

Zugang, Nachnutzung und Reproduzierbarkeit. Anmerkungen zur künftigen Ausrichtung einer wissenschaftsadäquaten Informationsinfrastruktur



Dr. Johannes Fournier

Deutsche Forschungsgemeinschaft
Gruppe „Wissenschaftliche Literatur-
versorgungs- und Informationssysteme“ (LIS)
Kennedyallee 40
D-53170 Bonn
E-Mail: Johannes.Fournier@dfg.de

Vor dem Hintergrund aktueller Entwicklungen in der Informationsinfrastruktur formuliert der Beitrag Empfehlungen, deren Umsetzung den offenen Zugang und die möglichst umfassende Nachnutzung von Forschungsergebnissen gewährleisten soll. Dazu wird dargestellt, welche rechtlichen, technischen, organisatorischen und finanziellen Vorkehrungen getroffen werden müssen, um eine umfassende Nutzung von herkömmlichen Publikationen, von Forschungsdaten oder von retrospektiv erstellten Digitalisaten zu ermöglichen. Schlüsselwörter: Informationsinfrastruktur; Open Access und Open Data; Nachnutzbarkeit

Access, Re-Use and Reproducibility. Notes for Future Direction of an Appropriate Science Information Structure

This contribution formulates recommendations that aim at ensuring open access and the comprehensive re-use of research results by taking the current developments in the information infrastructure into account. It describes legal, technical, organizational and financial provisions necessary to allow for the extensive re-use of traditional publications, research data or retrospectively digitized information.

Keywords: Information infrastructure; open access and open data; re-usability

0 Vorbemerkung

Die folgenden Ausführungen basieren auf einer Stellungnahme, die der Autor für eine Anhörung der Projektgruppe „Bildung und Forschung“ der Enquete-Kommission „Internet und digitale Gesellschaft“ des Deutschen Bundestags am 7. November 2011 angefertigt hat.¹ Diese Stellungnahme verfolgt die Zielsetzung, aus Sicht der Forschung und der

¹ S. unter http://www.bundestag.de/internetenquete/dokumentation/Bildung_und_Forschung/PGBuF_2011-10-27_Projektgruppe_Bildung_und_Forschung_laedt_Experten_ein/index.jsp.

Forschungsförderung Anforderungen an eine zeitgemäße Informationsinfrastruktur zu formulieren, die Wissenschaftler und Wissenschaftlerinnen in ihrer Forschung effizient unterstützt. Dafür wird zum einen auf einschlägige Positionspapiere rekurriert, die gerade in jüngster Zeit Vorschläge für einen effizienten Ausbau einer forschungsbezogenen Informationsinfrastruktur unterbreitet haben. Zum anderen basiert die Stellungnahme auf Erfahrungen und Aktivitäten der Gruppe „Wissenschaftliche Literaturversorgungs- und Informationssysteme“ (LIS) der Deutschen Forschungsgemeinschaft (DFG). Die von dieser Gruppe betreuten Programme und Projekte zielen auf den Aufbau einer optimalen Informationsinfrastruktur für die deutsche Wissenschaft ab. Schwerpunkte der DFG-Förderung sind dabei die überregionale Lizenzierung digitaler Inhalte, die Erstellung, Bereitstellung und Verbreitung elektronischer Publikationen im Open Access, die Retrodigitalisierung wissenschaftsrelevanter Kernbestände aus Bibliotheken und Archiven, die Entwicklung Virtueller Forschungsumgebungen und der Aufbau einer datenbezogenen Infrastruktur. In der Umsetzung ihrer Förderschwerpunkte arbeitet die Gruppe LIS im nationalen Rahmen eng mit den in der Allianz der Wissenschaftsorganisationen zusammengeschlossenen Partnern der DFG, im europäischen Netzwerk „Knowledge Exchange“ eng mit Infrastruktur-Fördereinrichtungen aus Dänemark, den Niederlanden und dem Vereinigten Königreich zusammen, mit denen die DFG die Vision teilt, wissenschaftliche Inhalte über das Internet für Nutzer frei verfügbar bereitzustellen.²

1 Kurzcharakteristik der die Informationsinfrastruktur aktuell prägenden Entwicklungen

Schon seit Mitte der 1990er Jahre ist eine deutliche Zunahme digitaler, für das wissenschaftliche Arbeiten hoch relevanter Informationen zu konstatieren. In analoger Form vorliegende Wissensbestände werden durch retrospektive Digitalisierung über das Internet verfügbar gemacht. Zugleich wächst die Menge genuin digital produzierter Informationen exponentiell an. Wissenschaftler und Wissenschaftlerinnen nutzen diese digitalen Bestände nicht nur für die eigene Lektüre, sondern auch um Fragestellungen zu lösen, deren Beantwortung eine über das Internet organisierte, koopera-

² Die Förderprogramme und Aktivitäten der Gruppe LIS sind auf den Internetseiten der DFG unter <http://www.dfg.de/lis> erläutert, die Schwerpunkte der Allianz-Initiative „Digitale Information“ unter <http://www.allianz-initiative.de/> beschrieben und die Arbeiten des Netzwerks „Knowledge Exchange“ unter <http://www.knowledge-exchange.info/> ausführlich dokumentiert.

tive Forschung erfordert sowie die Analyse und Auswertung digitaler Information mit Hilfe Computer-unterstützter Verfahren voraussetzt. Schließlich werden die Ergebnisse dieser Forschung gegebenenfalls in neuartigen Formen der wissenschaftlichen Publikation präsentiert.³

Insbesondere das Aufkommen sogenannter Virtueller Forschungsumgebungen belegt den Trend zu Computer-unterstützten, Internet-basierten, kollaborativen Arbeitsweisen. Virtuelle Forschungsumgebungen zielen auf eine Unterstützung des gesamten, über verteilte Datennetze organisierten wissenschaftlichen Arbeitsprozesses ab, von der Erhebung und Bereitstellung der Daten über deren Analyse und Auswertung bis hin zur Publikation. Zur Auswertung der Fülle digitaler Informationen wird dabei auf automatisierte Verfahren wie das Erkennen spezifischer Muster z.B. in Forschungsdaten, Texten oder Bildern sowie linguistische Analysen und die Feststellung statistischer Verteilungen gesetzt, wobei derartige Verfahren („*text mining*“ und „*data mining*“) keineswegs auf die Natur-, Lebens- und Technikwissenschaften beschränkt sind, sondern auch in den Sozial- und Geisteswissenschaften zur Anwendung kommen. Die Bedeutung einer digitalen Informationsinfrastruktur auch für diese Fächer belegt nicht nur die aktuelle Diskussion darüber, wie grundlegende physische Infrastrukturen der geisteswissenschaftlichen Forschung als digitale Bibliotheken, Archive und Sammlungen optimal verfügbar gemacht werden können, sondern auch der Einbezug geisteswissenschaftlicher Vorhaben in ESFRI, das European Strategy Forum on Research Infrastructures (z.B. mit CLARIN und DARIAH).⁴

Mit der stärkeren Verbreitung digitaler Informationen ändert sich allmählich auch die Darstellung von Forschungsergebnissen. Der wissenschaftliche (Zeitschriften-) Artikel und die Monographie in ihrer traditionellen Form sind nach wie vor die Größen, auf denen wissenschaftliche Karrieren gründen. Doch werden elektronische Publikationen schon heute mit einer Vielzahl höchst unterschiedlicher Inhalte angereichert. In Publikationen, die mit Forschungsdaten, Simulationen, Quellentexten, multimedialen Inhalten und vielem anderen vernetzt sind („*enhanced publications*“), weicht die streng linear und textbasierte Form der wissenschaftlichen Darstellung anderen Präsentationsformen. Dabei ermöglicht die vielfältige Vernetzung der angereicherten Publikation nicht nur, Forschungsergebnisse leichter zu finden und besser nachvollziehbar zu machen, sondern sie führt

auch zu neuen Fragen und somit zu neuen Erkenntnissen.⁵ Freilich wird die umfassende Nutzung solcher wissenschaftlicher Präsentationen, deren angereicherte Inhalte wie Daten, Digitalisate oder wissenschaftliche Artikel aus heterogenen Quellen stammen, häufig die Publikation auf (Verlags-) übergreifenden Rechercheplattformen ohne technischen Schutzmaßnahmen erfordern.

Mit dem Schlagwort der „*reproducible research*“ wird schließlich die Idee bezeichnet, dass Veröffentlichungen künftig nur noch als eine Art Verweis auf die grundlegenden, „*eigentlichen*“ Forschungsergebnisse – auf Daten, Methoden und Werkzeuge – dienen werden, die in für andere Wissenschaftler nachnutzbaren Formen bereitgestellt werden.⁶ Solches Vorgehen gewährleistet nicht nur maximale Transparenz, sondern erhöht die Effizienz und Innovationskraft von Forschung. Umgekehrt setzt eine streng reproduzierbare Forschung umfassende Rechtseinräumungen für die rechtlich statthafte Nachnutzung von Daten, Werkzeugen und Methoden voraus.

Aus dem Vorhergehenden folgt, dass die effiziente und innovative Kommunikation einer global agierenden Wissenschaft auf die möglichst offene Bereitstellung und die umfassende Nachnutzung digitaler Inhalte angewiesen ist. Mit der Forderung nach und der Umsetzung von *Open Access* und *Open Data* soll der für Nutzer und Nutzerinnen barrierefreie Zugang zu Forschungsergebnissen gewährleistet werden. Allerdings ist nicht nur der freie Zugang entscheidend, sondern auch die Möglichkeit, Forschungsergebnisse Dritter in der eigenen Forschung ohne rechtliche und technische Beschränkungen produktiv nachnutzen zu können („*re-use*“), worunter insbesondere die Nachnutzung von Forschungsergebnissen durch Anwendung Computer-unterstützter Verfahren fällt. Damit verlässlich dokumentiert werden kann, auf welchen und auf wessen Forschungsergebnissen die eigene Arbeit basiert, ist es schließlich erforderlich, dass digitale Daten, Quellen und Publikationen ebenso wie Autoren, Arbeitsgruppen und Institutionen über eindeutige, nicht wandelbare „*Kennungen*“ identifiziert werden können („*persistent identification*“). Dies ist nicht nur Voraussetzung einer effizienten Verknüpfung digitaler Information, sondern gewährleistet auch, dass Forschungsleistungen ihren Urhebern zugeschrieben werden und diese somit ihre wissenschaftliche Reputation vermehren können.

Die zuvor skizzierten Entwicklungen eröffnen der Wissenschaft Chancen, stellen sie aber auch vor grundlegende Herausforderungen. Die größte Chance ist gewiss darin zu sehen, dass die weltweite, grenzüberschreitende Kooperation in der Wissenschaft erleichtert wird. Mit dem Zugriff auf digitale Daten können E-Science und E-Humanities neue Methoden entwickeln. Der möglichst freie Zugang

3 Das im April 2011 vorgelegte Gesamtkonzept der Kommission „Zukunft der Informationsinfrastruktur“ stellt die grundlegenden, aktuellen Entwicklungen der Informationsinfrastruktur in acht ausgewählten Themenfeldern dar, wobei zu jedem Themenfeld systematisch sowohl die nationalen als auch die internationalen Entwicklungen beschrieben und analysiert werden. Als Themen werden (1) Lizenzierung, (2) Hosting und Langzeitarchivierung, (3) Nichttextuelle Materialien, (4) Kulturelles Erbe und Retrodigitalisierung, (5) Virtuelle Forschungsumgebungen, (6) Open Access, (7) Forschungsdaten und (8) Informationskompetenz und Ausbildung behandelt. Das Konzept ist verfügbar unter <http://www.leibniz-gemeinschaft.de/?nid=infrastr> und wird hier im Folgenden zitiert als „KII-Gesamtkonzept“.

4 Vgl. dazu die Ausführungen des BMBF zur nationalen und internationalen Strukturbildung in den Geistes-, Sozial- und Kulturwissenschaften unter <http://www.bmbf.de/de/4817.php> vom 28. Juli 2011.

5 Unter <http://www.surffoundation.nl/en/themas/openonderzoek/verrijktepublicaties/Pages/default.aspx> wird das Konzept der „Enhanced Publications“ ausführlich erläutert; außerdem sind dort Informationen zu Projekten aus fünf unterschiedlichen Wissenschaftsbereichen zu finden, in denen „Enhanced Publications“ erstellt und analysiert wurden.

6 S. die zu diesem Thema unter <http://www.stanford.edu/~vcs/AAAS2011> zugänglichen Vorträge.

trägt nicht nur zu einer weltweiten Sichtbarkeit von Forschungsergebnissen und somit zu einer Effizienzsteigerung des Forschungsdiskurses bei, sondern erleichtert auch die produktive Nachnutzung von Forschungsergebnissen durch den kommerziellen Sektor, insbesondere durch KMU.⁷ Dass gerade Open Access eine echte und auch für Industrienationen gewinnbringende Zusammenarbeit mit der Wissenschaft in Entwicklungsländern fördert, sei an zwei Beispielen kurz illustriert: Mit der Überführung der vom German Institute for Global Area Studies (GIGA) herausgegebenen regionalwissenschaftlichen Zeitschriften zu Afrika, Lateinamerika, Südostasien und China in den Open Access wurde der wissenschaftlichen Diskurs nicht nur *über*, sondern *mit* den jeweiligen Zielländern auf Augenhöhe erst ermöglicht.⁸ Und die Open-Access-Zeitschriftenplattform „Bioline International“ zeigt, wie die medizinische, ökonomische oder sozialwissenschaftliche Forschung in westlichen Ländern davon profitiert, wenn wissenschaftliche Erkenntnisse aus den Entwicklungsländern selbst Open-Access-verfügbar gemacht werden.⁹ Schließlich ermöglicht die Digitalisierung unikatlicher Materialien den virtuellen Zugriff auf Informationen durch eine weit höhere Zahl von Nutzern. Die Anwendung Computer-basierter Methoden auf eine bislang unvorstellbar große Menge an digitalen Volltexten und Daten kann zu neuen Fragestellungen und damit neuen Erkenntnissen führen, zumal Suchen, die gleichzeitig systematisch auf Publikationen und diesen zugrunde liegenden Forschungsdaten zugreifen, weitaus effizienter sein werden als rein textorientierte Retrieval-Methoden.

Die eigentliche Herausforderung für eine offene Wissenschaft besteht in der Notwendigkeit, die prinzipielle Nachnutzbarkeit von Forschungsergebnissen umfassend abzusichern. Dies setzt neben klaren (Lizenz-) rechtlichen Vorgaben insbesondere eindeutige und nachvollziehbare Regelungen und Verabredungen in technisch-organisatorischen Bereichen voraus. Dazu gehören z.B. die Entwicklung und Einhaltung technischer und gegebenenfalls Disziplin-spezifischer Standards, die Festlegung und Organisation der persistenten Identifizierung, der Auf- und Ausbau sowie die interoperable Vernetzung von Repositorien für langzeitverfügbare Forschungsinhalte sowie die Professionalisierung der Ausbildung. In Bezug auf die umfassende digitale Bereitstellung von Texten, Objekten und Sammlungen gilt es ferner, Mechanismen zu entwickeln und umzusetzen, nach denen eine von wissenschaftlich-fachlichen Prioritäten geleitete Auswahl der zu digitalisierenden Artefakte möglich

ist. Die schiere Menge potenziellen Digitalisierungsgutes nämlich verbietet dessen vollständige Digitalisierung. Eine besondere Herausforderung stellt schließlich eine auskömmliche Finanzierung dar, da Mittel nicht nur für den – häufig über Projekte realisierten – Auf- und Ausbau der Informationsinfrastruktur, sondern auch für deren langfristigen Betrieb bereitgestellt werden müssen. Insbesondere um zentralisierte, einer globalen Community dienende Infrastrukturen und Services nachhaltig abzusichern, müssen Wege gefunden werden, wie diese auch über föderale und nationale Grenzen hinweg gemeinsam finanziert werden können, ohne den für Nutzer entgeltfreien Zugang zu Inhalten und Diensten zu beeinträchtigen.

2 Rahmenbedingungen für eine effiziente Wissenschaftskommunikation

Wie oben schon angeführt, ist das Kernelement für die Ausgestaltung einer effizienten Informationsinfrastruktur das umfassende Absichern der Nachnutzbarkeit von Forschungsergebnissen. Deren Nachnutzung kann prinzipiell auf verschiedene Weise gestattet werden, z.B. über Lizenzen und Verträge, über besondere Privilegierung des wissenschaftlichen Gebrauchs in Form von Schrankenregelungen oder über eine Bereitstellung im Open Access mit entsprechender Nutzungseinräumung. Welcher dieser Ansätze am ehesten greift, hängt gegebenenfalls auch davon ab, ob der Zugriff auf und die Nachnutzung von traditionellen Publikationen, Forschungsdaten oder retrodigitalisierten Materialien intendiert ist. Dieser Differenzierung entsprechend werden im Folgenden – an unterschiedliche Akteure gerichtete – Vorschläge und Empfehlungen ausgeführt, deren Aufgreifen den offenen Zugang und die Nachnutzung von Forschungsergebnissen unterstützen wird. Dabei ist mit zu bedenken, dass eine produktive Nachnutzung von Forschungsergebnissen häufig die aktive Zustimmung derjenigen voraussetzt, die diese Ergebnisse produziert haben. Deshalb sind gleichzeitig Anreize zu entwickeln und umzusetzen, die zur offenen Bereitstellung eigener Forschungsergebnisse motivieren, indem dies zur Steigerung der wissenschaftlichen Reputation beiträgt.

2.1 Offener Zugang und Nachnutzung herkömmlicher Publikationen

Die nachfolgenden Empfehlungen beruhen auf der seit geraumer Zeit geführten Diskussion um Open Access, in der jüngst zunehmend deutlicher beobachtet werden kann, dass Appelle für eine rechtlich abgesicherte Nachnutzbarkeit von Veröffentlichungen die grundlegende Forderung nach dem für Nutzer und Nutzerinnen entgeltfreien Zugriff auf Publikationen ergänzen. In ähnlicher Weise gilt die Aufmerksamkeit zunehmend der Frage, wie der Umstieg auf Open Access vom dominanten Subskriptionsmodell her gestaltet werden kann und muss. Dies sind die Parameter für die folgenden Ausführungen.

7 S. z.B. Ware, Mark: Access by UK small and medium-sized enterprises to professional and academic information. Research report to the Publishing Research Consortium. August 2009, zugänglich unter <http://publishingresearch.net/SMEaccess.htm>; Swan, Alma: Study on the availability of UK academic „grey literature“ to UK SMEs. Report to the JISC Scholarly Communications Group. December 2008, zugänglich unter <http://eprints.ecs.soton.ac.uk/17667/>.

8 Vgl. Hoffmann, Bert: ‚Harvards‘ and ‚Have-nots‘ on a level playing field: Open Access as a publication model for contemporary area studies. In: Afrika Spectrum 43 (2008), S. 303-308.

9 S. unter <http://www.bioline.org.br/> und vgl. die Kurzdarstellung unter <http://www.oastories.org/2011/09/ukcanadabrazil-platform-bioline-international/>.

Unabdingbares Zweitveröffentlichungsrecht einführen

In vielen Wissenschaftsbereichen ist es üblich, dass Autoren sämtliche Verwertungsrechte an einem Zeitschriftenartikel vertraglich exklusiv an einen Verlag übertragen. Eine anderweitige Verbreitung des Artikels ist dem Autor selbst somit nicht mehr möglich. Da Autoren zudem die schwächere Vertragspartei sind, werden sie in aller Regel davor zurückscheuen, darauf zu insistieren, die Verlagsverträge um Regelungen zur anderweitigen Nutzung ihrer Werke modifizieren zu lassen. Die Einführung eines gesetzlichen Zweitveröffentlichungsrechts sollte daher wissenschaftlichen Autoren die Möglichkeit eröffnen, ihre eigenen Beiträge aus wissenschaftlichen Zeitschriften und Sammelbänden über Open-Access-Repositorien für den entgeltfreien Zugriff verfügbar zu machen, auch wenn sie zuvor Verlagen exklusive Verwertungsrechte übertragen haben. Auf diese Weise sollten Autoren selbst über den höheren Grad der Sichtbarkeit ihrer eigenen Forschungsergebnisse entscheiden können. Durch die Einführung einer Sperrfrist, vor deren Ablauf die Beiträge nicht über Open-Access-Repositorien verfügbar gemacht werden dürfen, können zugleich die berechtigten kommerziellen Interessen der Verlagswirtschaft geschützt werden. Konkrete Vorstellungen für die Ausgestaltung einer derartigen Regelung sind, auf Basis einer bereits 2006 durch den Deutschen Bundesrat vorgeschlagenen Formulierung, in jüngster Zeit mehrfach unterbreitet worden, sowohl von der Allianz der Wissenschaftsorganisationen als auch von Seiten verschiedener im Deutschen Bundestag verteilter Fraktionen (SPD, Die Linke, Bündnis90/Die Grünen).¹⁰

Nachnutzbarkeit von Forschungsergebnissen über Lizenzen sichtbar machen

In zunehmendem Maße werden Forschungsergebnisse nicht nur gelesen, sondern mit Hilfe computerisierter Verfahren weiter prozessiert. Die Anwendung derartiger Verfahren setzt in vielen Fällen voraus, dass entsprechende Nutzungen rechtlich gestattet wurden. Um die umfassende Nachnutzung von Publikationen rechtlich abzusichern, sollte daher die Verwendung von Lizenzen empfohlen werden, mit denen explizit festgelegt ist, zu welchen Nutzungszwecken veröffentlichte Forschungsergebnisse verfügbar sind. Die Verwendung z.B. der Creative-Commons-Lizenzen durch Autoren und Verlage bringt den Vorteil, dass die möglichen Nutzungsweisen in einer auch für juristische Laien gut verständlichen Form verdeutlicht werden können. Die Vergabe entsprechender Lizenzen könnte sogar zu einem Qualitätsmerkmal werden, wenn Förderorganisationen und Forschungsinstitute die Finanzierung von Open-Access-Publikationen im Hinblick darauf ausgestalten, wie klar und umfassend die mit einer Open-Access-Publikation verbundenen Rechte für Zugriff und produktive Nachnutzung geregelt sind.

Den Übergang zu Open Access vom Subskriptionsmodell her gestalten

Das Subskriptionsmodell bestimmt derzeit die wissenschaftliche Literaturversorgung; erst 10 bis 20 Prozent aller Publikationen sind im Open Access verfügbar.¹¹ Der Umstieg auf Open Access muss daher auch vom Subskriptionsmodell her gedacht werden. Um die Ablösung des noch dominanten Subskriptionsmodells durch Open Access zu unterstützen, sollten künftig in die Verträge für die Lizenzierung elektronischer Zeitschriften klare und eindeutige Regelungen zu einem verbesserten Zugang und zur Nutzbarkeit von Publikationen aufgenommen werden.

Ein Beispiel für derartige Regelungen liefern die sogenannten „Allianz-Lizenzen“.¹² Diese Lizenzen ermöglichen nicht nur die Nutzung der lizenzierten Produkte für die üblichen bibliothekarischen Dienste, sondern z.B. auch die Einbindung der lizenzierten Produkte in digitale Semesterapparate oder Virtuelle Forschungsumgebungen. Außerdem beinhalten die Lizenzen eine Open-Access-Komponente: Autoren aus lizenznehmenden Einrichtungen, die Artikel in Allianz-lizenzierten Zeitschriften veröffentlicht haben, dürfen diese Artikel auch über Open-Access-Repositorien verfügbar machen, und zwar in der Regel in Form des Verlags-PDFs. Das gleiche Recht gilt für die Einrichtungen, an denen die Autoren tätig sind.

Verlässliche Kostenabschätzungen ermöglichen, um Publikationsfonds aufzubauen und ausreichend auszustatten

Auch in einem Open-Access-Modell müssen die für Erstellung und Verbreitung von Publikationen anfallenden Kosten – häufig in Form sogenannter Publikationsgebühren – bedient werden. Um entsprechende Mittel zu budgetieren, müssten Forschungsinstitutionen in der Lage sein, sich ein verlässliches Bild der Publikationsaktivitäten ihrer Wissenschaftler und der damit verbundenen Ausgaben zu beschaffen. Die Erfahrungen mit den Anträgen im noch jungen, erst 2009 aufgelegten Förderprogramm „Open Access Publizieren“ der DFG zeigen jedoch, dass viele Universitäten noch eher vage Vorstellungen darüber haben, wie viele Publikationen im Lauf eines Jahres von Universitätsangehörigen vorgelegt werden. Die *Einführung eines verlässlichen Publikationsmanagements* wäre ein erster Schritt dazu, künftig belastbare Aussagen über die zur Finanzierung des Open-Access-Publizierens erforderlichen Mittel treffen zu können.

Dabei ist zu beachten, dass die Ermittlung allein der Anzahl von Publikationen nicht ausreicht, um Ausgaben für das Open-Access-Publizieren zu errechnen. Denn nicht jede Open-Access-Publikation muss auch von Angehörigen der Einrichtung finanziert werden: Zum einen sind Publikationen von mehreren Autoren zu beachten, die aus unterschiedlichen Einrichtungen kommen und bei denen gegebenenfalls der an einer anderen Einrichtung arbeitende „*corresponding author*“

¹⁰ Die Vorstellungen der Allianz sind auf den Internet-Seiten der Allianz-Initiative „Digitale Information“ ausführlich in mehreren Unterlagen dokumentiert, s. unter <http://www.allianz-initiative.de/>.

¹¹ Björk, Bo-Christer et al. (2010): Open Access to the Scientific Journal Literature: Situation 2009. In: PLoS ONE 5(6) e11273, s. unter <http://www.plosone.org/article/info:doi/10.1371/journal.pone.0011273>.

¹² Die „Grundsätze für den Erwerb überregionaler, von der DFG geförderter Lizenzen“ sind zugänglich unter http://www.dfg.de/formulare/12_18/12_18.pdf.

für die Open-Access-Publikationsgebühren aufkommt. Zum anderen erheben viele Open-Access-Zeitschriften keine Publikationsgebühren, sondern tragen sich – etwa durch institutionelle Finanzierung – anderweitig.

Solange die Einrichtungen nicht selbst in der Lage sind, die für sie anfallenden Publikationskosten exakt zu beziffern, bleibt nur der Verweis auf Schätzungen. Seit geraumer Zeit wird insbesondere auf die Aussage des britischen Wellcome Trust verwiesen, nach der lediglich 1 bis 2 Prozent des Forschungsbudgets erforderlich wären, um sämtliche aus der Förderung des Wellcome Trust resultierende Publikationen durch Zahlung einer Gebühr im Open Access zugänglich zu machen. Nach mündlicher Auskunft von Dr. Ralf Schimmer, Leiter des Bereichs Informationsmanagement der Max Planck Digital Library, wäre diese Annahme auch für die Max-Planck-Gesellschaft durchaus realistisch.

Die im Rahmen der Kommission „Zukunft der Informationsinfrastruktur“ mit dem Thema Open Access befasste Arbeitsgruppe hat einige Zahlen zur (Projekt-) Finanzierung von Open-Access-Aktivitäten zusammengetragen. In Bezug auf die Höhe gegebenenfalls anzusetzender Publikationsgebühren hat die Gruppe lediglich darauf verwiesen, dass der „*Journal Citation Report*“ jährlich ca. 80.000 Zeitschriftenartikel aus Deutschland nachweise. Wie viele Gelder vonnöten wären, um diese Artikel im Open Access verfügbar zu machen, hänge dann von Annahmen über die durchschnittliche Höhe von Publikationsgebühren ab.¹³

Da Kosten auch Nutzen gegenüberstehen, sei an dieser Stelle auf die Arbeiten des australischen Wirtschaftswissenschaftlers John Houghton verwiesen, der genaue Analysen dazu vorgelegt hat, welche „*costs and benefits*“ mit einer Hinwendung zu Open Access (in seinen verschiedenen Spielarten) verbunden wären: Die für unterschiedliche Länder – eine im Rahmen eines DFG-geförderten Projekts für Deutschland beauftragte Analyse wird Anfang 2012 veröffentlicht werden – vorgenommene Modellierung zeigt stets, dass der verbesserte Zugang und die durch Open Access bedingte Effizienzsteigerung der Forschung zu substantiellen Gewinnen („*benefits*“) führt.¹⁴

Subskriptionsetat und Open-Access-Publikationsfonds korrelieren

Die Bereitstellung von Forschungsergebnissen wird heute im Wesentlichen aus dem Subskriptionsetat gezahlt. In einem Open-Access-Szenario würde die Informationsversorgung aus einem Publikationsfonds finanziert. Um den Übergang zu gestalten, müssen daher Wege gefunden werden, wie der Subskriptionsetat (bzw. Teile dieses Etats) in ein Budget umgeschichtet werden kann, aus dem künftig die Finanzierung von Open-Access-Publikationen geleistet wird. Hier

13 KII-Gesamtkonzept (Anm. 3), B103. Zum Vergleich: Nach Ausweis der deutschen Bibliotheksstatistik haben die wissenschaftlichen Bibliotheken im Jahr 2010 Ausgaben in Höhe von 322 Mio. Euro für die Erwerbung getätigt, s. unter <http://www.hbz-nrw.de/dokumentencenter/produkte/dbs/aktuell/auswertungen/gesamt/>.

14 Vgl. Houghton, John: Open Access – What are the economic benefits? A comparison of the United Kingdom, the Netherlands and Denmark (2009), s. unter <http://www.knowledge-exchange.info/Default.aspx?ID=316>.

besteht zudem das Problem, dass zur Finanzierung von Publikationskosten vorhandene Budgets von unterschiedlichen Organisationseinheiten – z.B. von der Bibliothek und einem Fachbereich – verwaltet werden. Deshalb sind auch Lösungen dafür zu suchen, wie die insgesamt zur Finanzierung der Informationsversorgung beitragenden Mittel korreliert werden können und müssen.

Interoperabilität von Repositorien ausbauen

Damit Nutzer die weltweit über viele unterschiedliche Repositorien verfügbaren Informationen komfortabel nutzen können, muss es möglich sein, die Inhalte von Repositorien über standardisierte Schnittstellen austauschen und auf die Information in aggregierter Form zugreifen zu können. Dies ermöglicht die Einhaltung von Richtlinien zur interoperablen Vernetzung von Repositorien. Allerdings fokussieren die bisher verfügbaren Richtlinien auf Texte. Da in Repositorien zunehmend neue Inhaltstypen wie multimediale Angebote, Datensätze, Software und Computerprogramme oder Simulationen archiviert werden, müssen diese Richtlinien überarbeitet werden, damit ein interoperabler Austausch auch neuer Inhaltstypen ermöglicht wird. Zugleich muss sichergestellt werden, dass Repositorien sich nahtlos in andere Teile der entstehenden Informationsinfrastruktur wie Virtuelle Forschungsumgebungen, GRIDs oder Forschungsinformationssysteme integrieren lassen.¹⁵ Hier kommt der persistenten Identifikation von Personen, Institutionen, Objekten und Daten eine wesentliche Rolle für die Vernetzung der Angebote zu, die nur bei eindeutiger Referenzierung gelingen kann. Schließlich muss Interoperabilität auch gewährleisten, dass auf der Basis von Repositorien entwickelte Dienste ohne Brüche in den Arbeitsablauf von Wissenschaftlern eingebettet werden können.¹⁶

2.2 Data Sharing als Imperativ internationaler Forschung

In vielen Wissenschaftsbereichen beruht der Erkenntnisfortschritt auf sorgfältig erhobenen und analysierten Forschungsdaten. Werden diese Daten auch Dritten verfügbar gemacht, erhöht dies die Transparenz der Forschung, denn so können gegebenenfalls andere Wissenschaftler und Wissenschaftlerinnen die aus den Daten resultierenden Publikationen besser bewerten. Da die Erhebung von Forschungsdaten zudem mit durchaus hohen Kosten verbunden sein kann, gebietet auch die ökonomische Vernunft, Daten möglichst in einer Form

15 Vorschläge für die interoperable Verzahnung von Repositorien und Forschungsinformationssystemen erarbeitet das im Rahmen von Knowledge Exchange durchgeführte „CRIS/OAR Interoperability Project“, s. unter <http://www.knowledge-exchange.info/Default.aspx?ID=340>.

16 Für welche Themengebiete Interoperabilität stärker gestaltet werden muss, hat die „Confederation of Open Access Repositories“ (COAR) in einem Arbeitspapier vom Juli 2011 ausführlich beschrieben, s. unter http://www.coar-repositories.org/files/COAR_Interoperability_Briefing.pdf. Zugleich hat COAR die Repositorien-Betreiber und -Entwickler dazu eingeladen, Ideen zur Lösung der adressierten Felder zu unterbreiten, s. unter <http://www.coar-repositories.org/working-groups/repository-interoperability/coar-interoperability-project/>.

bereitzustellen, die deren produktive Nachnutzung durch Dritte ermöglicht. Dazu müssen jedoch besondere rechtliche, technische, organisatorische und finanzielle Voraussetzungen beachtet oder gar geschaffen werden.

Internationale Nachnutzbarkeit von Forschungsdaten (rechtlich) absichern

Gerade die datenbezogene Infrastruktur muss konsequent international gedacht werden. Wissenschaftler auf der ganzen Welt bearbeiten Fragestellungen, deren Beantwortung oft einen Zugriff auf Daten erfordert, die häufig in gänzlich anderen Forschungskontexten erhoben worden sind. Somit sind Rahmenbedingungen zu schaffen, die einen Austausch von Forschungsdaten über nationale Grenzen hinweg ermöglichen. Obgleich die rohen Daten bloße Fakten sind und somit grundsätzlich nicht als geistiges Eigentum geschützt werden, greift nach deren systematischer Zusammenstellung eine Vielzahl rechtlicher Regelungen, die den Austausch und die Nachnutzung von Daten – insbesondere im internationalen Kontext – erheblich erschweren können. Die Komplexität der Materie zeigt ein für das Netzwerk „Knowledge Exchange“ angefertigter Bericht zum rechtlichen Status von Forschungsdaten, der nicht nur die Herausforderungen beschreibt, sondern auch einige, im Folgenden knapp skizzierte Lösungsvorschläge unterbreitet.¹⁷

Schranken zur Datennutzung in Forschung und Lehre einführen

Da die Verwendung von Forschungsdaten nicht allein urheberrechtlichen Bestimmungen und der EU-Datenbank-Richtlinie unterliegt, sondern in einzelnen Ländern weitere, den Datenaustausch erschwerende Regeln existieren, ist es äußerst diffizil, Bestimmungen zur Nachnutzung von Forschungsdaten über Lizenzverträge zu regeln. Die Ausarbeitung solcher Verträge wäre auch deshalb besonders schwierig, weil sowohl Regelungen des Landes, in dem die nachzunutzenden Daten gespeichert sind, als auch Bestimmungen aus dem Land der intendierten Nutzung greifen. Damit sich die Vielfalt dieser Regelungen nicht nachteilig auf eine produktive Nachnutzung von Forschungsdaten auswirkt, sollten gesetzliche Ausnahmebestimmungen („Schrankenbestimmungen“) eingeführt werden, die eine dezidierte Nachnutzung von Daten für Wissenschaft und Forschung – auch über Ländergrenzen hinweg – erlauben.

Möglichkeiten zur Nachnutzung präzise definieren

Forschungsförderer könnten Wissenschaftler dazu verpflichten, in Projekten erhobene Daten nur unter bestimmten Lizenzen bereit zu stellen, die eine Nachnutzung der Daten zu klar definierten Bestimmungen ermöglichen. Um Wissenschaftlern eine Nachnutzung von Daten zu ermöglichen, ohne dass diese Klagen oder rechtliche Beeinträchtigungen fürchten müssen, sollten auch Verhaltensnormen („Codes of

Conduct“) entwickelt werden, in denen z.B. klar definiert ist, wie die ursprüngliche „Autorschaft“ an nachgenutzten Daten gekennzeichnet wird, ob und unter welchen Bedingungen Forschungsdaten auch zu kommerziellen Zwecken nachgenutzt werden können und zu welchen Nutzungszwecken diese verfügbar sind. Eine weitere Möglichkeit, die Nachnutzung rechtlich abzusichern, besteht darin, dass die „Autoren“ der Forschungsdaten dezidiert darauf verzichten („waiver“), eventuell an ihren Daten haftende Rechte geltend zu machen, wenn diese in Wissenschaft und Forschung durch Dritte wieder verwendet werden. Derartige Verhaltensnormen könnten im Sinne von Verträgen zwischen den Datenproduzenten bzw. einem Repository und den Nutzern der Forschungsdaten interpretiert werden.

„Community-Dienste“ finanziell absichern

Viele für die Informationsversorgung wichtige „Community Dienste“ haben sich aus lokalen Strukturen entwickelt und werden daher von der den Dienst betreibenden Institution getragen, die für die Community relevante Informationen unentgeltlich bereitstellt. Je stärker die gerade auch internationale Community einen solchen Dienst in Anspruch nimmt, desto höher sind die hierfür erforderlichen Kosten. Hier sind Lösungen dafür zu suchen, wie gegebenenfalls auch externe Nutzer an den Betriebskosten beteiligt werden können, ohne dass der freie Zugang zu den für die Community wichtigen Informationen beschnitten wird. Das erfordert gegebenenfalls auch Lösungen dafür, wie eine Mitfinanzierung von im Ausland betriebenen Infrastrukturen aus deutschen Steuergeldern ermöglicht werden kann, sofern diese Infrastrukturen für die deutsche Wissenschaft relevant sind.

In diesem Zusammenhang ist vor allem darauf hinzuweisen, dass das Einspeisen von Forschungsdaten in Repositorien ebenso wie die laufende Pflege von Datenrepositorien z.T. hohe (Betriebs-) Kosten verursachen kann. Damit der Träger eines für die Community wichtigen, zentralen Repositoriums seinen Aufgaben in der Kuratierung von Forschungsdaten angemessen nachkommen kann, sollten Forschungsprojekte daher über Mittel („Datenpauschale“) verfügen, aus denen die beim Betreiber des Datenrepositoriums anfallenden Kosten bedient werden können. Für eine nachhaltige Datenpflege müssen nach aktuellen Schätzungen 5 bis 10 Prozent der Forschungskosten eines Projekts angesetzt werden.¹⁸

Grundsätze zum Umgang mit Forschungsdaten durchsetzen

Wesentliche Voraussetzung für eine prinzipielle Nachnutzbarkeit von Forschungsdaten ist, dass Daten sorgfältig und nach gegebenenfalls Disziplin-spezifischen Standards beschrieben und dokumentiert sind. Bei der Erzeugung und Verwaltung von Forschungsdaten sind technische und organisatorische Anforderungen zu berücksichtigen, deren Einhaltung für die prinzipielle Nachnutzbarkeit entscheidend ist. Die Bedeutung dieser Anforderungen sollte in „Grundsätzen

¹⁷ de Cock Buning, Madeleine et al.: The legal status of research data in the Knowledge Exchange partner countries. Oktober 2011, s. insbesondere die Empfehlungen auf S. 48-58, auf denen die folgenden Ausführungen basieren, verfügbar unter <http://www.knowledge-exchange.info/Default.aspx?ID=461>.

¹⁸ Vgl. KII-Gesamtkonzept (Anm. 3), B121f.

zum Umgang mit Forschungsdaten¹⁹ dokumentiert und deren Befolgen gegebenenfalls verpflichtend werden. Ein erster Schritt, um die Bedeutung der technisch-organisatorischen Rahmenbedingungen bewusst werden zu lassen, sollte die dezidierte Aufforderung sein, in Anträgen auf Forschungsgeldern Maßnahmen zu beschreiben, durch die eine nachhaltige Sicherung und gegebenenfalls erneute Nutzung der Daten gewährleistet wird.²⁰

Standards entwickeln, Orientierung bieten

Im eigentlichen Infrastruktur-Bereich ist des Weiteren die Entwicklung Disziplin-spezifischer Standards zur Beschreibung von Forschungsdaten erforderlich, außerdem der Auf- und Ausbau von Datenrepositorien, da eine lokale und quasi „private“ Speicherung von Forschungsdaten deren Nutzbarkeit nicht ermöglicht. Anzueraten ist daher eine Speicherung in überregionalen Datenzentren, die nicht nur über qualifiziertes Personal und unterstützende Dienstleistungen verfügen, sondern auch die langfristige Verfügbarkeit der Daten gewährleisten. Des Weiteren sind Verzeichnisse vonnöten, die Wissenschaftlern rasch eine Orientierung darüber ermöglichen, in welche Repositorien sie Daten aus ihren eigenen Forschungen (mit unterschiedlichen Berechtigungen für den Zugriff auf und die weitere Verwendung der Daten) einstellen können.²¹

2.3 Digitalisate – Forschungsdaten für die historischen Wissenschaften

In gewisser Weise lassen sich retrospektiv erstellte Digitalisate als Sonderfall von Forschungsdaten ansehen, denn digital verfügbare Texte, Bilder oder Objekte sind die Daten insbesondere – doch nicht allein – für historisch arbeitende Wissenschaften. Auch hier ermöglicht und erfordert eine über Datennetze betriebene Wissenschaftskommunikation neuartige Zugänge und Verwendungsweisen, die in der Ausgestaltung der Informationsinfrastruktur konsequent unterstützt werden müssen.

Eindeutige Regelungen zur Nutzung retrospektiv erstellter Digitalisate treffen

Eine wissenschaftliche Nutzung von Digitalisaten beinhaltet weit mehr als das Anschauen, Lesen und gegebenenfalls Herunterladen einzelner digitaler Dokumente. Weitere Nutzungsformen sind z.B. der Zugriff auf sehr große Textmengen über Metadaten- und Volltextsuche, die quantitative

Analyse von digitalen Editionen und Textkorpora mit Hilfe Computer-basierter Verfahren, die maschinelle Erkennung von Mustern in digitalisierten Bildern, Musik- oder Filmdateien oder der Export vollständiger digitaler Sammlungen in Virtuelle Forschungsumgebungen. Um die Vielfalt derartiger Nutzungen zu gestatten, sollten auch retrodigitalisierte Materialien unter Lizenzen bereitgestellt werden, die ihre Verfügbarkeit für Forschung und Lehre über den rein lesenden Zugriff hinaus regeln. Diese Forderung sollte nicht nur für öffentlich geförderte Vorhaben, sondern auch für Digitalisierungen im Rahmen von Public-Private-Partnerships greifen, wie sie z.B. im Rahmen der „Deutschen Digitalen Bibliothek“ angestrebt werden. Optimal wäre die Bereitstellung der Digitalisate mit strukturellen und semantischen Erschließungsdaten (z.B. in Form des XML-Markups), um Forschungsfragen auf neue Art kollaborativ angehen zu können.

Rechtsrahmen zur Digitalisierung verwaister Werke schaffen

Für einen großen Teil (wissenschaftlicher) Werke des 20. Jahrhunderts ist heute nicht geklärt, wer die Verwertungsrechte an diesen Werken besitzt. Damit diese, als „verwaist“ bezeichneten Werke überhaupt in digitaler Form verfügbar gemacht werden können, sind Regelungen zu implementieren, nach denen die Digitalisierung verwaister Werke zulässig ist. Derartige Regelungen sollten insbesondere im Hinblick darauf geprüft werden, ob sie für die Massendigitalisierung tauglich sind. Dazu könnte eine weitere urheberrechtliche Schrankenbestimmung implementiert werden, nach der Bibliotheken, Archiven und Museen die Vervielfältigung und öffentliche Zugänglichmachung verwaister Werke zu nicht-kommerziellen Zwecken gestattet würde.²² Allerdings würde dies eine entsprechende Erweiterung der europäischen Richtlinie zum Urheberrecht voraussetzen.

Grundfinanzierung durch Bund und Länder sicherstellen

Damit die Erstellung digitaler Inhalte zügig voranschreitet, müssen Bibliotheken und Archive von ihren Trägern so ausgestattet werden, dass sie ihre Bestände – gegebenenfalls in Public-Private-Partnerships – digitalisieren können. Für die kommenden Jahre dürften hier Mittel in Höhe von ca. 30 Mio. Euro pro Jahr erforderlich sein.²³ Über DFG-geförderte Projekte sind in den letzten 15 Jahren zahlreiche Bestände digitalisiert worden. Die DFG wird sich auch weiterhin in diesem Bereich einsetzen, wenn es darum geht, Digitalisierungen für neue Materialgattungen von dezidiertem Forschungsrelevanz anzustoßen. Allerdings zielt die Förderung der DFG nicht auf Vollständigkeit ab, sondern darauf, einen wissenschaftsrelevanten Kernbestand digital anbieten zu können und strukturbildende Maßnahmen anzustoßen. Somit kann die DFG-Förderung lediglich ergänzend zu der notwendig grundständigen Digitalisierung durch Bund und Länder hinzu treten.

19 Als Beispiel seien die von der Allianz der deutschen Wissenschaftsorganisationen verabschiedeten Grundsätze genannt, s. unter <http://www.allianzinitiative.de/de/handlungsfelder/forschungsdaten/grundsätze/>.

20 Wie andere Förderorganisationen fordert auch die DFG im „Leitfaden für die Antragstellung“ dezidiert dazu auf, den Umgang mit im Projekt erzielten Forschungsdaten zu beschreiben, s. DFG-Merkblatt 54.01 unter http://www.dfg.de/formulare/54_01/54_01_de.pdf, S. 6.

21 Für Infrastrukturen von europäischer Bedeutung arbeitet das EU-geförderte Projekt MERIL (Mapping of the European Research Infrastructure Landscape) an derartiger Verzeichnung, s. unter <http://www.esf.org/activities/science-policy/research-infrastructures/meril-mapping-of-the-european-research-infrastructure-landscape.html>.

22 Ein ausformulierter Gesetzesvorschlag für eine derartige Schranke findet sich in einem Papier der Allianz der Wissenschaftsorganisationen zur Neuregelung des Urheberrechts vom 9. Juli 2010, s. unter http://www.allianzinitiative.de/fileadmin/user_upload/Home/Desiderate_fuer_Dritten_Korb_UrhG.pdf, S. 6f.

23 S. KII-Gesamtkonzept (Anm. 3), S. 38.

Digitalisierung von Objekt-bezogenen Sammlungen mitdenken

Bislang wurden insbesondere Texte in handschriftlicher und gedruckter Form digitalisiert. Um die digitale Transformation umfassend zu gestalten, müssen verstärkt auch Objekte aus wissenschaftlichen Sammlungen digitalisiert werden. Dies hat nicht nur bedeutende finanzielle Implikationen – die DFG hat 2011 erstmals zur Antragstellung in diesem Bereich eingeladen, woraufhin 110 Anträge auf Erschließung und Digitalisierung wissenschaftlicher Sammlungen mit einem Antragsvolumen von ca. 31 Mio. Euro vorgelegt wurden –, sondern erfordert auch eine wissenschaftsgeleitete Entwicklung geeigneter Auswahlkriterien, damit priorisiert werden kann, welche Sammlungen zum breiten Nutzen der Forschung vordringlich zu digitalisieren sind.

Konsequente Anreicherung von Digitalisaten ermöglichen

Erst die Erschließung von Digitalisaten durch strukturierte, nach internationalen Standards erstellte Metadaten ermöglicht gezieltes Auffinden, Nachnutzen und Interpretieren von Digitalisaten und deren Langzeitarchivierung. Strukturdaten, Normdaten und die Anreicherung mit weiteren Kontextinformationen ermöglichen die umfassende Verknüpfung eines digitalen Objekts mit anderen Objekten („*contextual linking*“), schaffen so Mehrwerte und lassen Digitalisate häufig erst für einen größeren Kreis von Wissenschaftlern nutzbar werden. Das Potenzial derartiger Verknüpfungen kann nur dann ausgeschöpft werden, wenn möglichst viele Metadaten und im Volltext vorliegende Dokumente frei verfügbar und beliebig miteinander kombinierbar sind.²⁴ Durch die Weiterentwicklung automatisierter Erschließungs- und Indexierungsverfahren unter Einbezug von Normdateien sollten daher die Möglichkeiten zu einer (halb-) automatisierten Anreicherung von Digitalisaten gezielt ausgebaut werden.

3 Ausblick

In den vorangehenden Abschnitten wurde insbesondere ausgeführt, über welcherlei Vorkehrungen eine produktive Nachnutzung von Forschungsergebnissen unterschiedlicher Art abgesichert werden sollte. Die weitere Ausgestaltung der die Forschung unterstützenden Informationsinfrastruktur wird allerdings nicht allein durch Maßnahmen zur Absicherung von Zugang und Nachnutzung bestimmt werden. Daher soll ein letzter Abschnitt einige Entwicklungen reflektieren, die aktuell ebenfalls zu beobachten sind und mit hoher Wahrscheinlichkeit so fortgeschrieben werden, dass auch diese eine deutliche Auswirkung auf die Gestalt der künftigen Informationsinfrastruktur haben werden.

Im Bereich traditioneller Publikationen werden die Repositorien-Infrastruktur weiter ausgebaut, deren Interoperabilität erhöht und deutlich mehr Inhalte über Repositorien verfügbar gemacht werden, wobei eine deutliche Zunahme von Publikationen, deren Text durch Datensätze, Animatio-

nen, multimediale Anwendungen, Software o.ä. angereichert sein wird, außer Frage stehen dürfte.²⁵ Zudem ist es durchaus realistisch, dass Inhalte – wie es in großem Maßstab im EU-geförderten Projekt „*Publishing and the Ecology of European Research*“ (PEER) praktiziert wurde²⁶ – direkt von Verlagen in Repositorien eingespielt werden. Es muss betont werden, dass der gezielte Ausbau der Repositorien-Infrastruktur auch dann erfolgen wird, wenn sich der aktuell zu beobachtende Trend einer stärkeren Fokussierung auf den Goldenen Weg des Open Access fortsetzen sollte: Da eine originäre Open-Access-Publikation mit der Erlaubnis verbunden ist, die Publikation in jeder geeigneten Weise weiter zu verbreiten, werden sich Universitäten und Forschungsinstitute zunehmend darum bemühen, die Publikationen ihrer Wissenschaftler zu sammeln und über eigene Angebote verfügbar zu machen, auch um auf diese Weise ihr Forschungsprofil optimal sichtbar zu machen. Zugleich ist zu berücksichtigen, dass in vielen Projekten derzeit darauf hin gearbeitet wird, auf Repositorien basierende Dienste in den wissenschaftlichen Arbeitsablauf zu integrieren.²⁷ Dies wird in vielen Fällen zwangsläufig dazu führen, dass die Publikation als Ergebnis dieser Arbeit zumindest in einer Vorversion über ein Repository verfügbar ist. Dabei werden die Beziehungen zwischen unterschiedlichen Fassungen einer Veröffentlichung durch genaue Kennzeichnung von Versionen verdeutlicht werden. Das ist schon deshalb vonnöten, um aggregierte Information z.B. über die Zitation einer wissenschaftlichen Arbeit zu bieten.

Im Bereich der Forschungsdaten werden sich Repositorien für die unterschiedlichen Disziplinen etablieren. Nicht nur Förderorganisationen, sondern auch wissenschaftliche Zeitschriften werden in „*data policies*“ Auflagen spezifizieren, in welchen Repositorien Forschungsdaten, die in einem Projekt generiert wurden oder die einer Publikation zugrunde liegen, in einer Form archiviert werden müssen, die einen Zugriff auf diese Daten erlaubt. Damit wird nicht nur der Forderung Rechnung getragen, Forschungsergebnisse transparent nachvollziehbar zu machen, sondern auch – wo dies rechtlich möglich ist – eine produktive Nachnutzung der Daten für neue Forschung unterstützt. Allerdings ist darauf zu achten, dass Forschungsdaten als Rohstoff des wissenschaftlichen Arbeitens nicht auf solche Weise kommerzialisiert werden, dass eine unentgeltliche Weiterverarbeitung der Daten in der Wissenschaft nicht mehr möglich ist.

Mit dem Schlagwort der Virtuellen Forschungsumgebung wird eine Entwicklung benannt, die derzeit nicht nur

25 Während im Jahr 2003 erst 7 Prozent der publizierten Zeitschriftenartikel derartige Anreicherungen aufwiesen, traf dies im Jahr 2009 schon auf 25 Prozent der Artikel zu, vgl. van der Graaf, Maurits; Waaijers, Leo: A Surfboard for Riding the Wave. Towards a four country action programme on research data. November 2011, S. 15, s. unter <http://www.knowledge-exchange.info/Default.aspx?ID=469>.

26 Nähere Informationen hierzu auf der Projekt-Homepage unter <http://www.peerproject.eu/>.

27 Dass diese Entwicklung essentiell für den Aufbau von Repositorien ist, wird ausführlich dargelegt in: The Research Library's Role in Digital Repository Services. Final Report of the ARL Digital Repository Issues Task Force. January 2009, <http://www.arl.org/bm~doc/repository-services-report.pdf>, S. 18f.

im nationalen, sondern auch im internationalen Kontext intensiv diskutiert wird und die in besonderer Weise zeigt, warum die Möglichkeit zur Nachnutzung von Forschungsdaten von zunehmendem Interesse ist. Im Gesamtkonzept der Kommission „Zukunft der Informationsinfrastruktur“ wird die Vision angeführt, dass bis zum Jahr 2020 Virtuelle Forschungsumgebungen möglichst für alle Wissenschaftsgebiete und Forschungsfelder etabliert werden.²⁸ Während heute durch Förderung des BMBF und der DFG die Entwicklung von insgesamt 33 Virtuellen Forschungsumgebungen angeschoben werden konnte, geht die KII-Vision von um die 240 Virtuellen Forschungsumgebungen aus.²⁹ Dabei ist jedoch zu berücksichtigen, dass die anfänglich zu tätigen Investitionen – die vom BMBF geförderten, sehr breit aufgestellten Projekte förderten die Entwicklung einer Virtuellen Forschungsumgebung mit durchschnittlich 5 Mio. Euro, die DFG-geförderten, auf engere Fragestellungen zugeschnittenen Virtuellen Forschungsumgebungen wurden mit Summen zwischen 300.000 und 500.000 Euro pro Vorhaben unterstützt – deutlich reduziert werden können, da künftige Virtuelle Forschungsumgebungen explizit modular aufgebaut sein werden. Somit eröffnet sich die Möglichkeit, dass die für neue Virtuelle Forschungsumgebungen erforderliche Kernarchitektur aus Dienstleistungen und Werkzeugen auf – gegebenenfalls anzupassenden – nachgenutzten Modulen aufsetzt, während lediglich spezifische fachliche Erweiterungen vollständig neu entwickelt werden müssen. Dieser Ansatz setzt allerdings voraus, dass die Landschaft existierender Virtueller Forschungsumgebungen in einer Weise kartiert wird, die nicht nur über vorhandene Virtuelle Forschungsumgebungen informiert, sondern aufschlüsselt, welche Komponenten der bestehenden Infrastruktur nachgenutzt werden können. Nicht gesichert ist allerdings bislang der verlässliche Betrieb der über Projektförderung aufgebauten Virtuellen Forschungsumgebungen. Hier muss sich noch erweisen, wer die geeigneten Träger für die neuen Strukturen sein können, wobei noch ungeklärte Fragen etwa bezüglich der Organisations- und Rechtsformen gelöst werden müssen.³⁰

Da Wissenschaftler dafür „belohnt“ werden wollen, wenn sie ihre Forschungsergebnisse in nachnutzbarer Form verfügbar machen, müssen die bislang in Ansätzen verbreiteten Lösungen zur eindeutigen Identifikation von Autoren weiter entwickelt und umfassend genutzt werden. Dazu werden die bereits heute genutzten Identifikationssysteme darauf hin-

arbeiten, ihre Informationen interoperabel austauschen zu können. In ähnlicher Weise werden Normdaten zur eindeutigen Identifizierung von Objekten verabredet werden. Unter den sich allmählich entwickelnden Verfahren, mit denen die Produzenten von Datensätzen wissenschaftliche Reputation erwerben können, dürften sogenannte „*data journals*“ eine besondere Rolle spielen. In solchen Zeitschriften werden Datensätze nach einem rigorosen Peer Review publiziert, so dass diese wie traditionelle Veröffentlichungen zitiert werden können.

Während der Auf- und Ausbau von Text- und Datenrepositorien unter Beteiligung auch kommerzieller Anbieter voranschreiten wird, wird ein gewissermaßen gegenläufiger Trend sich fortsetzen, der eine zunehmende Selbstorganisation der Wissenschaft in Bezug auf die Wissenschaftskommunikation fortschreibt. Dies betrifft zum einen den Aufbau von Verlags-unabhängigen Open-Access-Zeitschriften, die in Eigenregie von Wissenschaftlern und Wissenschaftlerinnen herausgegeben und produziert werden. Diese Entwicklung wird durch die gute Verfügbarkeit von effizienten und leicht bedienbaren Open-Source-Publikationswerkzeugen wie dem *Open Journal Systems*³¹ begünstigt. Zum anderen wird die auf Publikationen basierende Kommunikation in erheblichem Umfang durch den informellen Austausch von Wissenschaftlern flankiert werden, wie er nicht nur über Blogs, sondern auch in sozialen Wissenschaftsnetzwerken wie etwa *ResearchGate*³² stattfindet. Dies wird schon deshalb geschehen, weil die Altersgruppe der unter 30-Jährigen alltäglich soziale Netzwerke nutzt, um Wissen und Informationen auszutauschen und zu pflegen. Dieses Verhalten wird auch künftige berufliche Netzwerke prägen, in denen „vernetzte Individuen“ zusammen arbeiten.

Gerade die zunehmende informelle Kommunikation wird zur Entwicklung neuer Mechanismen der Reputationszuweisung führen, die die ja bislang weitestgehend auf den etablierten Strukturen des formalen Publizierens beruhenden Mechanismen ergänzen und womöglich in Teilbereichen ablösen werden. Nicht zuletzt hier zeigt sich, dass die weitere Entwicklung von Informationsinfrastrukturen nicht losgelöst vom umfassenderen „Ökosystem“ der Forschung betrachtet werden kann.³³

28 S. KII-Gesamtkonzept (Anm. 3), S. 39.

29 KII-Gesamtkonzept (Anm. 3), S. B82.

30 KII-Gesamtkonzept (Anm. 3), S. B83.

31 S. unter <http://pkp.sfu.ca/?q=ojs>.

32 S. unter <http://www.researchgate.net/>.

33 An dieser Stelle gilt mein ausdrücklicher Dank den Kolleginnen und Kollegen der Gruppe LIS, deren Einschätzung und Ratschläge eine große Hilfe beim Ausarbeiten der von der Projektgruppe „Bildung und Forschung“ erbetenen Stellungnahme waren.