

# Virtuelle Umgebungen

NORBERT LOSSAU

## Virtuelle Forschungsumgebungen und die Rolle von Bibliotheken

Die Entwicklung digitaler Informationsinfrastrukturen wird das noch junge Jahrzehnt in der Wissenschaft und bei Infrastrukturanbietern prägen. Ob in den Forschungen zu Klima, Ernährung, Gesundheit, Migration, Religionen und Kulturen, Ökonomie oder Ökologie: kooperatives Arbeiten in virtuellen Forschungsumgebungen wird sich über die kommenden fünf bis zehn Jahre in allen Wissenschaftsdisziplinen als Normalverfahren etabliert haben, nicht mehr wegzudenken wie die E-Mail oder das Internet im Alltag des Forschers.

Bibliotheken sind daher gut beraten, sich mit virtuellen Forschungsumgebungen, ihren Potentialen und Problemen, insbesondere aber auch mit Fragen der praktischen Implementierung auseinanderzusetzen. Die Chance für Bibliotheken ist eine sichere Zukunft im digitalen Zeitalter. Bibliothekare können über die Disziplingrenzen hinweg ihre genuinen Aufgaben der Erschließung, Verwaltung und Sicherung der Zugänglichkeit zum Wissen in die Zeit des Internets übersetzen und zugleich eine wesentliche Rolle in Forschungsvorhaben einnehmen.

»Bibliothekare in die Forschung!«, so könnte plakativ das Motto für Bibliotheken zu virtuellen Forschungsumgebungen lauten. Es sollte uns nicht erschrecken, sondern vielmehr Ansporn sein, unsere Grund-Kompetenzen und Erfahrungen in die für viele Wissenschaftler und Bibliothekare gleichermaßen neue, digitale Welt der kollaborativen Forschung einzubringen.

The beginning decade will be shaped by the development of digital information infrastructures in the sciences and the providers of this infrastructure. Whether one is dealing with research on climate, nutrition, health, migration, religion and culture, economics or ecology, cooperative endeavours in virtual research environments will establish themselves as the norm in the coming five to ten years, and become as entrenched as the use of email and internet in the everyday life of a researcher.

Hence libraries will be well-advised to deal with the issues of the virtual research environment, its potentials and problems, and, most particularly, with the issues of practical implementation. This is the chance for libraries to secure their future in the digital age. Transcending the borders between disciplines, librarians will be able to apply themselves to their genuine tasks of cataloguing, administering and safekeeping the accessibility of knowledge in the age of internet and at the same time take on an essential role in the research process.

»Libraries Go Research!« could become the catchy motto for libraries in regard to virtual research environments and should not be frightening but, rather, an incentive for bringing our basic competencies and experience to the new digital world of collaborative research, which is new to many researchers and librarians alike.

### EINFÜHRUNG UND DEFINITION

»Virtuelle Forschungsumgebungen (VFU, engl. virtual research environment, VRE) stehen für flexible Infrastrukturen, die es Forschern erlauben, die Potentiale elektronischer Medien für das kollaborative Arbeiten zu nutzen und daraus auch neue Forschungsmethoden und Gegenstände zu entwickeln. (...) Technisch betrachtet bestehen sie (meist) aus einer Kernarchitektur mit allgemeinen Dienstleistungen und Werkzeugen (Medienspeicher, Rechenressourcen, Kommunikationsmittel etc.), an die Umgebungen und Module für einzelne For-

schungsgruppen mit spezifischen Konfigurationen und Erweiterungen angeschlossen werden können.«

So lautet der Beginn der Definition einer »Virtuellen Forschungsumgebung«, wie sie von der entsprechenden Arbeitsgruppe im Kontext der »Kommission Zukunft der Informationsinfrastruktur« (KI) erarbeitet wurde.<sup>1</sup> Ergänzend wird hier die Definition der Arbeitsgruppe der Allianz der Forschungsorganisationen angeführt:

»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 potenziell 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. Virtuelle Forschungsumgebungen sind wesentliche Komponenten moderner Forschungsinfrastrukturen.«<sup>2</sup>

Strukturell betrachtet bestehen virtuelle Forschungsumgebungen aus drei Grundbausteinen:

- Informationsobjekte (generell alle Formen von digitalen Informationsobjekten, insbesondere Forschungsdaten, Literatur, Digitalisate, Sammlungsobjekte, nichttextuelle Materialien)
- Werkzeuge (Anwendungs-Software zur Analyse, Sequenzierung, Visualisierung, zum Data-Mining, Hyperlinking etc.)
- Systeme (Hardware und Betriebssoftware, Rechner, Speicher, Netze, Wikis, Videokonferenzen u. a.)

Ihre Einsatzmöglichkeiten umfassen alle Arbeitsprozesse der Forschung und unterscheiden sich damit von anderen Informationsinfrastrukturen:

»Virtuelle Forschungsumgebungen decken in der Regel die wesentlichen Arbeitsprozesse in der Forschung ab: Diese umfassen die Recherche und Informationsbeschaffung, die Literatur- und Forschungsdatenverwaltung, die Bearbeitung von Literatur- und Forschungsdaten durch Annotieren, Sequenzierung, Analysieren etc., die Kommunikation von Forschungszwischenständen und die Produktion wissenschaftlichen Outputs bis hin zur Publikation von Forschungsergebnissen. Virtuelle Forschungsumgebungen können je nach fachspezifischen und individuellen Charakteristika strukturell sehr weit ausdifferenziert sein. Sie unterscheiden sich mit ihrem intensiven Bezug zur Arbeitspraxis der Forschung



Norbert Lossau

grundlegend von anderen, eher technologiebezogenen bzw. der Grundversorgung dienenden Infrastrukturen wie beispielsweise der Lizenzierung, dem Hosting oder der Archivierung, integrieren diese aber nach Möglichkeit«.<sup>3</sup>

Für die Wissenschaft eröffnen virtuelle Forschungsumgebungen neue Möglichkeiten und Methoden des Arbeitens und der Wissensgenerierung:

»Vielfalt und Komplexität virtueller Forschungsumgebungen bieten neue Chancen für den wissenschaftlichen Arbeitsprozess und stellen zugleich neue Anforderungen an die Kooperationsbereitschaft der Akteure, die Organisation, Finanzierung sowie den Technologieeinsatz. Virtuelle Forschungsumgebungen bieten ein hohes Innovationspotenzial. Sie fördern die direkte und globale Zusammenarbeit mehrerer Forschender und Forschungsgruppen und damit einen inter- und transdisziplinären Forschungsansatz. Einzelne Forschende werden bei der Entwicklung neuer wissenschaftlicher Fragestellungen, neuer Arbeitsweisen und Methoden angeregt und unterstützt. Ergebnisse aus solchen oft neuartigen transdisziplinären Untersuchungsparadigmen begünstigen die wissenschaftliche Ausdifferenzierung auf allen Ebenen«.<sup>4</sup>

Virtuelle Forschungsumgebungen sind komplexe Konstruktionen und, aus infrastruktureller Sicht, an der Schnittstelle zu anderen Themengebieten angesiedelt. Dazu gehören insbesondere Infrastrukturen für Forschungsdaten, Open Access, Digitalisierung, digitale Langzeitarchivierung und elektronisches Publizieren.

Der Trend hin zu virtuellen Forschungsumgebungen wird durch wissenschaftspolitische Diskussionen und Empfehlungen verstärkt. Genannt seien beispielhaft die Forderung des Wissenschaftsrats an die Union der Akademien nach verstärkter Digitalisierung und Sichtbarkeit im Internet<sup>5</sup> und die Empfehlungen des Wissenschaftsrats zu Forschungsinfrastrukturen in den Geistes- und Sozialwissenschaften.<sup>6</sup>

Die ersten europäischen Förderprogramme zum Aufbau »Virtueller Forschungsumgebungen« wurden in Großbritannien durch den Infrastrukturförderer JISC im Jahr 2004 aufgelegt.<sup>7</sup> Die Hauptakteure in Europa sind heute die Europäische Kommission,<sup>8</sup> unter den Mitgliedsstaaten insbesondere Großbritannien und Deutschland, außerhalb Europas die USA, Australien und Japan. Einen aktuellen und guten Überblick über die internationale Entwicklung liefert die im Januar 2010 im Auftrag von JISC herausgegebene Studie zu »Virtual Research Environments«.<sup>9</sup> Hauptförderer in Deutschland zu virtuellen Forschungsumgebungen sind die Deutsche Forschungsgemeinschaft<sup>10</sup> und das Bundesministerium für Bildung und Forschung.<sup>11</sup>

## VIRTUELLE FORSCHUNGsumGEBUNGEN IN DER PRAXIS

In den Natur- und Lebenswissenschaften spielen virtuelle Forschungsumgebungen in unterschiedlicher Ausprägung bereits seit mehreren Jahren eine wichtige Rolle. Dabei hat sich die strukturierte Bereitstellung von Forschungsdaten auch unterschiedlicher Quellen in einem Netzwerk zum Normalfall entwickelt. In der Teilchenphysik greifen mehr als 10.000 Forscher weltweit auf die Daten des Large Hadron Collider (LHC) zu.<sup>12</sup> In der Astronomie erhält die Riesenteleskopanlage ESO in Paranal, Chile, jährlich ca. 2.000 Forschungsanträge. In der Grippeforschung ist GISAID im Zusammenhang mit der Vogelgrippe weltweit bekannt geworden als öffentlich zugängliche Datenbank, in die zahlreiche Labore Gen-Sequenzen als gemeinschaftliche Forschungsdatenbasis einspeisen. Ob in den Forschungen zu Klima, Ernährung, Gesundheit, Migration, Religionen und Kulturen, Ökonomie oder Ökologie: kooperatives Arbeiten in virtuellen Forschungsumgebungen wird sich bis spätestens zum Jahre 2020 in allen Wissenschaftsdisziplinen als Normalverfahren etabliert haben, nicht mehr wegzudenken aus dem Alltag des Forschers wie die E-Mail oder die Internet-Nutzung.

Ein sehr frühes Beispiel für eine virtuelle Forschungsumgebung in den textbasierten geisteswissenschaftlichen Disziplinen stellt das BMBF-geförderte Projekt TextGrid dar. Der Fokus liegt auf den Fachdisziplinen der Philologen, Linguisten, Musikwissenschaftler und Kunsthistoriker. Das TextGridLab, ein intuitiv bedienbarer Einstiegspunkt, bietet Zugriff auf fachwissenschaftliche Werkzeuge, Services und Inhalte. Das TextGridRep, ein fachwissenschaftliches Langzeitarchiv, eingebettet in eine Grid-Infrastruktur, garantiert die langfristige Verfügbarkeit und Zugänglichkeit der geisteswissenschaftlichen Forschungsdaten.<sup>13</sup>

Auch in den Akademien sind virtuelle Forschungsumgebungen zum Thema geworden. »Blumenbach-Online« ist ein kürzlich gestartetes Langzeitprojekt der Akademie der Wissenschaften zu Göttingen und im Bereich der Edition angesiedelt.<sup>14</sup> Johann Friedrich Blumenbach (1752–1840), Professor der Medizin und Naturgeschichte an der Universität Göttingen, ist von entscheidender Bedeutung für das Verständnis der Wissenschaftskultur Europas im späten 18. und im frühen 19. Jahrhundert. Ein wissenschaftshistorisches Projekt im klassischen Verständnis hätte sich dem Wissenschaftler über seine Schriften genähert, diese ggf. digitalisiert, inhaltlich erschlossen und die Forschungsergebnisse einschließlich der Faksimilierung ausgewählter Schriften in einer (Print-)Edition publiziert.

Blumenbach-Online weist in der Zielsetzung darüber hinaus, indem es zu den Schriften naturhistorische Objekte seiner Sammlungen in den Blick nimmt. Tatsächlich wäre der Wissenschaftler Blumenbach nur unvollständig rezipiert, würde man die Vielfalt der von ihm gesammelten und beschriebenen Objekte unberücksichtigt lassen. Zu weit über 1.000 geowissenschaftlichen Blumenbachiana (Fossilien, Mineralien, Gesteine, Meteoriten u. a.) ist in Göttingen weiteres Material in den Sammlungen der Mediziner / Anatomen (Schädelsammlung), der Anthropologen, der Botaniker, der Zoologen und der Ethnologen vorhanden. Das Projekt erstellt eine virtuelle Forschungsumgebung zu Johann Friedrich Blumenbach mit einer Quellenbasis aus Primär- und Sekundärschriften sowie Objekten, die online zugänglich und für Forschungen nutzbar sein werden. Schriftzeugnisse und Objekte werden systematisch miteinander verknüpft<sup>15</sup> und diese Verknüpfungen über die Nutzeroberfläche visualisiert. Ausgewählte thematische Zugänge führen den interessierten Nutzer zu für Blumenbach relevanten Themen wie Forschungsreisen und -reisenden, die deutsche Aufklärung usw. Damit entsteht eine Modell-Edition neuen Typs, die Wissenschaftshistorie in einer ungewohnten Breite und Tiefe über das Internet publiziert.

Die Rezeptionsgeschichtliche, biographische, soziokulturelle und -politische Erforschung Blumenbachs ist natürlich ein Hauptzweck des Göttinger Akademieprojekts. Zugleich wird aber eine weltweit sichtbare virtuelle Forschungsumgebung zu dem deutschen Mediziner und Naturhistoriker geschaffen, an die einzelne Forscher und Forschergruppen »andocken« können. Blumenbach-Online wird sich somit nicht einfach in die Präsentation berühmter Wissenschaftler einreihen, wie wir sie von Darwin, Newton oder Linné her kennen.<sup>16</sup> Durch den Einsatz spezieller Software, die insbesondere auf den Entwicklungen von TextGrid<sup>17</sup> aufsetzt, werden zugleich neue Wege eröffnet, die nachfolgende Forscher beschreiten können. Dr. Heiko Weber, Koordinator des Blumenbach-Online-Projekts, beschreibt anschaulich das Potential eines Programms, des sog. Synopsen-Tools, mit dem in der Differenzialdiagnose Konsistenzen und Inkonsistenzen, Kontinuitäten und Diskontinuitäten in der Genese von Blumenbachs Texten schnell aufzuweisen sind, die dann als Ausgangspunkt für die eigentliche wissenschaftliche Forschung dienen können. Das Tool zeigt also nicht das »Warum eines Sachverhaltes«, sondern nur, dass »Etwas ein Sachverhalt ist«. Das aber so komfortabel, flexibel und schnell, wie es keine typographische Edition (auf Grund der in der Materialität liegenden Ursachen und Limitierungen) leisten könnte. Die Mitarbeiter des

Blumenbach-Online-Projekts sind somit zugleich Nutzer der zusammengestellten Quellenbasis und werden selber zum Anbieter für andere Forscher. Jürgen Renn vom MPI für Wissenschaftsgeschichte in Berlin bezeichnet diese »Rekursivität« als charakteristisches Merkmal von virtuellen Forschungsumgebungen.<sup>18</sup>

## **VIRTUELLE FORSCHUNG- UMGEBUNGEN, BIBLIOTHEKEN UND FACHDISZIPLINEN**

Das Verständnis der Rolle von Bibliotheken in virtuellen Forschungsumgebungen ist derzeit noch überwiegend geprägt vom generellen Verständnis einer Bibliothek in der jeweiligen Fachdisziplin. Generalisierend lässt sich feststellen, dass die Geistes- und Kulturwissenschaften Bibliothekare als natürliche Partner auch in der digitalen Welt verstehen, während Natur- und Lebenswissenschaften eher ein traditionelles, literatur-zentriertes Bild von Bibliotheken pflegen. Für die Geistes- und Kulturwissenschaften sind Bibliotheken traditionell die »Labore«, Bibliotheksbestände liefern zugleich Forschungsdaten. Im Zuge der Digitalisierung haben Bibliothekare und Wissenschaftler seit den 1990er Jahren gemeinsam ihre ersten Schritte in die digitale Welt gemacht. In den Natur- und Lebenswissenschaften hatten Bibliotheken in den Laboren und Forschungsprozessen hingegen keinen Platz und mit Forschungsdaten kamen Bibliothekare erst über die Publikationen in Berührung. Dieses Verständnis resultiert in unterschiedlichen Kooperationsmustern zwischen Bibliotheken und Fachwissenschaften im Bereich virtueller Forschungsumgebungen, wie auch in den folgenden Aufgabenbereichen erkennbar wird.

## **NEUE AUFGABENGEBIETE FÜR BIBLIOTHEKEN**

Im Folgenden werden wesentliche Aufgabengebiete skizziert, mit denen Bibliotheken infrastrukturseitig den Aufbau und Betrieb virtueller Forschungsumgebungen unterstützen können. Nicht alle Gebiete (wie elektronisches Publizieren oder Repositorien) sind per se neu für Bibliotheken. In der Systemumgebung einer virtuellen Forschungsumgebung können sie jedoch zusätzliche Anforderungen erfahren, wie die weitere Beschreibung zeigen wird.

## **KONZEPTIONIERUNG UND ANWEN- DUNG VON METADATENSCHEMATA, SICHERUNG DER INTEROPERABILITÄT UND STANDARDISIERUNG**

Literatur, Forschungsdaten und Wissensobjekte bilden das materielle Fundament für die Forschung. Sie werden themenbezogen in virtuellen Forschungsumge-

bungen zusammengespielt und nach verschiedenen Fragestellungen verknüpft. Daten unterschiedlicher virtueller Forschungsumgebungen können zugleich unter neuen, übergeordneten Fragestellungen in föderierten Datennetzwerken zusammengeführt werden.

Grundlage sowohl der Daten-Zusammenführung als auch der Bearbeitung sind ihre formale Beschreibung und inhaltlich-semantiche Erschließung. Diese basieren auf Metadatenschemata. Neue Metadatenschemata müssen entwickelt werden, die sukzessive das gesamte Spektrum der Wissensobjekte und ihre Verknüpfung beschreiben, wie Text-Text-, Text-Bild- oder Text-Forschungsdaten-Relationen. Metadatenschemata ermöglichen die auch automatisierte Verarbeitung der Wissensressourcen in virtuellen Forschungsumgebungen und bieten damit erst die Basis für Software-Werkzeuge. Ohne die Definition und Erstellung von Metadatenschemata wären IT-gestützte, forschungsrelevante Prozesse wie die Recherche und das Data-Mining, das Annotieren, Sequenzieren, Visualisieren etc. nicht möglich. Oder kurz gefasst: jede Funktion in einer Benutzeroberfläche erfordert eine Entsprechung im Datenmodell von Metadatenschemata. Im Folgenden sind einige Beispiele für Metadatenschemata aufgeführt.<sup>19</sup>

- LIDO (Lightweight Information Describing Objects),<sup>20</sup> wurde als Austauschformat für Metadaten aus dem Museumsbereich entwickelt.
- OAI-ORE (OAI Object Reuse and Exchange),<sup>21</sup> Standard für die Beschreibung und den Austausch von komplexen Ressourcen wie der Verknüpfung von Publikationen mit Objekten oder Forschungsdaten.
- Dryad Metadata Application Profile,<sup>22</sup> Kern-Format für die Beschreibung von Daten und Veröffentlichungen im DRYAD-Projekt mit dem Forschungsschwerpunkt Evolution und Ökologie.

Zusätzlich ermöglichen prozessbezogene Metadaten die Rekonstruktion der Produktionsbedingungen von Forschungsdaten inkl. ihres Entstehungsworkflows. Technische Metadaten sollten alle relevanten Informationen für die digitale Langzeitverfügbarkeit und -archivierung mit sich führen.

Zentrale Elemente in Metadatenschemata sind sog. Identifier, die Informationseinheiten eindeutig identifizieren und sicherstellen, dass ein Wissensobjekt dauerhaft und unabhängig vom physischen Speicherort eines Wissensobjekts gefunden und genutzt werden kann. Hier hat die Technische Informationsbibliothek (TIB) Hannover mit Anschubförderung der DFG erfolgreich einen Service aufgebaut, der gegen Gebühr Forschungsdatensets mit sog. Digital Object Identifiern (DOIs)<sup>23</sup> im Katalog registriert. Die Gesellschaft für Wissenschaftliche Datenverarbeitung (GWVG) in Göt-

tingen betreibt unter anderem für die Max-Planck-Gesellschaft als Alternative zur DOI-Registrierung einen Service, der auf dem ursprünglich in den USA entwickelten sog. Handle-System beruht.

Die Erfassung der Metadaten sollte, wo immer möglich, automatisiert erfolgen. Eine vollständige Automatisierung dürfte derzeit aber noch die Ausnahme bilden. Für die formulargestützte Erfassung können Bibliothekare zum Einsatz kommen, analog zur Erschließung von Strukturen (Strukturdaten) in Digitalisierungsprojekten oder zur Beschreibung von Internetquellen in virtuellen Fachbibliotheken.

Die Bedeutung von Metadatenschemata wird, über die Geistes- und Gesellschaftswissenschaften hinaus, zunehmend auch in den Natur- und Lebenswissenschaften erkannt. In einigen Disziplinen wie der Klimaforschung, der Astronomie oder den Geowissenschaften haben sich Experten innerhalb der Fachgebiete der Thematik angenommen. In vielen anderen Disziplinen steht die Entwicklung von Metadatenschemata noch am Anfang, sowohl in den Geistes- und Kulturwissenschaften wie in natur- und lebenswissenschaftlichen Fachgebieten. Bibliotheken finden hier disziplinübergreifend neue Aufgaben in einem ureigenen Kompetenzfeld, der Beschreibung und Erschließung von Informationsressourcen. Da Metadaten-Beschreibungen und Forschungsfragestellungen in virtuellen Forschungsumgebungen eng zusammenhängen, ist die Zusammenarbeit mit Fachwissenschaftlern zwingend erforderlich. Diese wird dadurch befördert, dass im Zuge der fortschreitenden Entwicklung virtueller Forschungsumgebungen ein wachsendes Bewusstsein für die Notwendigkeit von Metadatenschemata und Standards zugunsten von Interoperabilität zu beobachten ist. Damit wächst, wenn auch unterschiedlich ausgeprägt in den Disziplinen, die Erkenntnis, dass angesichts der Komplexität der Thematik ein professionelles Informationsmanagement hilfreich sein kann. Dies schafft eine günstige Ausgangsposition für Bibliothekare, auf Fachwissenschaftler zuzugehen und die Partnerschaft beim Aufbau und Betrieb virtueller Forschungsumgebungen anzubieten. Voraussetzung hierfür ist natürlich zunächst der Eigenerwerb einschlägiger Kompetenzen durch die Bibliothekare selbst (dazu mehr an anderer Stelle). Die Möglichkeit der auch finanziellen Förderung des professionellen Informationsmanagements in etablierten und bei Forschern angesehenen Förderinstrumenten wie den Sonderforschungsbereichen der Deutschen Forschungsgemeinschaft<sup>24</sup> wirkt zusätzlich unterstützend für die neuen Partnerschaften. Bibliotheken und Rechenzentren werden in den Förderinformationen explizit als Partner genannt.<sup>25</sup>

## **REPOSITORIEN, DATENARCHIVE UND DAS MANAGEMENT VON INFORMATIONSSRESSOURCEN IN VIR- TUELLEN FORSCHUNGsumGEBUNGEN**

Die in virtuellen Forschungsumgebungen gesammelten Daten und Sekundärquellen müssen strukturiert verwaltet und für den Zugriff verlässlich aufbewahrt werden. Forscher sprechen hier von Archiven und betonen deren zentrale Bedeutung als Basis für alle weiteren Forschungsprozesse. Auf Infrastrukturseite wird auch von Repositorien gesprochen, deren Zahl sich innerhalb der letzten Jahre rasant entwickelt hat.<sup>26</sup>

Repositorien können als Bausteine in virtuellen Forschungsumgebungen für Disziplinen sowie zur Bearbeitung spezifischer Forschungsfragestellungen integriert sein, oder sie stehen außerhalb einzelner virtueller Forschungsumgebungen als übergreifende Informationsinfrastruktur für multiple Anwender bereit. Im letzteren Fall wird auch allgemein von Forschungsdaten-Infrastrukturen gesprochen, wie sie beispielsweise in den Geisteswissenschaften alleine in Deutschland mit weit mehr als 100 digitalisierten Quellensammlungen von Bibliotheken bereitgestellt werden.<sup>27</sup> In den Sozialwissenschaften zählen die Datenarchive von GESIS<sup>28</sup> in diese Kategorie und aus den Naturwissenschaften können das Deutsche Klimarechenzentrum mit seinen Datenarchiven<sup>29</sup> oder das NASA Astrophysics Data System (ADS)<sup>30</sup> als Beispiele angeführt werden. In beiden Fällen können Bibliotheken grundsätzlich als Dienstleister auftreten.

Im Weiteren werden Repositorien als Bausteine in virtuellen Forschungsumgebungen betrachtet. Dabei ist grundsätzlich zwischen Publikationen auf der einen und Forschungsdaten bzw. anderen Wissensobjekten auf der anderen Seite zu unterscheiden. Relativ unstrittig wird in allen Disziplinen die Verantwortung von Bibliotheken für die Verwaltung und Archivierung von Online-Publikationen gesehen. Für Forschungsdaten und andere digitale Objekte ergibt sich ein differenziertes Bild. In den Geistes- und Kulturwissenschaften ist die Rolle der Bibliotheken aus bekannten Gründen weitgehend unangefochten. Aber bereits in den Sozialwissenschaften haben sich, wie oben genannt, innerhalb des Faches Infrastrukturen für Datenarchive etabliert. In den Fachgebieten, wo solche Strukturen etabliert wurden, ist die Zurückhaltung Bibliotheken gegenüber spürbar. Dennoch werden Bibliotheken auch dort als Infrastrukturpartner angenommen, wenn sie konzeptionell und mit Kompetenz überzeugen.

Zu den Kompetenzen gehört auch die Kenntnis und gegebenenfalls eigene Erfahrung mit den Technologien für Repositorien und Datenarchive. Hier wurden in den letzten 10–15 Jahren Repositoriensysteme

wie Fedora, DSpace<sup>31</sup>, Opus<sup>32</sup> und ePrints<sup>33</sup> entwickelt, die insbesondere durch Bibliotheken zum Einsatz kommen. Fedora oder seine durch die Max Planck Digital Library (MPDL) mit dem FIZ Karlsruhe betriebene deutsche Weiterentwicklung eSciDoc<sup>34</sup> stehen, ebenso wie DSpace, grundsätzlich auch als Systeme für die Speicherung und Archivierung von Forschungsdaten und anderen digitalen Objekten zur Verfügung. Bibliotheken eröffnet sich mit der konkreten Implementierung von Metadatenschemata in Repositorien-Systemen methodologisch über Disziplinengrenzen hinweg die Perspektive, sich als kompetente Infrastrukturpartner in den virtuellen Forschungsumgebungen einzubringen.

## **DAS PUBLIZIEREN AUS VIRTUELLEN FORSCHUNGsumGEBUNGEN**

Das Beispiel von Blumenbach-Online zeigt die Herausforderung für die Wissenschaft, Forschungsergebnisse aus virtuellen Forschungsumgebungen zu publizieren. Der semantische Gehalt der Verknüpfung von Schriftzeugnissen mit Sammlungsobjekten ließe sich über die klassische Edition nur in reduzierter Form veröffentlichen. Hier besteht also der Bedarf an neuartigen Online-Editionen, eine Entwicklung, wie sie für Wörterbücher und Enzyklopädien bereits vor zehn Jahren eingesetzt hat. Die etablierten Veröffentlichungsformen (die Monographie, der Zeitschriftenartikel, Proceedings etc.) stehen also mit der Verbreitung virtueller Forschungsumgebungen auf dem Prüfstand. Es geht um die Bewahrung der Verknüpfung von Text mit unterschiedlichen Objekttypen und Formaten bis hin zur Software (z. B. für Simulationen). Zugleich sollen die einmal publizierten Forschungsergebnisse als Ausgangspunkt für die Bearbeitung weiterer Forschungsfragen genutzt werden, wie eingangs am Beispiel von Blumenbach-Online gezeigt wurde.

Bibliotheken haben im Bereich des elektronischen Publizierens in den letzten mehr als fünf Jahren Kenntnisse erworben, auf die für das Publizieren aus virtuellen Forschungsumgebungen aufgebaut werden kann. Zu den wesentlichen Bausteinen zählen auch hier Repositorien als Archive sowie die Anwendung von Metadaten. Letztere sind besonders wichtig, da sie über die persistente Identifizierung die dauerhafte Zitierbarkeit aller Komponenten einer Online-Publikation (Text, Forschungsdaten, Objekte) sicherstellen. Nur so ist die Akzeptanz neuer Publikationsformen zu gewinnen, wie bereits *Nature* in einschlägigen Beiträgen unterstrichen hat.<sup>35</sup>

## **VIRTUELLE FORSCHUNGsumGEBUNGEN UND IHRE VERNETZUNG**

Essentiell für das Arbeiten in virtuellen Forschungsum-

gebungen ist die Bereitschaft, die eigene Quellenbasis mit anderen zu teilen, im Interesse der Forschungsthematik Wissensbasen gemeinschaftlich zu erstellen. Diese Kooperation kann, je nach Arbeitsstand, in geschlossenen Gruppen stattfinden. Sie kann aber auch öffentlich sichtbar erfolgen, indem die Forschungsdaten und Quellen für jedermann zugänglich sind. Das Prinzip von Open Access ist eng mit der Verfügbarkeit und Vernetzbarkeit wissenschaftlicher Informationen in virtuellen Forschungsumgebungen verknüpft.<sup>36</sup>

Die kooperative Nutzung von Informationsressourcen war in den datenintensiven Wissenschaftsgebieten wie der Hochenergiephysik, der Klimaforschung oder Astronomie die wesentliche Triebkraft zur Etablierung virtueller Forschungsumgebungen. Heute wird die Bedeutung der Zusammenführung von Daten und Wissenssammlungen weitgehend anerkannt, aber in vielen wissenschaftlichen Communities nur ansatzweise umgesetzt.

Die Interoperabilität der einzelnen Datenquellen über die Standardisierung von Metadatenschemata ist, wie zuvor ausgeführt, eine entscheidende Voraussetzung für die Vernetzung virtueller Forschungsumgebungen. Für digitalisierte Sammlungen gibt es auf nationaler Ebene mit dem Aufbau der Deutschen Digitalen Bibliothek<sup>37</sup> und auf europäischer Ebene mit der Europeana<sup>38</sup> wichtige Ansätze zur Vernetzung. Einschränkend ist aber zumindest derzeit noch festzustellen, dass der Portalgedanke in diesen Initiativen vorherrscht, weniger der Austausch von Daten oder ihre externe Nutzung außerhalb der Portale in verschiedenen virtuellen Forschungsumgebungen. Für Publikationen gibt es hier vielversprechende Ansätze auf europäischer Ebene mit EU-Projekten wie DRIVER und OpenAIRE<sup>39</sup>, die sich explizit als Informationsinfrastrukturen mit offenem Datenaustausch verstehen und technisch entsprechend konzipiert sind. Für die Forschungsinfrastrukturen im europäischen ESFRI-Prozess wie DARIAH<sup>40</sup> und CLARIN<sup>41</sup> sind vergleichbare Konzepte vorgesehen.

Bibliotheken sind an einigen dieser Initiativen bereits beteiligt, teilweise federführend.<sup>42</sup> Es gilt nun, im Interesse der Unterstützung virtueller Forschungsumgebungen, offene, serviceorientierte Infrastrukturen disziplinen- und institutionenübergreifend weiterzuentwickeln, mit aktiver Beteiligung von Bibliotheken.

## **DIE SICHERUNG DER DIGITALEN LANGZEITVERFÜGBARKEIT UND NACHHALTIGKEIT**

Digitale Langzeitarchivierung ist ein wesentlicher Baustein für virtuelle Forschungsumgebungen und zugleich eine Grundinfrastruktur für digitale Informa-

tionsressourcen außerhalb solcher Systemumgebungen. In den meisten Fachdisziplinen besteht ein Bedarf an digitaler Langzeitarchivierung. Die Erfahrung der letzten Jahre zeigt, dass die Sicherung der Langzeitverfügbarkeit von Daten insbesondere für Wissenschaftler aus den Geistes- und Sozialwissenschaften eine essentielle Voraussetzung ist, sich auf virtuelle Forschungsumgebungen überhaupt einzulassen. Sie gehen vom dauerhaften Wert ihrer Forschungsergebnisse aus und übertragen zu dessen Sicherung das bewährte Archivkonzept der Bibliothek auf die digitale Welt.

Für Natur-, Lebens- und Ingenieurwissenschaften hat sich Langzeitverfügbarkeit erst in den letzten Jahren zu einem relevanten Thema entwickelt, zusätzlich sensibilisiert durch die expliziten Hinweise der DFG auf die Erfordernisse eines langfristigen Datenmanagements. Im Sinne der guten wissenschaftlichen Praxis wird hier die Verfügbarkeit für mindestens zehn Jahre nach Abschluss eines Forschungsvorhabens verlangt. Darüber hinaus erkennen Wissenschaftler zunehmend den Mehrwert von Datenbanken und Datenarchiven als grundlegender Infrastruktur. Damit wächst auch die Einsicht in die Notwendigkeit zur Sicherung der Daten. Entsprechende Konzepte auf institutioneller Ebene fehlen aber zumeist.

Das Verständnis von Langzeitarchivierung und Langzeitverfügbarkeit ist insgesamt eher diffus und beschränkt sich zumeist auf die Datensicherung, im Idealfall gekoppelt mit einer Strategie zur Bitstream Preservation, der Minimalanforderung zur Überlieferung der Daten-Bit-Folgen. Richtlinien (Policies) und Workflows fehlen, ebenso wie Ansprechpartner vor Ort. Hier sind Bibliotheken gefordert, diese Lücke zu füllen, in enger Zusammenarbeit mit Rechenzentren.

Die Verantwortung für die digitale Langzeitarchivierung und damit die Sicherung des langfristigen Zugriffs auf die Inhalte virtueller Forschungsumgebungen wird, abgesehen von Fachgebieten mit etablierten Sicherungsstrukturen,<sup>43</sup> sowohl in den Geistes- als auch in den Naturwissenschaften in der Verantwortung von Bibliotheken und Rechenzentren gesehen. Dies schafft eine gute Grundlage für Bibliotheken, sich als Kompetenzzentrum an der eigenen Institution zu etablieren.

Auf nationaler Ebene wird das Thema Langzeitarchivierung durch das Kompetenznetzwerk NESTOR<sup>44</sup> fortentwickelt, allerdings mit einem Fokus auf Publikationen und digitalisierten Sammlungen des kulturellen Erbes. Das Projekt WissGrid<sup>45</sup> entwickelt disziplinübergreifend (Medizin, Klimaforschung, Teilchenphysik, Astronomie, Geisteswissenschaften) Langzeitarchivierungsdienste und ein gemeinsames Verständnis von

der nachhaltigen Sicherung und Nachnutzung von Forschungsdaten.

#### **DIE ENTWICKLUNG UND BEREITSTELLUNG VON WERKZEUGEN / ANWENDUNGS-SOFTWARE**

Werkzeuge stellen neben den Informationsressourcen die Kernbausteine einer virtuellen Forschungsumgebung dar. Die Software-Entwicklung ist nun kein genuines Aufgabengebiet von Bibliotheken. Mit dem Einzug der digitalen Bibliothek in den 1990er Jahren hat sich eine Reihe von Einrichtungen dennoch mit Erfolg diesem Aufgabengebiet gewidmet. Die Gründe hierfür waren sicherlich unterschiedlich. Zu nennen sind neue Services, für die Anwendungsprogramme auf dem begrenzten IT-Markt der Bibliotheken fehlten, sowie der Antrieb zur Innovation in der neuen, digitalen Welt.

Einige Software-Anwendungen in virtuellen Forschungsumgebungen stammen aus der Bibliothekswelt, z. B. die erwähnten Repositoriensysteme wie DSpace oder Fedora bzw. eSciDoc. Diese Werkzeuge sind generisch in dem Sinne, dass sie durch alle Fachgebiete eingesetzt werden können. Selten anzutreffen sind hingegen Bibliotheken als Entwickler von fachspezifischen Anwendungsprogrammen zur Bearbeitung von Forschungsdaten und Objekten. Die SUB Göttingen wurde hier bekannt als Koordinator des BMBF-geförderten Verbundprojekts TextGrid,<sup>46</sup> das in der Kooperation mit Fachwissenschaftlern und IT-Experten Software zur Analyse, Beschreibung und Erschließung von Textkorpora entwickelt und als Open-Source-Produkte zur Verfügung stellt.

Das Interesse der Wissenschaftler selbst an der Zusammenarbeit mit Bibliotheken bei der Software-Entwicklung ist sicherlich am stärksten in den Geisteswissenschaften ausgeprägt. In den Natur- und Lebenswissenschaften haben sich in einigen Fachgebieten, beispielsweise in der Medizin oder den Biowissenschaften, entsprechende Fachinformatiker etabliert, die bestimmte Aufgaben in virtuellen Forschungsumgebungen wahrnehmen.

Bibliotheken wie Rechenzentren sollten jedoch eine Aufgabe darin sehen, zumindest eine Grundkenntnis von Anwendungsprogrammen zur Bearbeitung von Daten und Informationsquellen zu erwerben. Als Anlaufstellen zur Beratung für virtuelle Forschungsumgebungen (wie im Weiteren ausgeführt) können sie den Blick von Wissenschaftlern auf Werkzeuge wie Software zur automatisierten Textanalyse oder Informationsextraktion<sup>47</sup> lenken, die über eine Fachdisziplin und einen Projektkontext hinaus einsetzbar sind.

#### **BERATUNG UND KOOPERATION BEI DER KONZEPTION UND IM AUFBAU VIRTUELLER FORSCHUNGSUMGEBUNGEN**

Der Trend zum Aufbau virtueller Forschungsumgebungen ist in vielen Wissenschaftsbereichen sichtbar. Damit sind zugleich Ansprechpartner für eine Erstberatung an Universitäten und bei den Forschungsorganisationen gefragt, die gemeinsam mit den Forschern ihre Fragestellungen analysieren und in den zu Beginn eingeführten drei Grundbausteinen von virtuellen Forschungsumgebungen – Informationsobjekte, Anwendungssoftware und Systeme – Lösungsansätze aufzeigen. Bibliotheken könnten über die Erstberatung hinaus gemeinsam mit Rechenzentren Expertentandems aufbauen, die auch aktiv an der Projektkonzeption und -durchführung mitwirken. Bereits mit der Einführung von Digitalisierungsprogrammen durch die DFG ab 1997 konnten Bibliotheken nachweisbar eine positive Vermittlerrolle zwischen Forschern und Technologie-Dienstleistern übernehmen. Für die virtuellen Forschungsumgebungen wird es auf die Akzeptanz der einzelnen Bibliothek und noch konkreter auf die Akzeptanz der agierenden Personen ankommen, inwieweit Forscherverbände eine Beratungsfunktion in Anspruch nehmen und Kooperationen in der Praxis leben. Auch hier wird es unterschiedliche Annäherungen in Institutionen und Fachdisziplinen geben.

#### **KOOPERATION VON BIBLIOTHEKEN MIT DEN FACHWISSENSCHAFTEN UND IT-INFRASTRUKTUREN**

Virtuelle Forschungsumgebungen müssen forschungs- und disziplingetrieben entwickelt werden. Das vertraute Handlungsmuster der Entwicklung eines Services für Wissenschaftler greift daher nicht. Stattdessen müssen Bibliothekare mit Wissenschaftlern und IT-Experten in einem gemeinsamen Team tätig werden, gesteuert durch die Fragestellungen der Forschung und begleitet von Vorschlägen von Bibliothekaren und IT-Experten. So entfaltet sich ein iterativer Prozess der Entwicklung, der die Perspektive der Forschung ebenso berücksichtigt wie die Perspektive von Infrastrukturdienstleistern mit dem Einsatz von Standards, Verlinkung mit externen Forschungsumgebungen, Betrieb des Systems nach Projektende etc. Weder sollten Wissenschaftler in Projekten sich langwierig in das Verständnis standardisierter Metadaten oder der digitalen Langzeitarchivierung einarbeiten, noch sollten Bibliothekare und Mitarbeiter von Rechenzentren eigenständig für Wissenschaftler Werkzeuge entwickeln in dem Selbstverständnis, schon zu wissen, »was Wissenschaftler benötigen«. Konkrete Ansätze für neue

Strukturen zur Unterstützung der beschriebenen Zusammenarbeit liegen vor, z. B. mit der Max Planck Digital Library oder dem Anfang 2010 gegründeten Göttinger Centre for Digital Humanities (GCDH).<sup>48</sup>

Bibliotheken benötigen kompetente Partner aus dem IT-Management. Dazu bieten sich Rechenzentren, überregionale Servicezentren<sup>49</sup> sowie bibliothekarische Verbundzentralen für bestimmte Aufgaben an. Bibliotheken und Institutionen des IT-Managements stehen damit gemeinschaftlich Fachwissenschaftlern als Partner für den Aufbau und nachhaltigen Betrieb virtueller Forschungsumgebungen zur Verfügung.

## **HERAUSFORDERUNGEN UND KONSEQUENZEN FÜR BIBLIOTHEKEN**

Virtuelle Forschungsumgebungen stellen große Herausforderungen für Bibliotheken dar. Ihr bisheriger Aufgabenschwerpunkt lag auf der »Bereitstellung« von Informationen, bevorzugt durch Recherche-Instrumente und Zugriffspfade, über Kataloge und Portale. Mit der Etablierung virtueller Forschungsumgebungen rückt die Unterstützung der Auswertung und Bearbeitung von Informationsressourcen mit Software-Applikationen außerhalb der eigentlichen Bibliothekssystemumgebung durch die Wissenschaftler in den Fokus. Bibliothekare werden zu Akteuren im eigentlichen Forschungsprozess. Damit setzt sich ein Trend der Annäherung von Bibliotheken an die Forschung fort, der bereits mit dem Aufkommen des elektronischen Publizierens und der Diskussion um Open Access zu beobachten war, und sich unter anderem in verstärkten Kontakten von Bibliothekaren zu Wissenschaftlern und vice versa gezeigt hat. In Projekten zu virtuellen Forschungsumgebungen arbeiten Mitarbeiter der Bibliothek in gemeinschaftlichen Forschungs- und Arbeitsgruppen mit Fachwissenschaftlern und IT-Experten zusammen, wie die Beispiele von TextGrid, Blumenbach-Online und eine Reihe der DFG-geförderten Projekte zeigen.

Bibliotheken werden Stellung beziehen müssen, ob sie den Aufbau eines professionellen Informationsmanagements für Forschungsdaten verantwortlich mitgestalten und damit Verantwortung über die vertrauten Bibliotheksmaterialien hinaus übernehmen wollen. Eng damit verbunden ist die Einarbeitung in neue Metadatenschemata für Forschungsdaten, die signifikant von den etablierten Regelwerken in Bibliotheken abweichen.

Diese neuen Herausforderungen müssen auf verschiedenen Ebenen vorbereitet und begleitet werden. Im Mittelpunkt stehen Aspekte der Organisation von Bibliotheken und der Aus- bzw. Fortbildung von Bibliothekaren. Die Gründung neuer Organisationseinheiten

zur Abdeckung neuer Aufgabengebiete und Tätigkeiten kann sinnvoll sein, um ein deutliches Signal an die Mitarbeiter bezüglich einer neuen Daueraufgabe zu setzen. An der SUB Göttingen wurde beispielsweise im Rahmen einer großen Organisationsentwicklung im Jahre 2008 die Gruppe »Metadaten und Datenkonversion« gegründet.<sup>50</sup> Die neue Gruppe bietet einen für alle Mitarbeiter sichtbaren »Brückenkopf« aus der Welt der RAK-Katalogisierung in die Welt moderner Metadatenschemata. Die Gruppe »Elektronisches Publizieren« wurde ebenfalls an der SUB neu ins Leben gerufen.<sup>51</sup> Die beiden Gruppen wurden gezielt in der Abteilung »Informationsversorgung Zentral« angesiedelt, einer integrierten Erwerbungs- und Katalogisierungsabteilung, die sich, ganz allgemein gesprochen, um »Inhalte« kümmert. Damit wird es Mitarbeitern einer klassischen Organisationseinheit erleichtert, den Schritt in ein neues Aufgabengebiet zu wagen, da sie den vertrauten Abteilungsrahmen nicht verlassen. In dieser Abteilung können auch Forschungsdaten ihren Platz finden. Die Abteilung »Forschung und Entwicklung« der SUB Göttingen, die in dieser Form in Deutschland wohl nur noch bei der Max-Planck-Gesellschaft mit der Max Planck Digital Library zu finden ist, hat die wichtige Funktion der Exploration neuer Aufgabengebiete wie virtuelle Forschungsumgebungen, Forschungsdaten oder Digital Humanities. Sie hat daher schon vor einigen Jahren den Weg der Bibliothek »in die Forschung« mit vorbereitet. Das eingangs genannte Motto »Bibliothekare in die Forschung« ist auch wörtlich zu verstehen. Bibliothekare, die in einem der o. g. neuen Aufgabengebiete aktiv werden, könnten zumindest einen Teil ihrer Arbeitszeit räumlich außerhalb der Bibliothek in Forschungsverbänden als Mitglied in den mehrfach beschriebenen Teams mit Fachwissenschaftlern und IT-Experten verbringen.<sup>52</sup>

Alle Vorstellungen und Planungen zur Rolle von Bibliothekaren in virtuellen Forschungsumgebungen werden nur dann in die Praxis umgesetzt werden können, wenn entsprechende Aus- und Fortbildungsangebote angeboten und genutzt werden. Die kommenden Jahre erfordern daher auf Seiten der Bibliotheken und Ausbildungseinrichtungen für Bibliothekare einige Kraftanstrengungen, um die Voraussetzungen hierfür zu schaffen. Die Überarbeitung der Curricula an den Hochschulen sollte konsequent und mit Unterstützung der Hochschulrektorenkonferenz vorangetrieben werden. Letztere will sich über die in 2010 eingerichtete Arbeitsgruppe »Digitale Information in Forschung und Lehre« aus Mitgliedern von Hochschulleitungen den Fragen der Aus- und Fortbildung für Wissenschaftler und Fachpersonal in diesem Themenbereich annehmen. Neben nationalen sollten unbedingt auch

internationale Angebote wesentlich stärker als bisher genutzt werden. Besonders sei hier auf die Aktivitäten der Organisation wissenschaftlicher europäischer Bibliotheken, LIBER<sup>53</sup>, und das qualitativ hochwertige Angebot der TICER »Digital Libraries Summer School«<sup>54</sup> verwiesen.

Bibliothekare haben bereits in den 1990er Jahren den Weg in die digitale Welt angetreten. Elektronische Zeitschriften, Datenbanken, E-Books (in der frühen, ersten Ausprägung), Fachportale (»Virtuelle Fachbibliotheken«) und die Digitalisierung von Sammlungen gehören heute zum Kerngeschäft von Bibliotheken, einschlägige Standards wurden etabliert. In den vergangenen fünf Jahren haben sich Dokumentenserver, auch Repositorien genannt, und Open Access im Repertoire des bibliothekarischen Sprachgebrauchs und Tätigkeitsspektrum ausgebreitet. Bibliotheken werden auch die neuen Aufgaben zu virtuellen Forschungsumgebungen, Forschungsdaten etc. erfolgreich annehmen.

Über einzelne Bibliotheken hinaus sind viele weitere offene Fragen zu klären. Dazu gehören die Organisation und Rechtsform von virtuellen Forschungsumgebungen über Institutionen- und Ländergrenzen hinweg, ihre, auch nachhaltige Finanzierung, ein differenziertes Rechtemanagement sowie die digita-

le Langzeitarchivierung. Diese und andere infrastrukturell bedeutsamen Fragen werden derzeit durch die bereits erwähnte »Kommission Zukunft der Informationsinfrastruktur« (KII) behandelt. Die dort tätigen Akteure sind sich einig, dass Bund und Länder in Deutschland diese übergreifenden Fragen der Informationsinfrastrukturen mit Priorität angehen müssen und fühlen sich durch die jüngsten »Übergreifenden Empfehlungen zur Informationsinfrastruktur« des Wissenschaftsrats<sup>55</sup> in ihrem Bemühen bestätigt.

<sup>1</sup> Aus dem derzeit noch unveröffentlichten Bericht der Arbeitsgruppe »Virtuelle Forschungsumgebungen« im Rahmen der von der Gemeinsamen Wissenschaftskonferenz von Bund und Ländern (GWK) initiierten KII, [www.wgl.de/?nid=infrastr&nidap=&print=0](http://www.wgl.de/?nid=infrastr&nidap=&print=0)

<sup>2</sup> Ebenfalls derzeit unveröffentlicht; [www.allianzinitiative.de/de/handlungsfelder/virtuelle\\_forschungsumgebungen/arbeitsgruppe/](http://www.allianzinitiative.de/de/handlungsfelder/virtuelle_forschungsumgebungen/arbeitsgruppe/)

<sup>3</sup> Wie Anm. 1.

<sup>4</sup> A.a.O.

<sup>5</sup> [www.wissenschaftsrat.de/aktuelles-presse/pressemitteilungen/2009/nummer-o8-vom-29-mai-2009/](http://www.wissenschaftsrat.de/aktuelles-presse/pressemitteilungen/2009/nummer-o8-vom-29-mai-2009/)

<sup>6</sup> [www.wissenschaftsrat.de/index.php?id=345&L=](http://www.wissenschaftsrat.de/index.php?id=345&L=)

<sup>7</sup> Joint Information Systems Committee, [www.jisc.ac.uk/whatwedo/programmes/vre.aspx](http://www.jisc.ac.uk/whatwedo/programmes/vre.aspx)

<sup>8</sup> Hier sind insbesondere die Projekte zu nennen, die vom European Strategy Forum on Research Infrastructures (ESFRI) empfohlen werden.

<sup>9</sup> Virtual Research Environment: Collaborative Landscape Study. A JISC funded project. Erarbeitet durch Annamaria Carusi und Torsten Reimer. Januar 2010, [www.jisc.ac.uk/publications/reports/2010/vre-landscapestudy.aspx](http://www.jisc.ac.uk/publications/reports/2010/vre-landscapestudy.aspx)

<sup>10</sup> Deutsche Forschungsgemeinschaft (DFG) Aktionslinie 13: »Virtuelle Forschungsumgebungen«, [www.dfg.de/foerderung/programme/infrastruktur/lis/lis\\_themenschwerpunkt\\_digitale\\_information/lis\\_informationsmanagement/index.html](http://www.dfg.de/foerderung/programme/infrastruktur/lis/lis_themenschwerpunkt_digitale_information/lis_informationsmanagement/index.html)

<sup>11</sup> Bundesministerium für Bildung und Forschung (BMBF), Projekte in der D-Grid-Initiative, [www.d-grid.org/](http://www.d-grid.org/)

<sup>12</sup> <http://lh.web.cern.ch/lhc/>

<sup>13</sup> Seit 2006 wird eine Internet-basierte Plattform aufgebaut, die Wissenschaftlern Werkzeuge und Dienste für die Auswertung von textbasierten Daten in unterschiedlichen digitalen Archiven bietet – unabhängig von Datenform, Softwareausstattung oder Standort, [www.textgrid.org](http://www.textgrid.org)

<sup>14</sup> [www.blumenbach-online.de/](http://www.blumenbach-online.de/)

<sup>15</sup> Z. B. mit Hilfe des TextGrid-Werkzeuges »Text-Bild-Link-Editor«, [www.textgrid.de/beta/text-bild-link-editor.html](http://www.textgrid.de/beta/text-bild-link-editor.html)

<sup>16</sup> [www.newtonproject.sussex.ac.uk/prism.php?id=44](http://www.newtonproject.sussex.ac.uk/prism.php?id=44); <http://darwin-online.org.uk/>; [www.linnaeus.ci8.net](http://www.linnaeus.ci8.net)

<sup>17</sup> A.a.O.

<sup>18</sup> Eingeführt durch Jürgen Renn, MPI Wissenschaftsgeschichte in die o.g. AG Virtuelle Forschungsumgebungen der KII.

<sup>19</sup> Dank für diese Hinweise geht an Frau Stefanie Rühle, Mitarbeiterin in der Gruppe »Metadaten« der SUB Göttingen und im DFG-Projekt »Kompetenzzentrum Interoperable Metadaten« (KIM)

<sup>20</sup> [http://cidoc.icom.museum/WG\\_Data\\_Harvesting\(en\)\(E1\).xml](http://cidoc.icom.museum/WG_Data_Harvesting(en)(E1).xml)

<sup>21</sup> [www.openarchives.org/ore/](http://www.openarchives.org/ore/)

<sup>22</sup> [https://www.nescent.org/wg\\_dryad/Metadata\\_Profile](https://www.nescent.org/wg_dryad/Metadata_Profile)

<sup>23</sup> Angeboten in Deutschland von der TIB Hannover im Rahmen der Organisation DataCite, <http://datacite.org/>

<sup>24</sup> Module und Programmelemente in Sonderforschungsbereichen, Informationsmanagement und Informationsinfrastruktur: [www.dfg.de/foerderung/programme/koordinierte\\_programme/sfb/module/modul\\_inf/index.html](http://www.dfg.de/foerderung/programme/koordinierte_programme/sfb/module/modul_inf/index.html)

<sup>25</sup> »Ein professionelles Management der Daten wird erwartet. Daher ist die systematische Zusammenarbeit des Sonderforschungsbereichs mit Informationseinrichtungen (z. B. Bibliotheken, Rechenzentren) zu sichern.«, a.a.O.

<sup>26</sup> Vgl. das Directory of Open Access Repositories, OpenDOAR, [www.opendoar.org/](http://www.opendoar.org/)

<sup>27</sup> Die umfassendste Übersicht in Deutschland bietet das Zentralverzeichnis Digitalisierter Drucke, ZVDD, [www.zvdd.de/](http://www.zvdd.de/)

<sup>28</sup> [www.gesis.org/dienstleistungen/forschungsdatenzentren/](http://www.gesis.org/dienstleistungen/forschungsdatenzentren/)

<sup>29</sup> [www.dkrz.de/dkrz/intro\\_s](http://www.dkrz.de/dkrz/intro_s)

<sup>30</sup> <http://adsabs.harvard.edu/>

<sup>31</sup> Fedora und DSpace haben sich in der Organisation DuraSpace zusammengeschlossen, <http://duraspace.org/>

<sup>32</sup> [www.bsz-bw.de/digitalebibliothek/opus.html](http://www.bsz-bw.de/digitalebibliothek/opus.html)

<sup>33</sup> [www.eprints.org/](http://www.eprints.org/)

<sup>34</sup> <https://www.escidoc.org/>

## LITERATUREMPFEHLUNGEN ZU »VIRTUELLEN FORSCHUNGSUMGEBUNGEN«

Virtual Research Environment: Collaborative Landscape Study. A JISC funded project. Erarbeitet durch Annamaria Carusi und Torsten Reimer. Januar 2010, [www.jisc.ac.uk/publications/reports/2010/vrelandscapestudy.aspx](http://www.jisc.ac.uk/publications/reports/2010/vrelandscapestudy.aspx) (mit vielen weiterführenden Hinweisen).

Neuroth, Heike; Jannidis, Fotis; Rapp, Andrea; Lohmeier, Felix: »Virtuelle Forschungsumgebungen für e-Humanities. Maßnahmen zur optimalen Unterstützung von Forschungsprozessen in den Geisteswissenschaften«, in: Bibliothek: Forschung und Praxis, 33, 2009, Nr. 2, S. 161–169.

Neuroth, Heike; Aschenbrenner, Andreas; Lohmeier, Felix: »e-Humanities – eine virtuelle Forschungsumgebung für die Geistes-, Kultur- und Sozialwissenschaften«, in: Bibliothek: Forschung und Praxis, 31, 2007, Nr. 3, S. 272–279.

## FÖRDERMÖGLICHKEITEN ZU VIRTUELLEN FORSCHUNGSUMGEBUNGEN IN DEUTSCHLAND

Deutsche Forschungsgemeinschaft (DFG) Aktionslinie 13: »Virtuelle Forschungsumgebungen«, [www.dfg.de/foerderung/programme/infrastruktur/lis/lis\\_themenschwerpunkt\\_digitale\\_information/lis\\_informationsmanagement/index.html](http://www.dfg.de/foerderung/programme/infrastruktur/lis/lis_themenschwerpunkt_digitale_information/lis_informationsmanagement/index.html)

Bundesministerium für Bildung und Forschung (BMBF), Projekte in der D-Grid-Initiative, [www.d-grid.org/](http://www.d-grid.org/)

<sup>35</sup> Z. B. in: *Nature Genetics*, Vol. 41, Number 10, October 2009 und Vol. 42, Number 9, September 2010, jeweils im »Editorial«.

<sup>36</sup> Zu dieser Feststellung gelangte auch die AG Virtuelle Forschungsumgebungen der KII, a. a. O. Ausnahmen sind beispielsweise in der Industrie-Auftragsforschung oder beim Umgang mit Datenschutzsensiblen Forschungsdaten anzutreffen.

<sup>37</sup> [www.deutsche-digitale-bibliothek.de/](http://www.deutsche-digitale-bibliothek.de/)

<sup>38</sup> [www.europeana.eu/](http://www.europeana.eu/)

<sup>39</sup> [www.driver-repository.eu/](http://www.driver-repository.eu/); [www.openaire.eu/](http://www.openaire.eu/)

<sup>40</sup> [www.dariah.eu/](http://www.dariah.eu/)

<sup>41</sup> [www.clarin.eu/](http://www.clarin.eu/)

<sup>42</sup> Die Deutsche Nationalbibliothek (DNB), die Bayerische Staatsbibliothek, die SUB Göttingen sowie die SLUB Dresden sind bibliothekarische Mitglieder im Kompetenznetzwerk der Deutschen Digitalen Bibliothek; die DNB ist federführend an der Entwicklung von Europeana beteiligt; die SUB Göttingen federführend in den laufenden Projekten OpenAIRE und DARIAH.

<sup>43</sup> Vgl. GESIS in den Sozialwissenschaften und das DKRZ für Klimadaten, a. a. O.

<sup>44</sup> [www.langzeitarchivierung.de/index.htm](http://www.langzeitarchivierung.de/index.htm)

<sup>45</sup> [www.wissgrid.de/workgroups/ap3.html](http://www.wissgrid.de/workgroups/ap3.html)

<sup>46</sup> A. a. O.

<sup>47</sup> Beispielhaft seien die Entwicklungen im Rahmen des BMBF-geförderten Projekts eAqua zur Informationsextraktion aus Quellen in den Altertumswissenschaften genannt, [www.eaqua.net/](http://www.eaqua.net/)

<sup>48</sup> Das GCDH wird von den fünf geistes- und gesellschaftswissenschaftlichen Fakultäten der Universität gemeinsam mit der Staats- und Universitätsbibliothek getragen mit dem Ziel, Services, Forschung und Lehre in den IT-gestützten Geisteswissenschaften integriert zu entwickeln und anzubieten.

<sup>49</sup> Z. B. das FIZ Karlsruhe.

<sup>50</sup> Beginnend mit der nicht RAK-basierten Erfassung von Struktur-Metadaten für Digitalisate und Beschreibungen von Internetquellen in virtuellen Fachbibliotheken ist hier auch personell die Expertise für das gemeinschaftlich mit der Deutschen Nationalbibliothek und der Max Planck Digital Library betriebene Kompetenzzentrum für Interoperable Metadaten (KIM) angesiedelt.

<sup>51</sup> Diese Gruppe fasst die Aktivitäten des Universitätsverlags, der elektronischen Dissertationen, zu GoeScholar als institutionellem Repository und zahlreiche einschlägige Projektaktivitäten zusammen.

<sup>52</sup> Für Mitarbeiter der Abteilung »Forschung und Entwicklung« der SUB wird diese Perspektive wohl noch in 2011 Realität werden, wenn sie in die geplanten Räumlichkeiten des Göttingen Centre for Digital Humanities (GCDH) mit einziehen werden, um dort an neuen Verbundprojekten wie DARIAH-DE oder Digital Humanities mitzuwirken.

<sup>53</sup> [www.libereurope.eu/](http://www.libereurope.eu/)

<sup>54</sup> [www.tilburguniversity.edu/research/institutes-and-research-groups/ticer/courses.html](http://www.tilburguniversity.edu/research/institutes-and-research-groups/ticer/courses.html)

<sup>55</sup> [www.wissenschaftsrat.de/index.php?id=344&L=](http://www.wissenschaftsrat.de/index.php?id=344&L=)

## DER VERFASSER

**Prof. Dr. Norbert Lossau** ist Direktor der Niedersächsischen Staats- und Universitätsbibliothek Göttingen, 37070 Göttingen, Tel.: 0551 – 5212, Mail: [lossau@sub.uni-goettingen.de](mailto:lossau@sub.uni-goettingen.de)

# Kommentar

GERHARD LAUER

## Kommentar aus Sicht eines Literaturwissenschaftlers

Virtuelle Forschungsumgebungen sind Modernisierungssphänomene. Wie alle Modernisierungen sind diese mit systemischen Irritationen etablierter Institutionen verbunden. Das gilt für die Bibliotheken, das gilt aber nicht weniger für die Wissenschaften. Und wie andere Modernisierungsprozesse auch, verläuft auch diese Modernisierung nicht linear und nicht in allen Bereichen mit derselben Geschwindigkeit. »Bibliothekare in die Forschung« – mit dieser Zuspitzung beschreibt der Beitrag von Norbert Lossau einen Modernisierungsschub, den virtuelle Forschungsumgebungen auf Bibliotheken haben und haben werden, ganz einfach weil sie die tradierten Funktionszuweisungen an die Bibliotheken verändern. Bibliotheken, darauf macht Norbert Lossau zu Recht aufmerksam, werden durch virtuellen Forschungsumgebungen zu Informationsinfrastrukturen, in denen Forschung stattfindet, genauer durch die vielfach erst Forschung möglich wird. Auch Bücher waren in dieser Hinsicht so etwas wie virtuelle Forschungsumgebungen, weil sie mit anderen Büchern in einer Systematik aufgestellt,

den Denk- und Arbeitsraum der Wissenschaften lange mitbestimmt haben. Aber die digitalen Forschungsumgebungen haben mehr als Bücher. Sie stellen neben die Bücher und Zeitschriften analytische Werkzeuge bereit und modellieren die Untersuchungsobjekte. Sie eröffnen neue Formen der wissenschaftlichen Zusammenarbeit, die bisher bekannte räumliche Grenzen weitgehend überwindet. Schließlich kümmern sie sich auch um die Verbreitung und langfristige Vorhaltung der Forschungsergebnisse. Wenn Bibliotheken Häuser sind, die Bücher in unterschiedliche Räume anordnen, dann sind virtuelle Forschungsumgebungen Städte, die Häuser mit unterschiedlichen Funktionen in einer Stadt zusammenstellen.

Solche Städte gibt es bereits, etwa in der Astro- oder Teilchenphysik. Hier sind unter der D-Grid-Initiative im AstroGrid oder teilweise schon in Fachrepositorien wie arXiv.org die Beobachtungsdaten, Publikationen und im Grid inzwischen auch die Tools für die Forschung zusammengeführt.<sup>1</sup> Der Umbau ist in diesen Fächern wie der Physik kein Zufall, sondern hängt



Gerhard Lauer

Foto: privat

mit einem Umstand zusammen, der gar nicht genug hervorgehoben werden kann: der Entstehung von datenintensiven Forschungen.<sup>2</sup> Sie sind ein Kind des digitalen Zeitalters. Die gewaltigen Datenmengen, die hier die Forschung ausmachen, waren bis vor kurzem noch kaum vorstellbar, noch weniger, wie sie als Infrastruktur überhaupt zu organisieren sind. Der Large Hadron Collider in Genf produziert im Jahr selbst dann, wenn der größte Teil seiner Daten gar nicht archiviert wird, immer noch derzeit zehn Petabyte Daten, das sind 20 Millionen CD übereinandergeschichtet, ein Datenberg höher als der Mont Blanc. DNA-Sequenzierautomaten, wie sie heute in den Lebenswissenschaften zum Einsatz kommen, analysieren Daten im Umfang von einem Terabyte pro Minute. Datenhäuser wie die GenBank führen solche Datenströme zusammen.<sup>3</sup> Facebook erzeugt ca. 30 Milliarden Postings im Monat, die als Quelle für sozialwissenschaftliche Daten prinzipiell von Interesse sind.<sup>4</sup> Und es bleibt nicht bei diesen Fächern stehen. In Disziplinen wie der Archäologie, deren Arbeitsalltag sich noch gestern nicht wesentlich von dem ihrer Kollegen vor einem Jahrhundert unterschieden hat, können jetzt ganze Grabungsgelände in 3D-Scannern aufgenommen und für die weitere Forschung zur Verfügung gestellt werden. In den Textwissenschaften entsteht mit TextGrid gerade die erste Grid-basierte virtuelle Forschungsumgebung in den Geisteswissenschaften, die korpusorientierte Forschung in einem ganz neuen Umfang möglich machen wird.<sup>5</sup> In der Linguistik entstehen Forschungsergebnisse, die auf der Untersuchung von mehr als tausend Jahren Sprachdaten beruhen,<sup>6</sup> und Google Books stellt mit Culturomics Tools für die Untersuchung von mehr als fünf Millionen Büchern zur Verfügung.<sup>7</sup> Manche wie Tony Hey sprechen vom vierten Paradigma der datenintensiven Forschung, dem die Zukunft gehöre,<sup>8</sup> andere wie der Chefredakteur des Magazins »Wired« gar vom Petabyte-Zeitalter, das Daten anstelle von Theorien setze.<sup>9</sup> Das mag übertrieben sein, aber mindestens einzelne Fächer haben bereits die Schwelle zu den virtuellen Forschungsumgebungen überschritten, die datenintensive Forschung möglich macht. Wenn der Sloan Digital Sky Survey demnächst seine 1998 begonnene Durchmusterung des Himmels abgeschlossen haben wird und das auf diesen Beobachtungsdaten aufruhende Virtual Observatory an den Start geht, dann werden nicht Teleskope auf den Himmel gerichtet, sondern Tools auf schier unvorstellbar große Mengen strukturierter Daten.<sup>10</sup> Damit wird Forschung erst in einer virtuellen Forschungsumgebung möglich.

Der Beitrag von Norbert Lossau macht darauf aufmerksam, dass dieser Umbau der Forschung durch virtuelle Forschungsumgebungen mehr als nur eine Erweiterung der Bibliotheken ist. Wenn Bibliotheken bis in die rechtlichen Dimensionen solche datenintensive Forschung ermöglichen sollen, das heißt dann auch Reputationsverteilung über Data-Sharing organisieren müssen,<sup>11</sup> werden sie selbst zu einem Teil der Forschung, weil an ihrer Strukturierung der Datenmengen und -typen wesentlich hängt, welche Forschung überhaupt möglich sein wird. Zu Recht macht er daher auf die besondere Rolle von Metadatenschemata aufmerksam, deren Entwicklung und Standardisierung Wissenschaftler und Bibliothekare zusammenführt, oder sollte ich nicht genauer sagen: Informationswissenschaftlerinnen und -wissenschaftler. Denn mit den klassischen Aufgaben hat und wird das immer weniger zu tun haben. Informationswissenschaftliche Forschungsabteilungen, die datenintensive Forschungsinfrastrukturen entwickeln und pflegen können, werden an vielen Bibliotheken zum Alltag gehören. Umgekehrt müssen nicht nur Bibliothekare in die Forschung, sondern Forscherinnen und Forscher in die Informations-Infrastrukturen. Der Rollenwechsel findet auf beiden Seiten statt. Das gehört zu dieser digitalen Modernisierung, an deren Anfang wir gerade erst stehen.

#### Der Verfasser

**Prof. Dr. Gerhard Lauer**, Universität Göttingen, Seminar für Deutsche Philologie, Käte-Hamburger-Weg 3, 37073 Göttingen, Tel.: 0551 – 39 7526, Mail: gerhard.lauer@phil.uni-goettingen.de

<sup>1</sup> [www.d-grid.org](http://www.d-grid.org).

<sup>2</sup> Vgl. das Sonderheft von »Science« (2011), [www.sciencemag.org/special/data/](http://www.sciencemag.org/special/data/).

<sup>3</sup> [www.ncbi.nlm.nih.gov/genbank/](http://www.ncbi.nlm.nih.gov/genbank/).

<sup>4</sup> Z. B. D. Lazer, A. Pentland et al. (2009), Life in the network: the coming age of computational social science, *Science* 323 (5915), 721–723.

<sup>5</sup> [www.textgrid.org](http://www.textgrid.org).

<sup>6</sup> E. Lieberman et al. (2007), Quantifying the evolutionary dynamics of language, *Nature* 449, 713–716.

<sup>7</sup> [www.culturomics.org](http://www.culturomics.org).

<sup>8</sup> T. Hey et al. (2010), The Fourth paradigm. Data-Intensive Scientific Discovery, auch online: <http://research.microsoft.com/en-us/collaboration/fourthparadigm/>.

<sup>9</sup> C. Anderson (2008), The End of Theory: The Data Deluge Makes the Scientific Method Obsolete, *Wired*, auch online: [www.wired.com/science/discoveries/magazine/16-07/pb\\_theory](http://www.wired.com/science/discoveries/magazine/16-07/pb_theory).

<sup>10</sup> [www.sdss.org](http://www.sdss.org).

<sup>11</sup> B. Nelson (2009), Empty Archives, *Nature*, [www.nature.com/news/2009/090909/pdf/461160a.pdf](http://www.nature.com/news/2009/090909/pdf/461160a.pdf)