds, Mitgliedschaften, ggf. auch Freikauf von kommerziellen Publikationen) zur komplementären Finanzierung zu den Förderungen anderer Organisationen, wie z. B. der DFG, vor.

Die Mittelbewirtschaftung für Projekte des Landes zur Förderung des Umstiegs auf Open Access könnte – analog zur Vorgehensweise der DFG – beim Konsortium BW angesiedelt werden.

<sup>4</sup> zur Übersicht Ansprechpartner des Konsortiums <http://www.konsortium-bw.de/a020/start> (abgerufen am 01.03.2013)

<sup>5</sup> Besonders extreme Beispiele der Preissteigerung sind (dargestellt am Preis der Universität Freiburg):

· Beck-Online (Jura): von 6.336 Euro im Jahr 2005 auf 16.660 Euro im Jahr 2012 (160 % in 7 Jahren)

· REAXYS (Chemie): von 20.210 Euro im Jahr 2005 auf 42.138 Euro im Jahr 2012 (100 % in 7 Jahren)

<sup>6</sup> Science, Technology, Medicine, also Naturwissenschaften, Ingenieurwissenschaften, Medizin)

<sup>7</sup> Arbeitskreis der Leiter der wissenschaftlichen Rechenzentren von Baden-Württemberg (ALWR-BW), Strategiepapier „IT an den Universitäten Baden-Württembergs“, August 2010, S.7

<sup>8</sup> Vgl. Kommission Zukunft der Informationsinfrastruktur: Gesamtkonzept für die Informationsinfrastruktur in Deutschland, Seite B19

<sup>9</sup> [http://www.dfg.de/foerderung/info\\_wissenschaft/archiv/2012/info\\_wissenschaft\\_12\\_55/index.html](http://www.dfg.de/foerderung/info_wissenschaft/archiv/2012/info_wissenschaft_12_55/index.html) (abgerufen am 01.09.2013),

[http://www.dfg.de/foerderung/info\\_wissenschaft/info\\_wissenschaft\\_13\\_49/](http://www.dfg.de/foerderung/info_wissenschaft/info_wissenschaft_13_49/) (abgerufen am 01.09.2013)

<sup>10</sup> [http://www.dfg.de/formulare/12\\_18/12\\_18.pdf](http://www.dfg.de/formulare/12_18/12_18.pdf) (abgerufen am 01.03.2013)

<sup>11</sup> [http://www.dfg.de/formulare/12\\_18/12\\_18.pdf](http://www.dfg.de/formulare/12_18/12_18.pdf) (abgerufen am 01.03.2013)

<sup>12</sup> Die Arbeitsgruppe ReDI beteiligt sich an der Ausschreibung der DFG zur Entwicklung und des dauerhaften Betriebs zum Themenfeld 3 – Electronic Resource Management – das Management digitaler Publikationen (ERM) gemeinsam mit dem Hessischen Verbund (HeBIS), und den Verbundzentralen von Nordrhein-Westfalen (HBZ) und Niedersachsen (VZG).

<sup>13</sup> ReDI hat in den Jahren 2011 und 2012 gemeinsam mit dem Hessischen Verbund (HeBIS) ein Statistikserversystem entwickelt, das die Nutzungsdaten nach standardisierten Verfahren von den Anbietern/Verlagen abrufen und zur Auswertung aufbereitet.

<sup>14</sup> Verwaltungs- und Betriebskosten ergeben sich aus den Personalkosten der Universität Freiburg für den Serverbetrieb, den Sachkosten für den Serverbetrieb und den Projektkosten (ab 2010).

<sup>15</sup> COUNTER steht für Counting Online Usage of Networked Electronic Resources und ist ein internationales, von Bibliotheken, Konsortien, Händlern und Verlagen gemeinsam getragenes Projekt ([www.projectcounter.org](http://www.projectcounter.org)), das mittlerweile den Goldstandard für die Nutzungsmessung von digitalen Inhalten darstellt.



# AG 2 · Digitalisierung

## INHALT

|   |    |
|---|----|
| 1. Nutzen der Digitalisierung für die Forschung .....   | 49 |
| 2. Inhaltliche Eingrenzung .....  | 49 |
| 3. Spartenübergreifende Digitalisierung .....   | 50 |
| 4. Bestehende Verpflichtungen des Landes im Bereich Retrodigitalisierung .....  | 50 |
| 5. Status quo der Retrodigitalisierung in den Archiven, Bibliotheken, Museen,<br>universitären Sammlungen und Filmeinrichtungen im Geschäftsbereich des Wissenschaftsministeriums ..... | 50 |
| 6. Zukünftige Digitalisierungsstrategie .....   | 52 |
| 6.1 Spartenspezifische Digitalisierungszentren .....  | 52 |
| 6.2 Scan-Dienstleistungen .....   | 52 |
| 6.3 Beratungsdienstleistungen .....   | 53 |
| 6.4 Kriterien bei der Mittelvergabe .....   | 54 |
| 6.5 Regionale und überregionale Koordinierung von Digitalisierungsprojekten .....   | 54 |
| 7. Derzeit vorhandene Mittel für die Digitalisierung .....  | 54 |
| 8. Mittelbedarf .....   | 55 |
| 9. Public-Private-Partnership-Projekte .....  | 55 |
| 10. Nutzung und Verwertung .....  | 56 |
| 11. Evaluierung der Digitalisierungsstrategie .....   | 56 |
| Anhang .....  | 56 |

## MITGLIEDER DER ARBEITSGRUPPE

**DR. URSULA BERNHARDT**

Ministerium für Wissenschaft, Forschung und Kunst  
Baden-Württemberg (Vorsitz)

**PETER CASTELLAZ**

Ministerium für Wissenschaft, Forschung und Kunst  
Baden-Württemberg

**DR. JUTTA DRESCH**

Badisches Landesmuseum

**DR. HANNSJÖRG KOWARK**

Württembergische Landesbibliothek

**DR. GERALD MAIER**

Landesarchiv Baden-Württemberg

**DR. VEIT PROBST**

Bibliothek, Universität Heidelberg

**PROF. DR. DIETER SPECK**

Universitätsarchiv/Uniseum, Universität Freiburg

**STEFAN WOLF**

Bibliotheksservice-Zentrum Baden-Württemberg, Konstanz

**DR. REINER ZIEGLER**

Haus des Dokumentarfilms Stuttgart

## 1 · NUTZEN DER DIGITALISIERUNG FÜR DIE FORSCHUNG

Die digitale Transformation bzw. Retrodigitalisierung für die Forschung relevanter analoger Bestände der wissenschaftlichen und kulturellen Überlieferung gewinnt seit Jahren insbesondere für die historisch arbeitenden Wissenschaftler al-

ler Disziplinen zunehmend an Bedeutung: Digitalisate als Teil Virtueller Forschungsumgebungen ermöglichen es Wissenschaft und Forschung, zeit- und ortsunabhängig mit dem kulturellen historischen Erbe zu arbeiten und dank der internationalen Vernetzung dieses Erbe immer wieder neu zu kontextualisieren.

*„Die Qualität der Forschungsergebnisse würde sich dadurch [i. e. Ausweitung der Digitalisierung von Archivgut] weiter verbessern lassen. Die Nutzung von digitalisiertem Archivgut wäre im Interesse der Qualitätssicherung und weiteren Profilierung der universitären Forschung und Lehre nicht nur am Wissenschaftsstandort Baden-Württemberg, sondern insgesamt in der globalisierten Wissenslandschaft wünschenswert, ja letztlich unverzichtbar.“*

*Angela Borgstedt (Universität Mannheim) in einer gutachterlichen Stellungnahme für die AG Digitalisierung der AG E-Science vom 29.8.2012.*

Digitalisate sind ein wesentlicher Beitrag zur Beschleunigung der Forschung, zur Setzung neuer Forschungsimpulse und zur Erschließung des Materials. Bislang unbekanntes oder nur schwer zu-

gängliches Material kann der Forschung sichtbar und zugänglich gemacht werden. Die Ersparnis von Reise- und Kopierkosten ist ein ebenso zu begrüßender Nebeneffekt wie die Schonung der Originalbestände.

*Dies ermöglicht die Einsichtnahme von Quellen, die in früheren Zeiten vielleicht nicht berücksichtigt worden wären oder von deren Existenz man nur durch Zufall erfahren hätte.“*

*Christine Sievert: Wissenschaftliches Arbeiten im Zeitalter digitaler Quellen, in ZfBB 59 (2012), S. 158-165, 158*

Förderlich für die starke Zunahme der Digitalisierungsprojekte der letzten Jahre war die Schaffung der DFG-Praxisregeln aus dem Jahr 2009<sup>16</sup>, die für einheitliche technische Standards sorgen, so dass Austausch, Weitergabe und Nachweis der digitalisierten Dokumente und Objekte über definierte technische Schnittstellen problemlos

möglich ist. Eine Aktualisierung der DFG-Praxisregeln ist 2013 erfolgt. Die Bedeutung der Retrodigitalisierung für eine innovative Informationsinfrastruktur wurde von der Deutschen Forschungsgemeinschaft, dem Wissenschaftsrat und der Kommission Zukunft der Informationsinfrastruktur wiederholt hervorgehoben<sup>17</sup> und mit Empfehlungen zur finanziellen Förderung verbunden.

## 2 · INHALTLICHE EINGRENZUNG

Angeht die Fülle forschungsrelevanter, nachgefragter Dokumente und Objekte in den Archiven, Bibliotheken, Museen, universitären Sammlungen und Filmeinrichtungen<sup>18</sup> im Geschäftsbereich des Ministeriums für Wissen-

schaft, Forschung und Kunst Baden-Württemberg sowie aus urheberrechtlichen Gründen konzentrieren sich die spartenübergreifende Digitalisierungsstrategie und das Digitalisierungsprogramm des Landes vorerst auf das kulturelle historische Erbe im urheberrechtsfreien Raum.

### 3 · SPARTENÜBERGREIFENDE DIGITALISIERUNG

Der spartenübergreifende Ansatz der Digitalisierungsstrategie der Landesbibliothek und der Landesarchivverwaltung der Forschung des Landes ergibt sich aus den veränderten Informationsanfor-

*„Mittelfristig werden aber alle Wissenschaftler und Wissenschaftlerinnen erwarten – seien es Geistes-, Sozial- oder Naturwissenschaftler, dass sie Unterstützung beim digitalen Zugang zu Publikationen, bei der fachlichen Erschließung von Informationsräumen sowie der Verwaltung und Vernetzung von Forschungsdaten und elektronischen Lehrmodulen erhalten. Voraussetzung dafür werden überregional integrierte digitale Informationsumgebungen sein, die standortübergreifend den Zugriff auf ein breites Spektrum aktueller Forschungsliteratur, digitalisierter Forschungsprimärdaten (z. B. historische Bibliotheksbestände, Archivalien, Objekte etc.), E-Learning Materialien und virtuelle Forschungsverbünde ermöglichen. Daher müssen sich die heute noch weitgehend getrennt operierenden Bibliotheken, Archive und Fachinformationseinrichtungen mit überregionaler Ausstrahlung zu einem kohärenten Gesamtsystem der digitalen Informationsversorgung für die Wissenschaft vernetzen. ... Bibliotheken, Archive und Museen werden zu Trägern des E-Science Gedankens.“*

*Hans-Jochen Schiewer (Universität Freiburg) in einer gutachterlichen Stellungnahme für die AG Digitalisierung der AG E-Science am 10.9.2012*

### 4 · BESTEHENDE VERPFLICHTUNGEN DES LANDES IM BEREICH RETRODIGITALISIERUNG

Da derzeit noch zu wenig Zeugnisse des kulturellen historischen Erbes digital im Netz zur Verfügung stehen, haben sich Bund und Länder entschlossen, mit der Deutschen Digitalen Bibliothek (DDB) eine spartenübergreifende Internetplattform für Digitalisate von Werken aller Art (Bücher, Bilder, Archivalien, Noten, Musikstücke, Filme, 3D-Aufnahmen von Skulpturen und Kulturdenkmälern) aus unterschiedlichen Sparten einzurichten, die in die Europäische Digitale Bibliothek (Europeana) eingebunden werden soll. Als Vertragspartner der DDB hat sich das Land Baden-Württemberg vertraglich verpflichtet, für die in seiner Trägerschaft befindlichen Einrichtungen die entsprechenden Rahmenbedingun-

gen zu schaffen, die die Produktion von E-Content und dessen Speicherung ermöglichen.<sup>19</sup>

Die Digitalisate, welche Archive, Bibliotheken, Museen und universitäre Sammlungen im Geschäftsbereich des Wissenschaftsministeriums sowie das Haus des Dokumentarfilms produzieren und auf ihren Webseiten bzw. in regionalen Portalen wie z. B. dem landeskundlichen Informationssystem LEO-BW ([www.leo-bw.de](http://www.leo-bw.de)) anbieten, können dank einheitlicher Schnittstellen mehrheitlich auch der DDB zur Verfügung gestellt werden.

### 5 · STATUS QUO DER RETRODIGITALISIERUNG IN DEN ARCHIVEN, BIBLIOTHEKEN, MUSEEN, UNIVERSITÄREN SAMMLUNGEN UND FILMEINRICHTUNGEN IM GESCHÄFTSBEREICH DES WISSENSCHAFTS MINISTERIUMS

Baden-Württemberg verfügt im Bereich der Digitalisierung bereits über beträchtliches Know-how und kann bei der Entwick-

lung eines Digitalisierungsprogramms auf ein großes Wissenspotenzial zurückgreifen. Langjährige

praktische Erfahrungen haben insbesondere das Landesarchiv Baden-Württemberg und die Universitätsbibliothek Heidelberg. Die beiden Landesbibliotheken und die Universitätsbibliotheken Freiburg und Tübingen sowie die staatlichen Museen sind in diesem Bereich ebenfalls seit Jahren aktiv. Diese Einrichtungen verfügen mehrheitlich über Digitalisierungswerkstätten, deren Ausstattung jedoch noch nicht überall hinreichend ist.

In den Universitätsarchiven und den universitären Sammlungen spielte das Thema Digitalisierung insbesondere wegen der fehlenden technischen Ausstattung bislang eine untergeordnete Rolle, so dass hier erheblicher Nachholbedarf besteht. Die 4.200 digitalisierten Filme der Landesfilmsammlung werden in absehbarer Zeit im Netz zur Verfügung stehen.

Im Oktober 2012 konnten die genannten Einrichtungen im Netz 7,7 Mio. Images anbieten, die sich spartenweise wie folgt verteilen:

| EINRICHTUNGEN   | IMAGES           |
|---|------------------|
| 5 Altbestandsbibliotheken (Württembergische Landesbibliothek, Badische Landesbibliothek, Universitätsbibliotheken Heidelberg, Freiburg, Tübingen) | 5.100.000        |
| Landesarchiv Baden-Württemberg  | 1.975.123        |
| Universitäre Sammlungen   | 600.000          |
| Staatliche Museen   | 32.500           |
| Universitätsarchive   | —                |
| <b>Summe</b>  | <b>7.707.623</b> |

Die Analyse des Ist-Zustands im Land belegt aber auch, welcher Handlungsbedarf insgesamt noch besteht: Aktuell sind von den stark nachge-

fragten Beständen des kulturellen historischen Erbes in den genannten Einrichtungen noch rd. 205 Mio. Images zu produzieren, die sich spartenweise wie folgt verteilen:

| EINRICHTUNGEN                        | IMAGES                                  |
|--------------------------------------|---|
| Landesarchiv Baden-Württemberg       | 88.000.000                              |
| Universitäts- und Landesbibliotheken | 42.303.000                              |
| Staatliche Museen*                   | 65.200.000                              |
| Universitäre Sammlungen              | 6.300.000                               |
| Universitätsarchive                  | 4.000.000                               |
| Landesfilmsammlung                   | 2.800 Filme                             |
| <b>Summe</b>                         | <b>205.803.000 Images + 2.800 Filme</b> |

\* Die hohe Imagezahl bei den Staatlichen Museen ist bedingt durch die Vielzahl an dreidimensionalen Objekten, die 3D-Digitalisierungen erforderlich machen.

Wie verhältnismäßig gering die Zahl des noch zu produzierenden E-Contents im Verhältnis zum Ge-

samtbestand der Einrichtungen ist, illustrieren exemplarisch die Zahlen des Landesarchivs: 88 Mio. Images entsprechen nur 7,43 % des

Gesamtbestands, der sich derzeit auf 144 lfd. Kilometer Archivgut beläuft. Die ständig steigenden Zugriffszahlen auf die bereits jetzt im Netz angebotenen Digitalisate der genannten Einrichtungen sind deutlicher Beleg für den weltweiten Bedarf an digitalen Kultur-

und Wissenschaftsinformationen. Das Landesarchiv Baden-Württemberg verzeichnete zuletzt jährliche Zugriffszahlen auf Images i. H. von 3,75 Mio., die Universitäts- und Landesbibliotheken i. H. von 7,6 Mio.

## 6 · ZUKÜNFTIGE DIGITALISIERUNGSSTRATEGIE

### 6.1 SPARTENSPEZIFISCHE DIGITALISIERUNGSZENTREN

Um von den kleinteiligen Digitalisierungsprojekten im Land zu einer großflächigen Struktur mit koordinierten Digitalisierungsvorhaben zu kommen, sollen Digitalisierungskompetenzzentren geschaffen

bzw. ausgebaut werden, ausreichende Speicherkapazitäten bereitgestellt und Mittel für die Bildung des E-Contents zur Verfügung gestellt werden.

*„Wenn nun über Heidelberg hinaus ein Digitalisierungsschwerpunkt in Baden-Württemberg eingerichtet würde, gäbe das ohne Zweifel vielfältige neue Impulse für die Mittelalterforschung. Angesichts der reichen einschlägigen Bestände in unseren Bibliotheken (vor allem natürlich in Karlsruhe und Stuttgart sowie in den Universitätsbibliotheken, aber durchaus auch in kleineren Bibliotheken und Archiven wie z. B. dem Zentralarchiv Neuenstein Hohenlohe) würde ein solches Engagement in hohem Maße und mit langer Wirkung gewinnbringend sein für die Forschungslandschaft, die Bibliotheken und Archive sowie für das kulturelle Erbe Baden-Württembergs.“*

*Ludger Lieb (Universität Heidelberg) in einer gutachterlichen Stellungnahme für die AG Digitalisierung der AG E-Science vom 9.7.2012.*

Die Digitalisierungsstrategie des Landes sieht aus Gründen der Praktikabilität, der Wirtschaftlichkeit und mit Rücksicht auf die unterschiedlichen Materialien sowie den Transport wertvoller Kulturgüter eine dezentrale spartenspezifische Struktur an Digitalisierungszentren vor. Diese Digitalisierungs- und Kompetenzzentren

sollen gleichzeitig als Dienstleister fungieren. Spartenübergreifend können jedoch Synergieeffekte bei der Digitalisierung von Objekten erzielt werden, für die vor Ort Scanner gemeinsam genutzt werden könnten und deren Transport nicht aufwendig ist.

### 6.2 SCAN-DIENSTLEISTUNGEN

Nach Aufrüstung der vorhandenen Digitalisierungswerkstätten im Landesarchiv BW, in den fünf Altbestandsbibliotheken, in den staatlichen Museen und in der Landesfilmsammlung sollen diese als Digitalisierungszentren Scan-Dienstleistungen gegen Entgelt anbieten:

Freiburg, Tübingen, Badische Landesbibliothek, Württembergische Landesbibliothek) sollen insbesondere für andere Bibliotheken Scanaufträge erledigen. Die universitären Sammlungen sollen den Digitalisierungszentren materialspezifisch zugeordnet werden. Im Hinblick auf die zahlreichen dreidimensionalen Objekte in den Museen und universitären Sammlungen ergibt sich ein besonderer Digitalisierungsaufwand sowohl in Bezug auf die Technik als auch auf den Zeitaufwand. Da dreidimensionale Museumsobjekte aus konservatorischen Gründen und aus Sicherheitsgründen nicht be-

Das Landesarchiv BW, das auf ein verteiltes Digitalisierungszentrum an den Standorten Ludwigsburg, Karlsruhe und Stuttgart mit einer zentralen Koordinierungsstelle setzt, soll diese Dienstleistung bei speziellen Projekten oder Sonderformaten für Universitätsarchive, Landes- und Universitätsbibliotheken und Landesmuseen erbringen. Die fünf Altbestandsbibliotheken (Universitätsbibliotheken Heidelberg,

liebig durch das Land transportiert werden können, soll eine mobile 3D-Werkstatt aufgebaut werden. Für die Digitalisierung von Spezialformaten ist in Ergänzung zu einer entsprechenden technischen Grundausstattung eine Kooperation mit den Bibliotheken und dem Landesarchiv BW denkbar. Die Landesfilmsammlung im Haus des Dokumentarfilms soll Dienstleistungen beim Überspielen und Digitalisieren von Filmkassetten übernehmen.

### 6.3 BERATUNGSDIENSTLEISTUNGEN

Die Kompetenzzentren sollen bei entsprechender personeller Ausstattung Beratungsdienstleistungen erbringen.

#### **Beratungsdienstleistungen des Landesarchivs BW für Universitätsarchive, Kommunalarchive (Kreis-, Stadt- und Gemeindearchive) und sonstige nichtstaatliche Archive:**

- Beratung bei der Antragstellung für Fördermittel
- Beratung bei Ausschreibungen und der Auftragsvergabe an externe Dienstleister, Beratung bei der technischen Ausstattung und in Fragen der Geschäftsprozesse
- Abstimmung und Beratung bei Metadatenstandards für die Online-Präsentation im Kontext der Erschließung
- Unterstützung bei der Anbindung an regionale, nationale und internationale Nachweis- und Präsentationssysteme (LEO-BW, DDB, „Archivportal-D“, Europeana).

#### **Beratungsdienstleistungen der fünf Altbestandsbibliotheken für andere Bibliotheken:**

- Inhaltliche Beratung zum Digitalisierungsbedarf, Unterstützung bei der Bestandsanalyse, Vermeidung von Doppeldigitalisierungen
- Technische Beratung bei der Geräteauswahl
- Aufsetzen und Optimieren der Geschäftsprozesse
- Abstimmung von Standards für technische Verfahren. Standardisierung der Datenformate
- Unterstützung bei Anbindung an regionale, nationale und internationale Nachweis- und Präsentationssysteme (LEO-BW, DDB, ZVDD, Europeana)
- Unterstützung und Beratung bei Beantragung von Projektgeldern und Fördermitteln
- Beratung bei der Auswahl von Dienstleistern

Im Zentrum soll die Beratung und Unterstützung bei der Durchführung eigener Projekte sowie bei der Vergabe von Projekten an externe Firmen stehen. Das Angebot richtet sich ausdrücklich auch an Bibliotheken in kommunaler, kirchlicher und privater Trägerschaft, die über bedeutende historische Sammlungen verfügen.

#### **Beratungsdienstleistungen des Bibliotheksservice-Zentrums Baden-Württemberg:**

- Datentechnische Unterstützung bei der Auswahl und Erschließung geeigneter Sammlungen
- Vorgaben für Werkstätten und Dienstleister, besonders im Museumsbereich
- Verarbeiten von Digitalisaten in automatisierten Workflows, besonders im Museumsbereich
- Langzeitarchivierung der Digitalisate
- Präsentationswerkzeuge und Lieferung in überregionale Portale

Für Archive wird das BSZ bei speziellen Projekten in Absprache mit dem Landesarchiv Baden-Württemberg technisch unterstützend tätig.

#### **Beratungsdienstleistungen der Landesfilmsammlung:**

- Beratung zu allen Fragen der Filmdigitalisierung

#### **6.4 KRITERIEN BEI DER MITTELVERGABE**

Zu einer Digitalisierungsstrategie gehören einheitliche Kriterien bei der Mittelvergabe. Diese sind:

- Nachweis der Forschungsrelevanz
- Nachweis der Metadaten der zu digitalisierenden Objekte
- Nachweis, dass die fraglichen Objekte nicht an anderer Stelle digitalisiert worden sind
- Freigabe der Digitalisate im Sinne des Open Access
- Anwendung der einschlägigen technischen Standards (Maßgebend sind derzeit die DFG-Praxisregeln „Digitalisierung“ in ihrer jeweils aktuellen Fassung)
- Die Rechte Dritter dürfen nicht verletzt werden.
- Nachweis über die Nachhaltigkeit bzw. Persistenz der Online-Präsentation und die Sicherung der zu erstellenden Digitalisate

#### **6.5 REGIONALE UND ÜBERREGIONALE KOORDINIERUNG VON DIGITALISIERUNGSPROJEKTEN**

Die landeseigene Digitalisierungsstrategie muss überregionale Digitalisierungsstrategien und -projekte berücksichtigen. Die Strategie der DFG zielt mit deren Förderprogramm „Erschließung und Digitalisierung“ auf die digitale Transformation und Bereitstellung herausragender und für die Forschung überregional bedeutender Bestände der wissenschaftlichen und kulturellen Überlieferung ab. Dabei wird sich der Schwerpunkt der Förderaktivitäten nach und nach von den Drucken hin zu historischen Zeitungen<sup>20</sup> und unikalem Material (Archivgut, mit-

telalterliche Handschriften, dreidimensionale Objekte) verschieben.<sup>21</sup> Für die Deutsche Digitale Bibliothek gibt es bislang noch keine zwischen Bund und Ländern abgestimmte Strategie für die Koordinierung des E-Contents. Die Abstimmung neuer forschungsrelevanter Digitalisierungsprojekte innerhalb des Landes und mit überregionalen Aktivitäten wird Aufgabe der Digitalisierungs- und Kompetenzzentren sein.

### **7 · DERZEIT VORHANDENE MITTEL FÜR DIE DIGITALISIERUNG**

Das Wissenschaftsministerium stellt den Bibliotheken und Archiven seit 2011 zentrale Digitalisierungsmittel zur Verfügung: Zum einen über die neue Förderlinie der Stiftung Kulturgut Baden-Württemberg, die jährlich durchschnittlich 250.000

Euro bereitstellen kann, zum anderen über ein Projekt für die fünf Altbestandsbibliotheken („Tresor - Schätze der baden-württembergischen Bibliotheken“), in dessen erster Projektphase

(1/2011-12/2012) 500.000 Euro bewilligt wurden. Mittel in gleicher Höhe wurden für den Doppelhaushalt 2013/2014 angemeldet. Für die Inventarisierung und Digitalisierung wurden den staatlichen Museen seit einigen Jahren jährlich rd. 350.000 Euro bereitgestellt. Die geplante Auf-

stockung dieser Mittel wird aufgrund der Haushaltskonsolidierungszwänge voraussichtlich nicht realisiert werden können. Zu den Ist-Ausgaben vgl. den Anhang. Diese Landesmittel und die Förderprogramme der DFG reichen nicht aus, um in einem überschaubaren Zeitraum den beschriebenen Digitalisierungsbedarf bei den genannten Einrichtungen zu decken.

## 8 · MITTELBEDARF

Digitalisierung im größeren Stil ist erst möglich, wenn die richtigen finanziellen, organisatorischen und technischen Rahmenbedingungen geschaffen werden. Der größte Teil der Kosten entsteht nicht bei der eigentlichen Image-Digitalisierung, sondern bei der Aufbereitung der Digitalisate für die Präsentation und bei der Qualitätssicherung.

Erst die notwendige Erschließung und Strukturierung, die Aufbereitung und Verwaltung sowie die Langzeitarchivierung der Digitalisate machen diese zu einem zentralen Bestandteil der Informationsinfrastruktur. Der zusätzliche Mittelbedarf für die Archive, Bibliotheken, Museen, universitären Sammlungen im Geschäftsbereich des Wissenschaftsministeriums sowie die Landesfilmsammlung teilt sich in drei Kategorien:

| KATEGORIE   | KOSTEN EURO | ANMERKUNG                                |
|---|-------------|--|
| Zusätzliche technische Ausstattung  | 1.274.550   | Einmalige Kosten                         |
| Personalkosten für Beratungsdienste der Digitalisierungszentren und die mobile 3D-Werkstatt | 458.450     | Jährliche Kosten                         |
| E-Content-Produktion und Speicherplatz für 2 Mio. Images Jahresproduktion                   | 2.000.000   | Jährliche Kosten: variabel <sup>22</sup> |

## 9 · PUBLIC-PRIVATE-PARTNERSHIP-PROJEKTE

PPP-Projekte sollten nur dann ins Auge gefasst werden, wenn eine vollständige Transparenz bezüglich der Auswahlkriterien der zu digitalisierenden Objekte und der Nachnutzungsrechte besteht und

die Qualität der Digitalisate den Anforderungen der Forschung genügt. Alle entsprechend geförderten Digitalisate müssen im Sinne des Open Access dauerhaft zur Verfügung stehen.

## 10 · NUTZUNG UND VERWERTUNG

Im Gegensatz zur kommerziellen Nutzung muss der lesende Zugriff auf den von der öffentlichen Hand geförderten E-Content und dessen Nachnutzung im Sinne des Open Access für die Forschung bzw. für die private Nutzung kostenlos und rechtssicher sein. Für die kommerzielle Verwertung von Digitalisaten gelten die Gebühren der jeweiligen Gebührenordnungen.

## 11 · EVALUIERUNG DER DIGITALISIERUNGSSTRATEGIE

Die spartenübergreifende Digitalisierungsstrategie des Landes sollte im Hinblick auf das Nutzerverhalten und die Forschungsbedürfnisse in regelmäßigen Abständen evaluiert und bei Bedarf angepasst werden. Über die Verteilung zentraler Digitalisierungsmittel sollte ein Beirat mit Vertretern der einzelnen Sparten entscheiden.

## ANHANG

### Ist-Ausgaben des Ministeriums für Wissenschaft, Forschung und Kunst Baden-Württemberg für Digitalisierungsmaßnahmen 2011-2012:

| JAHR         | AUSGABEN IN EURO |
|--------------|------------------|
| 2011         | 920.633          |
| 2012         | 422.922          |
| <b>Summe</b> | <b>1.343.555</b> |

### Zusätzlicher Ausstattungsbedarf für die Digitalisierungszentren:

| EINRICHTUNGEN                                  | KOSTEN IN EURO   |
|--|------------------|
| Landesarchiv Baden-Württemberg                 | 585.000          |
| 5 Altbestandsbibliotheken                      | 42.000           |
| Staatliche Museen                              | 482.850          |
| Landesfilmsammlung im Haus des Dokumentarfilms | 164.700          |
| <b>Summe</b>                                   | <b>1.274.550</b> |

Der Bedarf wurde im Detail erhoben, wird hier jedoch nur summarisch wiedergegeben.

### Bedarf zusätzlicher Stellen für Beratungsdienste der Digitalisierungs- und Kompetenzzentren und die mobile 3D-Werkstatt:

| EINRICHTUNGEN              | STELLEN              | KOSTEN (TV-L GEMÄSS RICHTWERT 2013/14) IN EURO |
|----------------------------|----------------------|--|
| Landesarchiv BW            | 1 hD E 13, 1 gD E 9  | 115.400  |
| 5 Altbestandsbibliotheken  | 5 x ½ gD E 9         | 137.250  |
| Bibliotheksservice-Zentrum | 1 hD E 13, ½ gD E 12 | 96.800   |
| Staatliche Museen          | 2 gD E 9             | 109.000  |
| <b>Summe</b>               |                      | <b>458.450</b>                                 |

### Kosten für jährliche E-Content-Produktion mit Speicherplatz:

| EINRICHTUNGEN        | IMAGEZAHL                          | IMAGEKOSTEN IN EURO | TB           | TB-KOSTEN IN EURO | KOSTEN INSG. IN EURO |
|----------------------|------------------------------------|---------------------|--------------|-------------------|----------------------|
| Landesarchiv BW      | 1.000.000                          | 1.000.000           | 40           | 14.400            | 1.014.400            |
| Universitäts- und LB | 1.000.000                          | 780.000             | 40           | 14.400            | 794.400              |
| Universitätsarchive  | 50.000                             | 50.000              | 2            | 720               | 50.720               |
| Staatliche Museen    | 20.000                             | 40.000              | 1            | 360               | 40.360               |
| Univ. Sammlungen     | 2.000                              | 2.000               | 0,1          | 36                | 2.036                |
| Landesfilmsammlung   | 800 Objekte                        | 200.000             | 30           | 10.800            | 210.800              |
| <b>Summe</b>         | <b>2.072.000 +<br/>800 Objekte</b> | <b>2.072.000</b>    | <b>113,1</b> | <b>40.716</b>     | <b>2.112.716</b>     |

Bei den Speicherplatzkosten ist zu beachten, dass die Zahl der TB Jahr für Jahr kumuliert.

<sup>16</sup> [http://www.dfg.de/download/pdf/foerderung/programme/lis/praxisregeln\\_digitalisierung.pdf](http://www.dfg.de/download/pdf/foerderung/programme/lis/praxisregeln_digitalisierung.pdf).

<sup>17</sup> Zuletzt: DFG-Positionspapier „Die digitale Transformation weiter gestalten – der Beitrag der Deutschen Forschungsgemeinschaft zu einer innovativen Informationsinfrastruktur für die Forschung“, Bonn 2012; Wissenschaftsrat: „Empfehlungen zur Weiterentwicklung der wissenschaftlichen Informationsinfrastrukturen in Deutschland bis 2020“ (Drs. 2359-12), Köln 2012; Kommission Zukunft der Informationsinfrastruktur im Auftrag der GWK des Bundes und der Länder: „Gesamtkonzept für die Informationsinfrastruktur in Deutschland“, Berlin 2011. In diesem Sinne auch die Handlungsempfehlungen der Enquete-Kommission des Deutschen Bundestags „Internet und digitale Gesellschaft/ Projektgruppe „Bildung und Forschung“ vom 23.10.2012 (Drs. 17(18)321).

<sup>18</sup> Rüdiger Steinmetz (Universität Leipzig): „Zum einen werden durch die digitale Bereitstellung von historischen und kulturell wichtigen filmischen Dokumenten fiktionaler und vor allem non fiktionaler Art im Internet sicher Forschungen angeregt und intensiviert werden... Und zum anderen wird die Schul- und Hochschul-Lehre intensiv befruchtet“ in einer gutachterlichen Stellungnahme für die AG Digitalisierung der AG E-Science vom 31.8.2012.

<sup>19</sup> Das Verwaltungs- und Finanzierungsabkommen über die DDB wurde nach der Ermächtigung durch das Bundeskabinett und die Ministerpräsidentenkonferenz von Baden-Württemberg am 4.1.2010 unterzeichnet.

<sup>20</sup> Edgar Wolfrum (Universität Heidelberg): „Im internationalen Vergleich wurden regionale und lokale Zeitungen bislang in Deutschland nicht ausreichend in digitalisierter Form zugänglich gemacht. Die in einzelnen Fällen vorliegenden Sammlungen auf Mikroformen sind ein wichtiger Bestandteil zur Sicherung des Quellenbestands. Ihre Nutzung ist jedoch mit Einschränkungen verbunden.“ Auszug aus seiner gutachterlichen Stellungnahme für die AG Digitalisierung der AG E-Science vom 3.7.2012.

<sup>21</sup> DFG-Positionspapier „Die digitale Transformation weiter gestalten“, a.a.O., S. 10

<sup>22</sup> Materialspezifische Durchschnittspreise: 0,78 Euro pro Image für Bibliotheksgut, 1 Euro pro Image für Archivgut, 2 Euro pro Image für Museumsgut.

## AG 3 · Open Access

Neue Rahmenbedingungen für wissenschaftliches Publizieren in Open Access  
an den Hochschulen in  
Baden-Württemberg

## INHALT

|  |    |
|--|----|
| 1. Programmkonzeption .....  | 61 |
| 1.1 Ausgangslage .....   | 61 |
| 1.2 Handlungsfelder .....  | 62 |
| 2. BW-Open Access Förderprogramm .....                             | 63 |
| 2.1 Ebene Hochschule .....   | 64 |
| 2.2 Ebene MWK .....  | 64 |
| Zu 1: Programmkonzeption   |    |
| Anhang 1: Beschaffungsetats der Hochschulbibliotheken .....        | 65 |
| Anhang 2: Abbestellungen .....                                     | 65 |
| Anhang 3: Nutzerperspektiven .....                                 | 67 |
| Anhang 4: Füllstand von Repositorien in Baden-Württemberg .....    | 69 |
| Anhang 5: Technische Infrastrukturen .....                         | 70 |
| Anhang 6: Rechtsrahmen von Open Access .....                       | 71 |
| Zu 2: BW-Open Access Förderprogramm                                |    |
| Anhang 7: Pilotprojekt zu hybriden Open Access Zeitschriften ..... | 73 |
| Anhang 8: Meilensteine .....                                       | 74 |
| Anhang 9: Literaturverzeichnis .....                               | 75 |

## MITGLIEDER DER ARBEITSGRUPPE

### **DR. CHRISTOPH BRUCH**

Helmholtz Open Access Koordinationsbüro,  
Alfred-Wegener-Institut für Polar- und Meeresforschung in  
der Helmholtz-Gemeinschaft, Bremerhaven

### **PROF. DR. THOMAS DREIER**

Zentrum für Angewandte Rechtswissenschaft am KIT

### **PETRA HÄTSCHER**

Bibliothek, Universität Konstanz

### **DR. MARION MALLMANN-BIEHLER**

Bibliotheksservice-Zentrum Baden-Württemberg,  
Konstanz

### **PROF. DR. HERBERT MÜTHER**

Prorektor, Institut für Theoretische Physik,  
Universität Tübingen

### **DR. THOMAS PFLÜGER**

Ministerium für Wissenschaft, Forschung und Kunst  
Baden-Württemberg (Vorsitz)

### **PROF. DR. HANS-JOCHEN SCHIEWER**

Rektor, Universität Freiburg

### **DR. CHRISTIANE SPARY**

Bibliothek, Pädagogische Hochschule Ludwigsburg

### **PROF. DR. HEINER STUCKENSCHMIDT**

CIO, Institut für Informatik, Universität Mannheim

## 1 · PROGRAMMKONZEPTION

### 1.1 AUSGANGSLAGE

Die Informationstechnik einschließlich des Internet ermöglicht es, Forschungsergebnisse (Forschungsdaten und wissenschaftliche Publikationen) weltweit zugänglich und nachnutzbar zu machen. Die mit dem wissenschaftlichen Publikationswesen, dem Fokus der Open Access AG, verknüpften traditionellen Geschäftsmodelle stehen der Nutzung dieser neuen Möglichkeiten zum Nachteil von Wissenschaft, Wirtschaft und Gesellschaft zum Teil entgegen.

Die Wissenschaftspolitik des Landes Baden-Württemberg hält es für geboten, das wissenschaftliche Publikationsregime so zu gestalten, dass

- die für Bildung und Wissenschaft aufgewendeten Steuermittel größtmöglichen Nutzen für Wissenschaft, Wirtschaft und Gesellschaft erzielen,
- die Sichtbarkeit der wissenschaftlichen Forschung in Baden-Württemberg weltweit erhöht wird und damit
- die Wettbewerbsfähigkeit des Landes nachhaltig gesichert und gesteigert wird.

Dafür sollen jedenfalls diejenigen wissenschaftlichen Publikationen, die aus vom Land Baden-Württemberg (ko-)finanzierter Forschung stammen, entsprechend dem Open Access Prinzip<sup>23</sup> ab dem Zeitpunkt ihrer Erstveröffentlichung (goldener Weg) oder nach einer Embargofrist von in aller Regel 6 Monaten ab der Erstveröffentlichung (grüner Weg) in elektronischer Form weltweit kostenfrei zugänglich sein.

Das Land Baden-Württemberg wäre damit das erste deutsche Bundesland, das solche Zielvorgaben entwickelt und würde damit eine bundesweite Vorreiterrolle einnehmen. International betrachtet ordnet

sich die von Baden-Württemberg angestrebte Politik in eine breite Bewegung in der Wissenschaft ein. Dies wird beispielsweise durch die Liste der Unterzeichner der Berliner Erklärung über offenen Zugang zu wissenschaftlichem Wissen<sup>24</sup> deutlich. Diese erhebt im Übrigen, dass es sich bei dem Open Access-Paradigma in erster Linie um ein aus der scientific Community – durch die sog. „Zeitschriftenkrise“ lediglich beschleunigtes – kommendes alternatives Publikationskonzept handelt, dessen Ziel in erster Linie darin besteht, die möglichst breite Verfügbarkeit und Nachnutzbarkeit von Forschungsergebnissen<sup>25</sup> zu gewährleisten. Dort wurde als Herausforderung für das gesamte Wissenschaftssystem erkannt, dass wissenschaftlicher Fortschritt überwiegend aus Steuermitteln finanziert wird, seine Verbreitung über Publikationen aber kommerzialisiert ist und – bei hoher Marktmacht der Anbieter – sich der Erwerb zu angemessenen Bedingungen immer schwieriger gestaltet.

Auf transnationaler Ebene sind innerhalb Europas die Europäische Kommission<sup>26</sup> und die Regierung von Großbritannien bei der Entwicklung und Implementation Open Access-Politik<sup>27</sup> in der Wissenschaft, aber auch Science Europe<sup>28</sup> und die LERU<sup>29</sup> als wichtige Treiber zu nennen. Die für eine optimale Nutzung von Forschungsergebnissen notwendige Umgestaltung des wissenschaftlichen Publikationswesens bedarf internationaler Zusammenarbeit. Die Entscheidung der Landesregierung von Baden-Württemberg hat deshalb über das Land hinaus Bedeutung, da sie einen wesentlichen Beitrag dazu leisten würde, das Postulat der „fünften Grundfreiheit“ der EU – der Förderung des freien Verkehrs von Wissen – mit Leben zu erfüllen<sup>30</sup>.

Die Landesregierung wird sich weiter dafür einsetzen, im Urheberrechtsgesetz ein unabdingbares Zweitveröffentlichungsrecht für öffentlich geförderte wissenschaftliche Autoren einzufügen. Sie orientiert sich damit an der von Abgeordneten aller im Bundestag vertretenen Parteien in der Projektgruppe Bildung und Forschung der Enquete Kommission Internet und digitale Gesellschaft des Deutschen Bundestages einstimmig beschlossenen

Empfehlung<sup>31</sup>, Open Access und Zweitveröffentlichungsrecht umzusetzen, sowie an den entsprechenden Forderungen der Kultusministerkonferenz<sup>32</sup> und der Allianz der deutschen Wissenschaftsorganisationen<sup>33</sup>.

Die Open Access Politik des Landes Baden-Württemberg ist darauf fokussiert, im Land die Voraussetzungen zu schaffen, den notwendigen internationalen Umgestaltungsprozess in herausgehobener Stellung zu ermöglichen, zu fördern und aktiv mitzugestalten, um so auch die Wahrnehmung der Forschungsergebnisse aus Baden-Württemberg international zu erhöhen. Die unter den Stichworten „grüner/goldener Weg“ definierten Ansätze für die Implementierung von Open Access an Hochschulen und Forschungseinrichtungen werden vorliegend im Sinne weitgehender Komplementarität verstanden, ohne dass damit weitere Ansätze ausgeschlossen sind. Welche Wege die Hochschulen präferieren, ist auch eine Frage der hochschulautonomen Profilbildung und Schwerpunktsetzung.

Das vom Land Baden-Württemberg entwickelte Open Access Förderprogramm kann dabei auf vielfältige, teilweise auch von den großen Drittmittelgebern geförderte Aktivitäten an Hochschulen, außeruniversitären Forschungseinrichtungen sowie einiger Verlage im Lande aufbauen. Eine stärkere Akzentuierung der Informationsversorgung an Hochschulen über Open Access wird allerdings mit

zusätzlichen Kosten für die Umsteuerung verbunden sein. Im Übrigen ist davon auszugehen, dass die Hochschulbudgets für die Literaturversorgung im Rahmen der Zuwachsraten der letzten Jahre weiter in gewissem Umfang wachsen werden<sup>34</sup>. Insgesamt zeigt sich bereits, dass gerade auch national und international agierende Verlagshäuser zunehmend Open Access als Publikationsform akzeptieren und integrieren. Somit bleiben – unbeschadet eines auch über Open Access zunehmenden Wettbewerbs auf Seite der Anbieter – wirtschaftliche Belange nationaler und internationaler Wissenschaftsverlage grundsätzlich unberührt.

Mit einem BW-Open Access-Förderprogramm soll ein markanter Beitrag zur nachhaltigen Sicherung der Informationsinfrastrukturen an den Hochschulen des Landes und der entsprechenden Umgestaltung des internationalen wissenschaftlichen Publikationswesens geleistet werden. Im Fokus steht das Ziel, dass Wissen und Innovation, die in gemeinnützig organisierten Wissenschaftsorganisationen erarbeitet werden, auch „barrierefrei“ zur Nachnutzung durch Wissenschaft, Wirtschaft und Gesellschaft zur Verfügung stehen.

## 1.2 HANDLUNGSFELDER

Im Förderprogramm sind Aktivitäten auf den folgenden drei Handlungsfeldern vorgesehen:

### 1) Auf- und Ausbau einer Infrastruktur für die elektronische wissenschaftliche Kommunikation (grüner Weg)

Die technischen Infrastrukturen an den Hochschulen für elektronisches Publizieren müssen weiter ausgebaut werden. Im Bereich des wissenschaftlichen Publikationswesens kommt dem „grünen Weg“ kurz- und mittelfristig eine besondere Bedeutung zu, da in vielen Bereichen neue Open Access Zeitschriften noch nicht das Renommee der traditionellen Zeitschriften erreicht haben und Angebote zur Open Access Stellung einzelner Artikel in ansonsten traditionell subskriptionsbasierten Zeitschriften (hybride Open Access Zeitschriften) in ihrer bisherigen Ausprägung in der Wissenschaft noch zu wenig Akzeptanz finden. Für die internationale wissenschaftliche Informationsversorgung sowohl mit Publikationen als auch mit Forschungsdaten haben inzwischen einzelne Repositorien eine

herausragende Bedeutung erlangt<sup>35</sup>. Daher sollte das Land Baden-Württemberg Infrastrukturen mit dem Potenzial, eine herausragende Bedeutung für die internationale wissenschaftliche Informationsversorgung in einem besonders kompetitiven Forschungsfeld zu erlangen, gesondert als „Open Access Leuchtturm“ fördern.

## **2) Aufbau von Fonds für die Finanzierung von Open Access Publikationen (goldener Weg)**

Komplementär hierzu soll die Transformation von Subskription zu Open Access weiterentwickelt werden. Konkret werden die folgenden Aktivitäten gefördert:

- a) Entwicklung von Kriterien für die Übernahme von Publikationsgebühren für Zeitschriftenaufsätze und Buchveröffentlichungen in Open Access,
- b) stufenweiser Aufbau von Fonds für die Finanzierung von Publikationsgebühren,
- c) Erhöhung der Steuerungsfähigkeit des Landes und der Hochschulen bei den Ausgaben für die wissenschaftliche Kommunikation, speziell für die Informationsversorgung, durch die valide und zeitnahe Erfassung einschlägiger Daten.

## **3) Zukunftsfähige Geschäftsmodelle im Land Baden-Württemberg**

Open Access („grün“ und „gold“) eröffnet ferner Möglichkeiten, mit Wissenschaftsverlagen und den Trägern verlagsungebundener Publikationsformen zukunftsfähige Geschäftsmodelle wie folgt zu entwickeln:

- a) Unterstützung von Open Access kompatiblen Geschäftsmodellen und des Umstiegs auf solche durch Verhandlungen mit den am wissenschaftlichen Publikationswesen beteiligten Akteuren,
- b) Förderung von Wissenschaftlerinnen und Wissenschaftlern sowie wissenschaftlichen Gesellschaften beim Aufbau qualitativ hochwertiger Open Access Zeitschriften sowie bei der Umstellung bestehender Zeitschriften auf Open Access,
- c) Pilotprojekte über den Aufbau von Open Access Geschäftsmodellen für Monografien.

Die beschriebenen Handlungsfelder weisen Schnittstellen zu den in den Arbeitsgruppen „Virtuelle Forschungsumgebung“, „Forschungsdatenmanagement“ sowie „Lizenzierung“ auf, die in der Projektsteuerung auch unter Nutzeraspekten operationalisiert werden müssen. Auf die grundsätzliche Relevanz für das „E-Learning“ wird verwiesen<sup>36</sup>.

## **2 · BW-OPEN ACCESS FÖRDERPROGRAMM<sup>37</sup>**

Zur beschleunigten und nachhaltigen Implementierung des Open Access-Paradigmas an den Hochschulen in Baden-Württemberg kann ein zielführendes Förderprogramm bereits auf bestehende Ansätze in den Hochschulen aufbauen; dieser Prozess muss jedoch durch gezielte Fördermaßnahmen des Landes ergänzt werden. Dabei kommt es auf Basis des

Auf- und Ausbaus der erforderlichen technischen Infrastrukturen auf eine intelligente Verknüpfung „weicher“ und „harter“ Förderfaktoren an. Zu ersteren zählen – auf Ebene der Hochschulen – die Verankerung von Open Access auf Ebene der jeweiligen Rektoratspolicy und – auf Ebene des Landes – das beharrliche Eintreten für eine Verbesserung der rechtlichen Rahmenbedingungen für Open Access. Als Beispiele für den zweiten Bereich sind einschlägige Re-

gelungen in den Berufungsvereinbarungen mit dem wissenschaftlichen Personal sowie entsprechende Anpassungen der Förderbedingungen des Landes in der Projektforschung zu nennen. Diese Maßnahmen erfordern keine zusätzlichen finanziellen Aufwendungen und können rasch umgesetzt werden.

Herausfordernd und in besonderem Maße erfolgskritisch dürfte es sein, ein Anreizsystem zu entwickeln, das zielführend und in mittelfristiger Perspektive den Anteil Open Access-publizierter Forschungsergebnisse in allen Wissenschaftsdisziplinen signifikant erhöht. Dieses Anreizsystem muss bei Hochschulen und den wis-

senchaftlich Publizierenden gleichermaßen ansetzen und ist ohne konditioniert und temporär zur Verfügung gestellte zusätzliche Mittel im System nicht zu haben. Ferner kommt es darauf an, das Anreizsystem so auszugestalten, dass die Schnittstellen zu den Virtuellen Forschungsumgebungen und des Forschungsdatenmanagements reibungslos gestaltet werden.

Unter Zugrundelegung dieser Maßstäbe schlägt die AG der Landesregierung folgendes Förderprogramm Open Access-BW vor:

## 2.1 EBENE HOCHSCHULE

- Open Access ist als wichtiger Teil der Informationsinfrastruktur (IIS) Thema der Rektoratspolitik und wird auf Rektoratebene entsprechend vertreten.
- Open Access wird in Berufungsverhandlungen wie folgt gestärkt: Berufene werden grundsätzlich verpflichtet,
  - sich das Recht der Zweitveröffentlichung bei allen Publikationen in Periodika vorzubehalten,
  - ihre Publikationen in der jeweiligen Hochschulbibliographie nachzuweisen und
  - einen Teil der Budgetzusagen für Open Access-Publikationen zu verwenden.
- Einführung von Publikationsfonds (Übergang von Subskription zu Open Access „gold“).
- Verknüpfung Open Access-Repositoryn mit Forschungsdaten/Virtueller Forschungsumgebung (Schnittstellen, Kompatibilität eal.).
- Implementierung von Qualitätssicherungsmaßnahmen (z. B. DINI-Zertifika, Data Seal of Approval<sup>38</sup>, DIN-31644 / ISO-16363<sup>39</sup>) und Monitoring bei Repositoryn (grüner Weg).

## 2.2 EBENE MWK

- Open Access wird im Landeshochschulgesetz als Leitbild wissenschaftlichen Publizierens abgesichert.
- In allen Förderprogrammen des MWK wird Open Access für die Publikation der Projekt- bzw. Forschungsergebnisse verpflichtend (Open Access-Auflage in der Förderzusage).
- Finanzierung des Personalaufwandes für eine Beratungsstelle für Open Access und Lizenzierung<sup>40</sup>.
- Zur Förderung des Übergangs von Subskription zu Open Access „gold“ wird für den Zeitraum von 5 Jahren vom MWK ein zentraler Publikationsfonds aufgelegt.
- Monitoring des neuen Publikationsmanagements (Aktivitäten, Kosten, Umsetzung)
- Pilotprojekte von Hochschulen mit Wissenschaftsverlagen/Fachgesellschaften
  - Open Access-Journals (auch hybrid) – vgl. Anhang 7
  - Publikationen von Monografien in Open Access
- Förderung von standortübergreifender Infrastrukturprojekten an Hochschulen zur Vernetzung von Publika-

onen, Forschungsdaten und Virtuellen Forschungsumgebungen in Open Access-Repositorien (Open Access Leuchtturmprojekte), einschließlich Web 2.0-Technologien und Linked Data.

- Überregional: Das Land setzt sich für bessere rechtliche Rahmenbedingungen für Open Access sowie für ein wissenschafts- und bildungsfreundlicheres Urheberrecht ein.

## ZU 1: PROGRAMMKONZEPTION

### ANHANG 1: BESCHAFFUNGSETATS DER HOCHSCHULBIBLIOTHEKEN

Die Ausgaben für die Erwerbung in den wissenschaftlichen Bibliotheken (Universitätsbibliotheken, Hochschulbibliotheken, Fachbibliotheken und Regionalbibliotheken) in Deutschland lagen zuletzt bei knapp 300 Mio. Euro jährlich. Davon entfielen mit knapp 100 Mio. Euro etwa 1/3 auf Ausgaben für digitale bzw. elektronische Medien<sup>41</sup>. Von den Erwerbungsmitgliedern allein der Universitätsbibliotheken in Baden-Württemberg in Höhe von ca. 26 Mio. Euro entfallen sogar 10 Mio. Euro, also ca. 40 %, auf digitale bzw. elektronische Medien. Aufschlussreich ist der Blick auf die Entwicklung der Erwerbungsmitgliedern der wissenschaftlichen Bibliotheken von 2003 bis 2010. Lagen sie in den Jahren von 2003 bis 2005 zwischen 220 bis 232 Mio. Euro, stiegen sie von 245 Mio. Euro (2006) und 276 Mio. Euro (2007)

auf 298 Mio. Euro in 2008 und haben sich seither auf diesem Level eingependelt<sup>42</sup>. Die Steigerungsrate von 2003 bis 2010 liegt also bei insgesamt 36 %. Im Zuge des weiteren Ausbaus der Kapazitäten an den Hochschulen kann auch in den nächsten Jahren mit vergleichbaren Steigerungsmitgliedern gerechnet werden, da in den zusätzlich von Bund und Ländern für die Finanzierung neuer Studienplätze zur Verfügung gestellten Budgets auch Infrastrukturmittel enthalten sind. Hinzu kommen noch die Vergütungen der Länder nach den einschlägigen Tantiemetatbeständen des UrhG, die für den gesamten Bereich der öffentlichen und wissenschaftlichen Bibliotheken bei über 20 Mio. Euro p.a. liegen. Davon entfallen allein über 17 Mio. Euro für die Verleih- und Vermiettantieme des § 27 UrhG. Im Ergebnis ist daher festzuhalten, dass in Deutschland allein die Länder als Träger der wissenschaftlichen Bibliotheken den Verlagen eine jährliche garantierte Nachfrage in Höhe von ca. 300 Mio. Euro eröffnen<sup>43</sup>. Hinzu kommen noch die von den großen öffentlichen Drittmittelgebern für den Bereich der Literaturversorgung zur Verfügung stehenden Mittel.

### ANHANG 2: ABBESTELLUNGEN

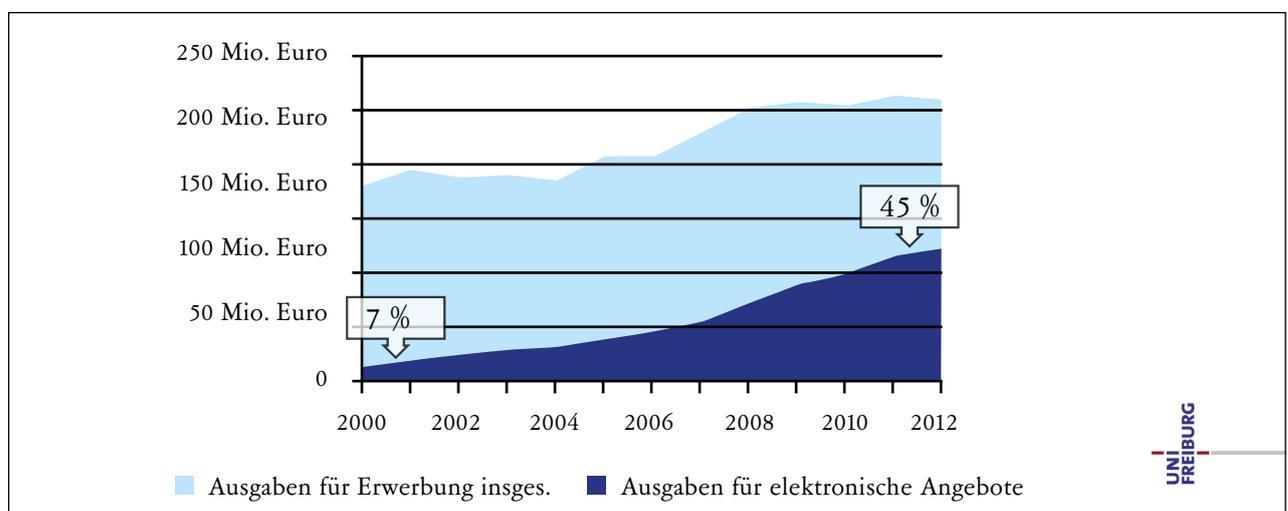


Abb. 4: Ausgaben für Literaturerwerbung bzw. für elektronische Angebote an den Universitätsbibliotheken in Deutschland 2000-2012 (Quelle: Deutsche Bibliotheksstatistik. Diagramm: Antje Kellersohn)

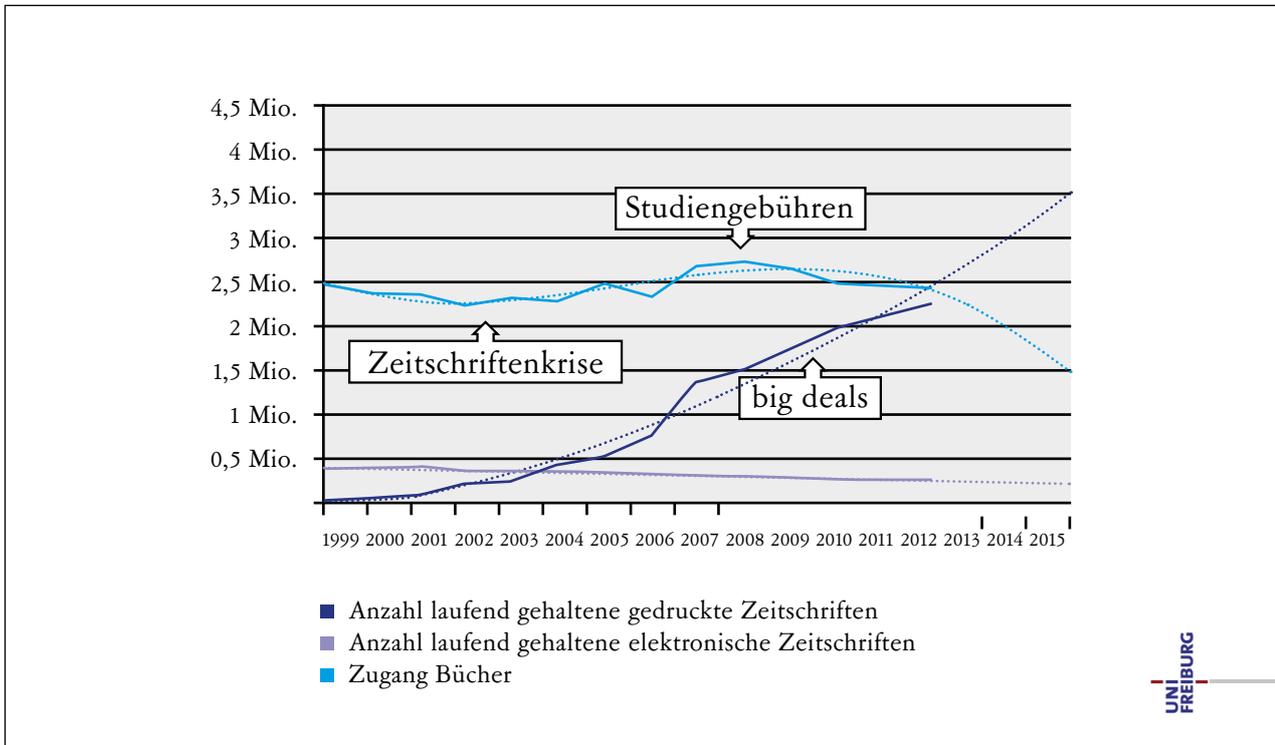


Abb. 5: Medienversorgung an den Universitätsbibliotheken in Deutschland: Zeitschriften und Monografien-Zugang der Universitätsbibliotheken in Deutschland. (Quelle: Deutsche Bibliotheksstatistik. Diagramm: Antje Kellersohn)

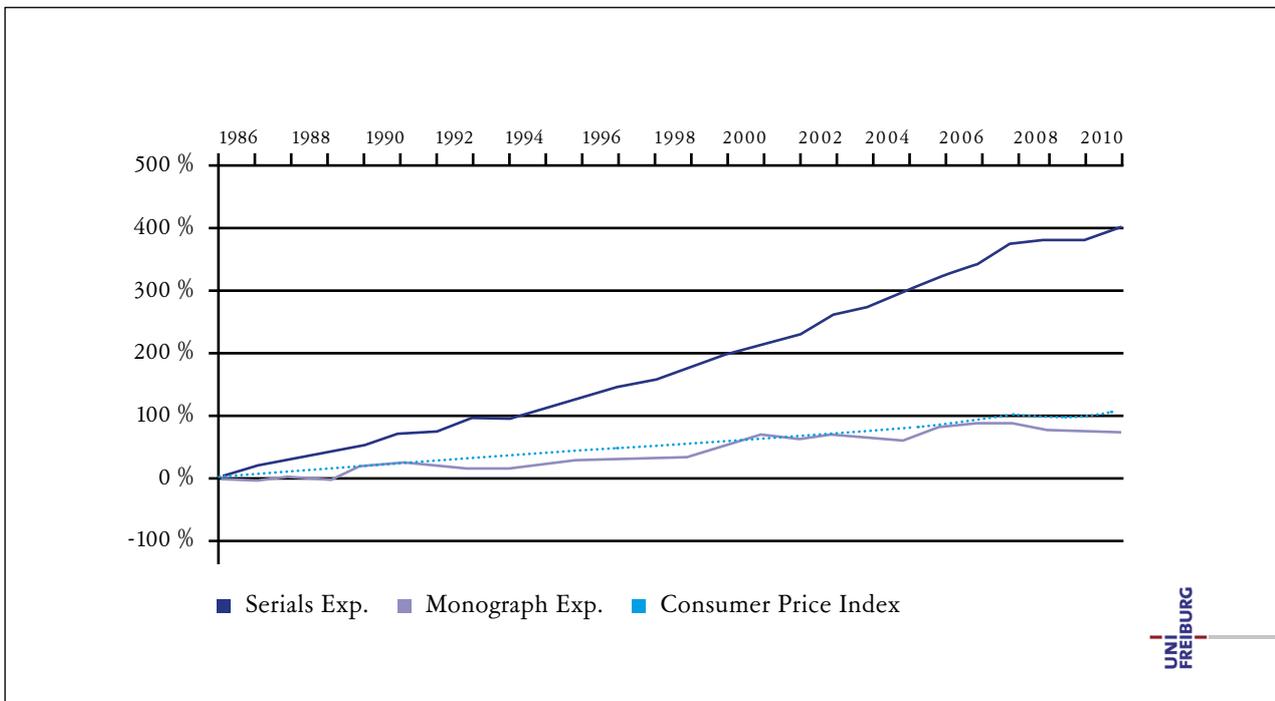


Abb. 6: Ausgabenentwicklung für Fachinformation der Bibliotheken in der Association of Research Libraries (ARL) 1986 bis 2011 für Monographien und Zeitschriften im Vergleich zur allgemeinen Preisentwicklung (US CPI) (Quelle: Kyrillidou/Shaneka (2011): ARL Statistics 2008-2009, Washington, DC, Association of Research Libraries. Diagramm: Antje Kellersohn)

In der ersten Folie ist die Verschiebung der Erwerbungssetats hin zu elektronischen Angeboten insgesamt dargestellt.

In der zweiten Folie geht es darum, wie sich die Zeitschriftenangebote durch den Kauf der großen Pakete insgesamt entwickelt haben. Die Titelzahlen für die elektronischen Ausgaben insgesamt steigen exponentiell an, während die der gedruckten Zeitschriften kontinuierlich sinken. Besonders beachtenswert ist die Entwicklung der Zugangszahlen für die Bücher mit einer Extrapolation bis 2015. Hier wird deutlich,

dass die steigenden Ausgaben für die Zeitschriften immer mehr die verfügbaren Mittel für den Monografienkauf „auffressen“. Das Polster, das ganz offensichtlich durch die Studiengebühren generiert wurde, ist inzwischen aufgebraucht. Es handelt sich um jeweils aufsummierte Werte für alle deutschen Universitätsbibliotheken. Dieser Trend wird immer stärker die Geisteswissenschaften treffen, die im Vergleich zu den STM-Fächern wenig Bedarf an Zeitschriften haben, aber besonders abhängig von einer guten Ausstattung an – in der Regel noch gedruckten – Monografien haben.

Die dritte Folie zeigt eine Übersicht aus den USA, die die überproportionale Preissteigerung bei den Zeitschriften verdeutlicht.

### ANHANG 3: NUTZERPERSPEKTIVEN

Open Access konnte sich inzwischen als neues Paradigma im Bereich des wissenschaftlichen Publizierens aufgrund des einhelligen inhaltlichen Commitments etablieren. Die Allianz der Wissenschaftsorganisationen sowie die Universitäten tragen die Idee und befördern die Etablierung. Doch sind weitere konkrete Schritte erforderlich, um eine nachhaltige Struktur zu erreichen. Um diesen Prozess in der operativen Umsetzung durch Fördermaßnahmen signifikant beschleunigen zu können, kommt es darauf an, die spezifischen Interessen der unterschiedlichsten Nutzer in den Blick zu nehmen. Dabei handelt es sich um so unterschiedliche Nutzergruppen wie die wissenschaftlichen Autoren und Rezipienten an den Hochschulen, die Hochschulleitungen, die einschlägigen Infrastruktureinrichtungen, sowie die interessierte Öffentlichkeit, Gesellschaft und Politik, aber auch innovierende Unternehmen.

Die mit Open Access jeweils verbundenen Erwartungen lassen sich wie folgt beschreiben:

- Die **wissenschaftlichen Autoren und Rezipienten** an den Hochschulen verbinden mit Open

Access die Erwartung der erhöhten Wahrnehmbarkeit ihrer Werke durch Wegfall von Lizenzschranken bzw. den schrankenlosen und möglichst umfassenden Zugriff auf qualitätsgesicherte (peer-reviewed) Forschungsergebnisse.

- Den **Hochschulleitungen** wird es neben der Erwartung lokaler Kostenoptimierung insbesondere um die bessere Informationsversorgung an ihrer Hochschule sowie um deren bessere Sichtbarkeit als Institution im weltweiten Wettbewerb gehen; dabei kann Open Access ein wichtiger Faktor des Umstiegs von der quantitativen zur qualitativen internen Mittelsteuerung sein.
- **Fakultäten, Institute und Lehrstühle**, die zum Teil zur Finanzierung der Informationsversorgung herangezogen werden bzw. auf deren finanzielle Ausstattung sich Kostensteigerungen und Verlagerungen auf der Ebene der Infrastruktureinrichtungen gegebenenfalls auswirken können, sind an Transparenz der Entscheidungen und einem Höchstmaß an einer kompetitiven Informationsversorgung interessiert.
- Den **Infrastruktureinrichtungen** – Hochschulbibliothek, Rechen- und Medienzentrum – kommt es auf die Verbesserung der Dienstleistung durch umfassende Versorgung sowie um eine transparentere Gestaltung der Kosten an.
- Der interessierten **Öffentlichkeit und Gesellschaft** liegen insbesondere bessere Möglichkeiten der Partizipation an For-

schungsergebnissen und die Transparenz von Forschung und Studium am Herzen.

- Insbesondere für *innovierende KMU* leistet eine rasche Diffusion öffentlich finanzierter Forschungsergebnisse einen wichtigen Beitrag für den Wissens- und Technologietransfer.
- **Regierung und Parlament** geht es um den optimalen Einsatz von Steuergeldern mit größtmöglichem Wirkungsgrad sowie die bestmögliche Platzierung des Wissenschafts- und Forschungsstandortes in einem föderal organisierten aber weltweit agierenden Hochschulsystem.

Aus diesen Erwartungen lassen sich etwa folgende Anforderungen ableiten:

1. Qualitätssicherung für Open Access-Publikationen:  
Open Access Zeitschriften und Publikationen müssen die gleichen Standards erfüllen wie subscriptionsbasierte elektronische oder gedruckte Zeitschriften und Publikationen. Sie erscheinen auf den gleichen Wegen wie Verlagspublikationen, Publikationen von Fachgesellschaften oder von Wissenschaftlerinitiativen. Die Qualitätssicherung erfolgt über Peer Review, Gutachtergremien und Editorial Bords. Bei Qualifikationsschriften erfolgt die Qualitätssicherung im jeweiligen Verfahren.
2. Wissenschaftsadäquate Such- und Mehrwertdienste:  
Wie bei subscriptionsbasierten Publikationen ist die Einbindung in bibliographische Datenbanken und Rechercheumgebungen des jeweiligen Wissenschaftsfaches zu gewährleisten; dasselbe gilt für Nutzungsstatistiken und bibliometrische Analysen.
3. Open Access über author fees und Subskriptionszahlungen:  
Mit author fees kann die Erstveröffentlichung ohne Embargofrist erfolgen. Im Rahmen von Subskriptionsmodellen sind unterschiedliche Varianten etwa wie folgt denkbar:
  - Alle Publikationen von Wissenschaftlern aus Baden-Württemberg werden bei entsprechenden Lizenzverträgen sofort Open Access gestellt.
  - Im Rahmen konventioneller Journals werden Publikationen von BW-Wissenschaftlern Open Access gestellt (author fees) mit der Folge, dass die gesamte Zeitschrift für die Wissen-

schaft in Baden-Württemberg freigeschaltet ist (Pilotprojekt hybride Publikation).

#### 4. Weitere Desiderate

- Repositorien werden inzwischen fast flächendeckend vorgehalten, jedoch sind die Breite des fachlichen Inhalts und der Füllstand der Sammlungen noch „unterkritisch“. Zu selten stellen Wissenschaftler Preprints, Projektberichte oder Aufsätze in die Repositorien ein.
- Die Hochschulrepositorien müssen in einem strukturell neu gedachten zentral (Metadaten) / dezentral (Inhalte) angelegten Konzept unter Einbeziehung nicht-textueller Materialien (vgl. AG 2 Forschungsdaten) organisiert werden, das die Materialfülle in allen Erscheinungsformen umfasst und die Langzeitarchivierung der Inhalte sicherstellt.
- Ferner stellt sich die Frage des Mehrwerts aggregierter Repositorien. Im Sinne der Auffindbarkeit im Retrievalprozess der Nutzergruppen sind fachlich oder sachlich-thematisch gebündelte Angebote lokalen Angeboten vorzuziehen. Für die kleineren Hochschulen ist dies besonders wichtig, da ihnen dadurch gewährleistet werden kann, in der elektronischen Informationswelt mit ihren Potenzialen gesehen und beachtet zu werden.
- Zunehmend ist festzustellen, dass nationale und internationale Fördergeber in ihren Ausschreibungen zwingend ausgearbeitete Open Access-Strategien als Förderungsbedingungen voraussetzen. Im Rahmen der den Hochschulen in Baden-Württemberg nach § 7 LHG obliegenden Struktur- und Entwicklungsplanung müssen daher auch Aussagen zur Open Access-Strategie der jeweiligen Hochschule eingefordert werden.

· Schließlich erscheint es aus Nutzersicht erforderlich, eine rechtliche Beratung zu allen Fragen im Open Access-Publikationsprozess zu etablieren. Hier könnte eine Beratungsstelle für Open Access und Lizenzierung einen wichtigen Beitrag zur Nachhaltigstellung von Open Access leisten. Die wissenschaftlichen Bibliotheken haben immer wieder Beratungsbedarf zu rechtlichen und organisatorischen Fragen im Kontext von Open Access geltend gemacht, die durch die bestehenden Angebote (Informationsplattform open-access.net, IUWIS, einschlägige Literatur u.a.m.) nicht abgedeckt werden. Inhaltlich geht es neben

der Beratung in rechtlichen Fragen auch um die Erarbeitung und Überarbeitung von Empfehlungen und Richtlinien, in enger Kooperation mit der Schwerpunktinitiative „Digitale Information“ der Allianz der deutschen Wissenschaftsorganisationen und den Hochschulen, und deren Einbettung in die Gesamtstrategie der Open Access-Policy des Landes. Die Einrichtung einer zentralen Beratungsstelle zur Sicherung von Synergien und Wissenstransfer, angesiedelt etwa bei einer einschlägig ausgewiesenen Bibliothek, könnte diese Lücke schließen und Open Access im Land Baden-Württemberg weiter befördern.

#### ANHANG 4: FÜLLSTAND VON REPOSITORIEN IN BADEN-WÜRTTEMBERG

Im Einsatz ist an den meisten Hochschulen das System „OPUS“, gehostet vom Bibliotheksservicezentrum Baden-Württemberg, Konstanz. Zumeist liegen die Dokumente auch auf den Servern des BSZ und

werden in einem Outsourcing-Modell nicht auf hochschuleigenen Servern gespeichert.

Die folgende Tabelle zeigt den Stand vom 21.05.2012 (Quelle: Auskunft BSZ)<sup>44</sup>:

|    |   |              |
|----|---|--------------|
| 1  | Hochschule Ludwigsburg                                      | 278          |
| 2  | Hochschule Nürtingen-Geislingen                             | 1.286        |
| 3  | Hochschule Aalen  | 24           |
| 4  | Hochschule der Medien, Stuttgart                            | 643          |
| 5  | PH Weingarten und Hochschule Ravensburg-Weingarten          | 56           |
| 6  | Hochschule Esslingen  | 54           |
| 7  | Hochschule Offenburg  | 108          |
| 8  | Hochschule Reutlingen                                       | 104          |
| 9  | Hochschule für Technik, Wirtschaft und Gestaltung, Konstanz | 107          |
| 10 | PH Freiburg   | 248          |
| 11 | PH Heidelberg   | 35           |
| 12 | PH Karlsruhe  | 5            |
| 13 | PH Ludwigsburg  | 33           |
| 14 | PH Schwäbisch-Gmünd   | 7            |
|    | <b>Summe</b>  | <b>2.988</b> |

|   |                           |               |
|---|---------------------------|---------------|
| 1 | Universität Freiburg      | 7.619         |
| 2 | Universität Heidelberg    | 12.426        |
| 3 | Universität Hohenheim     | 685           |
| 4 | Universität Konstanz      | 17.399        |
| 5 | KIT Karlsruhe             |               |
|   | KAROLA (KIT Campus Nord)  | 5.795         |
|   | EVA STAR (KIT Campus Süd) | 7.310         |
| 6 | Universität Mannheim      | 24.523        |
| 7 | Universität-Stuttgart     | 7.042         |
| 8 | Universität Tübingen      | 5.552         |
| 9 | Universität Ulm           | 2.609         |
|   | <b>Summe</b>              | <b>91.010</b> |

Daraus ergibt sich ein Gesamtfüllstand in BW in Höhe von ca. 95.000 Medien.

## ANHANG 5: TECHNISCHE INFRASTRUKTUREN

Open Access zielt primär auf Veränderungen in der Wissenschaftskultur und etablierter Prozesse der Informationsgewinnung und -beschaffung ab. Um die neu einzurichtenden Prozesse effizient unterstützen zu können und hierdurch deren Akzeptanz bei Wissenschaftlern zu steigern, sind – zusätzlich zu entsprechenden Reputationsmechanismen – geeignete technische Infrastrukturen unerlässlich. Es sollten daher Vorhaben unterstützt werden, die darauf abzielen, Open Access Lösungen als gleichberechtigten Bestandteil der wissenschaftlichen Informationsversorgung zu etablieren und in bestehende Infrastrukturen und Prozesse zu integrieren (Schnittstellen zu AG 4 und 5)<sup>45</sup>. Die gegebenen, unterschiedlichen Zielsetzungen der Repositorien nach fachlichen bzw. institutionellen Kriterien verhindern, dass eine mögliche Vollständigkeit im Bestand jedes einzelnen Repositoriums erreicht wird, und zeichnen daher für Wissenschaftler ein sehr unübersichtliches Bild der Repositorienlandschaft.

Auf der Ebene der technischen Infrastruktur sollte die Förderung daher auf die Integration unterschiedlicher, vorhandener Open Access Angebote durch den Austausch von Erschließungsdaten abzielen, so dass es unerheblich ist, ob der Wissenschaftler seine

Publikation in ein fachliches oder institutionelles Repositorium einstellt. In beiden Fällen soll der Nutzer in jedem Repositorium entsprechend dessen inhaltlichen Zwecks einen möglichst vollständigen Bestand vorfinden. Zur Integration der Repositorien in die von den Nutzern bei ihrer Arbeit verwendeten Infrastrukturen stellen die Repositorien geeignete abstrakte Schnittstellen bereit, sodass z. B. aus dem Repositorium generierte Literaturlisten dynamisch auf persönlichen Webseiten der Forscher eingebunden werden können. Solche Technologien sowie die Verwendung der Repositoriendaten als Linked Open Data erfordern allerdings die Verwendung standardisierter Metadaten und normierten Vokabulars. Technische Voraussetzungen und organisatorische Abläufe müssten ggf. angepasst werden. Insbesondere die Autorentifizierung und ihre datenschutzgerechte Ausgestaltung sind hier zentrale Aspekte. Die Zusammenführung von Informationen über Open Access Angebote bedingt stets die Zusammenarbeit un-

terschiedlicher Akteure. Es sollten daher vor allem Kooperationen von Informationsanbietern bzw. von Informationsanbietern mit den Wissenschaftlern gefördert werden, um zu verhindern, dass Insellösungen geschaffen bzw. am Bedarf der Nutzer vorbei geplant werden. Neben dem freien Zugriff auf Publikationen, welcher das Hauptinteresse des Open Access Konzeptes darstellt, entwickelt sich auch der freie Zugriff auf Forschungsdaten sowie deren Verknüpfung mit den zugehörigen Publikationen zunehmend zu einem Thema (Schnittstelle zu AG 4)<sup>46</sup>. Daher sollte die Zusammenarbeit von Open Access Anbietern mit Betreibern von Forschungsdaten-Repositoryn sowie die Koordination von Standardisierungsbemühungen auf Publikations- und Datenseite gefördert werden. In einer gemischt zentral-dezentral ausgerichteten Organisationsstruktur wären die Digitalisate dezentral und die Metadaten zentral zu verorten.

In den vergangenen Jahren sind in etlichen Fach-Communities überregionale fachliche Open Access Repositoryn eingerichtet worden, die als Dienstleistungen für Wissenschaftler angeboten werden. Ein Beispiel ist das Social Science Open Access Repository SSOPEN ACCESSR, welches in den Jahren 2007 und 2008 im Rahmen eines von der DFG geförderten Projektes eingerichtet und seit 2008 vom Leibniz Institut für Sozialwissenschaften betrieben und weiterentwickelt wird. Ein Beispiel für ein internationales Preprint Repository im Fach Physik ist

„arXiv“, das seit 1991 betrieben wird und sich inzwischen weltweit zu dem zentralen Open Access Angebot des gesamten Faches entwickelt hat. Nachdem die Finanzierung über viele Jahre von einzelnen Institutionen (Los Alamos, aktuell die Bibliothek der Cornell University) getragen wurde, sucht man jetzt nach einer internationalen Basis der Finanzierung. An diesem Beispiel wird deutlich, dass die Entwicklung von fachspezifischen Open Access Repositoryn besonders effektiv gestaltet werden kann. Dennoch muss auch für eine nachhaltige Entwicklung dieser Plattformen eine Grundfinanzierung gesichert sein<sup>47</sup>. Ähnliche Lösungen existieren auch in anderen Bereichen. Zusätzlich besteht bereits eine Vielzahl institutioneller Repositoryn, die Publikationen einer bestimmten Forschungseinrichtung frei zugänglich anbieten. Die Universität Mannheim betreibt zum Beispiel den Mannheim Electronic Document Server (MADOC), der technische Berichte und Arbeitspapiere von Forschern der Universität frei zugänglich macht.

Aktuell besteht weniger ein Bedarf, die Einrichtung neuer Open Access Repositoryn zu fördern, zumal noch bestehende Lücken z. B. bei institutionellen Repositoryn durch das Hostingangebot des BSZ geschlossen werden können. Vielmehr besteht derzeit das Problem der Fragmentierung der Informationsangebote im Open Access Umfeld, da grundsätzlich für jede Publikation sowohl, u. U. verschiedene, fachliche Repositoryn als auch ein institutionelles Repository in Frage kommen und sich die Vorhaltung insofern unübersichtlich verteilt. Soweit die standardisierten Schnittstellen XMetaDissPlus der DNB sowie OAI bedient werden, ist immerhin eine einheitliche Auffind- und Zugreifbarkeit der Publikationen in Verbund- und Bibliothekskatalogen, überregionalen Diensten wie BASE, OAIster oder Google Scholar gewährleistet.

## ANHANG 6: RECHTSRAHMEN VON OPEN ACCESS<sup>48</sup>

Open Access-Strategien werfen in ihrem Kern eine Reihe urheberrechtlicher Fragestellungen auf. Je nachdem, ob der „goldene Weg“ oder der „grüne Weg“ gewählt wird, sind diese Fragestellungen zum

Teil identisch, zum Teil jedoch auch abweichend. Weiterhin ist beim „goldenen Weg“ zu unterscheiden, ob die Open Access-Veröffentlichung durch einen Verlag erfolgt oder von den Wissenschaftseinrichtungen selbst in Eigenregie übernommen wird.

### 1. „GOLDENER WEG“

„Golden“ Open Access erscheint aus urheberrechtlicher Sicht vergleichsweise unproblematisch. Bei diesem Geschäftsmodell lassen sich die Verlage von den Autoren der Beiträge – wie schon bisher beim klassischen Subskriptionsmodell – diejenigen Rechte einräumen, die sie benötigen, um die Artikel im Rahmen ihres digitalen Verlagsangebots online stellen zu können.

Allerdings ist dafür zu sorgen, dass die Autoren, die ihre Beiträge den Verlagen zur Veröffentlichung im Wege des „Golden Open Access“ überlassen und ihnen die dazu erforderlichen Rechte übertragen, diejenigen Rechte zurückbehalten, die zur Nachnutzung durch den Autor bzw. seine Institution erforderlich sind. Hier wäre zuvor zu definieren, welche Nachnutzungen in Betracht kommen und welche der in Betracht kommenden Nachnutzungen unverzichtbar sind. Eine solche Freiheit der Nachnutzung ließe sich ebenfalls durch eine gesetzliche Lösung absichern. Allerdings wäre eine gesetzliche Lösung nur auf Bundesebene durch entsprechende Änderung des UrhG möglich. Ohne eine solche Regelung müsste ein Katalog derjenigen Nachnutzungsrechte erstellt werden, die der einzelne Autor in seinem Vertrag mit dem Open Access-Verlag zurückbehalten soll/muss. Soweit „Golden Open Access“ nicht durch Verlage, sondern von den Bildungseinrichtungen selbst durchgeführt wird, lassen sich die Rechte einer Nachnutzung durch den einzelnen Urheber dagegen sehr viel einfacher auf vertraglichem Wege regeln.

Damit der Zugang kostenfrei ausgestaltet werden kann, müssen die Geldflüsse umstrukturiert werden: anstatt einer Zahlung allein durch die abnehmenden Bibliotheken (oder ggf. durch die aufgenommenen Autoren und die abnehmenden Bibliotheken) würden hier entweder die einzelnen Autoren selbst oder aber die Institutionen, denen diese angehören, für die mit der Veröffentlichung verbundenen Dienstleistungen der Verlage zahlen. Das Problem dabei ist aus der Sicht des Wissenschaftsbetriebes ein Zweifaches: Zum einen setzte dies eine vermutlich tiefgreifende Umstrukturierung interner Geldflüsse voraus. Zum anderen könnten die Verlage über den Preis, den ein Autor für die Veröffentlichung seines Beitrages zu zahlen hätte, ebenso wie schon bisher bei der Festsetzung der Subskriptionspreise auf die Ge-

samtkosten Einfluss nehmen, die der Wissenschaftsbetrieb für die Veröffentlichung von Forschungsergebnissen aufbringen müsste.

### 2. „GRÜNER WEG“

Urheberrechtlich komplexer ist die Situation, wenn der Autor die entsprechenden Beiträge selbst archiviert (auf einer individuellen Seite oder aber in einem institutionellen Repository). Hier ist im Wesentlichen zu unterscheiden zwischen Preprint- und Postprint-Veröffentlichungen.

Preprint Veröffentlichungen sind aus urheberrechtlicher Sicht zunächst unproblematisch, da und solange der Autor noch alle seine Urheberrechte hat. So stellt sich allein die Frage, ob die Verlage bereit sind, Artikel zur Veröffentlichung in ihren subskriptionspflichtigen Journals anzubieten, die bereits andernorts in einer Vorversion veröffentlicht worden sind. Wo dies für einen Verlag gleichwohl attraktiv erscheint, wird/muss er sich mit der nicht-exklusiven Rechteeinräumung zufrieden geben (was nicht ausschließt, dass der Verlag sich hier alle verbleibenden Rechte einräumen lässt, um sicherzustellen, dass derselbe Artikel nicht auch noch in einem anderen Journal erscheint).

Postprint Veröffentlichungen durch den Autor setzen hingegen voraus, dass der Autor die dazu erforderlichen Rechte nicht bereits zuvor dem Verlag eingeräumt hat. In aller Regel lassen sich die Verlage die Rechte jedoch umfassend einräumen (s. auch § 38 Abs. 1 Satz 1 des deutschen Urheberrechtsgesetzes (UrhG): „im Zweifel“ ausschließliches Nutzungsrecht des Herausgebers). Nach § 38 Abs. 1 Satz 2 UrhG darf der Urheber das Werk in diesem Fall zwar nach Ablauf eines Jahres seit Erscheinen anderweitig vervielfältigen und verbreiten, doch

gilt dies nur dann, wenn vertraglich nichts anderes vereinbart ist. Dieses Zweitveröffentlichungsrecht des Autors ist also abdingbar und es wird in den Verlagsverträgen auch regelmäßig abbedungen bzw. ausgeschlossen. Eine gesetzliche Regelung im Urhebergesetz, die den Autoren ein solches Zweitveröffentlichungsrecht gesetzlich garantieren würde, könnte nur vom Bundesgesetzgeber, nicht hingegen auf Landesebene vorgenommen werden.

Ein solches gesetzliches Zweitveröffentlichungsrecht stellt keine – international nur unter eingeschränkten Bedingungen mögliche – Beschränkung urheberrechtlicher Ausschließlichkeitsbefugnisse dar. Vielmehr würde sie den Urheber davor schützen, allzu viele Rechte pauschal an die Verlage abzutreten. Näherer Prüfung bedarf in diesem Zusammenhang wohl noch die Frage, ob bzw. inwieweit eine solche deutsche Lösung aus rechtlichen Gründen den Einsatz von Geolocation-Sperren erforderlich macht, mit denen sichergestellt würde, dass auf in deutschen Repositories vorgehaltene Postprint-Veröffentlichungen nur von Deutschland aus zugegriffen werden kann.

Kommt eine gesetzliche Lösung hinsichtlich eines Zweitverwertungsrechts nicht zustande, oder erfass-

te sie im Falle ihres Zustandekommens lediglich Verlagsverträge mit deutschen Verlegern, so ließe sich das Zweitverwertungsrecht wohl nur dann erreichen, wenn die Verlage dazu bewegt werden könnten, ihre bisherige Vertragspraxis zu ändern.

In wirtschaftlicher Hinsicht hätten die Verlage zwar wohl nichts gegen eine Selbstarchivierung einzelner Artikel durch individuelle Autoren, sofern dabei die Embargo-Fristen der Sherpa/Romeo-Liste<sup>49</sup> eingehalten werden. Die Liste ist allerdings nicht rechtsverbindlich. Die Verlage wehren sich jedoch strikt gegen eine umfassende Archivierung in Repositorien. Dabei gestehen sie zwar zu, dass mit Postprint Online-Veröffentlichungen kein nennenswerter Umsatz zu erzielen ist, sie sehen jedoch in umfassenden Postprint Repositorien eine Gefahr für die Wirtschaftlichkeit der Primärveröffentlichungen.

Schließlich kann aufgrund der grundgesetzlich garantierten Wissenschaftsfreiheit den einzelnen Autoren zumindest im Rahmen ihrer Grundausstattung keine rechtliche Verpflichtung auferlegt werden, ihre Aufsätze als Open Access-Publikationen zu veröffentlichen. Das betrifft den „goldenen“ wie „grünen“ Open Access gleichermaßen.

Anders sieht es hingegen bei der Vergabe von Fördermitteln aus. Hier kann in den Förderrichtlinien aus öffentlichen Geldern geförderter Projekte auf eine Open Access-Veröffentlichung hingewirkt werden.

## ZU 2: BW-OPEN ACCESS FÖRDERPROGRAMM

### ANHANG 7: PILOTPROJEKT ZU HYBRIDEN OPEN ACCESS ZEITSCHRIFTEN

Als hybride Open Access Zeitschriften werden Zeitschriften bezeichnet, deren Geschäftsmodell grundsätzlich auf dem traditionellen Subskriptionsprinzip beruhen, die jedoch ihren Autoren das Angebot machen, ihre Artikel durch die Zahlung einer Open Access Publikationsgebühr von der mit dem Subskriptionsprinzip verbundenen Zu-

gangsbarriere auszunehmen. Sie unterscheiden sich damit von genuinen Open Access Zeitschriften, deren Inhalte vollständig über ein Open Access kompatibles Geschäftsmodell finanziert und damit entsprechend zugänglich gemacht werden.

Alle großen Wissenschaftsverlage nutzen das hybride Geschäftsmodell. Das Angebot der Freistellung einzelner Artikel gegen die Zah-

lung einer Publikationsgebühr wird jedoch nur in sehr geringem Umfang angenommen. Kritiker werfen den Verlagen vor, sich ihre Inhalte mit diesem Geschäftsmodell doppelt bezahlen zu lassen (double dipping). Dem halten die Verlage entgegen, die aus den Open Access Publikationsgebühren erzielten Einnahmen würden bei der Preisgestaltung der Subskriptionsgebühren verrechnet. Diese Behauptung ist ohne einen vollständigen Einblick in die Bücher eines Verlages nicht überprüfbar. Solch eine Offenlegung hat bisher nicht stattgefunden, es gibt jedoch entsprechende Angebote.

Mit hybriden Geschäftsmodellen ist die Möglichkeit verbunden, Open Access kompatibles Publizieren unter Beibehaltung der bestehenden Publikationsgewohnheiten kurzfristig zu realisieren. Darüber hinaus besteht die Aussicht, mit Hilfe dieses Geschäftsmodell-Zeitschriften von einem subskriptionsbasierten zu einem Open Access kompatiblen Geschäftsmodell zu überführen. Um solch ein Projekt aus der Perspektive der Wissenschaftsorganisationen attraktiv zu machen, müssten zumindest folgende Bedingungen erfüllt werden:

- Der Verlag strebt die Umstellung des Geschäftsmodell an und implementiert dies nachprüfbar durch entsprechende Policies.
- Die Wissenschaftsorganisationen öffnen ihre Open Access Publikationsfonds für die Zeitschrift.
- Die Preisbildung der Zeitschrift ist überprüfbar.
- Die Höhe der Preise wird von beiden Seiten als angemessen bewertet.

Als Kandidaten für solch ein Projekt würden sich renommierte Zeitschriften anbieten, deren Autorentschaft im Wesentlichen aus Deutschland stammt, so dass die Umstellungsbedingungen weitgehend von deutschen Akteuren gestaltet werden können.

## ANHANG 8: MEILENSTEINE

Meilensteine für ein Fünfjahresprogramm:

- Leuchtturm (langfristige Finanzierung)
  - M01-M12: Prüfung in Frage kommender Projekte
  - M13-M18: Entscheidung, ob ein aussichtsreiches Projekt identifiziert werden konnte.
- Infrastruktur für die elektronische wissenschaftliche Kommunikation
  - Für Planung von Meilensteinen ist eine Abstimmung mit den AGs Forschungsdatenmanagement und Virtuelle Forschungsumgebung erforderlich.
- Open Access Publikationsfonds – goldener Weg
  - M01-M06 Erarbeitung von Kriterien zur Teilnahme am Programm für einen Publikationsfonds
  - M07-M09 Ausschreibung der Teilnahme
  - M10-M12 Entscheidung über Teilnahme
  - M13-M18 Erarbeitung von Kriterien für die Kostenübernahme
  - M13-M18 Aufbau des Monitoring
- Zukunftsfähige Geschäftsmodelle im Land Baden-Württemberg
  - M01-M06 Identifikation von Akteuren
  - M07-M12 Identifikation konkreter Projekte

## ANHANG 9: LITERATURVERZEICHNIS

- **AG OPEN ACCESS in der Allianz** „Open Access Positionen Prozesse Perspektiven“ (Oktober 2009)
- **WR** – Übergreifende Empfehlungen zu Informationsinfrastrukturen (28.01.2011)
- **GWK** – Gesamtkonzept für die Informationsinfrastruktur in Deutschland – KII (April 2011)
- **FAQ** – zu Open Access und Zweitveröffentlichungsrecht (**Allianz**) vom 30.06.2011
- Zur urheberrechtlichen Gestaltung von Repositorien (**IUWIS**-Projekt, 2011)
- **AG OPEN ACCESS in der Allianz** „Open Access Positionen Prozesse Perspektiven“ (Oktober 2009)
- **DFG** – Die digitale Transformation weiter gestalten (Juli 2012)
- **WR** – Weiterentwicklung IIS (13.07.2012)
- **EU** – Mitteilung vom 17. Juli 2012 – Verbesserung des Zugangs zu wissenschaftlichen Informationen – Steigerung der Wirkung öffentlicher Investitionen in die Forschung
- **EU** – Empfehlung der Kommission über den Zugang zu wissenschaftlichen Informationen und deren Bewahrung vom 17. Juli 2012 (2012/417/EU)

<sup>23</sup> Die Landesregierung orientiert ihr Verständnis der freien Zugänglichkeit / des Open Access an der Definition in der Berliner Erklärung über offenen Zugang zu wissenschaftlichem Wissen (<http://oa.mpg.de/lang/de/berlin-prozess/berliner-erklarung/>). Unter dem goldenen Weg wird Open Access für Erstveröffentlichungen verstanden, der grüne Weg bezieht sich in erster Linie auf Zweitveröffentlichungen.

<sup>24</sup> <http://oa.mpg.de/lang/de/berlin-prozess/signatoren>

<sup>25</sup> Der Begriff Forschungsergebnisse umfasst wissenschaftliche Publikationen und Forschungsdaten gleichermaßen.

<sup>26</sup> Neelie Kroos - EU-Kommissarin für Digitale Agenda - verkündete am 17. Juli 2012 das Ziel, dass bis 2016 der Großteil der mit EU-Mitteln finanzierten Forschungsergebnisse öffentlich und kostenfrei verfügbar sein müssen (SZ vom 19. Juli 2012).

<sup>27</sup> „Finch Report“ der „Working Group on Expanding Access to Published Research Findings“ vom 18. Juni 2012.

<sup>28</sup> Früher European Heads of Research Councils - EUROHORCS

<sup>29</sup> League of European Research Universities.

<sup>30</sup> Schlussfolgerungen des europäischen Rates vom 20. Mai 2008, Dokument Nr. 7651/1/08, Mitteilung der Kommission an das Europäische Parlament vom 17. Juli 2012 - Verbesserung des Zugangs zu wissenschaftlichen Informationen: Steigerung der Wirkung öffentlicher Investitionen in die Forschung (COM (2012) 401 final.) und Empfehlungen der Kommission vom 17. Juli 2012 (2012/417/EU).

<sup>31</sup> Beschluss vom 21.05.2012: „Im Bereich digitale Medien als Gegenstand von Forschung und Innovation diskutierten die Mitglieder unter anderem Handlungsempfehlungen zum Zweitveröffentlichungsrecht und zu Open Access. Die Projektgruppe empfiehlt übereinstimmend, Open Access im Wissenschaftsbereich umfassend zu unterstützen. Als wichtiges Signal bezeichnen es die Mitglieder beispielsweise, ein verbindliches Zweitveröffentlichungsrecht anzustreben, so dass in Periodika oder Sammelbänden publizierte wissenschaftliche Beiträge aus überwiegend öffentlich finanzierter Forschung im Internet frei und dauerhaft zugänglich gemacht werden könnten. Um Open Access im Wissenschaftsbereich zu fördern, so die Projektgruppe, müsse insbesondere auch die Veröffentlichung und Nachnutzbarkeit von Forschungsergebnissen umfassend abgesichert werden. Dazu sollen die rechtlichen Voraussetzungen für Open Access verbessert werden. In einer intensiven Diskussion einigten sich die Mitglieder auf eine Reihe gemeinsamer Empfehlungen.“ [http://www.bundestag.de/internetenquete/dokumentation/Bildung\\_und\\_Forschung/PGBuF\\_2012-05-21/PGBuF\\_2012-05-21\\_Bericht/index.jsp](http://www.bundestag.de/internetenquete/dokumentation/Bildung_und_Forschung/PGBuF_2012-05-21/PGBuF_2012-05-21_Bericht/index.jsp)

- <sup>32</sup> Positionspapier der KMK vom 8. Juni 2009, dem BMJ in Rahmen der Konsultationen für einen „Dritten Korb“ vorgelegt. Vgl. hierzu Pflüger, Positionen der Kultusministerkonferenz zum Dritten Gesetz zur Regelung des Urheberrechts in der Informationsgesellschaft - „Dritter Korb“, ZUM 2010, 938 (941).
- <sup>33</sup> Papier der Allianz vom 7. Juli 2010 „Neuregelung des Urheberrechts: Anliegen und Desiderate für einen Dritten Korb“.
- <sup>34</sup> Vgl. hierzu Pflüger, Die Befristung von § 52a UrhG - eine (un)endliche Geschichte?, ZUM 2012, S. 444 (447f.) und Anhang 1. Zum Themenbereich „Abbestellungen“ vgl. Anhang 2.
- <sup>35</sup> Beispiele hierfür sind Publikationsrepositorien wie arXiv oder PubMedCentral oder das Datenrepositorium PANGAEA.
- <sup>36</sup> In der Arbeitsgruppe „Open Access“ wurde auch die Einrichtung eines Open Access Repositoriums für Lehrmaterialien diskutiert. In einer solchen Sammlung sollten E-Learning Module gesammelt und veröffentlicht werden, deren Qualität durch ein Peer-Review Verfahren gesichert ist. Die Qualitätssicherung soll auch dazu dienen, dass die Autoren dieser Beiträge ihre Kompetenz im Bereich der Lehre dokumentieren können. Mittelfristig könnten insbesondere Bewerber für eine Professur von einem solchen Nachweis profitieren. Gleichzeitig könnte dadurch die Bedeutung der Lehre in Berufungsverfahren gestärkt werden. Die Arbeitsgruppe verfolgt diesen Ansatz hier nicht weiter, da dies den Rahmen des Auftrags sprengen würde.
- <sup>37</sup> Mit dem finanziellen Bedarf des Förderprogramms hat sich die AG nicht näher befasst. Legt man jedoch einen von der KMK in 2007 für die Anmeldung im Hochschulpakt ermittelten Finanzierungsrahmen für ein bundesweites Open Access-Infrastrukturprogramm in Höhe von 20 Mio. Euro p.a. zugrunde, entfielen auf BW in Anlehnung an den Königsteiner-Schlüssel ca. 2,7 Mio. Euro p.a..
- <sup>38</sup> <http://www.dataealofapproval.org/>
- <sup>39</sup> <http://www.trusteddigitalrepository.eu/Site/Trusted%20Digital%20Repository.html>
- <sup>40</sup> Könnte etwa über das Center of Digital Tradition (CODIGT, vormals: Kompetenzzentrum für kulturelle Überlieferung - digital Karlsruhe (KÜd-Ka)) oder die Geschäftsstelle des Konsortiums BW operationalisiert werden.
- <sup>41</sup> Deutsche Universitätszeitung (DUZ) 11/2010, S. 10f.
- <sup>42</sup> Deutsche Bibliotheksstatistik für das jeweilige Jahr, veröffentlicht vom HBZ Nordrhein-Westfalen, [http://www.hbz-nrw.de/dokumentencenter/produkte/dbs/aktuell/auswertungen/wb\\_gesamt\\_10.pdf](http://www.hbz-nrw.de/dokumentencenter/produkte/dbs/aktuell/auswertungen/wb_gesamt_10.pdf).
- <sup>43</sup> Dies entspricht etwas mehr als 3 % des Gesamtumsatzes auf dem Buchmarkt in Deutschland in Höhe von 9,7 Mrd. € in 2010, [www.boersenverein.de/de/158286](http://www.boersenverein.de/de/158286) (Abruf 13. April 2012).
- <sup>44</sup> Gezählt wurden Datensätze ohne Volltext, Datensätze mit Volltext ohne Zugänglichkeit, Datensätze mit Volltext mit Zugänglichkeit ohne Lizenz sowie Datensätze mit Volltext mit Zugänglichkeit mit Lizenz.
- <sup>45</sup> Zum Stand vgl. Anhang 4.
- <sup>46</sup> Die AG 4 (Forschungsdatenmanagement) behandelt die Einrichtung von Forschungsdatenrepositorien und die angesprochene Verknüpfung zu Publikationen. Der freie Zugriff auf Forschungsdatenbanken der Verlage ist ein weiteres Problem, dass nicht im Rahmen der AG Open Access behandelt wird. Die derzeitige Lizenzierungspraxis von Forschungsdatenbanken ist Gegenstand der AG 1 (Lizenzierung).
- <sup>47</sup> Empfehlung des Wissenschaftsrates 2359-12 „Empfehlungen zur Weiterentwicklung der wissenschaftlichen Informationsinfrastrukturen in Deutschland“, Seite 10-11.
- <sup>48</sup> Ergänzend wird auf die IUWIS-Publikation „Zur urheberrechtlichen Gestaltung von Repositorien – Handreichung für Universitäten, Forschungszentren und andere Bildungseinrichtungen“ verwiesen, [www.iuwis.de](http://www.iuwis.de).
- <sup>49</sup> <http://www.sherpa.ac.uk/romeo/>



# AG 4 · Forschungsdatenmanagement

## INHALT

|  |    |
|--|----|
| 1. Warum Management von Forschungsdaten? .....                             | 81 |
| 2. Definition .....  | 82 |
| 3. Nutzererwartungen .....   | 83 |
| 4. Nationale und europäische Empfehlungen und Positionspapiere .....       | 85 |
| 5. Status quo in Baden-Württemberg .....                                   | 86 |
| 6. Maßnahmen für die Hochschulen in Baden-Württemberg .....                | 88 |
| 6.1 Weiterentwicklung der technischen Infrastruktur .....                  | 88 |
| 6.2 Integration in die Lehre .....   | 88 |
| 6.3 Förderprogramm zum Infrastrukturaufbau .....                           | 88 |
| 6.4 Förderprogramm Data Life Cycle Labs .....                              | 90 |
| 6.5 Einrichtung einer landesweiten Koordination .....                      | 90 |
| Projekte und Best Practices  |    |
| Anhang 1: Nationale Projekte mit baden-württembergischer Beteiligung ..... | 90 |
| Anhang 2: Europäische Infrastrukturprojekte .....                          | 91 |
| Anhang 3: Globale Projekte .....   | 92 |
| Anhang 4: Forschungsdatenmanagement bei GESIS .....                        | 92 |

## MITGLIEDER DER ARBEITSGRUPPE

**DR. CHRISTIAN HÄNGER**

Bibliothek, Universität Mannheim

**BRIGITTE HAUSSTEIN**

GESIS – Leibniz-Institut für Sozialwissenschaften

**PROF. MICHAEL HEBGEN**

Rechenzentrum, Universität Heidelberg

**DR. MARKUS KLEIN**

Ministerium für Wissenschaft, Forschung und Kunst  
Baden-Württemberg (Vorsitz)

**MATTHIAS RAZUM**

FIZ Karlsruhe - Leibniz-Institut für  
Informationsinfrastruktur

**PETER REMPIS**

Bibliothek, Universität Tübingen

**PROF. DR. GERHARD SCHNEIDER**

Rechenzentrum, Universität Freiburg

**DR. MATTHIAS SCHULZE**

Bibliothek, Universität Stuttgart

**PROF. DR. ACHIM STREIT**

Steinbuch Computing Center, KIT

**PROF. DR. VOLKER WULFMEYER**

Institut für Physik und Meteorologie,  
Universität Hohenheim

## 1 · WARUM MANAGEMENT VON FORSCHUNGSDATEN?

Digitale Werkzeuge haben die Methoden der Wissenschaft revolutioniert: Die automatisierte Erzeugung und Analyse von Daten in vielen Bereichen der Wissenschaft hat sehr stark zugenommen, so dass sich die Methoden der „data-driven science“ – teilweise mit extrem großen Datenmengen („big data“) – in allen Wissenschaftsdisziplinen etabliert haben. Beispiele sind die empirischen Sozial- und Wirtschaftswissenschaften, die Daten in Umfragen und Beobachtungen erheben und sie statistisch analysieren. Natur- und Geowissenschaften vergleichen Daten, die in Experimenten, punktuellen Messungen, Monitorings oder Satellitenaufnahmen gemessen wurden, mit simulierten Daten einer modellhaften Rekonstruktion der Wirklichkeit. Medizin und Biologie erzeugen mit bildgebenden Verfahren und Gensequenzen riesige Datenmengen, die teilweise persönlichkeits- und datenschutzrelevant sind. Sprachwissenschaften analysieren Korpora verschriftlichter Sprache und Streams gesprochener Sprache. Die steigenden Anforderungen guter wissenschaftlicher Praxis und der Qualitätssicherung ihrer Ergebnisse verlangen eine zuverlässige Bewahrung der Primärdaten nach festgelegten (Mindest-)Standards. Idealerweise werden die Primärdaten, wenn keine berechtigten Schutzinteressen dem entgegenstehen, auf Open Access Repositorien veröffentlicht, damit Dritte die Ergebnisse sowie die zugrunde liegende wissenschaftliche Methodik und Argumentation nachvollziehen können.

Häufig sind die zugrundeliegenden Daten unwiederbringlich in dem Sinne, dass (i.) sie zeitabhängige bspw. soziale, klimatologische oder individuelle medizinische Zustände repräsentieren, die im Nachhinein nicht mehr beobachtet und gemessen werden können, oder dass (ii.) ihre erneute Gewinnung unangemessen aufwendig ist. Wissenschaft, die

Phänomene beschreibt, ihre Bedingungen analysiert und kausale Zusammenhänge erforscht, ist auf solche Daten angewiesen, nicht nur im Rahmen der Fragestellungen, für die sie erhoben wurden, sondern auch im Sinne von Nachnutzung, Analyse und Deutung in neuen, auch transdisziplinären Kontexten. Die Bewahrung der Daten erfordert im digitalen Zeitalter Design und Umsetzung der Prozesskette des Datenmanagements. Während digitale Daten der Apollo-Mission heute teilweise nicht mehr lesbar sind, verfügten schon die Babylonier über astronomische Aufzeichnungen von sechs Jahrhunderten. Das Auffinden von Daten setzt in dem von hoher Fluktuation geprägten Hochschulkosmos gesicherte Prozesse bei der Ablage der Daten voraus: Angabe standardisierter Metadaten und Qualitätskriterien, Beachtung rechtlicher Veröffentlichungsbedingungen, Speicherung über eine definierte Zeitspanne und Interpretierbarkeit der Daten.

Daher ist es das Bestreben des Ministeriums für Wissenschaft, Forschung und Kunst Baden-Württemberg (MWK) – gemeinsam mit den Universitäten, weiteren Forschungseinrichtungen und Forschungsverbänden – eine Forschungsdaten-Strategie zu entwickeln und alle Wissenschaftlerinnen und Wissenschaftler, die im Geschäftsbereich des MWK arbeiten, mit dem Zugriff auf die notwendigen Prozesse und die darunter liegende technische Infrastruktur zu versorgen. Der Wert wissenschaftlicher Daten im Sinne von Wissensrohstoff muss der Wissenschaft erhalten bleiben. Dadurch soll die strategische Wettbewerbsfähigkeit der Hochschulen und Forschungseinrichtungen bei der Einwerbung von Forschungsmitteln nachhaltig weiterentwickelt und der Wissenschaftsstandort Baden-Württemberg gestärkt werden.

## 2 · DEFINITION

Die DFG definiert Forschungsdaten als „digitale und elektronisch speicherbare Daten [...], die im Zuge eines wissenschaftlichen Vorhabens z. B. durch Quellenforschungen, Experimente, Messungen, Erhebungen oder Befragungen entstehen“<sup>50</sup>. Sie reichen von strukturierten Massendaten bis hin zu unstrukturierten (d.h. nur begrenzt maschinell auswertbaren) Daten unterschiedlichsten Formats und Inhalts. Forschungsdaten können auch Prozessdaten wie Bearbeitungsschritte, in einem Forschungsvorhaben entwickelte Algorithmen und Programme sein. Schließlich können auch Bücher und andere Schriften, Bilder und andere nicht-textuelle Materialien mittels Digitalisierung zu Forschungsdaten werden und müssen entsprechend gemanagt werden<sup>51</sup>, vgl. E-Science-AG 2 Digitalisierung.

Der Open Access-Gedanke bedeutet im Zusammenhang mit Forschungsdaten einen Paradigmenwechsel in der Wissenschaftskultur. Die Berliner Erklärung über offenen Zugang zu wissenschaftlichem Wissen dokumentiert den Willen der Unterzeichner, nicht nur wissenschaftliche Ergebnisse in textueller Form, sondern auch Forschungsdaten Open Access zu publizieren<sup>52</sup>, vgl. E-Science-AG 3 Open Access. Dies setzt nicht nur das dafür notwendige Datenmanagement voraus, sondern auch die Etablierung der neuen Publikationsform Forschungsdaten, die erst in wenigen Communities wie Geowissenschaften und Astronomie vorangeschritten ist. In anderen Disziplinen, erst recht wenn sie mit personenbezogenen Daten operieren, müssen Formate und Anreizsysteme erst entstehen. Die Veröffentlichung erfolgt in Repositorien oder Datenbanken, zu denen idealerweise ein einheitlicher strukturierter Zugriff erfolgt und die Basis für darauf aufbauende Dienste legt wie z. B. Serviceorientierte Architekturen für Web-Portale, Apps, Synchronisationsdienste, Virtuelle Forschungsumgebungen.

Forschungsdaten unterliegen typischerweise einem Lebenszyklus, der viele Schritte des Anreicherns und Aussortierens enthält: von Rohdaten über geprüfte Mess- und Erhebungsergebnisse, verschiedene Stufen der Aufbereitung und Analyse bis hin zu Datensätzen in archivierungsfähiger Form. Analog wächst die Nutzerdomäne von

persönlicher Nutzung über Nutzung in der Institution bis idealerweise offener Kenntnisnahme und Nachnutzung. Während temporäre Zwischenergebnisse verworfen werden, müssen Primärdaten und finale Ergebnisse, die Grundlage wissenschaftlicher Publikationen sind, im Sinne „guter wissenschaftlicher Praxis“ und nach den Richtlinien der Fördermittelgeber nach definierten Standards mit Metadaten dokumentiert und nachhaltig gesichert werden.<sup>53</sup> Umfang, Formate, Konsistenz, Qualität, Veröffentlichungsform und Lizenzen (z. B. cc-by), Aufbewahrungsdauer u.a. werden im Idealfall von den Communities, Forschungs- und Fördereinrichtungen festgelegt. Dies kann durch die Entwicklung koordinierter FDM-Richtlinien und gemeinsamer Best Practice Empfehlungen auf Basis landesweit verfügbarer und für die Forscher leicht nutzbarer Infrastruktur geschehen. Die Herausforderung besteht dabei in der Gratwanderung der genügenden Berücksichtigung disziplinspezifischer Anforderungen und dem Setzen allgemeingültiger Regeln, gerade auch im Hinblick auf die gewünschte und zunehmende Interdisziplinarität zwischen Natur-, Geistes- und Lebenswissenschaften. Die gleiche Balance zwischen disziplinspezifischen und -übergreifenden Anforderungen muss bei der Entwicklung von Metadatenschemata für Forschungsdaten gefunden werden, da neben bibliographischen und technischen Attributen möglicherweise auch disziplinspezifische Attribute für die Beschreibung von Inhalt, Kontext und Provenienz der Daten erhoben werden müssen. Um Metadaten von gleichbleibend hoher Qualität zu erhalten, die Voraussetzung für die Auffindbarkeit und Nachnutzung von Forschungsdaten sind, muss der Übergang von der persönlichen in die offene Domäne durch einen Workflow unterstützt werden, der die administrativen Anforderun-

gen an die Forschenden durch Beratung, Betreuung und Übernahme von organisatorischen Aufgaben minimiert. Der Zugang der Wissenschaftler auf die Forschungsdaten muss technisch und organisatorisch

so konzipiert sein, dass die Nutzer über benutzerfreundliche Portale eine Meta-Suche in mehreren Fach- und Daten-Repositoryen durchführen. Voraussetzung dafür sind standardisierte und vernetzte Metadaten in diesen Repositoryen.

### 3 · NUTZERERWARTUNGEN

In allen Wissenschaftsdisziplinen steigt die Menge der Daten, die durch Experimente, Beobachtungen oder Simulationen produziert werden, rasant an. Ohne diese Daten ist Spitzenforschung in vielen Disziplinen nicht mehr vorstellbar. Einige exemplarische Beispiele zeigen mögliche Sichten und Bedürfnisse der Wissenschaftler auf:

- a) Hochenergiephysik: Die Physik hat mit die längste Erfahrung und die größten Bedürfnisse an riesigen Datenmengen. Hier werden nicht mehr Rohdaten gesammelt, sondern bereits im Detektor Daten ausgefiltert, nach denen man nicht sucht. Die Daten selbst liegen in einer mehrschichtigen Tier-Architektur vor, z. B. auf dem Tier-1-Speicher beim KIT. Physiker können sich von dort Working-Sets herunterladen und unter Nutzung von High Performance Clustern analysieren. Eine lokale Sicherung der Daten ist nicht nötig, da dank der performanten Infrastruktur Primärdaten zentral gehalten werden können. So werden lokal nur Ergebnisse gespeichert.
- b) Systembiologie: Hier fallen sehr große Datenmengen von Experimenten mit elektronischen Mikroskopen, DNA-Sequenzierern usw. an, die abgenommen und gesichert werden müssen. Danach sind die Daten über die verschiedenen Experimente zu normalisieren und so aufzubereiten, dass man Suchen und Working-Sets (10-100 TB) bilden kann.
- c) Geowissenschaften zeichnen sich durch enges Zusammenwirken verschiedener Disziplinen und Messmethoden (in-situ, Fernerkundung, Chemie, Biologie, Physik, usw.) aus, die permanent weltweit Daten erheben und austauschen. Ohne die Festlegung auf gemeinsame Datenzentren, -strukturen und -formate auf der Basis international vereinbarter Standards (z. B. Metadaten, Fehlertoleranzangaben, Open Source Software) wäre die Entwicklung globaler Klimamodelle und Studien wie der aktuelle IPCC-Report, in dem die Ergebnisse von 15 Modellen verglichen werden, unmöglich.
- d) GESIS ist der Zusammenschluss von drei Einrichtungen der Sozialwissenschaften, um über ein Portal einen neuen Dienst anzubieten. Bei der Suche in den Datenbanken stößt man auf einen interessanten Fall und kann dann die Datensätze sowie die dazu verwendeten Erhebungs- und Analysemethoden (einschließlich der zum Analysezeitpunkt verwendeten statistischen Software) erhalten.
- e) Das Institut für Deutsche Sprache (IDS) baut in einem Netzwerk mit anderen Institutionen ein Zentrum für die Bereitstellung digitaler Ressourcen für die sprachwissenschaftliche Forschung auf und erarbeitet Konventionen für die Speicherung und Verarbeitung germanistischer Forschungsdaten. Forschende können Analysen auf verschiedenen Korpora veröffentlichter Sprache (z. B. Zeitungen, Web-2.0-Streams) durchführen.
- f) In den Sonderforschungsbereichen der DFG werden Kompetenzen verschiedener Disziplinen und Fakultäten zusammengeführt. Die interdisziplinäre Zusammenarbeit unter einem ordnenden Aspekt bringt eine kreative Forschungsarbeit und Erkenntnisprünge hervor, die den Horizont eines einzelnen Lehrstuhls übersteigen. Nicht umsonst fördert die DFG mit

den „INF-Projekten“ Maßnahmen zur Kollaboration der Forschenden und zur Sicherung ihrer Ergebnisse und Daten. Sonderforschungsbereiche in Baden-Württemberg sollen die besten Voraussetzungen haben, sich um die zusätzlichen Mittel für ein INF-Projekt erfolgreich zu bewerben.

Auch unabhängig von den INF-Projekten müssen die Hochschulen bei Antragstellung die Compliance-Vorgaben von DFG, BMBF, EU usw. erfüllen.

In Anlehnung an die vier großen Handlungsfelder, die das Gesamtkonzept der KII unterscheidet, kann man den resultierenden Bedarf der Forschung strukturieren:

- 1) **Sicherung:** Für Massendaten wird eine Kaskade von Datenspeichern vorgehalten, die die Daten von Experimenten performant abnimmt und zentral sichert. Für Einzeldaten wie Massendaten stehen Workflows bereit, die die Daten teil- bzw. vollautomatisch mit ersten Metadaten anreichern und rudimentär erschließen. Die Zugriffsrechte sind in der ersten Phase noch sehr restriktiv.
- 2) **Erschließung und Bereitstellung:** Die Erschließung der Daten für Dritte liegt i.d.R. nicht im primären Fokus der Urheber. Daher muss die Erschließung organisatorisch und technisch durch die zuständigen Einrichtungen unterstützt werden. Die Daten müssen so dokumentiert sein, dass alle Nutzer Gegenstand, Provenienz, Methodik, Semantik, Fehlertoleranzen und Qualität zuverlässig beurteilen und die Datenformate weiterverarbeiten können. Der fachspezifische Teil der Metadaten muss aufgrund der Interdisziplinarität der Forschung über größere Communities abgestimmt werden. Die Daten müssen so dokumentiert sein, dass der Nutzer verstehen und nachvollziehen kann, wie und wofür sie erzeugt wurden. Erst dann kann man davon ausgehen, dass auch andere Nutzer Interesse haben, die Daten für ihre Projekte zu verwenden. Die Vergabe von permanenten Identifiern sichert dem Urheber die Unversehrtheit und Zitierfähigkeit seiner Daten.
- 3) **Nutzung:** Die Schnittstelle zu dritten Nutzern muss unter Gesichtspunkten der Usability entwickelt werden. Die Verknüpfung zwischen Daten und textuellen Publikationen wird zen-

trale Bedeutung gewinnen. Die Suche in den Metadaten muss über mehrere Repositorien unterstützt werden und zur Auswahl von Working-Sets führen. Im Hinblick auf die Nachnutzer müssen Datenformate gewählt werden, deren Strukturen vollständig offengelegt sind, damit sie von zukünftigen Nutzern interpretiert werden können. Proprietäre Formate stellen ein Risiko für die spätere Interpretierbarkeit der Daten dar. Falls sie unumgänglich sind, müssen Konvertierungsroutinen und Umgebungen (Emulatoren), wie sie z. B. im Landesprojekt „Funktionale Langzeitarchivierung“ (bwFLA) untersucht werden, vorgehalten und gesichert werden.

Ein Datenschutzkonzept ist notwendig. Insbesondere für personenbezogene Daten aus Medizin, Psychologie usw. müssen sichere Verfahren der Anonymisierung und Codierung angewendet werden, die sowohl die Belange der untersuchten Personen als auch der (wechselnden) Forscher berücksichtigt. Die technische Steuerung der Zugriffsrechte kann über verteilte Mechanismen zur Authentifizierung und zur Autorisierung (wie z. B. das verteilte Identity Management in Baden-Württemberg bwIDM) vereinfacht werden, um die Daten komfortabel an allen Universitäten verfügbar zu machen.

- 4) **Langfristige Bewahrung und Pflege:** Die langfristige physikalische Bewahrung und Pflege der „Bitstreams“ ist eine Aufgabe des Landesdatenspeichers bwLSDF. Der Anstieg des Datenvolumens ist nach wie vor ungebrochen exponentiell. Daher müssen Aufbewahrungsfristen bzw. Lösungsbedingungen vorab geklärt sein wie auch die Trägerschaft der damit verbundenen Kosten. Die potenziellen Interessen zukünftiger Forschung müssen von den Communities antizipiert werden.

#### 4 · NATIONALE UND EUROPÄISCHE EMPFEHLUNGEN UND POSITIONSPAPIERE

Im April 2011 wurde das „Gesamtkonzept für die Informationsinfrastruktur in Deutschland“ der **Kommission Zukunft der Informationsinfrastruktur (KII)** im Auftrag der GWK vorgelegt.<sup>54</sup> Für die Handlungsfelder Forschungsdaten und Langzeitarchivierung empfiehlt das Papier

- die Entwicklung von Richtlinien zum Datenmanagement, indem ein Datenmanagementplan mit klaren Verantwortlichkeiten erstellt wird,
- die Etablierung von Anreizsystemen für die Wissenschaftler zur Publikation ihrer Forschungsdaten,
- Förderprogramme zur Entwicklung von Best Practice Vorgehensmodellen,
- die Bereitstellung von Mitteln für den Aufbau und die Verstetigung von disziplingetriebenen, nutzerorientierten, international vernetzten Forschungsdaten-Infrastrukturen mit angemessenen Service Levels (Verfügbarkeitszielen),
- die Festlegung verbindlicher Standards und Verfahren mit Nachweissystemen, persistenten Identifiern, Schnittstellen usw.,
- die konzeptionelle Einbindung aller Datentypen und -formate in die Infrastruktur für Langzeitarchivierung,
- die Mitwirkung bei der rechtlichen Absicherung,
- die Verankerung in der Lehre: Schlüsselqualifikationen in der Breite und Bildungswege zum Daten-Kurator im Speziellen.

Der **Wissenschaftsrat** schließt sich im Juli 2012 mit seinen „Empfehlungen zur Weiterentwicklung der wissenschaftlichen Informationsinfrastrukturen in Deutschland bis 2020“<sup>55</sup> den KII-Empfehlungen an. Sie empfehlen darüber hinaus, dass

- Infrastrukturen konsortial und spartenübergrei-

fend weiterentwickelt werden,

- für die wissenschaftliche und nicht-kommerzielle Nutzung der Infrastrukturen keine oder geringe Gebühren von den Wissenschaftlern erhoben werden,
- die Communities Qualitätskriterien für die Datenpublikation und verbindliche Entscheidungskriterien für die Langzeitarchivierung entwickeln,
- Forschungsdaten open access veröffentlicht werden,
- in die Budgetierung von Forschungsprojekten der Aufwand für die Langzeitarchivierung kalkuliert wird,
- die Langzeitarchivierung verlagsgebundener E-Journals auch für den Fall gesichert ist, dass ein Verlag seine Geschäftstätigkeit einstellt.

Die **DFG** sieht in ihrem im Juli 2012 vorgelegten Positionspapier „Die digitale Transformation weiter gestalten – Der Beitrag der DFG zu einer innovativen Informationsinfrastruktur für die Forschung“<sup>56</sup> ein sehr unterschiedliches Engagement der Wissenschaftsdisziplinen im Informationsmanagement von Forschungsdaten. In jedem Fall stellt es „neue Anforderungen an eine intensivere Kooperation zwischen Informationsinfrastrukturakteuren sowie Wissenschaftlerinnen und Wissenschaftlern“. Die DFG versteht Forschungsdaten sehr breit und zählt auch Informationen zu Objekten und wissenschaftlichen Sammlungen dazu. Offene Fragen sieht sie insbesondere in Organisationsformen und Verantwortlichkeiten, rechtlichen Rahmenbedingungen und der Sicherstellung von Nachhaltigkeit. Auf der Grundlage eines Rundgesprächs im Juni 2012 und der Erfahrung aus der laufenden Projektförderung wurde ein entsprechendes Programm 2013 aufgelegt.<sup>57</sup>

Auf **EU-Ebene** wird das Thema Forschungsdaten u.a. in folgenden Papieren aufgegriffen:

- Riding the wave – How Europe can gain from the rising tide of scientific data. Final report of the High Level Expert Group on Scientific Data. A submission to the European Commission,

October 2010.<sup>58</sup> Dort werden u.a. die Herausforderungen der Metadaten, der Interoperabilität und der Semantik im interdisziplinären Kontext, die Heterogenität und die Vertrauenswürdigkeit von Daten genannt.

- Mitteilung der Kommission an das Europäische Parlament, den Rat, den Europäischen Wirtschafts- und Sozialausschuss und den Ausschuss der Regionen zur Verbesserung des Zugangs zu wissenschaftlichen Informationen: Steigerung der Wirkung öffentlicher Investitionen in die Forschung, Juli 2012.<sup>59</sup> Darin wird die Open Access Veröffentlichung von Forschungsdaten empfohlen.

- e-IRG „Blue Paper“ on Data Management der e-Infrastructure Reflection Group, October 2012.<sup>60</sup> Darin wird eine integrierte e-Infrastruktur gefordert, die Computing Ressourcen und Management von Forschungsdaten einschließt. Letztere sollen insbesondere mittels ESFRI aufgebaut werden.

## 5 · STATUS QUO IN BADEN-WÜRTTEMBERG

Für das Hochleistungsrechnen an den Universitäten baut das Land Baden-Württemberg in den Jahren 2011-2016 die Ebenen bzw. Leistungsstufen 2 und 3 der „HPC-Leistungspyramide“ als Landesstrategie **bwHPC** aus. In diesem von der DFG positiv begutachteten Umsetzungskonzept werden die Vorhaben „Forschungshochleistungsrechner am KIT“ (ForHLR) als System der Ebene 2 sowie „Forschungs-Cluster“ (bwForCluster) und „HPC-Versorgungssystem des Landes Baden-Württemberg“ (bwUniCluster) auf der Ebene 3 beschrieben. Die Systeme der Ebene 3 werden zentralisiert und landesweit genutzt: bwUniCluster als Grundversorgungssystem und bwForCluster für ausgewiesene Fachwissenschaften.

Insbesondere auf der Ebene 3 beruht die Neuordnung auf den Möglichkeiten eines föderativen Betriebs- und Nutzerunterstützungskonzepts, welches sich auf den Erfahrungen im Grid-Umfeld, bwGRiD und den zugehörigen Landesprojekten, abstützt. Mit der Aufteilung in Versorgungs- und Forschungscluster wird der Bedeutung der Ebene 3 der Betriebsmittelhierarchie als Einstiegsebene für Computational Science and Engineering sowie als „HPC Enabler“ Rechnung getragen: der zu erwartende Leistungsanstieg auf den Ebenen 1 und 2 erfordert einen Anstieg der Leistungsfähigkeit auch auf Ebene 3 bei gleichzeitigem Angebot von einfachem Zugang, Methodenentwicklung und Anwendungsdomänen-bezogener Unterstützung der Nutzer.

Die Large-Scale Data Facility (**LSDF**) stellt Hardwareressourcen zur Speicherung, Archivierung und Analyse großskaliger Daten aus prinzipiell allen Wissenschaftsbereichen zur Verfügung. Gleichzeitig dient die LSDF als zentraler Datenspeicher für Baden-Württemberg, wobei das Kostenverrechnungsmodell spartenspezifisch geregelt ist. Derzeit sind Daten aus der Systembiologie, der Klima- und Umweltforschung von Synchrotronstrahlungsquellen, aus der Biophysik, den Geistes- und Lebenswissenschaften sowie der Geophysik in der LSDF gespeichert. Die Kombination aus Größe der Daten und Anforderung hinsichtlich des Datenmanagements und der -analyse stellen dabei die größten Herausforderungen dar. Darüber hinaus werden Datendienste zum Data-Browsing, Zugang, Transfer und Umgang mit Metadaten sowie dem datenintensiven Computing entwickelt und zur Verfügung gestellt. Mit einer initialen Förderung durch das MWK stehen aktuell 2,6 PetaByte (PB) Online-Festplatten-Speicher sowie entsprechende Archivierungskapazitäten, 464 Datenanalyse-CPU-Kerne sowie Hochgeschwindigkeits-Netzwerkverbindungen (100 Gb/s) zu ausgewählten Datenquellen am KIT zur Verfügung, hinzu kommen weitere 5 PB am Bioquant

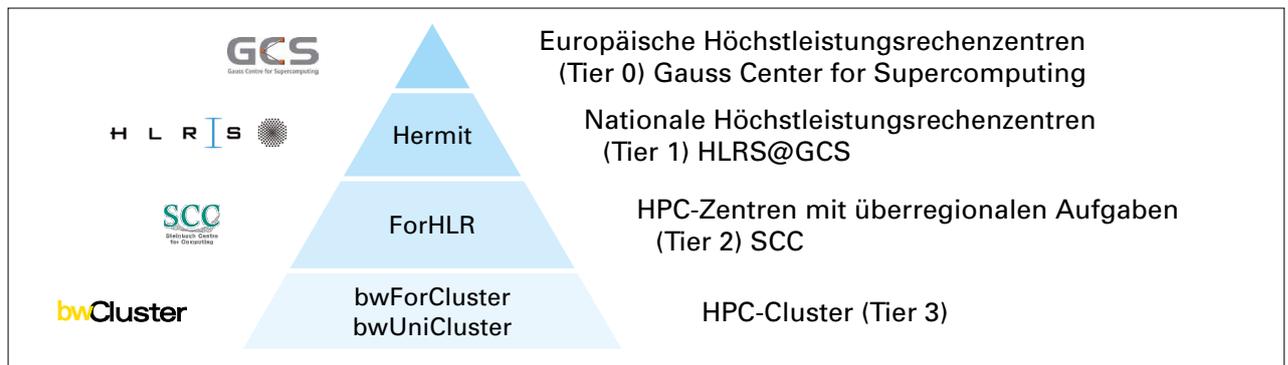


Abb. 7: Leistungspyramide der Hochleistungsrechner in Baden-Württemberg (bwHPC)

Zentrum der Universität Heidelberg. Eine weitere Finanzierung für ein kontinuierliches Anwachsen auf über 10 PB gemäß den Anforderungen der Nutzer steht zur Verfügung. Die Anbindung der LSDF an die HPC-Systeme am KIT bietet in Zukunft noch breitere und tiefere Möglichkeiten der Hochleistungs-Datenanalyse auch mit semantischen Methoden.

Mit dem Begleitprojekt **bwLSDF** fördert das MWK die Erweiterung der LSDF als zentralen Datenspeicher für die baden-württembergischen Hochschulen. Gemeinsam arbeiten die beteiligten Universitäten Stuttgart, Heidelberg, Mannheim und Ulm sowie das KIT an der Flexibilisierung der Zugriffsverfahren als Überlaufspeicher sowie an der Ausweitung der Funktionalität für allgemeine Speicherbedarfe, wie z. B. des Dienstes **bwSync&Share** zur Synchronisation und zum Austausch von Dateien, der – anders als Dropbox® – die Anforderungen an Datensicherheit und Datenschutz adäquat erfüllt. Darüber hinaus wird eine BW Storage Cloud entwickelt, um neben einer Policy-basierten Verteilung des Storage über mehrere Standorte auch hybride Cloud Lösungen zu etablieren, die auf sicherheitstechnischer Ebene mit föderativer Authentifizierung und Autorisierung (bwIDM) ausgestattet ist. Zur Archivierung komplexer digitaler Objekte fördert das MWK das Projekt **bwFLA** (Funktionale Langzeit-Archivierung) an der Universität Freiburg in

Zusammenarbeit mit der Universität Ulm, dem Center of Digital Tradition (CODIGT)<sup>61</sup>, dem Landesarchiv und dem Bibliotheks-Service-Zentrum. Dabei konzentriert sich das Vorhaben auf die Entwicklung von Methoden zur Unterstützung der Arbeitsabläufe bei der Einstellung (Ingest) der Daten in das Archiv und der langfristigen Zugriffssicherung (Access). Komplexe Objekte sind dadurch gekennzeichnet, dass sie nicht verlustfrei in dokumentierte Formate (z. B. PDF/A) konvertiert werden können oder eine spezielle originäre Laufzeitumgebung (z. B. spezielle Programme/Versionen/Betriebssysteme) benötigen, damit der Inhalt dynamisch interpretiert werden kann (z. B. Excel-Tabellen mit Makros). Für den späteren Zugriff auf die archivierten Daten kann die originalgetreue Ablaufumgebung emuliert werden. Beispiele komplexer Objekte sind wissenschaftliche Primärdaten sowie ihre (teilweise) selbst entwickelten Arbeitsumgebungen („scientific desktops“), die den technisch-wissenschaftlichen Prozess bis zum publizierfähigen Resultat repräsentieren, ferner Datenbanken oder Verwaltungsvorgänge. Die Entwicklung wird für exemplarische Klassen auf Basis der Schnittstellen und Komponenten des PLANETS<sup>62</sup> Frameworks durchgeführt.

Darüber hinaus wird am **CODIGT** in Kooperation mit dem KIT ein digitales Archiv für komplexe digitale Objekte aus Forschung und Kunst konzipiert, ein Lab zur Entwicklung von Archivierungs-, Migrations- und Emulationsverfahren aufgebaut sowie die Beratung zur Erhaltungsplanung (Preservation Planning), einschließlich der juristischen Aspekte, eingerichtet (IQF-Projekt). Die beteiligten Institutionen sind Mitglieder bei nector, dem deutschen Kompetenznetzwerk zur digitalen Langzeitarchivierung.<sup>63</sup>

## 6 · MASSNAHMEN FÜR DIE HOCHSCHULEN IN BADEN-WÜRTTEMBERG

### 6.1 WEITERENTWICKLUNG DER TECHNISCHEN INFRASTRUKTUR

Die technische Infrastruktur bzgl. Datenspeicher (bwLSDF, bwDATA), Datennetze (BelWue) und HPC-Ressourcen (bwHPC-Konzept) muss weiterhin nach dem Bedarf der Wissenschaft ausgebaut werden. Zu dieser Infrastruktur gehört auch eine Authentifizierung und Be-

rechtigungssteuerung auf Basis föderativer Strukturen der Universitäten (bwIDM). Diese Konzepte sind nicht Gegenstand dieses Papiers, aber Voraussetzung für die Einführung von Forschungsdatenmanagement.

### 6.2 INTEGRATION IN DIE LEHRE

Das Thema Forschungsdaten kommt i.d.R. gar nicht in der Lehre vor. Im Hinblick auf die Anforderungen in nächster Zukunft muss Datenmanagement Teil der notwendigen Ausbildung im Bereich Forschungsmethoden werden. Der sichere und nachvollziehbare Umgang mit Forschungsdaten über den gesamten Data Life Cycle wird zunehmend in allen Wissenschaftsbereichen als grundlegende Kompetenz benötigt. Lehrangebote sollten sich sowohl an eine breite Zielgruppe richten als auch eine Spezialisierung unterstützen. Beispiele sind:

- Disziplinübergreifende Kurse zur Vermittlung der Schlüsselqualifikation Informationskompetenz (inkl. Forschungsdaten-

management) – unterschiedlich auf den Qualifikationsstufen Bachelor, Master, Promotion

- Seminar in Masterstudiengängen: z. B. „(big) data methods“, Datenanalyse
- Zur Spezialisierung Kurse im Bereich Data Scientist / Data Librarian im Rahmen etablierter Studiengänge oder als Masterstudiengang für Berufserfahrene (z. B. an der Hector School am KIT)
- Weiterbildung für Berufstätige in berufsbegleitenden E-Learning-Kursen.

### 6.3 FÖRDERPROGRAMM ZUM INFRASTRUKTURAUFBAU

Ziel ist der Aufbau eines Forschungsdaten-Repositorys. Vorab ist wie auch bei Repositorien für textuelle Publikationen die Zielgruppe (Institution, Community) zu bestimmen und zu klären, wie die Repositorien vernetzt bzw. in eine übergreifende Struktur integriert werden.

#### Aspekte des Förderprogramms:

##### 6.3.1 Metadatenschema und Erschließung

- Strukturierung der Metadaten in einen allgemeinen und einen fachspezifischen Teil
- Harmonisierung der Metadaten über die verschiedenen Disziplinen hinweg (Interdisziplinäre Forschung)
- Nutzung nationaler und internationaler Metadaten-Standards
- Bei Experimentdaten sollte beim Laden eine halbautomatische Erzeugung von Metadaten entwickelt werden (Laborgeräte, Konfigurationen, Kalibrierung, ...) sowie Workflows, die die Aufnahme von Metadaten begünstigen. In einigen Disziplinen sollten die Metadaten teilweise auch Prozessinformationen (Auswertungsschritte, -software) abbilden. Hier besteht eine Schnittstelle zu Virtuelle Forschungsumgebungen.
- Erschließung durch automatisierte Einbindung in die Bibliothekskataloge und Erweiterung der Research Discovery Systems auf Forschungsdaten

### 6.3.2 Persistente Adressierung

Der persistent Identifier (PID) stellt die dauerhafte Erreichbarkeit von Objekten sicher. Hierfür haben sich verschiedene Standards etabliert (im Wesentlichen URN<sup>®64</sup>, DOI<sup>®65</sup> und handle<sup>®66</sup>), die von der Art der Daten (abgeschlossene oder kumulierende Daten) abhängen und häufig fachspezifisch sind.

- Information, Beratung, PID-Service
- Know-how Aufbau in Bibliotheken und bei Wissenschaftlern (Wissensplattform)

### 6.3.3 Qualitätssicherung und Datenschutz

- Die inhaltliche Prüfung kann nur durch die Community erfolgen! Die Communities können mit Best Practice Beispielen unterstützt werden.
- Die zentrale Redaktion kann Metadaten formal prüfen und den Bezug zu Normdaten herstellen.
- Empfehlung für die Aufbewahrungsdauer aus wissenschaftlicher Sicht. Ggf. müssen dafür Policies der Universitäten entwickelt und vereinbart werden.
- Datenschutzkonzept, das die Belange der betroffenen Personen und der Forscher berücksichtigt. Lizenzen, ggf. Restriktionen zur Datennutzung

### 6.3.4 Open Access für Forschungsdaten

- Entwicklung von Rahmenbedingungen – gemeinsam mit den Communities – für die Publikation von Forschungsdaten als Open Access. Dazu gehören die Definition „publikationswürdiger“ Primärdaten, Lizenzmodelle, Anreizsysteme und Mechanismen zur Verknüpfung von Text- und Datenpublikationen
- Entwicklung von Verfahren zur Bestimmung des wissenschaftlichen Impacts von Datenpublikationen
- Ziel ist die internationale Sichtbarkeit der Forschung in Baden-Württemberg, auch im Hinblick auf die Anbahnung von Forschungsk Kooperationen

### 6.3.5 Archivierung

- Entwicklung eines Archivierungsworkflows für typische Datenarten (Objektklassen, Entstehungskontexte)
- Umgang mit proprietären, nicht dokumentierten Dateiformaten
- Geschäftsmodell: Kosten setzen sich aus Storage (bwLSDF) und formatspezifischen Zusatzpauschalen (z. B. für Emulatoren) zusammen
- Wirtschaftliche und rechtliche Überlegungen (z. B. Haftung, Eigentümerschaft, Persönlichkeitsschutz)

### 6.3.6 Entwicklung von Geschäftsmodellen

- Dauerhafte Finanzierung von Betrieb und Entwicklung des Forschungsdaten-Repositoriums und der zugehörigen Services. Kostendeckungsanteile müssen durch Nutzer, Projekte, Communities, Einrichtungen usw. in Abhängigkeit von Aufbewahrungsdauer und Datenformaten erbracht werden.
- Verrechnungsmodell für die Nutzung des Speichers für die Langzeitarchivierung, insbesondere auch für die außeruniversitären Infrastruktureinrichtungen im Geschäftsbereich des Ministeriums für Wissenschaft, Forschung und Kunst sowie ggf. weitere Landeseinrichtungen
- Vorschlag für eine dauerhafte Struktur

#### 6.4 FÖRDERPROGRAMM DATA LIFE CYCLE LABS

Die **Data Life Cycle Labs** (DLCLs) des Projekts  **LSDMA** (siehe Anhang Nationale Projekte mit baden-württembergischer Beteiligung) sind auf die Forschungsbereiche der Helmholtz Gemeinschaft ausgerichtet. Es wird vorgeschlagen, in Baden Württemberg weitere DLCLs zu etablieren, welche sich an den wichtigsten Wissenschaftsrichtungen im Land orientieren. Sie können sowohl auf Wissenschafts-Communities, die keine eigenen Data Center haben und das Thema erst für sich entdecken, als auch gut organisierte Communities ausgerichtet werden. Die DLCLs setzen sich aus Wissenschaftlern der jeweiligen Fachbereiche und der Infrastrukturein-

richtungen verschiedener Universitäten zusammen. Sie sind Experten für Datenmanagement, -analyse und -migration. Durch ihre Arbeit an der Schnittstelle zwischen etablierten, standardisierten Datentechnologien und den spezifischen Anforderungen der Communities können sie allen Wissenschaftlergruppen an Forschungseinrichtungen des Landes bei deren Optimierung des jeweiligen Scientific Data Life Cycles unterstützen. Ferner stehen sie im engen Kontakt mit den internationalen Communities.

#### 6.5 EINRICHTUNG EINER LANDESWEITEN KOORDINATION

- Landesweite Clearing-Stelle für rechtliche Fragen im Umgang mit Forschungsdaten, Lizenzen und Publikationsformen. Sie verfolgt das Ziel, dass die Wissenschaft im rechtlichen Besitz ihrer Daten bleibt.
- Landesweite Beratungsstelle für Forschungsdaten- und Metadatenstandards, persistente Adressierung von Dokumenten, Verantwortung für die benötigten Wissensplattformen
- Landesweite Abstimmung und Umsetzung von Policies, Methoden und Verrechnungsmodelle für Langzeitarchivierung
- Entwicklung eines Zertifikats für Nachvollziehbarkeit und Reproduzierbarkeit von Forschungsdaten im Scientific Data Life Cycle in Kooperation mit der nationalen Ebene (DINI)

### PROJEKTE UND BEST PRACTICES

#### ANHANG 1: NATIONALE PROJEKTE MIT BADEN-WÜRTTEMBERGISCHER BETEILIGUNG

Seit 2012 unterstützt das Helmholtz-Projekt „Large Scale Data Management and Analysis“ ( **LSDMA**) Wissenschaftler verschiedener Disziplinen im Umgang mit sehr großen Datenmengen. LSDMA orientiert sich an den Datenlebenszyklen in den verschiedenen Wissenschaftsdisziplinen und zielt daher auf gemeinsame Forschungs- und Entwicklungsarbeiten im Kontext der sog. Data Life Cycle Labs (DLCL). Mit den DLCLs können die sich teils unterscheidenden Anforderungen der wissenschaftlichen Communities adressiert werden. In den DLCLs werden zunächst ausgewählte Themenbereiche adressiert, etwa Daten von Umweltsatelliten oder aus der Systembiologie. Die Datenexperten in den DLCLs sind Ansprechpartner für die Wissenschaftler in den jeweiligen Communities. In gemeinsamer Forschung und Entwicklung stellen sie den jeweiligen Data

Life Cycle auf und optimieren ihn in einem iterativen Prozess. Um methodenwissenschaftliche Forschungs- und Entwicklungsarbeiten kümmert sich ein verteiltes, horizontal arbeitendes Data Services Integration Team (DSIT), welches generische Datenservices für mehrere Communities entwickelt. Das Projekt LSDMA hat vier Partner in der Helmholtz-Gemeinschaft (DESY, GSI, FZ Jülich, KIT) und sieben außerhalb (HTW Berlin, DKRZ, TU Dresden, Uni Frankfurt, Uni Hamburg, Uni Heidelberg, Uni Ulm). Es ist für fünf Jahre angelegt und wird ab 2015 in die 3. Phase der Programmorientierten Förderung der Helmholtz-Gemeinschaft übergehen.

Baden-Württembergische Wissenschaftler sind an mehreren **DFG-Projekten** beteiligt, die exemplarisch Lösungen für spezifische Disziplinen oder allgemeine Fragestellungen entwickeln.

Die **KIT-Bibliothek** ist Partner im Projekt **re-3data**<sup>67</sup>, einem Projekt zum Aufbau eines disziplinübergreifenden globalen Registry of Research Data Repositories. Dazu gehört die Entwicklung von Auswahlkriterien und eines Metadatenschemas zur Beschreibung der Forschungsdaten-Repositories.

In dem Projekt **Radieschen**<sup>68</sup> erstellen Vertreter der **Helmholtz-Gemeinschaft**, der **Leibniz-Gemeinschaft**, des World Data Center System, der Max-Planck-Gesellschaft und der **Universitäten** eine Roadmap mit Handlungsempfehlungen für eine disziplinübergreifende Infrastruktur. Sie identifiziert und behandelt die Anforderungen an generische Komponenten einer Infrastruktur und die Vernetzung mit disziplinspezifischen Komponenten. Die Analyse basiert auf einer Bestandsaufnahme mit bestehenden und neuen Projekten sowie Maßnahmen zur Community-Bildung. Zentrale Dimensionen der Analyse sind Technik, Organisation und Kosten und die Untersuchung von Querschnittsthemen.

**GESIS** und **Universität Mannheim** erstellen Methoden zur „Integration von Forschungsdaten

und Literatur in den Sozialwissenschaften“ (InFoLiS). Mittels Text Mining werden die notwendigen Informationen aus den Daten extrahiert. Die Verknüpfung von Daten- und Textpublikation sowie die Einspeisung in das Recherchesystem „Primo“ sollen maschinell erfolgen.

Das **Institut für deutsche Sprache** baut in einem sechsjährigen Projekt ein **Zentrum für germanistische Forschungsprimärdaten** durch Zusammenführung ihrer umfangreichen Datensammlung in einem Portal auf. In der zweiten Phase werden auch Repositorien und Korpora anderer Forschungseinrichtungen verknüpft. Das Zentrum bündelt dann linguistische Korpora geschriebener Sprache, Sammlungen gesprochener Sprache, insbesondere transkribierte und ggf. annotierte Audio- und/oder Videodaten, sowie elektronische Wörterbücher und lexikalische Informationssysteme. Das **GESIS** – Leibniz-Institut für Sozialwissenschaften betreibt seit 2010 die von der DFG geförderte Datenregistrierungsagentur für Sozial- und Wirtschaftsdaten **dalra**<sup>69</sup>. dalra stellt datenhaltenden und -produzierenden Organisationen einen nicht-kommerziellen Dienst zur Registrierung von Forschungsdaten bereit. Diesen Daten werden eindeutige „Persistent Identifiers“ (PID) in Form von DOI®-Namen (Digital Object Identifiers) zugewiesen. GESIS kooperiert dabei mit der internationalen Vereinigung DataCite, deren Ziel die Förderung der Zitierfähigkeit und Referenzierbarkeit von Forschungsdaten ist. Der Registrierungsservice ist eine infrastrukturelle Daueraufgabe, bei der GESIS die bestehende Erfahrung im Rahmen der Langzeitsicherung von sozialwissenschaftlichen Forschungsdaten und bei der Bereitstellung zugehöriger Metadaten einbringt. (siehe Best Practice Beispiel im Anhang)

## ANHANG 2: EUROPÄISCHE INFRASTRUKTURPROJEKTE

Auf europäischer Ebene beteiligt sich GESIS aktiv an zwei ESFRI-Infrastrukturprojekten, die derzeit als dauerhafte europäische Forschungsinfrastrukturen implementiert werden. So nimmt GESIS eine führende Rolle innerhalb der Initiative der europäischen Datenarchive zur Gründung einer europäi-

schen Dateninfrastruktur (CESSDA) ein und beteiligt sich beim ESS (European Social Survey). Mit der Anerkennung dieser Forschungsverbände als European Research Infrastructure Consortium (ERIC) wird eine klare rechtliche Struktur auf europäischer Ebene geschaffen. Gegenwärtig kooperieren CESSDA und ESS mit den ESFRI-ERICs der Virtuellen Forschungsumgebungen CLA-

RIN und DARIAH im Rahmen des EU-Projektes DASISH. Das von der europäischen Kommission mit 9,3 Mio. Euro geförderte Projekt „EUDAT – Building a Pan-European Collaborative Data Infrastructure“ hat zum Ziel, eine kosteneffiziente und hochqualitative, kollaborative Daten-Infrastruktur in Europa aufzubauen, die die Anforderungen der unterschiedlichen Wissenschaftsdisziplinen von den Kapazitäten und Fähigkeiten auf flexible und dauerhafte Art und Weise erfüllt. Dazu haben sich 25 Partner aus 13 Ländern zusammengeschlossen, darunter Universität Tübingen und KIT. Die kollaborative Daten-Infrastruktur fokussiert auf die

Entwicklung und Bereitstellung von Diensten zur Föderierung existierender Ressourcen, die von den Partnern eingebracht werden. Gemeinsame Datendienste bzgl. Langzeitaufbewahrung mit bit-stream Persistenz, Konsistenz und Integrität, Persistent Identifier (PID) zum dauerhaften, eindeutigen Referenzieren von Daten, entsprechende Upload/Download Dienste für Nutzer sowie eine föderierte Sicherheits-Infrastruktur zur Authentifizierung und Autorisation stehen dabei im Vordergrund.

### ANHANG 3: GLOBALE PROJEKTE

Mit GridKa betreibt das Steinbuch Centre for Computing am KIT eines von 11 sog. Tier-1 Daten- und Rechenzentren in der weltweit, föderierten Infrastruktur (WLCG = Worldwide LHC Computing Grid) zur Speicherung und Prozessierung der Daten des LHC-Beschleunigers am CERN in Genf. Als eine von drei Säulen hat das LHC Computing maßgeblich zur ersten formalen Bestätigung der Existenz des sog. Higgs Bosons beigetragen, welches am 4. Juli 2012 weltweit für Aufsehen sorgte. An GridKa werden alle vier LHC-Experimente ATLAS, CMS, LHCb und ALICE unterstützt sowie noch weitere Experimente der Hochenergiephysik. GridKa stellt als community-spezifisches IT-Großgerät mit rund 12.500 CPU-Kernen, 11 PB an online-Speicher mit Festplatten und 17 PB Archivkapazität auf Band – das entspricht zusammen über 5 Millionen DVDs – rund 15 % der gesamten Tier-1 Ressourcen weltweit den LHC-Experimenten zur Verfügung. Mittels zahlreicher Hochgeschwindigkeits-Weitverkehrs-Netzwerkverbindung ist GridKa einerseits mit der Datenquelle am CERN verbunden,

andererseits sind über 20 Tier-2 Zentren in 6 europäischen Ländern angeschlossen und erhalten somit ihre Daten über GridKa. Wissenschaftler nutzen in diesen Ländern GridKa zu Datenanalyse und -archivierung sowie für Simulationsrechnungen. Im Jahr 2011 lag der gesamte eingehende Datenverkehr bei rund 18 PB, die Menge der ausgehenden Daten bei 80 PB. Neben dem reinen Betrieb des Großgerätes besteht ein zweiter wesentlicher Teil in der Erbringung von Grid-Diensten zum Betrieb der weltweit-föderierten WLCG-Infrastruktur. Diese Grid-Dienste zeichnen sich durch hohe Leistungsfähigkeit, Zuverlässigkeit, Verfügbarkeit und Betriebsstabilität aus, ohne die die kontinuierlich erzeugten Daten des LHC nicht in der Welt verteilt werden könnten. GridKa erzielt hier im direkten Vergleich mit anderen Tier-1 Zentren weltweit stets Spitzenplätze.

### ANHANG 4: FORSCHUNGSDATENMANAGEMENT BEI GESIS

Das GESIS – Leibniz Institut für Sozialwissenschaften – stellt Forschungsdaten, vorwiegend aus nationalen und international-vergleichenden Umfragen für die Sekundärnutzung bereit. Die Studien werden gemäß klar definierten methodisch-technischen Anforderungen akquiriert und bedarfsorientiert aufbereitet, do-

kumentiert, langfristig gesichert und der wissenschaftlich interessierten Öffentlichkeit zugänglich gemacht.

## DATENBESTAND UND ZUGANG

Der Datenbestand umfasst gegenwärtig ca. 5.100 Studien. Dabei handelt es sich zum größten Teil um Mikrodaten der Umfrageforschung sowie um historische Zeitreihen und Aggregatdaten. Alle Daten im Datenangebot von GESIS erhalten persistente Identifikatoren (DOI-Namen) und werden über Portale (auch mit der Möglichkeit zu einfachen Online-Analysen) oder ftp zum Download angeboten sowie auf Offline-Medien bereitgestellt. Zur Identifikation relevanter Daten werden verschiedene Re-

cherchesysteme angeboten. So liegen zu jeder Studie detaillierte und standardisierte deskriptive und strukturelle Metadaten vor, die sich über den Datenbestandskatalog (DBK<sup>70</sup>) online durchsuchen lassen. Für die wichtigsten Studien werden Recherchemöglichkeiten bis auf die Ebene einzelner Variablen angeboten. Um zukünftig auch die Zugänglichkeit sensitiver Daten für wissenschaftliche Forschungszwecke systematisch zu verbessern, wurde von GESIS ein Secure Data Center aufgebaut, das seit Mitte 2013 in einer ersten Ausbaustufe für schutzwürdige Datenbestände Vor-Ort-Zugänge an den GESIS Standorten ermöglicht.

## DATENAUFNAHME

Forschungsdaten werden von GESIS entweder gezielt akquiriert oder aber wie in der Mehrzahl der Fälle dem Archiv von Primärforschern zur Archivierung angeboten bzw. die Übergabe erfolgt routinemäßig auf der Grundlage längerfristiger Kooperationen. Darüber hinaus ist es durch ein neues Angebot namens DATORIUM (<http://datorium.gesis.org/>) seit Januar 2014 möglich, dass Nutzer die Publikation und Distribution von Daten weitgehend eigenständig durchführen können. Auf der Basis von Kooperationsvereinbarungen zu anderen sozial- und wirtschaftswissenschaftlichen Forschungszentren sowie universitären Sonderforschungsbereichen bietet GESIS die Möglichkeit der Langzeitarchivierung der im Rahmen dieser Einrichtungen erhobenen Daten.

gesis.org/) seit Januar 2014 möglich, dass Nutzer die Publikation und Distribution von Daten weitgehend eigenständig durchführen können. Auf der Basis von Kooperationsvereinbarungen zu anderen sozial- und wirtschaftswissenschaftlichen Forschungszentren sowie universitären Sonderforschungsbereichen bietet GESIS die Möglichkeit der Langzeitarchivierung der im Rahmen dieser Einrichtungen erhobenen Daten.

## MINDESTANFORDERUNGEN

Aufzunehmende Studien müssen bestimmten formalen und technischen Anforderungen entsprechen. Insbesondere müssen neben den Datensätzen selbst auch alle für eine Sekundarnutzung notwendigen Materialien vorhanden sein und an GESIS übergeben werden (Erhebungs- bzw. Messinstrumente, Codeplane, Methodenberichte etc.). Diese den Datensatz begleitenden Materialien und Dokumente bilden auch die Grundlage für die vom Datenarchiv im Rahmen der Archivierung erzeugten standardisierten Metadaten. Die Daten selbst müssen nicht unbedingt in einem bestimmten Format vorliegen (solange sie lesbar sind), sondern werden im Zuge der Archivierung in langfristsicherungsfähige Formate überführt und für den Datenservice wiederum

in von den Nutzern gewünschten Formaten bereitgestellt. Vor der endgültigen Aufnahme von Daten in das GESIS Datenarchiv wird im Rahmen des Ingest-Prozesses eine Eingangskontrolle durchgeführt, die sich auf die Erhebungsinstrumente, die Dokumentation und Daten bezieht und sicherstellt, dass die Informationen vollständig und lesbar sind und mit den Daten übereinstimmen. Die Kontrolle der Daten reicht von der Überprüfung der Anzahl der Fälle, dem Abgleich der Variablen mit dem Erhebungsinstrument bis hin zu Plausibilitäts- und Konsistenzkontrollen sowie der Überprüfung hinsichtlich datenschutzrechtlicher Aspekte. Eine Übernahme der Daten in das Langzeitarchiv erfolgt erst, wenn alle im Rahmen der Eingangskontrolle festgestellten Probleme bzw. Fehler behoben werden konnten (zumindest soweit, dass eine Nutzung der Daten durch Dritte möglich und in der Folge eine Langzeitarchivierung auch sinnvoll ist). In jüngerer Zeit ist GESIS dazu übergegangen

die Aufnahme in das Langzeitarchiv durch Archivierungsverträge mit dem Datengeber auch rechtlich zu formalisieren. Dort werden

insbesondere Fragen zu Verfügungs- und Nutzungsrechten sowie zum Datenschutz geklärt.

### **DATENVOLUMEN**

Das insgesamt in den Sozialwissenschaften anfallende Datenvolumen ist derzeit schwer quantifizierbar. Allgemein lässt sich sagen, dass die numerischen Datensätze der empirischen Sozialforschung tendenziell eher klein sind. So summieren sich bspw. die bei GESIS vorgehaltenen Bestände mit Daten und Dokumenten zu rd. 6.000 sozialwissenschaftlichen Studien, die von der physischen Speicherkapazität etwa 1 TB Speicherplatz benötigen und

sich auf etwa 600.000 Objekte verteilen. Selbst wenn man die verschiedenen on- und offline Spiegelungen bzw. Replikate im Datenarchiv von GESIS mitrechnet, ergeben sich gemessen an heute üblichen Speicherdimensionen keine generell schwer zu handhabenden Speichergrößen.

### **DATENFORMATE**

Quantitative Daten werden i.d.R. mit speziellen Statistiksoftwarepaketen verarbeitet und analysiert. In der akademischen Forschung sind dabei die Statistikpakete SPSS und Stata am weitesten verbreitet. Zunehmend gewinnt auch die Open Source Alternative „R“ an Bedeutung. In der kommerziellen Forschung spielt neben den zuvor genannten Paketen auch SAS eine wichtige Rolle. Dementsprechend werden quantitative Daten in der Regel in den meist proprietären Formaten dieser Softwarepakete erstellt, analysiert und weitergegeben. Bei den am meisten genutzten Formaten handelt es sich um proprietäre Binärformate – also um nicht oder nur eingeschränkt für die Langzeitarchivierung geeignete Formate. Im Rahmen der Langzeitarchivierung ist es daher erforderlich, diese in langfristsicherungsfähige Formate umzuwandeln. Dies kann jedoch i.d.R. ohne Informationsverlust bzw. unter Erhaltung der signifikanten Eigenschaften bewältigt werden (bspw. durch Überführung

der Daten in csv-Formate und der mit den Daten verbundenen internen Metadaten in textbasierte Formate, etwa DDI). Mit dem Bedeutungszuwachs der qualitativen Methoden bzw. von mixed-method Ansätzen und der generellen Tendenz zu komplexeren Studiendesigns sowie den bereits oben beschriebenen neuen Datenformen in den Sozialwissenschaften werden auch andere Formate (Audio, Bild, Video, Datenbanken, GIS) zunehmend wichtig und erfordern neue oder angepasste Strategien des Datenmanagements und der Langzeitarchivierung. Hier wird auch die Speichergröße eine wichtigere Rolle spielen, denn gerade eine Vielzahl von Audio-, Bild- und Videodateien erfordern andere Speicherdimensionen als rein quantitative Daten.

### **METADATEN**

Metadaten spielen eine entscheidende Rolle für die Nutzbarkeit sozialwissenschaftlicher Forschungsdaten. Im GESIS-Datenarchiv wird zur Erstellung der Metadaten der DDI-Standard<sup>71</sup> verwendet. Dabei wird gegenwärtig von der Version DDI-Codebook auf DDI-Lifecycle umgestellt. Inhaltliche, methodische und technische Charakteristika werden in einer Studienbeschreibung

zusammengefasst. Ausgewählte Studien durchlaufen einen intensiveren Dokumentationsprozess, für den spezielle Tools und Workflows entwickelt wurden. Die Dokumentation der Studien wird je nach Zielvorgabe erweitert um Metadaten sowohl auf Studien- als auch auf Datensatzebene (Standar-

disierung der Datensatzstruktur, Variablenamen/-label, etc.). Darüber hinaus erfolgt eine strukturierte Dokumentation des Fragebogens, ergänzt um Anmerkungen zur Datenqualität auf Variablenebene und weiteren Kontextinformationen. Dieser umfangreiche Bestand an strukturierten Metadaten wird dann einerseits über verschiedene (Meta)Datenportale und Recherchesysteme, wie etwa dem

Datenportal ZACAT oder dem Datenbestandskatalog DBK, zur Verfügung gestellt und dient andererseits aber auch dazu, begleitende Dokumentationen etwa in der Form von Codebüchern, Variablenreports oder Methodenberichten zu erstellen. Diese sind insbesondere bei komplexeren Studien oder Studienkollektionen neben den Daten selbst und den dazugehörigen Messinstrumenten häufig Voraussetzung für eine sekundäranalytische Nutzung durch Dritte.

<sup>50</sup> DFG-Ausschreibung „Informationsinfrastrukturen für Forschungsdaten“ (28.04.2010), [http://www.dfg.de/download/pdf/foerderung/programme/lis/ausschreibung\\_forschungsdaten\\_1001.pdf](http://www.dfg.de/download/pdf/foerderung/programme/lis/ausschreibung_forschungsdaten_1001.pdf)

<sup>51</sup> vgl. Gesamtkonzept der KII

<sup>52</sup> Berliner Erklärung vom 22. Oktober 2003: <http://oa.mpg.de/lang/de/berlin-prozess/berliner-erklarung/>

<sup>53</sup> z. B. DFG „Vorschläge zur Sicherung guter wissenschaftlicher Praxis“ (1997), darin Empfehlung 7: „Primärdaten als Grundlagen für Veröffentlichungen sollen auf haltbaren und gesicherten Trägern in der Institution, wo sie entstanden sind, für zehn Jahre aufbewahrt werden.“ Siehe auch: Empfehlungen zur gesicherten Aufbewahrung und Bereitstellung digitaler Forschungsprimärdaten (2009): [http://www.dfg.de/download/pdf/foerderung/programme/lis/ua\\_inf\\_empfehlungen\\_200901.pdf](http://www.dfg.de/download/pdf/foerderung/programme/lis/ua_inf_empfehlungen_200901.pdf); und: Antragsmuster für die Einrichtung eines Sonderforschungsbereichs (Stand: 06/2012): [http://www.dfg.de/formulare/60\\_100/60\\_100\\_de\\_rtf.rtf](http://www.dfg.de/formulare/60_100/60_100_de_rtf.rtf)

<sup>54</sup> <http://www.leibniz-gemeinschaft.de/?nid=infrastr&nidap=&print>, insbes. S. 62

<sup>55</sup> <http://www.wissenschaftsrat.de/download/archiv/2359-12.pdf>, insbes. S. 10 ff. und S. 53 ff.

<sup>56</sup> [http://www.dfg.de/download/pdf/foerderung/programme/lis/positionspapier\\_digitale\\_transformation.pdf](http://www.dfg.de/download/pdf/foerderung/programme/lis/positionspapier_digitale_transformation.pdf)

<sup>57</sup> [http://www.dfg.de/foerderung/programme/infrastruktur/lis/lis\\_foerderangebote/forschungsdaten/index.html](http://www.dfg.de/foerderung/programme/infrastruktur/lis/lis_foerderangebote/forschungsdaten/index.html)

<sup>58</sup> <http://cordis.europa.eu/fp7/ict/e-infrastructure/docs/hlg-sdi-report.pdf>

<sup>59</sup> [http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2012:194:0039:0043:DE:PDF, COM\(2912\)401 Juli 2012](http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2012:194:0039:0043:DE:PDF, COM(2912)401 Juli 2012).

<sup>60</sup> <http://www.e-irg.eu/publications/blue-papers.html>

<sup>61</sup> <http://codigt.de>, vormals: Kompetenzzentrum kulturelle Überlieferung - digital Karlsruhe (KÜdKa)

<sup>62</sup> Open Planets Foundation: <http://www.openplanetsfoundation.org/>

<sup>63</sup> <http://www.langzeitarchivierung.de/>

<sup>64</sup> Uniform Resource Name, in Deutschland von der Deutschen Nationalbibliothek vergeben: [http://www.dnb.de/DE/Netzpublikationen/URNService/urnservice\\_node.html](http://www.dnb.de/DE/Netzpublikationen/URNService/urnservice_node.html)

<sup>65</sup> Digital Object Identifiers des DataCite-Verbundes: <http://datacite.org/>

<sup>66</sup> handles werden vergeben von der Corporation for National Research Initiatives (USA) bzw. dem European Persistent Identifier Consortium (EPIC)-Konsortium <http://www.handle.net/>

<sup>67</sup> <http://www.re3data.org/>

<sup>68</sup> <http://www.forschungsdaten.org/uber-radieschen/>

<sup>69</sup> <http://www.da-ra.de/> da|ra wird von GESIS und ZBW – Leibniz-Informationszentrum Wirtschaft gemeinsam betrieben, siehe auch den Forschungsverbund der Leibniz-Institute „Science 2.0“: <http://www.leibniz-science20.de/>

<sup>70</sup> <http://www.gesis.org/unser-angebot/recherchieren/datenbestandskatalog/>

<sup>71</sup> Data Documentation Initiative: <http://www.ddialliance.org/>

# AG 5 · Virtuelle Forschungsumgebungen

## INHALT

|   |     |
|---|-----|
| 1. Vision Forschung 2025 .....  | 99  |
| 2. VFU in aller Munde – alle Herausforderungen gelöst? .....          | 100 |
| 2.1 Stimmen aus der Wissenschaft: „Implizites wird expliziert.“ ..... | 100 |
| 2.2 Sicht der Wissenschaftsorganisationen .....                       | 102 |
| 3. Definition VFU und Abgrenzung .....                                | 104 |
| 4. Nationale, europäische, internationale Entwicklungen .....         | 105 |
| 4.1 Deutschland .....   | 105 |
| 4.2 Baden-Württemberg .....   | 106 |
| 4.3 EU .....  | 106 |
| 4.4 Großbritannien .....  | 107 |
| 4.5 USA .....   | 107 |
| 4.6 Australien .....  | 108 |
| 5. Maßnahmen für die Hochschulen in Baden-Württemberg .....           | 108 |

## MITGLIEDER DER ARBEITSGRUPPE

**PROF. DR. ERHARD HINRICHS**

Seminar für Sprachwissenschaft, Universität Tübingen

**DR. MARKUS KLEIN**

Ministerium für Wissenschaft, Forschung und Kunst  
Baden-Württemberg (Vorsitz)

**DR. PETER LEINEN**

Rechenzentrum, Universität Mannheim

**DR. CLAUDIA PAULI**

Kommunikations- und Informationszentrum,  
Universität Ulm

**DR. EBERHARD PIETZSCH**

Bibliothek, Universität Tübingen

**PROF. DR. ERICH SCHELKLE**

Automotive Simulation Center Stuttgart

**FRANK SCHOLZE**

Bibliothek, KIT

**DR. KARL-HEINZ WEBER**

FIZ Karlsruhe - Leibniz-Institut für  
Informationsinfrastruktur

**DR. ANDREAS WITT**

Institut für Deutsche Sprache, Mannheim

## 1 · VISION FORSCHUNG 2025

Die „Wissenschaft der Zukunft“ wird im Vergleich zu der bisherigen Arbeitsweise zunehmend kooperativ, interdisziplinär, international und ortsunabhängig sein. Die Nutzung digitaler Ressourcen und Werkzeuge und die Verschmelzung digitaler und physischer Ressourcen werden zunehmen und Standortgrenzen verschwinden. Der Bedarf nach technischem Support auf Seiten der wissenschaftlichen Community wird perspektivisch deutlich ansteigen.

Kennzeichen der Wissenschaft der Zukunft werden in allen Wissenschaftsdisziplinen ein hohes Maß an kollaborativem Forschen anhand von allseitig anerkannten und standardisierten Referenzdaten sowie Analyse- und Visualisierungswerkzeugen sein, die sich in den Forschernetzwerken als Best Practices herausgebildet haben. Auch werden sich die Anforderungen zur Sicherstellung von Reproduzierbarkeit und Nachnutzbarkeit von Forschungsergebnissen verändern. Diese Daten und Werkzeuge sind in einem distribuierten Netzwerk von Kompetenzzentren – in der Regel in Form von Virtuellen Forschungsumgebungen – verfügbar. Solche Infrastrukturen unterliegen einem ständigen Wettbewerb um die Nutzer und entwickeln sich daher kontinuierlich weiter, so wie sich auch die Methoden und die Medien weiterentwickeln werden.

Weltweiter elektronischer Zugang zu Forschungsprimärdaten, Objekten und Forschungsanalysedaten prägt den Forschungsalltag bereits heute in vielen Disziplinen. In den nächsten Jahren wird sich die Nutzung der digitalen Informationsversorgung und -verarbeitung in Lehre und Forschung in allen Disziplinen intensivieren und dadurch die Methoden der Wissenschaften verändern. Die Communi-

ties werden sich verstärkt mit dem Potenzial digitaler Methoden und Virtueller Forschungsumgebungen (VFUen) auseinandersetzen und diese reflektieren. Neue Methoden und Arbeitsweisen müssen in die Ausbildung der Studierenden einfließen, um eine kontinuierliche Weiterentwicklung und Verstetigung zu erreichen. Die Verknüpfung digitaler Daten und Analysemethoden auf kollaborativ orientierten Plattformen mit offenen Schnittstellen befördert neue Formen transparenter Zusammenarbeit, die in einen Kulturwandel münden, der die Praxis und Arbeitsweise wissenschaftlicher Aktivität und ihrer Bewertung verändert.

Wissenschaftliche Kreativität erfordert Zeit, Muße, einfachen und unbeschränkten Zugang zu wissenschaftlichen Daten sowie den Dialog und die Zusammenarbeit mit Forschern – auch und gerade über traditionelle Fachgrenzen hinweg. Digitale wie nicht-digitale Methoden sind grundsätzlich gleich berechtigt. Forschungsinfrastrukturen und Virtuelle Forschungsumgebungen können kein Ersatz für ausreichende Forschungsfreiräume und kein Allheilmittel für die wissenschaftliche Kreativität sein, sie können aber den spielenden Umgang (im doppelten Sinne) mit Forschungsgegenständen ermöglichen und befördern.

Inhaltlich sollte eine VFU idealerweise alle Aspekte des Forschungsprozesses von der Projektidee über die Datensammlung und Auswertung bis zur Veröffentlichung und Nachnutzung berücksichtigen. Da die Wissenschaftsbereiche unterschiedliche Datenarten generieren und Methoden anwenden, sind die Bedarfe an unterstützenden Werkzeugen abgesehen von einigen grundlegenden Diensten disziplinspezifisch. In der Konsequenz werden Infrastruktureinrichtungen und Einrichtungen der Forschung sehr viel stärker kooperieren und teilweise örtliche oder überregionale Betriebs- und Geschäftsmodelle aushandeln müssen. In kooperativen Umsetzungsschritten muss die Bedeutung der Daten und Instrumente für den Wissenschaftsbereich spezifiziert und Lösungen für das Datenmanagement (Menge, Strukturierung, Metadaten, Sicherheit, Datenschutz) und gegen das Risiko „falscher“ Datenhaltung entwickelt werden. Dazu bedarf es einer Förderung der notwendigen Veränderungsprozesse an

den Universitäten des Landes innerhalb der anderen etablierten Förderinstrumente auf nationaler und europäischer Ebene. Dies betrifft auch die außeruniversitären Forschungseinrichtungen, wobei diese die Veränderungsprozesse bisher oft entschiedener vorangetrieben haben.

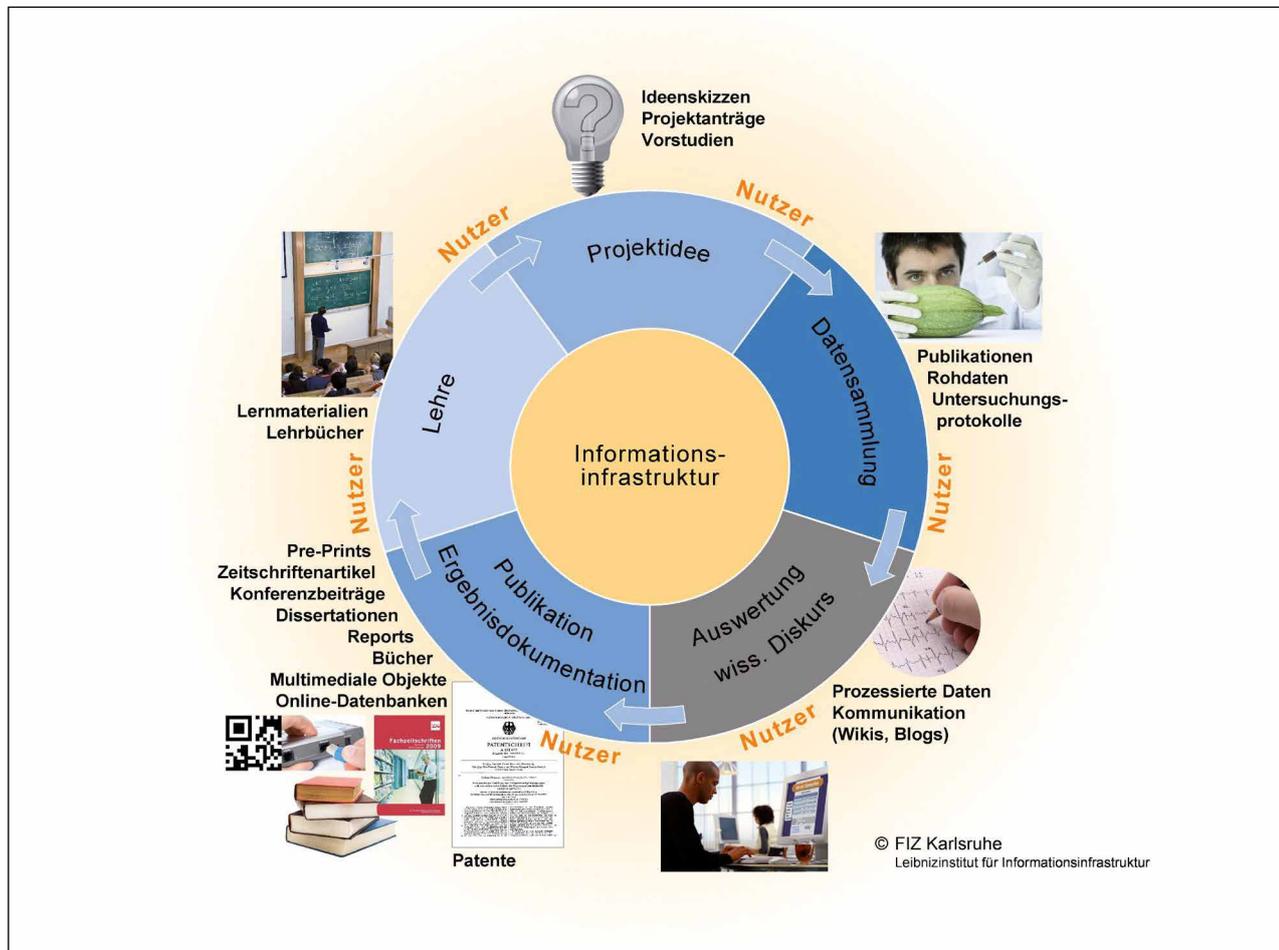


Abb. 8: Der wissenschaftliche Wertschöpfungsprozess (Quelle: FIZ Karlsruhe – Leibniz-Institut für Informationsinfrastruktur)

## 2 · VFU IN ALLER MUNDE – ALLE HERAUSFORDERUNGEN GELÖST?

### 2.1 STIMMEN AUS DER WISSENSCHAFT: „IMPLIZITES WIRD EXPLIZIERT.“

Die Mitglieder der AG haben Vertreter verschiedener Disziplinen nach Stellenwert und Nutzung der Virtuellen Forschungsumgebungen und nach ihren Erfahrungen und Erwartungen gefragt.

Dr. Thomas Rauch, Universität Tübingen, stellt dar, dass die Datenmengen in der Astronomie so exorbitant seien, dass ein Datentransport nahezu unmöglich wird. Daher müsse die Auswertung der Daten „vor Ort“ erfolgen, so dass nur extrahierte

Ergebnisse über die Netze transportiert werden. Voraussetzungen seien freier Zugang zu Daten und Software (Open Source) und deren zentrale Erfassung. „Wissenschaftliche Arbeit ist auf unserem Gebiet ohne VFU nahezu unmöglich. [...] Aufgrund der Entwicklung der Datenmengen [werden] die Energiekosten für Datenhaltung, Simulation etc. nicht mehr zu tragen sein.“

Die Geowissenschaften werden aus Sicht von Prof. Dr. Maik Thomas, GFZ Potsdam, durch interdisziplinäre Beobachtungs- und Modellierungsdaten getrieben, so dass internationaler und interdisziplinärer Austausch (tages-)aktueller Daten organisiert werden muss. Bei Verfügbarkeit würden VFUen intensiv genutzt, „aber der Aufwand für die Erstellung von VFUen ist gegenwärtig sehr hoch und die Wege zur Finanzierung vor allem des Routinebetriebs sind sehr begrenzt: nur große Zentren, oft kein Wettbewerb und geringes Innovationstempo.“ Er sieht wichtige Herausforderungen in der Entwicklung von Standards für Datenformate, Metadaten und Schnittstellen der VFU-Programmmodule. „Open Access von Daten ist essenziell für die Entfaltung der Wissenschaft.“

Das AWI kann auf eine sehr positive Entwicklung der Erdsystemdatenbank PANGAEA und dem World Data Center for Marine Environmental Sciences<sup>72</sup> von einer Informationsinfrastruktur hin zu einer Virtuellen Forschungsumgebung zurückblicken. Prof. Dr. Antje Boetius: „Schon jetzt gibt es z. B. eine Pipeline von der Probenahme (z. B. Forschungsschiffexpeditionen) samt Archivierung von nicht-digitalen Objekten (z. B. Meeresbodenproben, DNA Archive) bis hin zu den in Fachjournals veröffentlichten Daten, die mit eigener DOI versehen werden, inklusive eines nachhaltigen Metadaten-Konzeptes.“ Prof. Dr. Wolfgang Hiller meint, dass „der zeitliche Aufwand und die Kosten [...] gerade in Deutschland vielfach grob unterschätzt [wird], obwohl gerade die Effizienz und Produktivität von Grundlagenforschung stark gewinnen würde, wenn wir mehr in die Schaffung derartiger Strukturen investieren würden.“ Dr. Hans Grobe ergänzt: „Da die Komponenten der Forschungsumgebung ausschließlich von Wissenschaftlern

entwickelt wurden, folgen sie konsequent dem wissenschaftlichen Bedarf. Es soll aber nicht unerwähnt bleiben, dass sowohl die Akzeptanz als auch die Finanzierung dieser Strukturen noch weiterer Überzeugungsarbeit bedarf. Die Integration Virtueller Forschungsumgebungen in den wissenschaftlichen Alltag ist ein langer und stetiger Prozess. [...] Wesentlich bleiben auch in Zukunft die Wünsche der Wissenschaft gegenüber einer Installation vielfältigster technischer Möglichkeiten. Die Gefahr zu einer weiteren Komplexisierung liegt in der Natur der Sache. Auch führt die schnelle technische Entwicklung dazu, dass sich Arbeitsumgebungen ständig verändern; diesem kann und will der Nutzer nicht folgen.“

Die Systembiologie hat sich durch die Automatisierung von Synthese- und Analytik-Geräten sowie Gensequenzern zu einer sehr datenintensiven Disziplin entwickelt. Dadurch verschiebt sich die wissenschaftliche Leistung von der Datengewinnung zur weiterführenden Arbeit mit den Daten. Dominic Lütjohann, Doktorand am KIT, erwartet „eine Integrationsplattform, die den kompletten wissenschaftlichen Prozess abdeckt, sodass eine Forschergruppe im virtuellen Raum an einer gemeinsamen Problemlösung arbeiten kann, ohne physikalisch präsent sein zu müssen. [...] Durch eine im Gesamtprozess überwachte wissenschaftliche Arbeit werden Fehleinschätzungen vermieden, absolute Reproduzierbarkeit geschaffen und eine korrekte Datenerhebung in den wissenschaftlichen Medien gewährleistet.“

Für die theoretische Chemie sieht Prof. Dr. Axel Groß, Vizepräsident für Forschung und Informationstechnologie der Universität Ulm den Bedarf vordringlich im „standardisierte[n] Zugriff von überall her und zu jeder Zeit auf elektronische Ressourcen“ (Volltexte, Datenbanken, Preprints). Erst im Zusammenhang mit Großprojekten und der disziplin- und institutionsübergreifenden Nutzung von Forschungsdaten sieht er Potenziale für eine Forscher-Communities.

In der Philologie – stellvertretend für die Geisteswissenschaften – ist nach Meinung von Prof. Dr. Andrea Rapp, TU Darmstadt, „in der

Breite des Faches [...] das Thema VFU noch nicht rezipiert und angekommen,“ aber sie sieht große Chancen für die korpusbasiert und die historisch mit unikalenen Quellen in dislozierten Beständen arbeitenden Disziplinen. Es „besteht ein großes Potenzial, dadurch methodologisch neue Bereiche zu erschließen, z. B. durch die Verbindung von qualitativen und quantitativen Methoden oder durch neue Visualisierungsverfahren. [...] Implizites wird expliziert – was ich für eine der ganz großen Stärken computergestützter Arbeit halte.“ Prof. Dr. Fotis Jannidis schlägt eine Brücke zur Retro-Digitalisierung und digitalen Publikationsformen: „In dem Maße, in dem wir in unserem Arbeiten immer stärker auf digitale Quellen zurückgreifen, wächst die Abhängigkeit der Geisteswissenschaften von einer Umgebung, die ihnen das bietet, was sie gelernt haben zu erwarten: Nachhaltige Verfügbarkeit und gleichzeitig stetige Modernisierung. Ich glaube daher, dass VFUen zentral sind für die zukünftige Arbeit von Geis-

teswissenschaftlern, vergleichbar nur dem Status von Bibliotheken [...]“. Prof. Dr. Claudine Moulin, Universität Trier, stellt heraus: „Digitalität ist nicht nur ein wesentlicher Bestandteil unseres Alltags geworden, sondern verändert auch die Forschungsmethoden und -verfahren in den geisteswissenschaftlichen Disziplinen grundlegend.“ Sie warnt vor top-down-Initiativen. „Daher sollte man Virtuelle Forschungsinfrastrukturen – wie in dem Positionspapier der European Science Foundation erarbeitet – als komplexe Ökosysteme auffassen, die gleichermaßen der Interdisziplinarität, Mehrsprachigkeit und Polyfunktionalität ihrer Gegenstände Rechnung tragen müssen, um der Vielschichtigkeit geistes- und kulturwissenschaftlicher Fragestellungen begegnen zu können.“

## 2.2 SICHT DER WISSENSCHAFTSORGANISATIONEN

Der fundamentale Wandel der letzten Jahre in den Kommunikations- und Informationstechnologien hat gravierende Veränderungen des wissenschaftlichen Arbeitens nach sich gezogen. Forschungsergebnisse fallen zusehends in Gestalt digitaler Primärdaten an oder basieren zu erheblichen Teilen auf dieser empirischen Basis. Aggregationen von Forschungsdaten unterschiedlicher Provenienz in ihrer gemeinsamen Nutzung bergen – durch Anwendung eines geeigneten digitalen Werkzeugkastens – das Potenzial, unabhängig von ihrem ursprünglichen Erhebungszweck weiterführende Forschung zu beflügeln.<sup>73</sup> Verlust oder auch nur Nicht-Interpretierbarkeit von Daten entzieht der Wissenschaft somit die Grundlage ihrer Arbeit. Die Entwicklung einer leistungsfähigen Infrastruktur zur Erzeugung, Verwaltung, Recherche, Analyse, Verknüpfung, Publikation und Nachnutzung von Daten ist in der Konsequenz aus der Sicht der Wissenschaftler der entscheidende Schritt hin zu wissenschaftlicher Exzellenz. Wissenschaftler versprechen sich von VFUen mittels Vernetzung eine bisher ungekannte fachliche Qualität ihres Schaffens, von der im Endeffekt bessere Lösungen und auch neue Forschungsfragestellungen sowie neue Lehr- und Lernszenarien zu erwarten sind. So groß die Chan-

cen bei der Implementierung solcher Infrastrukturen auch sind, so hoch sind die Anforderungen der wissenschaftlichen Nutzer an die Stabilität ihres neuen Arbeitsplatzes.

Die im Rahmen öffentlich geförderter Projekte anfallenden Daten sollten idealerweise im Sinne des Open Access frei verfügbar, langfristig zugänglich und qualitätsgeprüft sein, um wissenschaftliche Forschung in ihrer Nachnutzung verifizierbar und rekapitulierbar zu erhalten. Aber auch die im Sinne guter wissenschaftlicher Praxis erhobenen Forderungen zum Schutz der persönlichen, rechtlichen und wissenschaftlichen Interessen von Nutzern und ggf. Probanden, die ein beschränktes Nutzungsverhältnis mit sich bringen, sollen durch den Einsatz von VFUen unterstützt werden. VFUen können hierfür eine technische Basis bilden, um langfristig den Zugriff auf diese Daten zu gewährleisten. Damit soll verhindert werden, die empirischen Grundlagen der

Forschung immer wieder aufs Neue in zeit- und kostenintensiver Arbeit wiederherstellen zu müssen.

Die Fachdisziplinen unterscheiden sich innerhalb einzelner wissenschaftlicher Communities bisweilen recht deutlich in ihrem Umgang mit Forschungsdaten. Nicht nur erheben sie unterschiedliche Datenformate, sie ziehen auch fachspezifische Werkzeuge zu ihrer Bearbeitung heran. VFUen sollten daher die nötige Flexibilität aufweisen, auf divergierende Nutzeranforderungen einzugehen bzw. sich daran anpassen zu lassen.

Vor diesem Hintergrund ist insbesondere bei Lösungen zur Langzeitarchivierung von Forschungsdaten in vernetzten Umgebungen die Verwendung von Standards in Format und die Anreicherung von Metadaten in der Dokumentation unabdingbar, sodass eine interdisziplinäre Nutzung ermöglicht wird.<sup>74</sup> (siehe auch den Beitrag der AG 4 Forschungsdatenmanagement)

Auf dem Weg, diesen Herausforderungen adäquate Antworten entgegenzustellen, sehen Wissenschaftsorganisationen Handlungsbedarf auf verschiedenen Themenfeldern und sprechen folgende Empfehlungen dazu aus<sup>75</sup>:

- Ein organisatorischer Rahmen (Betriebsmodelle) muss mittels Vernetzung der Akteure gefunden werden, der die VFU in eine stabile Struktur einbindet, Zuständigkeiten und Kompetenzen zuweist sowie ressourcensparend bündelt und Schnittstellen für Community- und Länder-übergreifende Aktivitäten bietet.
- Die VFUen tragenden Verbände sollten als Interessenvertreter für einzelne Disziplinen gegenüber größeren Ressourcenanbietern fungieren.
- Eine nachhaltige Finanzierung (Geschäftsmodelle) der Basisinfrastruktur der VFUen muss sichergestellt werden. Neuentwicklungen, um den Anschluss an technische Entwicklungen zu behalten, sollten mit Projektmitteln angeschoben werden. Hierfür ist mit Blick auf die föderale Verfasstheit Deutschlands ggf. eine Anpassung der nationalen Förderkultur erforderlich.
- Technische Lösungen zur Nachnutzbarkeit der Daten sowie Schnittstellen zu neuen Technologien und Community-Ansätzen müssen gefunden werden (Referenzarchitekturmodell, Austauschformate, Standards, Metadaten).
- Bei der Nutzung grundlegender Hardware-Komponenten der gemeinsamen Basisinfrastruktur werden Synergien zwischen den Nutzern verschiedener Einrichtungen gebildet. Voraussetzungen sind ein leistungsfähiges, an allen Orten nutzbares Daten-Netzwerk (BelWue, Eduroam, Zugriff über verschlüsselte DSL-, UMTS- und LTE-Verbindungen) sowie einheitliche Standards im Identity Management, um Urheber von Daten eindeutig identifizieren und Nutzungsrechte einrichten zu können.
- Um das Funktionieren der für eine VFU typischen Mischung aus sich stetig wandelnder Open Source, kommerzieller und individuell geschaffener Software zu gewährleisten, sollte jede VFU Lösungen zum Life-Cycle-Management beinhalten.
- Umgang mit und Mehrwert von VFUen ist in Zusammenarbeit von Lehrenden, VFU-Betreibern und E-Learning-Spezialisten in Lehrveranstaltungen und Schulungen zu vermitteln und in den Curricula zu verankern. Disziplinübergreifende Kompetenzen können im Rahmen von Kursen für Schlüsselqualifikationen vermittelt werden.
- Lösungen für offene Fragen des Urheber- und Persönlichkeitsrechts sind anzubieten und sichere Nutzungsszenarien für Hardware-Ressourcen, Netzwerkinfrastruktur, Software, Dienste und Daten müssen entwickelt werden. Für die Einbindung und Rechteverwaltung hochschulexterner Anbieter müssen Lösungen entworfen werden. Bedarfsgerechte Lizenzierungsmodelle (Musterlizenzen, Landes- und Nationallizenzen) sind zu entwickeln bzw. zu implementieren und technisch umzusetzen.

### 3 · DEFINITION VFU UND ABGRENZUNG

Virtuelle Forschungsumgebungen<sup>76</sup> sind Infrastrukturen, die Wissenschaftlerinnen und Wissenschaftlern eine kooperative, zeit- und ortsunabhängige Forschung ermöglichen. Neben allgemeinen digitalen Medien werden auch fachspezifische digitale und z. T. auch physikalische Umgebungen integriert und der Zugriff auf Forschungsinfrastrukturen wie Wissensressourcen (Datensammlungen, Archive, Datenbanken) und Großgeräte (Hoch- und Höchstleistungsrechner, Hochleistungsmikroskope, Gensequenzern) erleichtert und systematisiert. Ziel von VFUen ist es, alle Schritte der Forschung nachhaltig zu unterstützen (vgl. Abbildung 8).

Eine einheitliche und umfassende Dokumentation der Ergebnisse und systematische Erschließung von Quellen leistet einen wesentlichen Beitrag zur Qualitätssicherung, führt Forschende aber auch auf ganz neue Fragestellungen, die wiederum zur Weiterentwicklung der Umgebungen führen. Um den teilweise sehr fachspezifischen Anforderungen gerecht zu werden und die unterschiedlichen Forschungskulturen adäquat unterstützen zu können, muss der Aufbau von VFUen aus den Fach-Communities getrieben werden. Infrastruktureinrichtungen sollten sich der Aufgabe stellen, diese Entwicklungen aktiv zu unterstützen.

Die technische Grundlage bilden heute modular aufgebaute Softwaresysteme und leistungsfähige Kommunikationsnetzwerke. Zur flexiblen Vernetzung der einzelnen Module stehen Ansätze wie „Software as a Service“ (SaaS) zur Verfügung. Dies ermöglicht eine parallele Entwicklung sowie eine verteilte Realisierung und Datenhaltung, einen themenbezogenen Aufbau von Workflows, aber auch eine Nutzung durch andere Communities.

Ergänzt werden die fachspezifischen Werkzeuge in der Regel durch allgemein verfügbare Elemente wie z. B. Lösungen zur Kommunikation und Kollaboration, etwa zur gemeinsamen Bearbeitung von Dokumenten. Diese sind über standardisierte Schnittstellen anzubinden und ggf. an spezielle Bedürfnisse anzupassen. Die Entwicklung von anerkannten und nachhaltig verfügbaren, durch eine

Vielzahl von Wissenschaftscommunities nutzbaren Werkzeuge, wie sie z. B. im EU-Projekt EUDAT eingesetzt und gegenwärtig evaluiert werden, müssen in Zukunft verstärkt werden und beim zukünftigen Aufbau und der Weiterentwicklung von VFUen berücksichtigt werden. Die Balance zwischen fachspezifischen Applikationen, fächerübergreifenden Diensten und IT-Basiselementen zu finden, wird eine Aufgabe sein, der sich die Entwickler von VFUen stellen müssen.

Mit der Vernetzung von Forschung und Lehre wird auch der Umgang mit Virtuellen Forschungsumgebungen Teil der Ausbildung werden. Einzelne Aspekte vorhandener VFUen werden hierzu sicher in Lernmanagementsysteme integriert werden, dennoch werden letztgenannte eher eine Kommunikationsplattform zur Unterstützung der Lehre bleiben und sich von VFUen abgrenzen.

Die Standardisierung von Datenformaten und Schnittstellen ist eine wichtige Voraussetzung für einen dauerhaften Einsatz und Betrieb von VFUen sowie deren Anbindung an Datenarchive, Langzeitarchive und Bibliothekskataloge. Solche Formate werden zum Beispiel für die Repräsentation textbasierender Quellen von der Text Encoding Initiative (TEI) auf der Basis von XML oder für die herstellerunabhängige Darstellung von 3D-Daten von der Collaborative Design Activity Gruppe (COLLADA) entwickelt. Wissenschaftler arbeiten zum Beispiel im Projekt CLARIN an der Definition internationaler Standards (ISO, DIN, W3C).

Mit der zunehmenden weltweiten Vernetzung von Wissenschaftlern muss der Zugang und die Mög-

lichkeit der Partizipation an VFUen entsprechend offen gestaltet werden. Hierfür sind passende Governance-Strukturen aufzubauen. Zugriffe und Zugänge weltweiter Nutzer sind zu regeln, aber auch darüber hinaus gehende rechtliche Rahmenbedingungen festzulegen. Hier sind nicht nur die Wissenschaftsorganisationen gefragt, sondern zunehmend auch die politische Ebene, die die entsprechenden Rahmenbedingungen schaffen muss, damit VFUen das vorhandene Potenzial nutzen können. Mit der Anerkennung von Forschungsverbänden als „European Research Infrastructure Consortium“ (ERIC)<sup>77</sup>

besteht auf europäischer Ebene eine Möglichkeit, steuerliche Rahmenbedingungen klar zu regeln. Ebenso sind die rechtlichen Rahmenbedingungen für einen dauerhaften Betrieb und die Weiterentwicklung von VFUen zu regeln.

Neben den spezifischen Werkzeugen einer VFU werden Forscherinnen und Forscher auch in Zukunft mit Basisdiensten zur Informationsversorgung lokal vor Ort oder auf Landesebene zu versorgen sein. Mit einer zunehmenden Verbreitung von VFUen wird es hier teilweise Überschneidungen geben. Die Partizipation an VFUen setzt diese Basisdienste auch in Zukunft voraus.

## 4 · NATIONALE, EUROPÄISCHE, INTERNATIONALE ENTWICKLUNGEN

### 4.1 DEUTSCHLAND

Das BMBF hat ab 2005 im Rahmen der D-GRID-Initiative eine Reihe von Projekten gefördert, in denen themen- bzw. disziplinspezifische, auf Grid-Technologie basierende Infrastrukturen aufgebaut wurden. Sie ermöglichen den jeweiligen Communities die gemeinsame Nutzung von verteilten digitalen Ressourcen – z. B. Daten oder Rechenleistung – und schaffen damit wichtige Voraussetzungen für das Arbeiten in Virtuellen Forschungsumgebungen. Die DFG hat bisher in zwei Aufrufen (2008, 2009) 22 Projekte zum Aufbau von Virtuellen Forschungsumgebungen mit insgesamt 6,8 Mio. Euro gefördert. Ein dritter, Mitte 2011 publizierter Aufruf zielt auf den Übergang von VFUen aus der Aufbauphase in den Dauerbetrieb ab. Es sind zu diesem Aufruf 12 Anträge eingegangen.

In den vergangenen zwei bis drei Jahren haben sich insbesondere die von der GWK beauftragte „Kommission Zukunft der Informationsinfrastruktur“ (KII) sowie die Allianz-Initiative „Digitale Information“ neben anderen Themen mit VFUen befasst.

Die KII kommt dabei zu den Empfehlungen,

- Förderprogramme für Virtuelle Forschungsumgebungen auszuweiten und zusätzliche Finanzmittel für den dauerhaften Betrieb bereitzustellen,
- geeignete Koordinations- und Unterstützungsstrukturen zu schaffen (z. B. zur Vermeidung von Redundanzen bei der weiteren Entwicklung von VFUen),
- die Nachhaltigkeit und gemeinschaftliche Entwicklung durch enge Kooperation von Fachwissenschaftlern und Infrastrukturreinrichtungen zu sichern,
- die Nachnutzbarkeit durch Interoperabilität zu sichern,
- geeignete Organisationsstrukturen und Rechtsformen zu schaffen bzw. bei der Umsetzung in die Praxis zu unterstützen,
- Aus- und Weiterbildungsangebote auszubauen.

Der Wissenschaftsrat (WR) schließt sich in seinen im Juli 2012 veröffentlichten „Empfehlungen zur Weiterentwicklung der wissenschaftlichen Informationsinfrastrukturen in Deutschland bis 2020“ den Empfehlungen der KII zu den VFUen „uneingeschränkt“ an.

Die entsprechende AG VFU der Allianz-Initiative hat neben der Formulierung einer Definition einen „Leitfaden“ erstellt, der potenziellen

Nutzern und Entwicklern von VFUen Hinweise zu grundsätzlichen Fragestellungen gibt – etwa die Notwendigkeit der Beachtung von Standards (Interoperabilität), der (Nach-)Nutzung generischer Services, formale, rechtliche und organisatorische Rahmenbedingungen.

Ferner hat die Allianz-AG exemplarische VFU-Projekte unterschiedlicher Disziplinen auf ihre charakteristischen Eigenschaften

hin untersucht und wird solche Erhebungen in Kooperation mit DINI fortsetzen. Ferner ist künftig ein Leitfaden für den Betrieb von VFUen geplant sowie eine intensivere Untersuchung rechtlicher Fragestellungen, insbesondere hinsichtlich der für einen Dauerbetrieb geeigneten Organisationsformen.

#### 4.2 BADEN-WÜRTTEMBERG

Das Land Baden-Württemberg hat von 2009 bis 2012 zwei Projekte zur Entwicklung Virtueller Forschungsumgebungen gefördert: BW-eLabs und BW-eSci(T). Sie entstanden in enger Kooperation von Natur- bzw. Geisteswissenschaftlern und Informationsinfrastruktureinrichtungen der Universitäten Freiburg, Stuttgart und Tübingen und Entwicklern von FIZ Karlsruhe. Sie haben bestätigt, dass VFUen, die die methodische Varietät und die Produktivität der Wissenschaftler und die Qualität der Ergebnisse unterstützen und steigern, z. T. disziplinspezifische Anforderungen erfüllen müssen. Zum Beispiel müssen Kalibrierung und Datenausgang der Messgeräte interpretiert bzw. fachspezifische Analyse- und Bearbeitungswerkzeuge eingebunden werden. Daneben gibt es jedoch grundlegende verbindende Funktionen und Strukturen wie Rechteverwaltung (AAI, Single-Sign-On), sowie zentrale Datenhaltung in Repositorien und Datenbanken, die weiterentwickelt und in neue Anwendungen gebracht werden müssen. Seit 1987 sind die Hochschulen des Landes und weitere Bildungseinrichtungen im Datennetz „BelWue“ verbunden und in das Deut-

sche Forschungsnetz DFN eingebunden. Damit wird für den Datenverkehr zwischen den meisten Hochschulen eine hohe Übertragungsrate von 10 Gbit/s bereitgestellt. Auf dieser Basis wurden verschiedene Infrastrukturdienste entwickelt, die hochschulübergreifend nutzbar sind. Dazu gehört seit 2011 das zentrale Hosting des gemeinsamen lokalen Hochschulbibliotheksystems IBS|BW und die Entwicklung eines föderativen Identity Managements aller Universitäten (bwIDM). Letzteres wurde auf Basis der Erfahrungen, die in den o.g. VFU-Projekten hinsichtlich Rechteverwaltung gesammelt wurden, entwickelt und erlaubt eine einfache und komfortable Nutzung von Diensten, die nur an einer Universität der Föderation angeboten werden. Die dort entwickelten Funktionen sind Grundlage für die Rechteverwaltung bei neu entstehenden VFU- und Forschungsdaten-Diensten.

#### 4.3 EU

Die EU fördert(e) im Rahmen des 7. Forschungsrahmenprogramms im Bereich E-Infrastructures zehn Projekte eines Calls zu „Global Virtual Research Communities“. Bei den ESFRI-Projekten CLARIN und DARIAH handelt es sich um Infrastruktur-Projekte in den Geisteswissenschaften, in denen digitale Ressourcen und Tools einer größeren Community verfügbar gemacht werden. Sowohl CLARIN wie DARIAH befinden sich inzwischen in der sog. Umsetzungsphase mit jeweils nationaler Förderung, die in Deutschland durch das BMBF sowie mit

Begleitprojekten der Bundesländer geleistet wird. In Baden-Württemberg wurden im Rahmen von CLARIN D-SPIN (2009-2011) und CLARIN-D (seit 2011) die Universitäten Tübingen und Stuttgart sowie das IDS gefördert. An DARIAH ist u.a. das KIT beteiligt.

Da Virtuelle Forschungsumgebungen gerade die institutionenübergreifende – auch internationale –

Zusammenarbeit von Forschungsteams erleichtern, ist auf allen Ebenen möglichst frühzeitig auf Standardisierungen bei Zugriffsschnittstellen, Metadatenstandards etc. zu achten, um bei Bedarf und nach Möglichkeit schnell eine Vernetzung von digitalen Ressourcen erreichen zu können.

Um Insellösungen zu vermeiden, sollte sich jede Entwicklung von VFUen auf nationale (hier v.a. auf

die gerade vom Wissenschaftsrat vorbereitete national roadmap für Forschungsinfrastrukturen) und internationale Entwicklungen (ESFRI roadmap und ERIC) beziehen und sich solchen Initiativen durch eine finanzielle Beteiligung, bei der i.d.R. Bund und Länder gefordert sind, anschließen.

#### 4.4 GROSSBRITANNIEN

Das Joint Information Systems Committee (JISC) hat in Großbritannien seit 2004 in drei Phasen 28 Virtual Research Environments (VRE)-Projekte gefördert, die ein breites Spektrum von Fachdisziplinen umfassen. Die Förderung war in den drei Phasen jeweils auf verschiedene Schwerpunkte fokussiert: Entwicklung technischer Lösungen, Integrationsprojekte, Tools und Interoperabilität. Inzwi-

schon wird die Förderung vergleichbarer Projekte im Programm Research Infrastructure fortgesetzt.

In einer Arbeitsgruppe Virtual Research Environments (VRE) des Netzwerkes Knowledge Exchange (KE) tauschen sich von den Förderorganisationen JISC (UK), DEFF (DK), SURFfoundation (NL) und der DFG benannte Vertreter und Experten zu in den jeweiligen Ländern stattfindenden Aktivitäten zu VFU aus.

#### 4.5 USA

In den USA vollziehen sich unter der Bezeichnung „Cyberinfrastruktur“ ähnliche Entwicklungen. Die „National Middleware Initiative“ (NMI) der National Science Foundation (NSF) hat seit 2001 das Ziel, eine generische (Basis-) Infrastruktur aufzubauen, auf der dann disziplinnaher Anwendungen aufsetzen können. Es gibt in den USA zwar keine gezielte und koordinierte Förderung von Virtuellen Forschungsumgebungen, wohl aber einzelne, durch NSF (insbesondere „Science Gates“), Department of Energy (DOE) und dem National Institute of Health (NIH) geförderte Entwicklungen. Der Schwerpunkt lag bisher insgesamt eher auf der Förderung der Entwicklung von Basisdiensten. Seit 2008 werden im Programm DATANet der NSF Projekte gefördert, die die Entwicklung einer umfas-

senden digitalen Dateninfrastruktur zum Ziel haben. Insbesondere wird dabei die Bildung neuer Organisationsformen gefördert, die „library and archival sciences, cyberinfrastruktur, computer and information sciences, and domain science expertise“ integrieren. DATANet ist auf ein Volumen von 100 Mio. \$ in einem fünfjährigen Förderzeitraum angelegt. Im Bereich der geisteswissenschaftlichen Forschung verfolgt das Projekt Bamboo die Frage: „How can we advance arts and humanities research through the development of shared technology services?“ In der ersten Phase 2008-2010 wurde unter Beteiligung von 114 Institutionen der Bedarf der Wissenschaftler erhoben. In der aktuellen ersten Umsetzungsphase werden eine Infrastruktur aufgebaut und die ersten Dienste von den zehn Universitäten implementiert. Gefördert wird das Projekt durch die Mellon Foundation und Eigenmittel der Universitäten, bisher ca. 5 Mio. \$.

#### 4.6 AUSTRALIEN

In Australien erfolgt seit einigen Jahren eine koordinierte Förderung von „e-Research“, in der u.a. Grid-Entwicklungen und insbesondere der Aufbau einer Daten-Infrastruktur (durch ANDS)

und z. T. auch von VFUen gefördert werden, letztere auch innerhalb der Victorian e-Research Strategic Initiative (VeRSI).

### 5 · MASSNAHMEN FÜR DIE HOCHSCHULEN IN BADEN-WÜRTTEMBERG

Baden-Württemberg möchte – zur Umsetzung von Maßnahmen, wie sie auch im KII-Papier und den folgenden Empfehlungen des Wissenschaftsrats empfohlen wurden – eine führende Rolle bei der Entwicklung und Nutzung Virtueller Forschungsumgebungen einnehmen. Daraus folgt die Notwendigkeit, sich verstärkt mit anderen Ländern und dem Bund sowie der DFG abzustimmen.

len empfohlen. Ziel des Programms sind die Entwicklung, Anpassung, Verbreitung, Nutzung und der Betrieb von VFUen, ggf. inklusive der Ableitung generischer Komponenten. Dazu kann ggf. auch die Einrichtung eines Kompetenzzentrums beantragt werden. Weitere Ziele sind die Reflexion der methodischen Veränderung der Wissenschaften und die Integration in die Lehre auf der Basis erfolgreicher Anwendung von VFUen.

Außerdem wird die offene Ausschreibung eines Förderprogramms zur Unterstützung der Etablierung von VFUen an den Hochschu-

Drei Säulen der Förderung sollen in die Ausschreibung aufgenommen werden:

1. Identifizierung bestehender Infrastruktur und Bedarfe der Communities mit dem Ziel, dass Infrastruktureinrichtungen die Communities proaktiv beraten.
2. Anschub- und Kofinanzierung von Initiativen, die in nationale oder europäische Projekte münden. Ziel ist die Beteiligung an nationalen und internationalen Initiativen und die Profilbildung der Hochschulen. Dadurch werden Nuklei für die Entwicklung und Verbreitung weiterer VFUen geschaffen.
3. Schaffung von Voraussetzungen für die Weiterführung von Elementen Virtueller Forschungsumgebungen über Förderperioden hinaus: Entwicklung von Governance-Strukturen und Geschäftsmodellen, Überführung in Dauerbetrieb, Strukturwandel als Voraussetzung für Nachhaltigkeit.

Bedingungen für eine mögliche Förderung:

- a. Bewerbungen können sich Communities – über die Grenze der einzelnen Hochschule hinaus – im Verbund mit universitären oder außeruniversitären Infrastruktureinrichtungen. Die Unterstützung der Rektorate ist erforderlich. Dadurch wird lokale Hochschul-Infrastruktur eingebunden; Computing-, Storage- und Hosting-Ressourcen, bestehende VFUen und Services können zusammengeführt werden.
- b. Konzeptionelle Überlegungen für ein Forschungsdatenmanagement als zentrale Datenbasis für die VFU sollen vorhanden sein (s. a. Empfehlungen der AG 4 Forschungsdatenmanagement).
- c. Die Nutzung von AG 4 Forschungsdatenmanagement Standards und bisher entwickelter VFU-Komponenten wird vorausgesetzt.

d. Eine dauerhafte Wirkung eines VFU-Projekts setzt eine kritische Masse beteiligter Entwickler (oder anderer Experten) und beteiligter Einrichtungen voraus.

Aufgrund der verschiedenen Förderziele und der Inhomogenität der angesprochenen Communities muss das Förderprogramm hinreichend flexibel sein. Daher erscheint eine überlappungsfreie Aufspaltung in verschiedene Förderlinien nicht sinnvoll. Vorgeschlagen wird eine zeitlich offene Ausschreibung. Programmvolumen und Laufzeit müssen angemessen sein, um einen hinreichenden und sichtbaren Impact zu erzielen.

<sup>72</sup> <http://www.wdc-mare.org/>

<sup>73</sup> [http://www.ratswd.de/download/RatSWD\\_WP\\_2010/RatSWD\\_WP\\_156.pdf](http://www.ratswd.de/download/RatSWD_WP_2010/RatSWD_WP_156.pdf)

<sup>74</sup> [http://www.textgrid.de/fileadmin/TextGrid/konferenzen\\_vortraege/DH\\_Festakt/VRE-Schwiegelshohn-Goettingen.pdf](http://www.textgrid.de/fileadmin/TextGrid/konferenzen_vortraege/DH_Festakt/VRE-Schwiegelshohn-Goettingen.pdf),

<sup>75</sup> [http://www.allianzinitiative.de/fileadmin/hosting\\_studie\\_d.pdf](http://www.allianzinitiative.de/fileadmin/hosting_studie_d.pdf), 15ff.; [http://www.allianzinitiative.de/fileadmin/user\\_upload/KII\\_Gesamtkonzept.pdf](http://www.allianzinitiative.de/fileadmin/user_upload/KII_Gesamtkonzept.pdf), 44ff.; [http://www.wissgrid.de/publikationen/deliverables/wp1/WissGrid\\_AP1\\_D1-4\\_final\\_v2.pdf](http://www.wissgrid.de/publikationen/deliverables/wp1/WissGrid_AP1_D1-4_final_v2.pdf), 6ff.; [http://www.wissgrid.de/publikationen/deliverables/wp1/WissGrid\\_AP1\\_D15\\_v101\\_final.pdf](http://www.wissgrid.de/publikationen/deliverables/wp1/WissGrid_AP1_D15_v101_final.pdf), 35.

<sup>76</sup> Allianz-Initiative: „Eine Virtuelle Forschungsumgebung ist eine Arbeitsplattform, die eine kooperative Forschungstätigkeit durch mehrere Wissenschaftler an unterschiedlichen Orten zu gleicher Zeit ohne Einschränkungen ermöglicht. Inhaltlich unterstützt sie potentiell den gesamten Forschungsprozess – von der Erhebung, der Diskussion und weiteren Bearbeitung der Daten bis zur Publikation der Ergebnisse – während sie technologisch vor allem auf Softwarediensten und Kommunikationsnetzwerken basiert.“  
[http://www.allianzinitiative.de/de/handlungsfelder/virtuelle\\_forschungsumgebungen/](http://www.allianzinitiative.de/de/handlungsfelder/virtuelle_forschungsumgebungen/)

<sup>77</sup> CLARIN hat den Status eines ERIC. DARIAH ist in der Beantragungsphase.

# Mitglieder der E-Science-AG

## MITGLIEDER DER ARBEITSGRUPPE

### **DR. URSULA BERNHARDT**

Ministerium für Wissenschaft, Forschung und Kunst Baden-Württemberg  
Lenkungskreis, AG Digitalisierung

### **DR. CHRISTOPH BRUCH**

Helmholtz Open Access Koordinationsbüro, Stiftung Alfred-Wegener-Institut für  
Polar- und Meeresforschung in der Helmholtz-Gemeinschaft, Bremerhaven  
AG Open Access

### **PETER CASTELLAZ**

Ministerium für Wissenschaft, Forschung und Kunst Baden-Württemberg  
AG Digitalisierung

### **PROF. DR. THOMAS DREIER**

Zentrum für Angewandte Rechtswissenschaft (ZAR) am KIT  
AG Open Access

### **DR. JUTTA DRESCH**

Badisches Landesmuseum  
AG Digitalisierung

### **PROF. DR. OLAF GEFELLER**

Institut für Medizininformatik, Biometrie und Epidemiologie, Universität Erlangen-Nürnberg  
AG Lizenzierung

### **DR. CHRISTIAN HÄNGER**

Bibliothek, Universität Mannheim  
AG Forschungsdatenmanagement

### **PETRA HÄTSCHER**

Bibliothek, Universität Konstanz  
AG Open Access

### **BRIGITTE HAUSSTEIN**

GESIS – Leibniz-Institut für Sozialwissenschaften  
AG Forschungsdatenmanagement

**PROF. MICHAEL HEBGEN**

Rechenzentrum, Universität Heidelberg  
AG Forschungsdatenmanagement

**PROF. DR. ERHARD HINRICHS**

Seminar für Sprachwissenschaft, Universität Tübingen  
AG Virtuelle Forschungsumgebungen

**KARL-WILHELM HORSTMANN**

Kommunikations-, Informations- und Medienzentrum, Universität Hohenheim  
AG Lizenzierung

**DR. JOCHEN JOHANNSEN**

Badische Landesbibliothek Karlsruhe  
AG Lizenzierung

**WALTER KAAG**

Ministerium für Wissenschaft, Forschung und Kunst Baden-Württemberg  
Lenkungskreis

**DR. ANTJE KELLERSOHN**

Bibliothek, Universität Freiburg  
AG Lizenzierung

**DR. MARKUS KLEIN**

Ministerium für Wissenschaft, Forschung und Kunst Baden-Württemberg  
Lenkungskreis, AG Forschungsdatenmanagement,  
AG Virtuelle Forschungsumgebungen

**DR. HERIBERT KNORR**

Ministerium für Wissenschaft, Forschung und Kunst Baden-Württemberg  
Lenkungskreis (Vorsitz)

**PER KNUDSEN**

Bibliothek, Universität Mannheim  
AG Lizenzierung

**DR. HANNSJÖRG KOWARK**

Württembergische Landesbibliothek  
AG Digitalisierung

**DR. PETER LEINEN**

Rechenzentrum, Universität Mannheim  
AG Virtuelle Forschungsumgebungen

**DR. MARION MALLMANN-BIEHLER**

Bibliotheksservice-Zentrum Baden-Württemberg, Konstanz  
AG Open Access

**DR. GERALD MAIER**

Landesarchiv Baden-Württemberg  
AG Digitalisierung

**PROF. DR. HERBERT MÜTHER**

Prorektor / Institut für Theoretische Physik, Universität Tübingen  
AG Open Access

**DR. CLAUDIA PAULI**

Kommunikations- und Informationszentrum, Universität Ulm  
AG Virtuelle Forschungsumgebungen

**DR. THOMAS PFLÜGER**

Ministerium für Wissenschaft, Forschung und Kunst Baden-Württemberg  
Lenkungskreis, AG Open Access

**DR. EBERHARD PIETZSCH**

Bibliothek, Universität Tübingen  
AG Virtuelle Forschungsumgebungen

**DR. VEIT PROBST**

Bibliothek, Universität Heidelberg  
AG Digitalisierung

**MATTHIAS RAZUM**

FIZ Karlsruhe - Leibniz-Institut für Informationsstruktur  
AG Forschungsdatenmanagement

**PETER REMPIS**

Bibliothek, Universität Tübingen  
AG Forschungsdatenmanagement

**PROF. DR. ANDREAS REUTER**

Heidelberger Institut für Theoretische Studien  
Lenkungsreis

**ATO RUPPERT**

Bibliothek, Universität Freiburg  
AG Lizenzierung

**PROF. DR. ERICH SCHELKLE**

Automotive Simulation Center Stuttgart  
AG Virtuelle Forschungsumgebungen

**PROF. DR. HANS-JOCHEN SCHIEWER**

Rektor, Universität Freiburg  
AG Open Access

**PROF. DR. GERHARD SCHNEIDER**

Rechenzentrum, Universität Freiburg,  
Arbeitskreis der Leiter wissenschaftlicher Rechenzentren Baden-Württemberg  
Lenkungsreis, AG Forschungsdatenmanagement

**FRANK SCHOLZE**

Bibliothek, KIT  
AG Virtuelle Forschungsumgebungen

**ASTRID SCHÖNSTEIN**

Ministerium für Wissenschaft, Forschung und Kunst Baden-Württemberg  
Lenkungsreis, AG Lizenzierung

**DR. MATTHIAS SCHULZE**

Bibliothek, Universität Stuttgart  
AG Forschungsdatenmanagement

**DR. CHRISTIANE SPARY**

Bibliothek, Pädagogische Hochschule Ludwigsburg  
AG Lizenzierung, AG Open Access

**PROF. DR. DIETER SPECK**

Universitätsarchiv/Uniseum, Universität Freiburg  
AG Digitalisierung

**WALTER STEPHAN**

Bibliothek, Universität Stuttgart,  
AG der Direktoren wissenschaftlicher Bibliotheken Baden-Württemberg  
Lenkungsreis

**PROF. DR. ACHIM STREIT**

Steinbuch Computing Center, KIT  
AG Forschungsdatenmanagement

**PROF. DR. HEINER STUCKENSCHMIDT**

Institut für Informatik, Universität Mannheim  
AG Open Access

**PROF. DR. THOMAS WALTER**

Zentrum für Datenverarbeitung, Universität Tübingen  
AG Lizenzierung

**DR. KARL-HEINZ WEBER**

FIZ Karlsruhe - Leibniz-Institut für Informationsstruktur  
AG Virtuelle Forschungsumgebungen

**DR. ANDREAS WITT**

Institut für Deutsche Sprache, Mannheim  
AG Virtuelle Forschungsumgebungen

**STEFAN WOLF**

Bibliotheksservice-Zentrum Baden-Württemberg, Konstanz  
AG Digitalisierung

**PROF. DR. VOLKER WULFMEYER**

Institut für Physik und Meteorologie, Universität Hohenheim  
AG Forschungsdatenmanagement

**DR. REINER ZIEGLER**

Haus des Dokumentarfilms Stuttgart  
AG Digitalisierung

**HERAUSGEBER:**

Ministerium für Wissenschaft, Forschung und Kunst Baden-Württemberg  
[www.mwk.baden-wuerttemberg.de](http://www.mwk.baden-wuerttemberg.de)

**REDAKTIONELLE KOORDINATION:**

Markus Klein

**GESTALTUNG:**

Zimmermann Visuelle Kommunikation, Stuttgart

**DRUCK:**

Offizin Scheufele, Stuttgart

**DRUCKLEGUNG:**

Mai 2014

Diese Informationsschrift wird von der Landesregierung Baden-Württemberg im Rahmen ihrer verfassungsmäßigen Verpflichtung zur Unterrichtung der Öffentlichkeit herausgegeben. Sie darf weder von Parteien noch von deren Kandidaten oder Helfern während eines Wahlkampfes zum Zwecke der Wahlwerbung verwendet werden. Missbräuchlich sind insbesondere die Verteilung auf Wahlveranstaltungen, an Informationsständen der Parteien sowie das Einlegen, Aufdrucken oder Aufkleben parteipolitischer Informationen oder Werbemittel. Erlaubt ist es jedoch den Parteien, diese Informationsschrift zur Unterrichtung ihrer Mitglieder zu verwenden.



**Baden-Württemberg**

MINISTERIUM FÜR WISSENSCHAFT, FORSCHUNG UND KUNST