

von Nora Schmidt

Inhalt

1. Problemaufriss: Die Forschenden und ihre Daten
2. Relevanz für Bibliotheken: Expertisen und Chancen
3. Fragestellung und Methode: Die Praxis antwortet
4. Die Analyse: Das »volle Programm«
 - 4.1. Strategieentwicklung und Geschäftsplanung
data.bris an der University of Bristol
RoaDMap an der University of Leeds
ADMIRe an der University of Nottingham
 - 4.2. Policy
 - 4.3. Data Management Planning (DMP)
 - 4.4. Archivierung und Publikation
Orbital an der University of Lincoln
DaMaRO an der University of Oxford
 - 4.5. Training
Open Exeter an der University of Exeter
5. Bibliothekarische Aufgaben in den Projekten des Jisc-RDM-Programms 2011–13
6. Derivate

Zusammenfassung: Die Fördereinrichtung Jisc unterstützte 2011 bis 2013 27 britische Universitäten, Forschungsdatenmanagement-(RDM)-Services einzuführen oder auszubauen. Orientierung boten dabei die umfangreichen Materialien und Tools des Digital Curation Centers (DDC), die den gesamten Lebenszyklus der Daten bedienen. Auch die Struktur des vorliegenden Artikels folgt diesem Zyklus, um den State of the Art des RDM darzulegen, wie er von sechs näher zu beschreibenden Jisc-Projekten realisiert wurde. Ziel der Analyse war, anhand dieser Beispiele zu eruieren, welche Aufgaben von Bibliotheken übernommen werden. In erster Linie sind dies die Leitung des RDM-Services und die Durchführung von Trainingsmaßnahmen. Auch Repositorien und Metadaten werden meist hier gepflegt, wobei dies ebenso wie die Projektleitung und Bedarfsanalyse häufig auch an der IT angesiedelt ist. Die Policy-Entwicklung findet als weitere wichtige Bibliotheksaufgabe oft in Kooperation mit dem Forschungsservice statt, dem fast immer das Data Management Planning (DMP) obliegt.

Schlagwörter: Forschungsunterstützung, Forschungsdaten, Forschungsdatenmanage-

RESEARCH DATA MANAGEMENT AND LIBRARIES. LOCALISATION IN COOPERATIVE NETWORKS

Abstract: *The funder Jisc supported 27 british universities to establish or improve their research data management (RDM) services in 2011–2013. The comprehensive materials and tools by the Digital Curation Center (DDC), which handle the entire life cycle of the data, offered orientation. The structure of this article also follows this cycle to describe the state-of-the-art RDM as it was realized by six Jisc projects that will be focussed on. The aim of this analysis was to elicit from these examples which tasks are undertaken by libraries. Primarily, this are the overall management of RDM Services and the implementation of training. Repositories and metadata are also mostly curated here, while these tasks are often located at the IT services, as well as the project management and requirements analysis. The policy development is one more important task of the library, frequently held in cooperation with the research services, which are almost always in charge of data management planning (DMP).*

Keywords: *research support, research data, research data management, RDM, new library services, data librarian, academic library, research services, repository*

1. Problemaufriss: Die Forschenden und ihre Daten

Das PARSE-Insight-Projekt (FP7) fand in einer breit angelegten Umfrage heraus, dass 25% der Forschenden ihre Daten öffentlich zugänglich machen, während 11% die Zugänglichkeit auf ihre FachkollegInnen beschränken. Der größere Anteil, 58%, teilt Daten nur im Rahmen der eigenen Forschungsgruppe (Thaesis und Hoeven 2010). Die Hauptgründe sind fehlende Infrastrukturen, aber auch rechtliche Bedenken bzw. Hinderungsgründe sowie Missbrauchsängste und technische Fragen, die z. B. die Inkompatibilität der verwendeten Datenformate mit verbreiteten Analyseinstrumenten betreffen. In nur wenigen Fachbereichen sind bereits Standards etabliert, die den Austausch von Daten überhaupt ermöglichen. Ein Beispiel ist die *Systems Biology Markup Language* (siehe Hucka, Finney et al. 2003).

Eine andere Studie, in der v. a. UmweltwissenschaftlerInnen befragt wurden (Tenopir, Allard et al. 2011) – eine Disziplin, deren Fragestellungen

oftmals von hoher öffentlicher Relevanz sind – findet zwar als meistverbreitete Gründe für das Nichtzurverfügungstellen von Daten den Mangel an Zeit (54%) und Geld (40%), die Daten entsprechend aufzubereiten, aber auch hier sind Rechtsfragen (24%), Infrastruktur (24%) und das Nichtvorhandensein von Standards (20%) wichtige Punkte. Nur 15% geben an, die Veröffentlichung nicht zu wünschen.

Nur der Zugang zu den erhobenen und strukturierten Daten einerseits, und den analysierten Daten mitsamt den Analyseinstrumenten andererseits macht Forschungsergebnisse nachvollziehbar und damit verifizierbar. Der Öffentlichkeit und damit der Politik müsste es ein Anliegen sein, öffentlich finanzierte ebenso wie öffentlich angewandte Forschung, beispielsweise in der Medizin, verifizieren zu können. Wenigstens jedoch sollten Forschungsdaten sicher archiviert werden. Bislang lassen lediglich EU-Initiativen auf entsprechendes politisches Handeln hoffen. In einem Entwurf der Europäischen Kommission zu einem Handlungsplan im Rahmen des Anfang 2014 startenden Förderungsprogramms »Horizon 2020« wird die angedeutete Position vertreten (European Commission 2013).

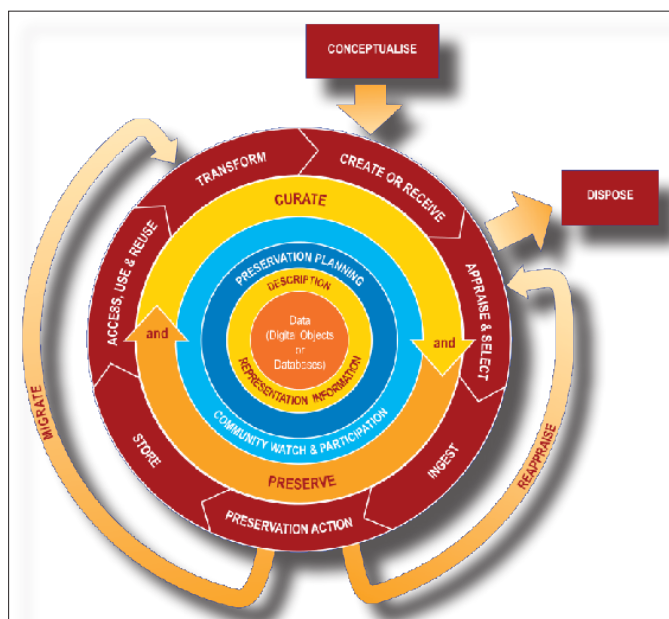


Abb. 1. Die Pflege der Daten über den gesamten Lebenszyklus, CC-BY: Digital Curation Centre (DCC), siehe <http://www.dcc.ac.uk/resources/curation-lifecycle-model>, zuletzt besucht am 27. Mai 2013.

Um entsprechende Services überhaupt entwickeln zu können, muss eine Vorstellung davon bestehen, wie Forschungsdaten »leben« (siehe Abbildung 1). Nur in einigen Fällen werden Daten vollkommen neu produziert – und dann muss die Erhebung bereits im Hinblick auf das Folgende gut geplant sein –, sondern bestehende Daten werden ausgewählt, ausgewertet und erweitert. Nicht nur für ein »Recycling«, auch für eine eventuell notwendige Entsorgung oder Überführung an andere Orte sollte es festgelegte Verfahren geben, denn mitunter handelt es sich um sensible Daten. In allen »Lebensphasen« der Daten sind Operationen notwendig, die je an diese Phasen angepasst unterschiedliche Formen annehmen können:

1. Die Pflege und Bewahrung der Daten,
2. die Beschreibung der Daten, die auch eine Beschreibung ihrer selbst beinhaltet, so dass beides, Daten und Beschreibung, theoretisch reproduzierbar werden,
3. die Einbeziehung und Beobachtung der relevanten Interessengruppen (siehe Abschnitt 3.) und
4. die administrative Planung aller Schritte.

Nachdem entsprechende Aktivitäten bislang v. a. von den Forschenden selbst initiiert wurden, bestehen in einigen Fachbereichen oder Themenfeldern seit vielen Jahren Datenzentren, die Forschungsdaten systematisch professionell archivieren und mitunter auch zugänglich machen, v. a. in den Geo- und Lebenswissenschaften. Klima- oder Genforschung ist vielfach nur möglich, weil viele Daten bereits vorhanden sind. Umgekehrt könnte eine einzelne Forschungsgruppe kaum für sich stehende Antworten auf Forschungsfragen dieser Disziplinen erarbeiten: Die Komplexität der Daten erfordert eine neuartige »wissenschaftliche Arbeitsteilung«.

Die große Mehrheit der Forschenden speichert jedoch Daten bevorzugt auf lokalen Festplatten oder den Servern ihrer Forschungseinrichtung. Beliebte sind auch USB-Sticks oder kommerzielle Cloud-Services wie Dropbox. Nur 20% – hier sind die genannten »Pioniere« mitgezählt – nutzen digitale Archive für Forschungsdaten (Thaesis und Hoeven 2010). Die »alternativen« Methoden haben eines gemeinsam: Beschrieben sind die Daten meist ausschließlich durch ihren Dateinamen; nähere Informationen über die ErzeugerInnen, die Erstellungszeit, den Umfang sowie inhaltliche Aspekte sind – wenn überhaupt für Außenstehende – lediglich über die Daten selbst zu erschließen. Falls also im Nachhinein doch Nachnutzungen ermöglicht werden sollen, ergeben sich enorme Hürden. Aufgrund der steigenden Be-

deutung von interdisziplinärer Forschung muss beim Erfassen von Metadaten von völlig anderen Nutzungsszenarien in der Zukunft ausgegangen werden. Verwendete Vokabulare – im Sinne eines allgemeinen Gebrauchs der Wissenschaftssprache – sollten keine vorausgesetzten innerdisziplinären Selbstverständlichkeiten implizieren. Solche Überlegungen kombiniert mit der bibliothekarischen Wahrheit, dass eine nachträgliche Korrektur von Metadaten tendenziell aufwändiger ist als ihre Erstellung, fundiert die vielfältigen Initiativen, die seit einigen Jahren an Entwürfen für internationale Metadaten-Standards arbeiten, allen voran DataCite.¹

2. Relevanz für Bibliotheken: Expertisen und Chancen

Auch unter den Bibliotheken gab es Pioniere: Die Idee hinter DataCite – Primärdaten zitierfähig und damit zu einer Reputationsquelle zu machen – wurde bereits Anfang 2002 unter maßgeblicher Beteiligung der TIB Hannover bei einer Sitzung des DFG-Bibliotheksausschusses präsentiert (Lautenschlager und Sens 2003). Die meisten anderen Bibliotheken entdecken hier jedoch erst allmählich ein Handlungsfeld bzw. erst jetzt sehen die anderen Akteure des Feldes in Bibliotheken wichtige Partnerinnen. So ist in nur einem der 48 Vorhaben im Rahmen des *European Strategy Forum on Research Infrastructures (ESFRI)* eine Bibliothek involviert. In allen Projekten geht es um die digitale Unterstützung des gesamten, kollaborativen Forschungsprozesses, der selbstverständlich starke disziplinäre Unterschiede aufweist. Ohne hier über die Gründe des Verzichts auf bibliothekarische Beteiligung spekulieren zu wollen, liegt auf der Hand, dass Kompetenzen in der Informationserschließung und -vermittlung, der Langzeitarchivierung unverzichtbar sind. In den Bibliotheken sind an Forschungseinrichtungen am ehesten ExpertInnen zu finden, die sich mit der Interoperabilität von Repositorien, der Sicherstellung von Integrität und Authentizität sowie der Erhöhung der Sichtbarkeit von Publikationen befasst haben – Wissen, dass sich leicht auf Forschungsdaten erweitern lässt – und muss, wenn Bibliotheken weiterhin als zentrale Orte für Informationskompetenz fungieren sollen. Sie können nicht dazu beitragen, den ForscherInnen die nötigen Ressourcen für die Aufbereitung der Daten zur Verfügung zu stellen. Aber können Sie den Aufwand für die Forschenden reduzieren, die sich Systeme wünschen, die so einfach zu bedienen sind wie eine *Dropbox*?² Sie sind weder ExpertInnen im Entwickeln solcher Systeme, noch kennen sie die Anforderungen der Fördernden und des Wissenschaftsmanagements gut genug, wohl aber

können sie Inhalte pflegen und markieren, so dass der Zugang zu ihnen jederzeit möglich ist. All dies gilt es beim Forschungsdatenmanagement zu vermitteln.

Selbstverständlich kann es angemessen sein, für neue Aufgaben neue Strukturen – Stellen und Abteilungen – zu schaffen, nur ist dies an Universitäten ein langer Prozess, in manchen Fällen aufgrund der Verpflichtung zum Stellenabbau auch unmöglich. Neue Stellen wie die des *data scientists* oder *data librarians* (siehe Pampel, Bertelmann et al. 2010) sind im Übrigen nur auf der Grundlage von klassischer wissenschaftlicher oder bibliothekarischer Tätigkeit denkbar.

Forschungsdatenmanagement bietet für wissenschaftliche Bibliotheken die Chance, bereits angebotene Services aus dem Bereich der Forschungsunterstützung für eine effektivere Nutzung besser zu verzahnen. So haben Bibliotheken mit der Etablierung von Repositorien für Publikationen am Ende des Forschungsprozesses angesetzt, der nächste logische Schritt – rückwärts – ist, sich um umfassende Forschungsdatenarchivierung zu bemühen. Gab es bislang nur schwache Anreize für die Nutzung von Publikationsservern, wird den Forschenden zunehmend bewusst, dass Ihre Daten schon für die eigene Nachnutzung sicherer archiviert und besser beschrieben werden müssen.³ Vorausgesetzt, die Verzahnung innerhalb der Software-Landschaft einer Institution oder eines Verbunds bietet entsprechende Verknüpfungen, so ist es von der Archivierung von Forschungsdaten zu jener von Publikationen nur ein kleiner Schritt. Auch sind die bislang verbreiteten Methoden zur Forschungsdatenarchivierung schlecht für immer größere Datenmengen geeignet. Fast die Hälfte der Forschenden müssen mehrere GB speichern; 6% sogar mehr als ein TB (Thaesis und Hoeven 2010). Dabei sind die Daten der Digital Humanities nicht zu unterschätzen: Ein 3D-Modell von einer Grabungsstätte ist im Normalfall mehrere TB groß.

Nichtzuletzt üben Forschungsfördernde Druck aus: So verpflichten sich die WissenschaftlerInnen im Rahmen von Projekten, die durch den Fonds zur Förderung der wissenschaftlichen Forschung (FWF) finanziert werden, auch ihre Daten Open Access verfügbar zu machen, ohne dass der FWF selbst dazu eine Infrastruktur zu Verfügung stellen würde.⁴ Schließlich ist auch nicht zu vernachlässigen, dass mittlerweile einige Zeitschriften ebenfalls die Zugänglichkeit der Daten fordern. Weiter unten wird darauf zurückzukommen sein (siehe S. 16).

3. Fragestellung und Methode: Die Praxis antwortet

Nachdem nun die Relevanz von Forschungsdatenmanagement als Arbeitsfeld von wissenschaftlichen Bibliotheken begründet wurde, soll es im Folgenden weniger um die Forderung nach Open Access für Daten gehen – dazu müssten z. B. auch Anonymisierungsmethoden für personenbezogene Daten und rechtliche Aspekte behandelt werden –, sondern die neuen bibliothekarischen Aufgaben sollen benannt werden. Welchen Beitrag können Bibliotheken zur Unterstützung der Forschenden während des gesamten Forschungsprozesses leisten, der sowohl die Erzeugung als auch die Verbreitung von strukturierten Daten und Wissen enthält? Immer mitlaufend sind 1) die durch die wissenschaftliche Kommunikation selbst-erzeugten, gegenwärtigen, disziplinspezifischen Regeln des wissenschaftlichen Arbeitens, 2) die Anforderungen der Fördernden und des Wissenschaftsmanagements und 3) die internationalen Standards für das Archivieren und Anbieten von Daten. Alle Gesichtspunkte erfordern vielfältige Kooperationen mit den jeweiligen AkteurlInnen dieser Felder.

Ziel dieser Untersuchung ist, anhand von Praxisbeispielen zu eruieren, welche Aufgaben innerhalb eines umfassenden Service-Portfolios (siehe Abb. 2) von Bibliotheken übernommen werden. Gleichzeitig sollen auch jene Aufgaben identifiziert werden, die andere Akteure effektiver werden bearbeiten können bzw. bereits erfüllen, so dass kein Bedarf besteht, dass Bibliotheken sich hier Verantwortung zuschreiben. Die Beispiele finden sich unter den insgesamt 27 britischen Projekten, die von der Förderungseinrichtung für Forschungsinfrastrukturen, Jisc (ursprünglich ein Akronym für *Joint Information Systems Committee*), im Rahmen des im März 2013 abgeschlossenen Programms *Managing Research Data* gefördert wurden. Das Programm deckt den gesamten »Lebenskreislauf der Daten« ab, weshalb es sich für die folgende Untersuchung so gut eignet.

Die Anwesenheit der Autorin beim Jisc Research Data Management Workshop am 25.–26.3.2013 in Birmingham, dem internen Abschluss-Workshop zum Förderprogramm, stellt den Anlass dieser Untersuchung dar. Die Verwertung der Aussagen einzelner Referierender ist jedoch nur möglich, wenn ihre Namen und ihre institutionellen Bezüge verschwiegen werden, denn der Workshop wurde unter der *Chatham House Rule*⁵ abgehalten. Es sollen daher im Folgenden die beim Workshop aufgenommenen Informationen dabei helfen, den auf den Websites der entsprechenden Institutionen zu findenden Informationen in einen impliziten Kontext zu stellen

und Relevanzen zuzuordnen. Hier wiederum wird sich ein großer Vorteil des Untersuchungsfeldes erweisen: Eine Förderung durch Jisc ist mit der Anforderung verbunden, eine Projektwebsite zu pflegen. Es wird empfohlen, jegliche Erträge öffentlich zugänglich zu machen. Für viele wichtige Dokumente wie Projektpläne und Workpackage-Listen stellt Jisc Vorlagen zur Verfügung, die gern benutzt werden. Diese Maßnahmen führen zu einer hervorragenden Vergleichbarkeit der Projekte, die jedoch wieder dadurch limitiert wird, dass die Organisationsstrukturen der einzelnen Universitäten sich unterscheiden, so dass von ähnlich betitelten Abteilungen ganz unterschiedliche Aufgaben übernommen werden. Die individuellen Kompetenzen der dort angesiedelten Mitarbeitenden lassen sich nicht von ihrer Stellenbezeichnung ablesen. Aus diesem Grund wurden auch stichprobenhaft *curricula vitae*, die häufig, nicht immer auf den Webseiten der Universitäten zu finden sind, in die Überlegungen einbezogen.

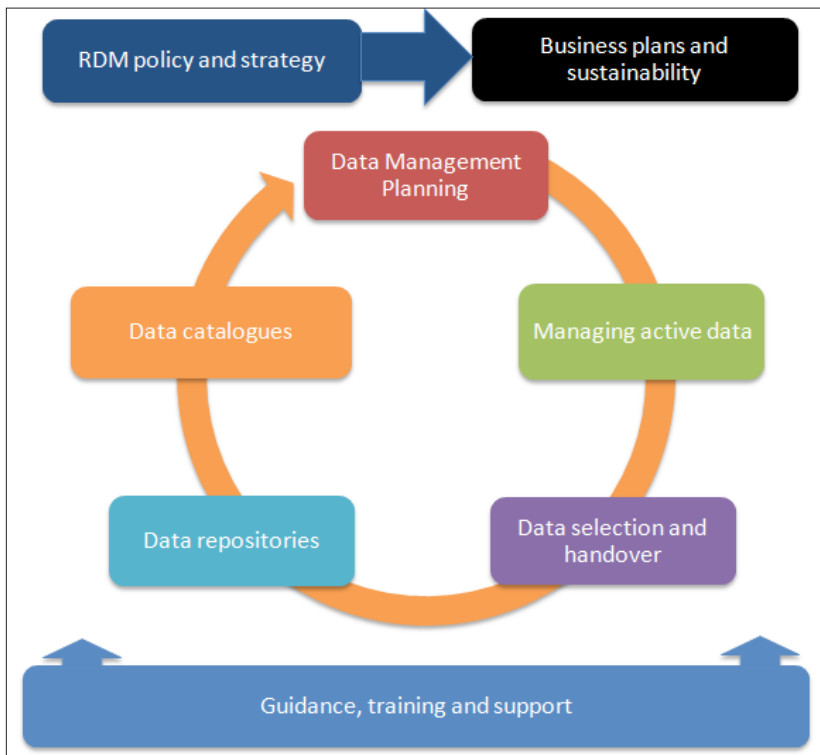


Abb. 2. Komponenten eines umfassenden Service-Portfolios, CC-BY: Jones, Pryor et al. 2013.

Die Struktur der Analyse durchbricht die Linearität eines gewöhnlichen Artikels durch eine Doppelgleisigkeit: Einerseits werden die einzelnen Aufgabenbereiche, die im Zuge der Entwicklung eines umfassenden Forschungsdatenmanagementservices unumgänglich abzudecken sind, nacheinander charakterisiert, andererseits die ausgewählten Projekte, teilweise über die Grenzen der entsprechenden Abschnitte hinaus vorgestellt.

4. Die Analyse: Das »volle Programm«

Jisc förderte bislang zwei Programme zum Forschungsdatenmanagement, mit einer Laufzeit von je knapp zwei Jahren, beginnend im Oktober 2009. Es geht dabei v. a. um ein Experimentieren, aber auch darum, sich als tragfähige Konzepte herausstellende Ansätze breit zu implementieren und zu etablieren. Jisc kooperiert mit vielen britischen und internationalen Partnern, wie der Europäischen Kommission oder dem Research Excellence Framework, der Bewertungsagentur für die Höhere Bildung in UK, um auf geteilte Ziele hinarbeiten zu können.

Das aktuelle Programm ist gemeinsam mit zwei anderen unter dem Dach des *Digital infrastructure: Research management* programme verortet worden und hat ein Gesamtvolumen von £ 4.600.000. Die Sparten des Managing Research Data Programmes decken gleichzeitig alle derzeit denkbaren Aspekte des Forschungsdatenmanagements ab:

1. Die Entwicklung von Services und Software, die Forschende bei der Erstellung eines Datenmanagementplans (DMP) unterstützen, wurde bei einer Laufzeit von sechs Monaten jeweils mit bis zu £ 100.000 gefördert. Bei vielen Forschungsfördernden, z. B. der Europäischen Kommission⁶, werden solche Pläne als Bestandteil von Forschungsanträgen gefordert (siehe Abschnitt 4.3.).
2. Für die Entwicklung von technischer Infrastruktur ist mit 18 Monaten zwar dreimal so viel Zeit veranschlagt, aber verhältnismäßig wenig Geld: Bei einer maximalen Förderung von £ 250.000 kann nur schwer neues Personal engagiert werden (zu technischen Aspekten siehe v. a. Abschnitt 4.4.).⁷
3. Unter dem Titel »Datenpublikation« fallen beispielsweise Projekte, die helfen, die Forschenden zu beraten, z. B., was die Wahl eines geeigneten Journals angeht, das Wert auf die Verifizierbarkeit von Forschungsergebnissen legt (mehr im Abschnitt 4.2.).

4. Schließlich wurden Trainingsprogramme für ForscherInnen, BibliothekarInnen und IT-Personal entwickelt, angefangen von Überlegungen zu Datenformaten bis hin zu Methoden, die Sichtbarkeit des publizierten Datensatzes zu erhöhen (Ausführliches in Abschnitt 4.5.). Fünf bereits im ersten Programm geförderte Projekte erhielten zusammen £ 300.000.

4.1. Strategieentwicklung und Geschäftsplanung

Jedes Projekt beginnt mit dem revidierbaren Entwurf einer Zielsetzung und einer Untersuchung der zu berücksichtigenden Bedingungen des Gelingens, insbesondere einer Bedarfsanalyse, einer Ressourcenplanung und einer Risikoanalyse. Am Ende dieses Prozesses steht idealerweise ein Geschäftsplan, der über einen Projektplan hinausgehend möglichst detailliert nicht nur kalkuliert, was die Projektdurchführung kosten wird, sondern v. a., wie der laufende Betrieb finanziert werden soll. Dazu muss skizziert werden, welcher Serviceumfang angeboten werden soll. Im Folgenden sollen keine Geschäftspläne verglichen werden – diese sind in aller Regel interne Papiere, sondern vielmehr die Aktivitäten, die diese zum Ziel haben. Darüber hinaus kann kurz nach Abschluss der Projektphase auch keine Aussage über sich bewährende Geschäftspläne getroffen werden.

Ist die Errichtung einer technischen Infrastruktur ein Projektziel, wird mit einer groben Vorstellung davon begonnen, was machbar und gewollt sein dürfte – meist vor dem Hintergrund der Beobachtung neuer Services an vergleichbaren Institutionen. Dem folgt eine Bedarfserhebung unter der anvisierten Zielgruppe, den zukünftigen Nutzenden des Services. Von den Ergebnissen der Umfragen anderer jedoch auf die eigene Institution zu schließen, ist einigermaßen riskant: An Institutionen bilden sich, so eine beim Jisc-Workshop breit unterstützte These, Mikrokulturen, vielleicht je weniger, desto höher die Fluktuation ist und damit Kulturen weitergetragen und adaptiert werden. Ein vielleicht noch wichtigerer Grund für die Durchführung einer Befragung an der eigenen Institution ist die – meist erstaunlich hohe – Rücklaufquote (siehe z. B. Kometa 2012), die verrät, dass es in der Zielgruppe durchaus Interesse am Thema gibt und somit die Wahrscheinlichkeit steigt, dass nach der Beteiligung an einer Umfrage auch Interesse am Ergebnis und den daraus entwickelten Aktivitäten bestehen wird.

Obwohl diese Station der Bedarfserhebung auf dem Weg zum Forschungsdatenmanagementservice für unsere Frage wenig relevant ist – für die

Anwendung statistischer Methoden findet sich unter dem universitären Servicepersonal im Normalfall eher zufällig die entsprechende Qualifikation⁸ – soll hier kurz erwähnt sein, dass es die Kenntnis der Grundlagen des Fragebogendesigns an dieser Stelle gar nicht braucht: Wiederum Jisc förderte bereits 2008 die Entwicklung eines *Data Audit Frameworks (DAF)*⁹ durch das *Humanities Advanced Technology and Information Institute (HATII)* an der *University of Glasgow* in Kooperation mit dem unter anderem dort verorteten DCC. Mittels des DAF wird eruiert, welche Praxen des Umgangs mit Forschungsdaten unter den Befragten herrschen. Es zielt darauf ab, Risiken zu erkennen und die Effizienz bei der Datenpflege zu steigern. Vergleicht man die Ergebnisse von Befragungen, denen die Anwendung der inhaltlichen Vorgaben und Werkzeuge aus diesem Rahmenwerk zugrunde lag, wie z. B. jenes, bereits oben zitiertes aus dem *iridium*-Projekt an der *University of Newcastle* (Kometa 2012) mit solchen, die auf die DAF-Verwendung scheinbar verzichtet haben, wie die *University of Oxford* beim Projekt *DaMaRO*,¹⁰ kommt man schnell zu dem Schluss, dass sie mit dem DAF an Aussagekraft gewinnen: Letzterem fehlt beispielsweise die entscheidende Frage nach den verwendeten Technologien zur kurz- und längerfristigen Datenspeicherung.

Ist dann ein Bild davon konstruiert, welche Bedürfnisse die Datenproduzierenden haben, kann die Zielsetzung konkretisiert und eine Ressourcenplanung begonnen werden: Kernstück eines Geschäftsplans. Dazu gehören nicht nur Überlegungen zu finanziellen Fragen, sondern es muss auch festgelegt werden, welche Expertisen und Autoritäten für den Erfolg des Projekts unverzichtbar, welche hilfreich sein werden. Die Einbeziehung aller relevanten Abteilungen der Institution – von der ersten Überlegung angefangen – ist entscheidend dafür, ob das Service und die damit verbundenen Praxen der Forschenden als für die gesamte Institution zentral und damit höchst relevant angesehen werden. In den meisten Jisc-Projekten entschied man sich dafür, der Bibliothek die Projektleitung zu übertragen. Es wäre durchaus denkbar, insbesondere bei den betont technischen Projekten, die Leitung bei der IT anzusiedeln oder, bei Projekten mit einem Schwerpunkt auf DMP, beim Forschungsservice. Weiters könnte für Trainingsprojekte auch die Personalentwicklung verpflichtet werden. Warum die Bibliothek? Auch wenn vielerorts eine mangelnde Anerkennung der Forschenden gegenüber BibliothekarInnen beklagt wird: Konzeptionell ist die Bibliothek die einzige Schnittstelle der nicht-wissenschaftlichen Organisationseinheiten einer Institution zu Forschenden, die nicht etwa durch eine inner- oder außeruniversitäre Bürokratie oder eine Bereitstellung mehr oder we-

niger inhaltsoffener Infrastruktur geschaffen wird, sondern ausschließlich über die Weitergabe von inhaltlichen Arbeitsgrundlagen eines jeden Forschenden. Sollte diese Schnittstelle auch in den meisten Fällen eine rein virtuelle, unbemerkte sein: Forschungsdatenmanagementservices könnten eine Chance sein, die Schnittstelle zu stärken, damit der Bibliothek die Anforderungen ihrer konkreten NutzerInnen besser bekannt werden und so beiderseits effektiver gearbeitet werden kann. Solche Services erreichen nämlich mitunter ein Komplexitätsniveau, das nur mit hohem Aufwand, wiederum beiderseits, ohne Interaktion unter Anwesenden vermittelt werden kann. Anhand eines Beispiels aus dem Jisc-Programm soll nun gezeigt werden, welche personellen Ressourcen für groß angelegte Projekte in Anspruch genommen werden.

data.bris an der University of Bristol

Das aus dem Projekt *data.bris*¹¹ an der *University of Bristol*¹² resultierende Service befindet sich derzeit in der auf 2,5 Jahre angesetzten Ausbauphase.¹³ Währenddessen hat das Service-Team nach außen sichtbar drei Mitglieder: einen Bibliothekar, zwei IT-Fachleute. Das Team kooperiert mit der Abteilung *Research and Enterprise Development (RED)*, dem Forschungsservice. Der Projektzeitpunkt war heikel, denn gleichzeitig sollte das neue Forschungsdokumentationssystem (*Pure*) eingeführt werden. Da man auch hierüber Daten hochladen und in einem damit verknüpften Repositorium archivieren kann, stellt sich schnell die Frage, ob nicht eine Eingabemaske genügt – und eine Präsentationsansicht. Forschungsdaten liegen jedoch häufig als viele Hundert Dateien in verschachtelten Ordnerstrukturen vor, die in der Präsentationsansicht auch zu durchblättern sein müssen, weil es sonst erforderlich wäre, riesige Datenmengen herunterzuladen, bevor man eine einzige Datei betrachten kann. *Pure* zeigte sich für diese Anwendung als ungeeignet. Im anderen Fall hätte dies enorme Auswirkungen auf den Fortgang des Projekts gehabt!

Das *data.bris*-Projekt war in der Forschungs- und Entwicklungsabteilung des Rechenzentrums (*IT Services R&D / ILRT*) lokalisiert. Von dieser Abteilung waren sechs MitarbeiterInnen im Projektteam involviert. Ebenfalls beteiligt waren zwei MitarbeiterInnen des Zentrums für Hochleistungsrechnen und ein Bibliothekar, der nun auch das Service betreut. Das Bild verschiebt sich etwas, wenn deutlich wird, dass die Stelle des Projektmanagers zwar am Rechenzentrum verankert ist, er aber als Konservator hochqualifiziert für die Pflege digitaler Sammlungen ist. Folgende Aufgaben

wurden von diesen beiden Mitarbeitern übernommen: die Entwicklung des Metadatensets, das Training und das Mitentwerfen der Policy und des Geschäftsplans.

Eine elfköpfige Steuerungsgruppe beriet das Projektteam. Darunter war die Pro-Dekanin für Forschung in den Geisteswissenschaften – die *Faculty of Arts* fungierte als Pilotpartner des Projekts. Hier entwickelte Policies, Tools, Workflows und Trainingsmaßnahmen sollten nach der erfolgreichen Erprobung auf die Bedürfnisse anderer Fakultäten angepasst werden. Außer der Pro-Dekanin berieten folgende das Projekt im Rahmen der Steuerungsgruppe: der Vize-Rektor für Forschung, der Zuständige für die Entwicklung des akademischen Personals, eine MitarbeiterIn des Forschungsservice und ein Mitglied der Bibliotheksleitung, dort zuständig für Informationsmanagement und die Entwicklung der Bibliothek. Die übrigen sind IT-Spezialisten. Geleitet wurde die Steuerungsgruppe vom Vorsitzenden des *Research Data Storage and Management Board*, dem eine Reihe von RepräsentantInnen der Fakultäten angehören. Das 13köpfige Board berät die Universität und auch das Projekt in disziplinspezifischen Belangen der Forschungsdatenspeicherung. Da es bereits im Februar 2011 eingesetzt wurde, ist seine Gründung offenbar unabhängig vom *data.bris*-Projekt gewesen.

Abgesehen vom beispielhaften phasenhaften Aufbau ist hier die IT-Dominanz sowohl in der Steuerungsgruppe der Projekt-, als auch im Service-Team der Ausbauphase auffällig. Das zeigt, dass ein Großteil der Kernaktivitäten offenbar vor allem IT-Kompetenz erfordert und bibliothekarische Expertisen nur im Bedarfsfall eingeholt werden. Die Leitung des Services wird in der Ausbauphase jedoch auf einen Bibliothekar übertragen, was darauf hinweist, dass erst im laufenden Betrieb v. a. Querschnittskompetenzen erforderlich sind: Es muss zwischen den Forschenden, dem Träger, den Standards internationaler Datenpflege-ExpertInnen und den auf klar formulierte Anforderungen angewiesenen EntwicklerInnen und SystemadministratorInnen vermittelt werden, was mitunter Übersetzungsleistungen erfordert, für die technisch versierte BibliothekarInnen prädestiniert sind.

Weiterhin bemerkenswert an der personalen Konstellation im Rahmen von *data.bris* ist die breite Beteiligung von Rektorats- und Dekanatsmitgliedern, was wahrscheinlicher macht, dass das Projekt auch von der ProfessorInnenschaft und vom akademischen Mittelbau ernst genommen wird. Ebenso vorteilhaft erscheint die Einsetzung eines disziplinär paritätisch besetzten Beratungsgremiums als allererster Schritt. Es kann – schon aus

Zeitgründen – nicht Aufgabe des Servicepersonals sein, mittels Recherchen zu eruieren, welche disziplinspezifischen Infrastrukturen bereits extern verfügbar sind und genutzt werden und an welchen Stellen die Unterstützung der Institution überhaupt erwünscht ist. Ein solches Gremium kann dabei helfen, den Forschenden im Rahmen der Bedarfserhebung die richtigen Fragen zu stellen. Die in der *data.bris*- Projektdokumentation aufgeführten Beteiligungen um die Abteilungen Finanzen, Qualitätssicherung und Öffentlichkeitsarbeit punktuell, aber doch frühzeitig zu ergänzen, könnte das Erfolgspotential noch weiter steigern.

Neben den Personalkosten müssen v. a. die Kosten für Speicherplatz kalkuliert werden, wobei entsprechende Fragen im Rahmen der Bedarfserhebung sehr wertvoll sind. *data.bris* verwendete hierzu das bereits erwähnte Tool CARDIO (siehe En. 9). Im Rahmen der Implementierung des institutionellen Forschungsdaten-Repositorys wurden £ 2 Mio. in Speicherplatz investiert. Forschungsprojektleitende erhalten für ihr Projekt bis zu 5 TB kostenfrei, darüber hinaus wird empfohlen, die anfallenden Gebühren bereits im Forschungsprojktantrag zu veranschlagen.

RoadMAP an der University of Leeds

Nicht viele Universitäten können so hohe Beträge in Speicherplatz investieren, dessen Nutzung in der Zukunft weitestgehend ungewiss ist. So haben sich andere dazu entschlossen, die Einhaltung von Archivierungsanforderungen, das Durchführen regelmäßiger Backups etc. externen Dienstleistern zu überlassen. Das Angebot von Arkivum an die *University of Leeds*¹⁴ scheint hier attraktiv: 19 p pro GB und Jahr und die Integration mit einem selbstadministrierten ePrints-Repository.¹⁵ Selbstverständlich bietet der Markt viele alternative Anbieter, die unter Konkurrenzdruck stehen. Entscheidet man sich dafür, die Aufgabe der sicheren Archivierung an externe abzugeben, verschieben sich die Anforderungen an die Kompetenzen der ProjektmitarbeiterInnen im Vergleich zu *data.bris* etwas von der IT weg. Für *RoadMAP*¹⁶ sind öffentlich zwei Bibliotheksmitarbeiterinnen – Projektmanagerin und -beauftragte – als Ansprechpartnerinnen angegeben. Zum Projektteam gehören darüber hinaus aber drei IT-Fachleute, zwei weitere BibliothekarInnen, zwei MitarbeiterInnen des DCC und je eine/r aus der Personalentwicklung und dem Forschungsservice. Während die Projektadministration, Bedarfserhebung, Policyentwicklung und Training in erster Linie von der Projektmanagerin und der Projektbeauftragten erledigt werden, sind die Arbeitspakete zum *Data Management Planning*, zur

Softwareauswahl und Metadaten (sic!) sowie Speicherplatz primär in den Händen der IT-Fachleute.

Wie *data.bris* – und alle anderen auf diesen Aspekt hin untersuchte Projekte – verfügt auch *RoaDMap* über zwei steuernde und beratende Gremien, auch wenn die Aufgabenkonstellation sich etwas unterscheidet, ebenso die Zusammensetzung: Die Steuerungsgruppe bekleidet eher eine Beratungsfunktion bezüglich (Förder-)Policies und unterschiedlicher Fachkulturen, denn von den 13 Mitgliedern sind nur drei IT-Fachleute und zwei BibliothekarInnen. Die 15köpfige *Research Data Working Group* berät wiederum die Steuerungsgruppe bzw. arbeitet aus, was diese entscheidet und stellt eine Art Qualitätssicherungseinrichtung für das Projektteam dar. Es gibt personelle Überschneidungen, wobei in der *Research Data Working Group* anstelle der DekanInnen ProfessorInnen beteiligt sind. Auch hier überwiegen die IT-Fachleute, gefolgt von Bibliothek, Forschenden sowie VizerektorIn, Personalentwicklung und Forschungsservice.

ADMIRe an der University of Nottingham

Das *ADMIRe*-Projekt¹⁷ an der *University of Nottingham*¹⁸ entwickelte über die Standard-Dokumentation der Projektplanung hinaus selbst eine Kompetenzmatrix¹⁹ für die anfallenden Service-Aufgaben im laufenden Betrieb. Im Gegensatz zu den anderen untersuchten Projekten, wo bei vielen Aufgaben mehrere Abteilungen kooperierten, wurden hier nach einer Diskussion mit den Vertretern der drei in Frage kommenden Abteilungen²⁰ die Aufgaben eindeutig verteilt:

1. Forschungsservice und DoktorandInnenzentrum: Projekt- und Serviceleitung, Verwertung, Integrität, (Förder-)Policies, Finanzen.
2. Forschungs- und Lehrressourcen (= Bibliothek, inklusive Lehrtechnologien): Datenpflege (Metadaten), Repository Management, Informationskompetenz (Training), Rechtsfragen.
3. Rechenzentrum: Daten-Archivierung, also Speicherplatz, Übertragung, Migration, Sicherheit und Authentifizierung.

Die Veranschlagung eines Servicepersonalbedarfs von etwa fünf Vollzeitäquivalenten findet sich nicht nur in der Untersuchung mehrerer Universitäten von Laurian Williamson und Tom Parsons (2013), sondern wurde auch während des Workshops mehrfach bestätigt. Es kommt jedoch selbstverständlich darauf an, wie umfangreich die angebotenen Services sein sollen: Hier reicht das Spektrum vom blanken Angebot einer Archivierungsin-

frastruktur, wofür ein technisch versierter *Repository Manager* genügen dürfte, über ein größeres Angebot, das persönliche Beratung und Training sowie auch für das Bekanntmachen des Services enthält – dies ist das *service level*, das die meisten britischen Universitäten anstreben, bis hin zu einem maßgeschneiderten Angebot für einzelne Disziplinen, was mit der Multiplikation der fünf Vollzeitäquivalente mit der Anzahl der Fachbereiche einhergeht.

4.2. Policy

Das *ADMIRe*-Projekt bietet auch Erkenntnisse zum Thema RDM-Policy, die an der *University of Nottingham* seit 2013 als erstrebenswerte, über einige Zeit zu implementierende Leitlinien eingeführt wurden. Hier lag die Hauptverantwortung dafür, anders als bei *data.bris* und *RoadMAP*, beim Forschungsservice und nicht bei der Bibliothek. Das macht insofern Sinn, als das Forschungsservice die Forschenden bei der Förderantragstellung unterstützt und eine RDM-Policy sich primär daran orientieren sollte, welche Vorgaben diesbezüglich von den Fördernden kommen. Das Forschungsservice verfügt hier über die größte Erfahrung. Eine Policy soll Orientierung bieten, was am ehesten erreicht wird, wenn die vielfältigen Anforderungen der unterschiedlichen Fördernden zu einem Set an Anforderungen kompiliert werden, das allen Fördernden genüge tut. Darüber hinaus sollten auch Journal-Policies berücksichtigt werden, die mitunter Aussagen über Forschungsdaten enthalten. Für diese wiederum kann die Bibliothek Kompetenzen aufweisen, so sie für die Publikationsberatung zuständig sein sollte.

Dabei empfiehlt es sich, gemeinsame Anforderungen der Fördernden stärker, wenn auch nicht verpflichtend zu formulieren – die übrigen in Form von Empfehlungen. Wie bereits aus den vielfach mühsamen Prozessen der Einführung einer Open-Access-Policy oder gar eines -Mandats gelernt werden konnte, erhöht der Einsatz eines Duktus des Erwartens eher als jener der Verpflichtung die Wahrscheinlichkeit der Veröffentlichung einer Policy, zumindest in einem ersten Schritt. Durch eine Verbindung mit häufig schon lange Zeit geltenden Leitlinien zur guten wissenschaftlichen Praxis kann der Erwartung noch mehr Nachdruck verliehen werden.

Neben der Formulierung der Erwartung an die Datenproduzierenden ist jedoch der zweite wichtige Grund für eine RDM-Policy das Bekenntnis zur institutionellen Selbstverpflichtung, entsprechende Infrastrukturen anzubieten. Es kann daher einen positiven, motivierenden Einfluss auf den Projektverlauf haben, wenn die Policy als erstes Projektergebnis präsentiert wird. Der Nach-

teil dieses Vorgehens besteht daran, zu früh Erwartungen der Verfügbarkeit der Infrastrukturen zu wecken. Durch die Befragung der Forschenden sollte man jedoch in der Lage sein, einzuschätzen, ob eine früh veröffentlichte Policy von den Forschenden als Druckmittel, Vision oder leeres Versprechen angesehen werden wird. Das Einfließenlassen der Erkenntnisse über die lokale Forschungskultur in die Policy kann den Erfolg derselben begünstigen.

Ein weiteres Jisc-Projekt soll hier kurz vorgestellt werden, da es, ähnlich wie bei der Bedarfserhebung ein Werkzeug zur Verfügung stellt, das jeden Policy-Erarbeitenden unabhängig von Erfahrungen mit den forschungsdatenbezogenen Anforderungen von Journals machen könnte; für die Anforderungen von internationalen Forschungsfördernden gibt es SHERPA/JULIET²¹. Im Rahmen von *Journal Research Data Policy Bank (JoRD)*²², ebenfalls an der *University of Nottingham* verortet, wurden 371 Journals auf Daten-Policies untersucht, 230 davon hatten solche. Beim Versuch, diese auf ein Set von Einzelanforderungen herunter zu brechen, taten sich große Schwierigkeiten auf, denn ein gemeinsames Vokabular existiert hier genauso wenig wie eine konvergierende Tiefenschärfe: Einige Journals forderten nur sehr vage den Zugang zu Daten (welche? wann? wie lange? für wen? wie?), andere – wenige – waren sehr präzise, insbesondere über den Archivierungs-ort. Nur sechs Journals forderten Open-Access-Daten und nur 62 beanspruchten verpflichtende Geltung, 22 drohten andernfalls Konsequenzen an. Dieses Bild könnte zur Schlussfolgerung verleiten: Noch haben die Verlage diesen Themenkomplex keineswegs determiniert; es bestehen noch große Gestaltungsmöglichkeiten für Forschungsinstitutionen. Allerdings, so fand *JoRD* in Interviews heraus, sind sechs große Verlage derzeit entweder mit der Formulierung von allgemein oder nur für bestimmte Journals geltenden Policies befasst. Auch wenn bei den interviewten Verlagen, Bibliotheken und Forschenden großes Interesse an einer Datenbank besteht, die Daten-Policies, ähnlich wie *SHERPA-RoMEO* dies für Nutzungsrechte von Publikationen tut, knapp und übersichtlich darstellt, so war die Bereitschaft, dafür Ressourcen zur Verfügung zu stellen, eher verhalten. Nur die Verlage zeigten Bereitschaft, für dieses Service auch zu zahlen – um die Konkurrenz im Auge behalten zu können.

4.3. Data Management Planning (DMP)

Zur Planung des Forschungsdatenmanagements konkreter Projekte, was nicht nur dem Druck der Fördernden genüge tut, sondern auch als Selbstverpflichtung dient, mit unter großem Aufwand erzeugten Daten

entsprechend wohlüberlegt zu verfahren, ist die Anwendung des *DMP-Online-Tools*²³ an britischen Universitäten sehr verbreitet. Es ermöglicht das Erstellen eines DMP mit Hilfe von Templates, die sich je nach den Anforderungen der Fördernden oder der Institution unterscheiden.²⁴ DMPonline hat eine Reihe nützlicher Features wie die Möglichkeit der Öffnung des Zugang zu DMP-Entwürfen für eine kollaborative Bearbeitung oder der Erstellung einer Kopie der beim Fördernden eingereichten DMP-Version zur Bearbeitung während der Projektlaufzeit.²⁵ Allerdings wird der große Aufwand, DMP zu erstellen, Forschende so lange abschrecken, dies zu tun, wie Fördernde es nicht flächendeckend einfordern.

Die beiden kurzen Abschnitte über Policies und DMP sollen vor allem zeigen, wie sehr RDM-Services an den Interessen und Vorgaben von Fördernden und dem Wissenschaftsmanagement orientiert sein müssen. Wenn die lokal stärksten Förderinstitutionen noch keine RDM-Anforderungen veröffentlicht haben, gilt es für eine Forschungseinrichtung, sich mit diesen in Verbindung zu setzen, um nicht nach der Einführung einer RDM-Policy mit völlig anderen Anforderungen der Fördernden konfrontiert zu werden. Das externe Wissenschaftsmanagement wiederum muss daraufhin beobachtet werden, welche Relevanz die Wiederverwendungsraten der Daten von Forschenden bei der Beurteilung ihrer Leistungsfähigkeit haben. Wie schon bei der Diskussion um Open Access und die Orientierung an alternativen Metriken bei Berufungsverfahren, sollte eine Universität bei der Einführung von zentralen RDM-Services selbst beginnen, zusätzlich zu den etablierten Formen der Einflussbestimmung für die Arbeit einzelner Forscher auch andere Indikatoren einzubeziehen. Die Unterstützung der in Entwicklung begriffenen, neuartigen Einfluss-Indikatoren für Publikationen und Daten, die an der eigenen Einrichtung hergestellt werden, sollte bei der Implementierung von Publikationsservices höchste Priorität haben.

4.4. Archivierung und Publikation

Die Archivierung von Forschungsdaten erfolgt idealerweise mehrstufig und in disziplinären Kontexten. Forschende verhalten sich gegenüber ihrer Fachgemeinschaft im Normalfall loyaler als gegenüber ihrer Institution, insbesondere, weil sie die Institution eher wechseln als das Fachgebiet, aber auch, weil zentrale Services meist zwar Querschnittsanforderungen zu erfüllen bemüht sind, dies aber auch dazu führt, dass spezielle Anforderungen unbeachtet bleiben müssen. Wo es also bereits außer-institutionelle Infrastrukturen gibt, sollten diese auch genutzt und deren Ausbau

unterstützt werden. Das gilt im Übrigen auch für institutionelle Services, denn, um auf die Mehrstufigkeit zurückzukommen, für die Kurzzeitarchivierung von Daten liegt bereits Open-Source-Software vor, die im Falle von CKAN²⁶ für die Daten von (Regierungs)Organisationen von der Open Knowledge Foundation entwickelt wurde – dazu gleich mehr. Für die Langzeitarchivierung sollten Daten selektiert werden, denn alle produzierten Daten vorzuhalten, sogar zugänglich zu machen, erzeugt enorme Kosten, selbst bei sinkenden Preisen für Speichermedien. Forschende bei der Selektion zu unterstützen, ist eine wichtige Aufgabe im Rahmen von RDM-Services.²⁷ Technische Lösungen für die Langzeitarchivierung werden ebenfalls bereits vielerorts entwickelt, z. B. im Rahmen der nun vorzustellenden Jisc-Projekte.

Orbital an der University of Lincoln

Das *Orbital-Projekt*²⁸ an der *University of Lincoln*²⁹ ist stark auf technische Infrastrukturen fokussiert, mit dem interessanten Schwerpunkt auf der Kompatibilität mit den Anforderungen von Partnern aus der Industrie – der Pilot findet an der Fakultät für Ingenieurwesen statt. Verortet ist das Projekt am Zentrum für die Lehre, dem *Centre for Educational Research and Development*, an dem auch Forschungsprojekte zur Lehre stattfinden. Starke Unterstützung erfährt *Orbital* aus dem Forschungsservice, da dieser auch für die Pflege der Beziehungen zur Industrie verantwortlich ist. Das Team ist dreifach gruppierbar: fünf Personen informieren das Projekt über die Bedürfnisse und Kenntnisse aus ihrer *peer group* und geben Feedback, darunter sind zwei Forschende, ein Bibliothekar, ein Datenbankadministrator und ein Industriepartner. Die zweite Gruppe, geleitet von einem Bibliothekar, versorgt das Projekt mit dem nötigen Hintergrundwissen über vergleichbare Projekte und (Quasi)Standards. Die drei weiteren Mitglieder der Recherche-Gruppe kommen vorrangig aus dem Forschungsservice. Es bleibt die IT-Gruppe mit zwei Entwicklern und einem Sicherheitsexperten. Die Bibliothek spielt also bei der Projektdurchführung eine eher untergeordnete Rolle.

Ausgangspunkte waren ein bestehendes, institutionelles *EPrints-Repository* sowie ein selbstentwickeltes *Award Management System*, das Forschende verwenden müssen, wenn sie ein extern gefördertes Forschungsprojekt an der *Lincoln University* durchführen wollen. Dort sind also bereits viele Projektdaten hinterlegt. Es mussten Brücken geschlagen werden zwischen diesen bestehenden und mehreren neu zu entwickelnden Systemen: ein

Datenkurzzeitarchivierungs- und ein frei zugängliches Forschungsinformationssystem. Diese Brücke ist das nun gelaunchte *Researcher Dashboard*³⁰, über das darüber hinaus auch die Trainingsmodule, das Blogsystem und DMPonline erreichbar sind.

Die größten Energien des Projekts flossen in die Entwicklung der Lösung für die Kurzzeitarchivierung. Auf Grundlage des bereits erwähnten CKAN, das aufgrund seiner standardisierten Programmierschnittstellen mit jeder der verbreiteten Repository-Softwares zusammenarbeitet, sind alle erforderlichen Anpassungen aufgrund des modularen Aufbaus der Software nach den Erfahrungen des Projekts recht problemfrei implementierbar (Winn 2013). CKAN bringt Visualisierungsmöglichkeiten, Metriken und eine große EntwicklerInnen-Community bereits mit. Dennoch steckt die Anpassung an die RDM-Bedürfnisse noch in den Anfängen.

Auch das *data.bris*-Projekt hat sich für die Adaption von CKAN entschieden, hier jedoch nicht zur Kurzzeitarchivierung, sondern um einen öffentlichen Datenkatalog umzusetzen, der für jedes Datenset eine Seite vorhält, auf die ein DOI verweisen kann. Es handelt sich dabei also um ein Präsentations-Frontend für das Daten-Repository. Was den Daten-Upload angeht, so sei hier noch kurz erwähnt, dass es einige weniger, aber dafür nicht unbedeutende Sonderanforderungen an Daten- gegenüber Publikations-Repositories gibt: Neben den erwähnten komplexen Ordnerstrukturen innerhalb eines Sets kommt z. B. hinzu, dass das HTTP sich für die Übertragung großer Datenmengen schlecht eignet, v. a., weil eine unterbrochene Übertragung nicht fortgesetzt werden kann. *data.bris* setzt hier auf BitTorrent.

DaMaRO an der University of Oxford

In Oxford³¹ war aufgrund der starken Forschungsorientierung die Entwicklung eines umfassenden RDM-Services besonders angezeigt: Im Projekt *DaMaRO*³² wurde ebenfalls darauf geachtet, dass Forschende möglichst nur einmal dieselben Daten eingeben müssen (siehe auch En. 25). Nach und nach sollen verschiedene Module implementiert werden, die jeweils dieselben Daten in einen Kontext stellen, der wiederum andere Funktionalitäten bietet. Die Idee entstand offenbar bereits während der Durchführung eines Jisc-Projekts im Rahmen des ersten RDM-Programms. Die Ergebnisse – Software – wurden für die Community als *DataFlow-Projekt*³³ geöffnet. Dieses und einige andere frühere Projekte bilden die Grundlage

für *DaMaRO*, wobei v. a. zwei Module weiterentwickelt werden sollten: Die Idee hinter *DataStage* ist, dass Forschende es auf dem Institutsserver o. ä. installieren lassen und wie das allseits bekannte Netzlaufwerk verwenden können, ergänzt um ein Web-Interface mit der Eingabemöglichkeit für Metadaten. Ausgewählte Daten werden dann mit *DataBank*, einem Fedora-Repositorium mit SWORD2-Protokoll zur Übertragung der Datenpakete, per Mausklick langzeitarchiviert.

An der *University of Oxford* ist die Forschendenunterstützung während der einzelnen Prozesse des RDM-Zyklus ähnlich wie in Nottingham und Lincoln auf Bibliothek, Forschungsservice und Rechenzentrum aufgeteilt. Zusätzlich kann hier noch eine Rechtsberatung hinzugezogen werden. Insgesamt werden nicht weniger als elf Kontaktstellen angeboten.³⁴ Eine zentrale Servicestelle, an die man sich wenden kann, wenn z. B. eine allgemeine Beratung zum Daten-Lebenszyklus gewünscht ist, gibt es nicht, was auch dazu führt, dass es für einen umfassenden Support nötig wäre, sich an immer wieder andere Stellen zu wenden. Dies könnte einerseits eine Hürde darstellen, diese Services zu nutzen, andererseits ist davon auszugehen, dass die einzelnen Stellen dann über besonders detailliertes, weil spezialisiertes Wissen verfügen.

4.5. Training

Alle analysierten Projekte bieten umfangreiche, sehr ähnliche, weil sich an den DCC-Materialien orientierende Online-Informationen an. Im Rahmen des Jisc-Programms wurden auch nachnutzbare Trainingsmodule nicht nur für Forschende, sondern auch für BibliothekarInnen und IT-Personal entwickelt. *RDMRose*³⁵, ein Kollaborationsprojekt mehrerer Universitäten unter der Leitung der *University of Sheffield*, richtet sich beispielsweise v. a. an FachreferentInnen, deren Einsatz aufgrund der disziplinären Besonderheiten von RDM sehr sinnvoll ist. Darüber hinaus ist die Veranstaltung von Workshops auf Anfrage ein Standard-Angebot.

Open Exeter an der University of Exeter

Im Projekt *Open Exeter*³⁶ wurde kurzerhand die Entwicklung des RDM-Supports an der *University of Exeter*³⁷ mit dem Training von DoktorandInnen, die aufgrund ihrer MultiplikatorInnen-Funktionen häufig die Hauptzielgruppe sind, verbunden: Es wurden sieben Freiwilligenstellen für ein Jahr ausgeschrieben. Die Aufgaben der DoktorandInnen reichten von der Bedarfsana-

lyse über die Policyentwicklung bis zur Aktualisierung des Projekt-Blogs³⁸. Obwohl das Pensum bei mehreren Wochenstunden lag und es als Aufwandsentschädigung und gleichzeitig als Arbeitswerkzeug lediglich einen Tablet-Computer gab, erreichten 70 Bewerbungen die an der Bibliothek sitzende Projektleitung. Die DoktorandInnen betreuten unterschiedliche Themen- und Fachbereiche und lernten auf diese Weise viel voneinander. *Open Exeter* ist das einzige Projekt im Jisc-Programm, in dem die Bibliothek alle Schritte im Lebenszyklus der Daten abdeckt: Hier wurden die DoktorandInnen betreut und auch das Management des institutionellen Repositoriums, das für Forschungsdaten genutzt wird, ist hier verortet.

5. Bibliothekarische Aufgaben in den Projekten des Jisc-RDM-Programms 2011–13

Zusammenfassend wird hier in einer Tabelle dargestellt, für welche RDM-Services, die durch die Projekte realisiert wurden, von der Bibliothek primäre Verantwortung übernommen wurde. In manchen Fällen war dies anhand der Projekt- und Service-Informationen nicht leicht nachzuweisen, weshalb die Ergebnisse eher als Tendenzen verstanden werden sollen. In vielen Fällen wird auch in den dargestellten Teilbereichen abteilungsübergreifend gearbeitet. Mitunter ist wenigstens nach außen eine Serviceleitung gar nicht kommuniziert und es lässt sich lediglich anhand der Aktivitäten vermuten, dass eine Stelle Leitungscharakter hat.

	Projekt-admin.	Bedarfs-analyse	Policy	DMP	Meta-daten	Rep-Man	Training	Service
data.bris	-	-	x	-	x	x	x	x
RoaD-MAP	x	x	x	-	-	-	x	x
ADMIRe	-	-	-	-	x	x	x	x
Orbital	-	x	x	-	-	-	x	x
DaMaRO	x	-	-	-	x	x	-	-
Open-Exeter	x	x	x	x	x	x	x	x
Total = 6	3	3	4	1	4	4	5	5

Das Ergebnis ist eindeutig: Zu den primären Aufgaben der Bibliothek gehört offenbar die Leitung des umfassenden RDM-Services und die Versor-

gung der Forschenden und Mitarbeitenden mit dem nötigen Handwerkzeug. Darüber hinaus sind BibliothekarInnen gefragt bei der Entwicklung von Policies und Metadaten-Sets, wobei – vielleicht überraschend – diese Aufgabe mitunter auch von IT-Fachleuten übernommen wird. Dasselbe gilt für das Management von Repositorien. In der Hälfte der Fälle lag auch das Management der Projekte zur Entwicklung der RDM-Services sowie die Feststellung des Bedarfs und die Formulierung der daraus abzuleitenden Anforderungen bei der Bibliothek. Recht eindeutig Aufgabe des Forschungsservice ist nach den Erfahrungen aus den Jisc-Projekten das DMP.

6. Derivate

Andernorts wurde natürlich bereits die Formulierung der bibliothekarischen RDM-Aufgaben begonnen. Inwiefern jedoch die Ergebnisse dieser Untersuchung sich mit den Vorschlägen aus der noch recht überschaubaren Literatur.³⁹ deren ausführliche Diskussion an anderer Stelle erfolgen muss, decken, bleibt anderen Arbeiten vorbehalten. Da es sich bei dieser Literatur v. a. um theoretische Empfehlungen einerseits und um singuläre Projektpräsentationen andererseits handelt, erhält der vorliegende empirisch-vergleichende Ansatz eine gewisse Originalität.

Eine Evaluation des Jisc-Programms oder der RDM-Aktivitäten der erwähnten britischen Universitäten kann hier ebenfalls nicht stattfinden. Es soll nur daran erinnert werden, dass Jisc die einzelnen Projekte als Experimentierfelder versteht und davon ausgegangen werden kann, dass Modelle, die offensichtlich immer wieder von den einzelnen Institutionen übernommen wurden, nicht mehr als Prototypen, sondern als Serienprodukte gelten müssen. Das gilt weniger für die Konstruktion der vernetzten Systeme, die sich zu einer E-Science-Infrastruktur arrangieren, denn hier gibt es durchaus bedeutende Unterschiede, sondern für die Struktur des Service-Portfolios, DMP(online) und das Online-Trainingsangebot.

Inwiefern die institutionellen RDM-Policies sich unterscheiden, konnte ebenfalls nicht analysiert werden.

Die Breite der Jisc-RDM-Programme bietet hervorragendes Material für weitere vergleichende Untersuchungen. Möglicherweise war die RDM-Initiative für viele Universitäten die erste Erfahrung der ausgedehnten Interabteilungskooperation über drei und mehr Abteilungen hinweg. Die dabei höchstwahrscheinlich auftauchenden Missverständnisse und Konflikte

bezüglich Zuständigkeit und Kompetenz vor einer organisationssoziologischen Folie genauer zu untersuchen, könnte wertvolle Ergebnisse liefern, aus denen für solche zukünftig sicherlich häufiger vorkommenden Kooperationen gelernt werden kann.

Die Anwendung der Methodologie dieser Untersuchung war aufwändiges Experiment. Es wurden unzählige Projektdokumente gesichtet, die wie gesagt den Vorteil hatten, einer vergleichbaren Struktur zu folgen. Solche Daten aus einer Datenbank abrufen zu können – Stichwort Data-Mining – hätte unendlich viel Zeit und Mühe gespart.

Mag.^a Nora Schmidt
Universitätsbibliothek Wien, Open Access Office
E-Mail: nora.schmidt@univie.ac.at

Literatur

- Antell, Karen, Jody Bales Foote und Brian Shults (2013). »Dealing with Data: Science Librarians' Participation in Data Management at Association of Research Libraries Institutions«. In: *College & Research Libraries*. Preprint. url: <http://crl.acrl.org/content/early/2013/04/05/crl13-464> (besucht am 10. 06. 2013).
- European Commission (2013). *Research Data e-Infrastructures: Framework for Action in H2020*. url: https://ec.europa.eu/digital-agenda/sites/digital-agenda/files/data_einfra_h2020_fiches_online_consult.pdf (besucht am 16. 05. 2013).
- Hucka, M., A. Finney, H. M. Sauro, H. Bolouri, J. C. Doyle, H. Kitano et al. (2003). »The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models«. In: *Bioinformatics* 19.4, S. 524–531. doi: 10.1093/bioinformatics/btg015.
- Jones, Sarah, Graham Pryor und Angus Whyte (2013). *How to Develop Research Data Management Services – a guide for HEIs*. DCC How-to Guides. url: <http://www.dcc.ac.uk/resources/howguides/how-develop-rdm-services> (besucht am 27. 05. 2013).
- Kometa, Simon (2012). *Research data management requirements online survey report. iridium project output*. url: http://research.ncl.ac.uk/media/sites/researchwebsites/iridium/iridium_online_survey_report_17_8_2012_v2.1_SK.pdf (besucht am 02. 06. 2013).

- Lautenschlager, Michael und Irina Sens (2003). »Konzept zur Zitierfähigkeit Wissenschaftlicher Primärdaten«. In: *Information – Wissenschaft und Praxis* 54.8, S. 463–466.
- Lewis, Martin (2010). »Libraries and the Management of Research Data«. In: *Envisioning Future Academic Library Services: Initiatives, ideas and challenges*. Hrsg. von Sue McKnight. London: Facet Publishing.
- Martinez, Luis (2009). *Using the Data Audit Framework: An Oxford Case Study*. url: <http://www.discuk.org/docs/DAF-Oxford.pdf> (besucht am 02. 06. 2013).
- Pampel, Heinz, Roland Bertelmann und Hans-Christoph Hobohm (2010). »'Data Librarianship': Rollen, Aufgaben, Kompetenzen«. In: *Working Paper Series des Rates für Sozial- und Wirtschaftsdaten* 144. url: <http://hdl.handle.net/10419/43622>.
- Rice, Robin und Jeff Haywood (2011). »Research Data Management Initiatives at University of Edinburgh«. In: *The International Journal of Digital Curation* 6.2, S. 232–244. doi: 10.2218/ijdc.v6i2.199.
- Starr, Joan, Perry Willett, Lisa Federer, Claudia Horning und Mary Linn Bergstrom (2012). »A Collaborative Framework for Data Management Services: The Experience of the University of California«. In: *Journal of eScience Librarianship* 1.2. doi: 10.7191/jeslib.2012.1014.
- Tenopir, Carol, Suzie Allard, Kimberly Douglass, Arsev Umur Aydinoglu, LeiWu, Eleanor Read, Maribeth Manoff und Mike Frame (Juni 2011). »Data Sharing by Scientists: Practices and Perceptions«. In: *PLoS ONE* 6.6, e21101. doi: 10.1371/journal.pone.0021101.
- Thaesis und Jeffrey van der Hoeven (Juni 2010). *Insight into digital preservation of research output in Europe. Insight Report (D3.6)*. url: <http://www.parse-insight.eu/publications.php#d3-6> (besucht am 15. 05. 2013).
- Times Higher Education (Apr. 2012). »Wealth check: Financial data for UK higher education institutions, 2010–11«. In: *Times Higher Education*. url: http://www.timeshighereducation.co.uk/Journals/THE/THE/12_April_2012/attachments/financial_data.pdf (besucht am 03. 06. 2013).
- Whyte, Angus und Andrew Wilson (2010). *How to Appraise and Select Research Data for Curation. DCC How-to Guides*. url: <http://www.dcc.ac.uk/resources/how-guides/appraise-select-data> (besucht am 06. 06. 2013).
- Williamson, Laurian und Tom Parsons (2013). *UK HEIs RDM service models and skills to support Research Data Management*. url: http://admire.jiscinvolve.org/wp/files/2013/05/RDM_HumanInfrastructure.pdf (besucht am 04. 06. 2013).

- Wilson, James (2013). »University of Oxford Research Data Management Survey 2012. The Results«. In: *DaMaRO-Blog*. url: <http://blogs.it.ox.ac.uk/damaro/2013/01/03/university-of-oxford-research-data-management-survey-2012-the-results/>.
- Winn, Joss (2013). »Open data and the academy: an evaluation of CKAN for research data management«. In: *IASSIST, 28-31 May 2013, Cologne*. url: <http://eprints.lincoln.ac.uk/id/eprint/9778> (besucht am 08. 06. 2013).
- 1 Zur Dokumentation des aktuellen DataCite Metadata Schema v 2.2: <http://dx.doi.org/10.5438/0005>. Die große Vielfalt an disziplinären Standards hat das *Digital Curation Centre (DCC)* dokumentiert: <http://www.dcc.ac.uk/resources/metadata-standards>, zuletzt besucht am 16. Mai 2013.
 - 2 Siehe Cameron Neylon, Building the perfect data repository. . . or the one that might get used, in: Science in the Open. The online home of Cameron Neylon, 27. Oktober 2011, <http://cameronneylon.net/blog/building-the-perfect-data-repository-or-the-one-that-might-get-used/>, zuletzt besucht am 23. Mai 2013.
 - 3 Bei der Befragung von Tenopir et al. (2011) schätzten 41% der Forschenden ihre Beschreibungsmöglichkeiten für Daten negativ oder neutral ein.
 - 4 Siehe Open Access Policy bei FWF-Projekten, http://www.fwf.ac.at/de/public_relations/oai/index.html, zuletzt besucht am 23. Mai 2013.
 - 5 Chatham House Rule, Chatham House, <http://www.chathamhouse.org/about-us/chathamhouserule>, zuletzt besucht am 11. Juni 2013.
 - 6 Im Rahmen des FP7-Programms gilt Folgendes: »The grant agreement [. . .] shall require the submission to the Commission of a plan for the use and dissemination of foreground.«, siehe Regulation (EC) No 1906/2006, <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2006:391:0001:0018:EN:PDF>, zuletzt besucht am 2. Juni 2013.
 - 7 Zu den Förderungshöhen siehe Simon Hodson, JISCMRD Programme 2011–13. Grant Funding Call 07/11, Präsentationsfolien, Juli 2011, <http://www.jisc.ac.uk/media/documents/programmes/mrd/mrdbriefingday/Hodson-JISCMRD%2011-13%20Briefing%20Day.pptx>, zuletzt besucht am 15. Mai 2013, sowie den Call auf der Jisc-Homepage: http://www.jisc.ac.uk/fundingopportunities/funding_calls/2011/06/managingresearchdata.aspx.
 - 8 Gleiches gilt auch für die hier nicht näher zu betrachtende letzte Station bei der Entwicklung des Geschäftsmodells, nämlich der Risikoanalyse.

- 9 Data Asset Framework, Four steps to effective data management, <http://www.data-audit.eu/methodology.html>. Hier ist auch ein DAF-Online-Tool zugänglich, das mit CARDIO eine Erweiterung gefunden hat, siehe Collaborative Assessment of Research Data Infrastructure and Objectives (CARDIO), <http://cardio.dcc.ac.uk/>, beide zuletzt besucht am 3. Juni 2013. Beide Tools enthalten Online-Fragebogen, CARDIO ist jedoch umfangreicher und bietet erweiterte Funktionalitäten wie unterschiedliche NutzerInnenrollen.
- 10 Siehe Wilson 2013. Es wäre interessant, zu erfahren, warum man sich in Oxford, nach einem Fungieren als Pilot-Partner für das DAF, gegen seinen produktiven Einsatz entschieden hat, siehe Martinez 2009.
- 11 JISC data.bris Project, University of Bristol, <https://data.blogs.ilrt.org/jisc-project/>, zuletzt besucht am 3. Juni 2013.
- 12 Als »dritte Schiene« läuft hier in den Fußnoten der Vergleich jener Universitäten mit, deren Jisc-Projekte näher vorgestellt werden. Alle bieten ein breites disziplinäres Spektrum. Die Daten sollen abschätzen helfen, was Universitäten unter welchen Voraussetzungen in die Einführung von RDM-Services investieren. Basisdaten der University of Bristol 2011/12: 15.730 Studierende, davon 2.280 Postgraduierte. 5.720 Mitarbeitende, davon 1.940 Service-Mitarbeitende. Gesamtetat: £ 408,8 Mio. Für die Etats 2010/11 der im folgenden untersuchten Universitäten siehe Times Higher Education 2012, alle anderen Zahlen stammen von den jeweiligen Facts and figures-Seiten der entsprechenden Homepages.
- 13 Research Data Management, University of Bristol, <http://data.bris.ac.uk/>, zuletzt besucht am 3. Juni 2013.
- 14 Basisdaten: Mit 30.760 Studierenden, davon 6.940 Postgraduierten, ist die University of Leeds deutlich größer als die University of Bristol. Der Personalumfang steht aber in diesem Vergleich nicht proportional zu den Studierendenzahlen: 7.140. Gesamtetat: £ 547,3 Mio., also nicht sehr viel mehr als der University of Bristol zur Verfügung steht.
- 15 Tim Banks, Storage & Archive of Research Data, JISC RDM Workshop, 26. März 2013, http://prezi.com/yx-gcmxvzfd9/storage-archive-of-research-data/?auth_key=d1c460eb2b1489e55929755a41652dea75756819, zuletzt besucht am 3. Juni 2013.
- 16 Leeds Research Data Management Pilot. RoaDMAP, University of Leeds, <http://library.leeds.ac.uk/roadmap-project>, zuletzt besucht am 4. Juni 2013.
- 17 ADMIRe, University of Nottingham, <http://admire.jiscinvolve.org/>, zuletzt besucht am 4. Juni 2013.

- 18 Basisdaten: Die Universität betreibt zusätzliche Campus in China und Malaysia. Die hier angegebenen Zahlen betreffen den britischen Campus. Gesamtetat: £ 510,7 Mio. Mit 34.080 Studierenden, davon 6.140 Postgraduierten, und 7.100 Angestellten entspricht die University of Nottingham in Größenordnung und Haushalt der University of Leeds.
- 19 Siehe Williamson und Parsons 2013. Das Papier enthält auch eine Analyse der Servicepersonalplanungen einiger britischer Universitäten, die bereits 2009, meist im Rahmen des ersten Jisc-RDM-Förderprogramms mit der Entwicklung begonnen haben.
- 20 Zum Recherchezeitpunkt war die Trennung der Informationsservices im Gange, zu der vormals alle nachfolgend genannten Abteilungen gehörten, siehe Introduction to IS Separation, University of Nottingham, <http://www.nottingham.ac.uk/is/news/the-issue/isseparation.aspx>, zuletzt besucht am 4. Juni 2013.
- 21 SHERPA/JULIET, University of Nottingham, <http://www.sherpa.ac.uk/juliet/>, zuletzt besucht am 10. Juni 2013.
- 22 JoRD, <http://jordproject.wordpress.com/>, zuletzt besucht am 5. Juni 2013.
- 23 DMPOnline, DCC, <https://dmponline.dcc.ac.uk/>, steht unter der AGPL-Lizenz und kann bei GitHub heruntergeladen werden, <https://github.com/DigitalCurationCentre/DMPOnline>. In den USA wurde parallel und mittlerweile kooperativ das sehr ähnliche DMPTool entwickelt, <https://dmp.cdlib.org>, alle zuletzt besucht am 5. Juni 2013.
- 24 Es ist davon auszugehen, dass die Möglichkeit, in der DCC-Instanz eigene Templates zu erstellen, britischen Institutionen vorbehalten ist.
- 25 Die University of Oxford entwickelt DMPOnline ebenfalls im Rahmen des Jisc-Programms weiter, siehe <http://datamanagementplanning.wordpress.com/>, zuletzt besucht am 6. Juni 2013. Besonderes Augenmerk liegt dabei auf der Vermeidung von doppelten Eingaben der Projektdaten: DMPOnline soll mit dem Forschungsdokumentationssystem und mit einem Repositorium kommunizieren können, genauer soll eine CERIF- und DataCite-konforme Ontologie verwendet, RDF-Metadaten erzeugt und das Ganze mit der Datei für die Verwendung einer SWORD2-Schnittstelle verpackt werden.
- 26 CKAN, The open source data portal software, <http://ckan.org/>, zuletzt besucht am 6. Juni 2013.
- 27 Auch hierzu gibt es eine hilfreiche Anleitung vom DCC, siehe Whyte und Wilson 2010.
- 28 Orbital, <http://orbital.blogs.lincoln.ac.uk/>, zuletzt besucht am 6. Juni 2013.

- 29 Als eher kleine Universität mit gerade einmal 1.320 Postgraduierten bei 11.340 Studierenden und 1.250 Angestellten verfügt die University of Lincoln über einen entsprechend kleinen Gesamtetat von £ 91,3 Mio.
- 30 Researcher Dashboard, Lincoln University, <https://orbital.lincoln.ac.uk/>, zuletzt besucht am 8. Juni 2013.
- 31 Basisdaten: Mit 9.860 Postgraduierten bei 22.180 Studierenden ist der Anteil ersterer enorm. Die Zahl von 11.000 Angestellten spiegelt den Status als Forschungsuniversität. Gesamtetat: £ 919,6 Mio.
- 32 DaMaRO, <http://damaro.oucs.ox.ac.uk/>, zuletzt besucht am 6. Juni 2013.
- 33 DataFlow, <http://www.dataflow.ox.ac.uk/>, zuletzt besucht am 8. Juni 2013.
- 34 Points of Contact – Research Data Management, University Administration and Services (UAS), University of Oxford, <http://www.admin.ox.ac.uk/rdm/help/contact/#Documentation>, zuletzt besucht am 9. Juni 2013.
- 35 RDMRose, <http://rdmrose.group.shef.ac.uk/>, zuletzt besucht am 9. Juni 2013.
- 36 Open Exeter, University of Exeter, <http://as.exeter.ac.uk/library/resources/openaccess/openexeter/>, zuletzt besucht am 9. Juni 2013.
- 37 Basisdaten: Mit 3.452 Postgraduierten bei 18.540 Studierenden und 1.800 Angestellten ist die University of Exeter zwar etwas größer als die University of Bristol, hat aber deutlich weniger Personal und Budget zur Verfügung. Gesamtetat: £ 259,1 Mio.
- 38 Open Exeter: Human Factors in Research Data Management, University of Exeter, <http://blogs.exeter.ac.uk/openexeter/rdm/>, zuletzt besucht am 9. Juni 2013.
- 39 Siehe z. B. Lewis 2010, Rice und Haywood 2011, Starr, Willett et al. 2012, v. a. aber Antell, Foote et al. 2013 und die dort angegebene Literatur.