

Gruppendomäne

JOCHEN KLAR, HARRY ENKE

Forschungsdaten in der Gruppendomäne – Zwischen individuellen Anforderungen und übergreifenden Infrastrukturen

Foto: AIP / Kerstin Mork



Jochen Klar

Foto: AIP / Kerstin Mork



Harry Enke

Gerade die Forschungsdaten, die innerhalb von wissenschaftlicher Zusammenarbeit anfallen, werden mehr und mehr zur Hauptressource der wissenschaftlichen Arbeit. Für diese Daten der sog. Gruppendomäne stehen jedoch nur unzureichende Infrastrukturen bereit. Angeregt durch die Diskussion der letzten Jahre haben die Förderer, allen voran die Deutsche Forschungsgemeinschaft (DFG), verschiedene Förderprogramme ins Leben gerufen, um Forschungsprojekten ein adäquates Forschungsdatenmanagement zu ermöglichen. Es ist an der Zeit zu hinterfragen, inwieweit diese Maßnahmen Früchte tragen und wie sie weiterentwickelt werden können. In diesem Artikel wird insbesondere die Rolle der INF-Projekte in den Sonderforschungsbereichen (SFB) und Transregio der DFG, der DFG-Ausschreibung »Infrastrukturen für Forschungsdaten« vom April 2010, und von Virtuellen Forschungsumgebungen kritisch diskutiert.

Research data which have been collected in the course of scientific collaboration are becoming with increasing frequency the main resource of scientific research. For data of the so-called group domain, however, the requisite infrastructure is still inadequately developed. As a result of discussions in recent years, sponsoring organizations – first and foremost the German Research Foundation (DFG) – have initiated various programmes in order to provide research projects with adequate data management systems. It is now time to evaluate whether these measures have been fruitful and how they can be further developed. This article gives a critical analysis of the role of so-called INF projects, which aim to develop information infrastructure within the DFG's Transregional and Collaborative Research Centers (SFB), the DFG's funding programme »Information Infrastructures for Research Data« which was announced in April 2010, and the Virtual Research Environments.

EINLEITUNG

Schon seit einiger Zeit ist die Menge an Forschungsdaten, die in bestimmten *datenintensiven* Disziplinen erzeugt werden, so groß, dass ohne ein dezidiertes Datenmanagement, welches auf einer soliden organisatorischen und infrastrukturellen Grundlage steht, keine sinnvolle wissenschaftliche Tätigkeit mehr vorstellbar ist. Groß angelegte, internationale Experiment- oder Beobachtungskampagnen wenden meist signifikante Ressourcen für den professionellen Umgang mit Forschungsdaten auf. Zeitgleich mit dem eigentlichen Forschungsaufbau, sei es ein Experiment, ein Satellit oder ein Observatorium, entstehen IT-Kapazitäten an Rechenzentren und die damit verbundenen Organisationsstrukturen. Musterbeispiel ist hier der Large Hadron Collider (LHC) am CERN in Genf mit seinem Worldwide LHC Computing Grid (WLCG), in dem über eine speziell an die Bedürfnisse der Experimente am LHC ausgerichtete Hierarchie die Daten an mehr als 150 Rechenzentren zur Weiterverarbei-

tung verteilt werden. In der Radioastronomie findet sich mit dem Low Frequency Array (LOFAR) ein weiteres Beispiel für ein solches Netzwerk. Hier findet der Datentransport jedoch umgekehrt statt. Für LOFAR werden über 7.000 Antennen in derzeit 45 Stationen in ganz Europa verteilt und die von ihnen aufgezeichneten Daten über ein Netzwerk an einen Supercomputer in den Niederlanden transportiert, wo sie miteinander korreliert werden und so zu einem einzelnen virtuellen Teleskop werden. Hierdurch können dann radioastronomische Messungen mit bisher unerreichter Qualität von weit entfernten Galaxien, veränderlichen Sternen oder von Prozessen in unserer Sonne vorgenommen werden.

Die Quelldaten, die von Projekten dieser Größenordnung erzeugt werden, weisen in der Regel eine homogene Form auf, und die verwendeten Formate und Standards werden im Vorhinein festgelegt und adäquat dokumentiert. Dies dient zuallererst einer möglichst unproblematischen Nutzung der Daten durch die am Projekt direkt beteiligten Wissenschaftlerinnen und Wissenschaftler. Im Fall der vier LHC-Projekte sind dies über 10.000 Personen. Auch eine Nutzung der gewonnenen Daten durch nachfolgende Forschungsgenerationen wird erst möglich, wenn diese in angemessenem Rahmen archiviert werden. Hierfür ist jedoch ein professioneller Umgang mit den Daten in jedem Arbeitsschritt erforderlich, um eine Nachvollziehbarkeit auf wissenschaftlichem Niveau zu gewährleisten.

Im Allgemeinen lässt sich also feststellen, dass das Datenmanagement in diesen Großprojekten, zumindest was die erzeugten Primärdaten angeht, in der Regel auf einem adäquaten Niveau stattfindet. Für die Daten, die dann im Laufe der Analyseprozesse der Wissenschaftlerinnen und Wissenschaftler erzeugt werden, oft als Sekundärdaten bezeichnet, gilt dies jedoch nur eingeschränkt.¹ In jüngerer Zeit hat jedoch auch die Menge dieser Daten ein Ausmaß erreicht, bei dem es – auch aufgrund der Schwierigkeiten, solche Daten innerhalb der Zusammenarbeit auszutauschen – nicht mehr zweckmäßig ist, sie nur in der Verantwortung der einzelnen Forscherin bzw. des einzelnen Forschers zu belassen, sondern sie in ein dezidiertes Datenmanagement zu überführen. Die Heterogenität dieser Daten, was Formate und Standards angeht (ausgehend

Daten, die im Laufe der
Analyseprozesse erzeugt
werden

von der recht individuellen Arbeitsweise der Forscherinnen und Forscher), und die mangelnde Bereitschaft, signifikante Ressourcen für die Organisation dieser Daten bereit zu stellen, stellen das Datenmanagement in diesem Bereich noch vor signifikante Herausforderungen.

Das Thema Datenmanagement geht aber inzwischen weit über diese Großprojekte hinaus und betrifft immer mehr Wissenschaftlerinnen und Wissenschaftler. Mehr und mehr kleinere Gemeinschaftsprojekte erzeugen große Datenmengen, verfügen jedoch nicht über die Mittel, sich im Vorhinein so stark mit dem Datenmanagement zu beschäftigen, wie es bei den Großprojekten der Fall ist. Auch Projekte in vermeintlich weniger datenintensiven Wissenschaften haben inzwischen einen Output an Daten, der nicht mehr ohne Datenmanagement zu organisieren ist. Beispiele hierfür sind große Surveys in den Sozialwissenschaften oder Digitalisierungsprojekte in den Geisteswissenschaften. Nicht zuletzt stellen die diversen Policies seitens der Geldgeber wie beispielsweise *Die Grundsätze zur Sicherung guter wissenschaftlicher Praxis* der Deutschen Forschungsgemeinschaft (DFG) Ansprüche an den Umgang mit Forschungsdaten, die über den bisherigen Arbeitsalltag vieler Wissenschaftlerinnen und Wissenschaftler deutlich hinausgehen.

Seit ca. zehn Jahren wird über den Umgang mit digitalen Forschungsdaten, im Allgemeinen und auch speziell für die Daten aus kollaborativen Projekten, auf internationaler wie auch auf nationaler Ebene intensiv diskutiert. Hierbei stehen der Aufbau der dafür notwendigen Forschungsdateninfrastruktur, Standards, Langzeitarchivierung, Datenpublikation und diverse andere Aspekte auf der Agenda. Verschiedenste Konzepte sind dabei vorgeschlagen worden und eine Vielzahl von Projekten wurde und wird gefördert. In der Diskussion lassen sich verschiedene Akteure identifizieren, die sich aus unterschiedlichen Gründen und mit unterschiedlichen Hintergründen und Traditionen dem Thema annehmen:

- Zum einen sind es die beschriebenen Communities² selber, die an neuen Konzepten und Infrastrukturen zur Bewältigung der oben beschriebenen Datenmengen arbeiten. Gerade im Umfeld von größeren Kooperationen finden sich Projekte, die nicht von den eher abstrakten Idealen von Datenarchivierung und Wiederverwendung getrieben werden, sondern aus der konkreten Notwendigkeit eines Datenmanagements entstehen, ohne welches das Erreichen der Forschungsziele nicht möglich wäre.
- Dadurch, dass besonders Computersimulationen zu der Vermehrung der Menge an Forschungsda-

ten beigetragen haben, sind auch die akademischen Rechenzentren in der Diskussion stark vertreten. Ihre primären Aufgaben haben diese Institutionen zum einen im Hochleistungsrechnen und, insbesondere die universitären Rechenzentren, in der Bereitstellung von Infrastruktur. Durch ihre Erfahrung mit digitalen Daten und dadurch, dass immense digitale Speicherressourcen an den Rechenzentren vorhanden sind, bieten sie sich auch als Ressourcenprovider für Projekte zur Forschungsdateninfrastruktur an.

- Auch die klassischen Gedächtnisorganisationen wie Bibliotheken und Archive sind am Thema Forschungsdatenmanagement interessiert. Aufbauend auf eine jahrhundertealte Erfahrung im Aufbewahren und Kuratieren von analogen Forschungsdaten, versuchen sie dieses Aufgabenfeld für die Anforderungen des digitalen Zeitalters zu erweitern. Es verwundert daher nicht, dass die Gedächtnisorganisationen besonders im Bereich der Langzeitarchivierung, bei der Publikation von Forschungsdaten, ihrer Referenzier- und Zitierbarkeit und bei den Daten-Journals aktiv sind. Es besteht auch eine gewisse Überlappung zwischen dem freien Zugang zu Forschungsdaten und der Open-Access-Thematik mit all ihrer Komplexität.
- Die Förderinstitutionen haben den prinzipiellen Bedarf an einer modernen Forschungsdateninfrastruktur erkannt. Innerhalb der letzten Jahre wurde eine Reihe von Förderinstrumenten geschaffen, die den Aufbau einer zukunftsfähigen Forschungsdateninfrastruktur fördern soll. So dient beispielsweise das Programmelement »Informationsmanagement und Informationsinfrastruktur« (INF) in den Sonderforschungsbereichen (SFB) und Transregio der Deutschen Forschungsgemeinschaft (DFG) dazu, den Umgang mit den im SFB/Transregio anfallenden Daten zu unterstützen, erhaltenswerte Daten zu archivieren und nachfolgenden Forscherinnen und Forschern zugänglich zu machen. Auch die Ausschreibung »Infrastrukturen für Forschungsdaten« vom April 2010 des Bereiches Wissenschaftliche Literaturversorgung und Informationssysteme (LIS) der DFG förderte Projekte zum Forschungsdatenmanagement. Auf internationaler Ebene sind in diesem Kontext unter anderem die Aktivitäten von JISC in Großbritannien und der SURF-Foundation in den Niederlanden zu erwähnen.
- In den deutschen Wissenschaftsorganisationen wird das Thema intensiv diskutiert. In den letzten Jahren hat sich eine Vielzahl von Kommissionen und Arbeitskreisen gebildet, die auf verschie-

Projekte in vermeintlich weniger datenintensiven Wissenschaften

verschiedene Akteure

denen Ebenen eine Reihe von Konzepten vorgeschlagen haben. Beispiele hierfür sind die Schwerpunktinitiative *Digitale Information* der Allianz der deutschen Wissenschaftsorganisationen oder die Handlungsempfehlungen der Kommission Zukunft der Informationsinfrastruktur (KII) innerhalb der Leibniz-Gemeinschaft.

Trotz all dieser Bemühungen lässt sich zum gegenwärtigen Zeitpunkt, im Jahr 2013, jedoch noch nicht feststellen, dass sich eine Art von Konsens gebildet hätte. Insbesondere im Bereich der eingangs beschriebenen Kooperationen, seien es Großprojekte mit hunderten von Mitarbeiterinnen und Mitarbeitern oder kleinere Projekte, sind noch viele Fragen offen. Das von der DFG geförderte Projekt *Rahmenbedingungen einer disziplinübergreifenden Forschungsdateninfrastruktur (Radieschen)*, an dem auch die Autoren dieses Artikels beteiligt waren, untersuchte in den Jahren 2011 bis 2013 die deutsche Forschungsdatenlandschaft, um aus einer gründlichen Bestandsaufnahme Handlungsempfehlungen für eine disziplinübergreifende Weiterentwicklung der Forschungsdateninfrastruktur zu entwickeln. Methodisch wurden hierzu diverse Projekte, insbesondere aus den schon angesprochenen Programmen der DFG mithilfe von Experten-Interviews und Literaturrecherchen untersucht und bewertet. Im Folgenden werden nun einige der Ergebnisse von *Radieschen* vorgestellt. Der Schwerpunkt liegt hierbei auf den Erkenntnissen, die sich insbesondere für die Forschungsdateninfrastrukturen in wissenschaftlichen Gemeinschaftsprojekten ergaben.³

EIN ERWEITERTES DOMÄNENMODELL FÜR FORSCHUNGSDATEN

Bevor jedoch die bestehenden Fördermaßnahmen kritisch diskutiert werden können, ist es zweckmäßig, eine Struktur einzuführen, anhand derer die geförderten Projekte klassifiziert werden können. Im *Radieschen*-Projekt wurde hierzu ein spezielles Domänenmodell verwendet. Dieses Modell basiert seinerseits auf dem Domänenmodell von Treloar und Harboe-Ree⁴, das

aus dem Vergleich von thematischen Kontinua (die anhand von Aspekten wie der Nutzung von Metadaten oder dem Aufwand, der für Datenmanagement betrieben wird, ermittelt wurden) drei distinkte Domänen identifiziert hat. Während in der *Privaten Domäne* die Forscherin oder der Forscher noch vollständig selbst für ihre/seine Daten zuständig ist, arbeiten in der *Gruppendomäne* mehrere Wissenschaftlerinnen und Wissenschaftler an einem gemeinsamen Satz von Forschungsdaten. In der Regel ist dann auch ein dediziertes Datenmanagement, mit den dazugehörigen personellen Ressourcen, nötig. In der öffentlichen Domäne befinden sich schließlich die Daten, die (meist nach abgeschlossener Forschung) für die Öffentlichkeit und insbesondere andere Forscher zugänglich sind. Das Modell wird von Treloar und Harboe-Ree insbesondere dafür genutzt, Datenworkflows innerhalb einer Institution zu klassifizieren und zu modellieren. Die Domänen werden hierbei als Stationen eines Life-Cycles gesehen, den die Daten zeitlich nacheinander durchlaufen.

Für eine Charakterisierung der Zusammenhänge im Kontext einer ganzen Forschungsdisziplin oder gar bezüglich disziplinübergreifender Aspekte ist diese Sicht jedoch unzureichend. Anstatt eines temporalen Zusammenhangs der Domänen ist es vielmehr nötig, die einzelnen Domänen als nebeneinander stehende Bereiche eines komplexen Forschungsdatenworkflows zu sehen. Daten können ihn zwar von privater über Gruppen- bis hin zur öffentlichen Domäne durchlaufen, genauso ist es aber vorstellbar, dass Daten in der Gruppendomäne erzeugt werden und dann durch Verarbeitungsschritte in die private Domäne wandern und von dort, nach einiger Zeit, veröffentlicht werden. Um sowohl nicht-öffentliche Archivierung als auch eine Veröffentlichung unter bestimmten Bedingungen bzw. nur für eine eingegrenzte Zielgruppe zu berücksichtigen, wird die oben beschriebene öffentliche Domäne weiter in eine *dauerhafte Domäne* und eine *Zugangsdomäne* unterteilt. Es ist sinnvoll, die jeweiligen Übergänge zwischen den Domänen als die kritischen Stellen für spezifische infrastrukturelle Anforderun-

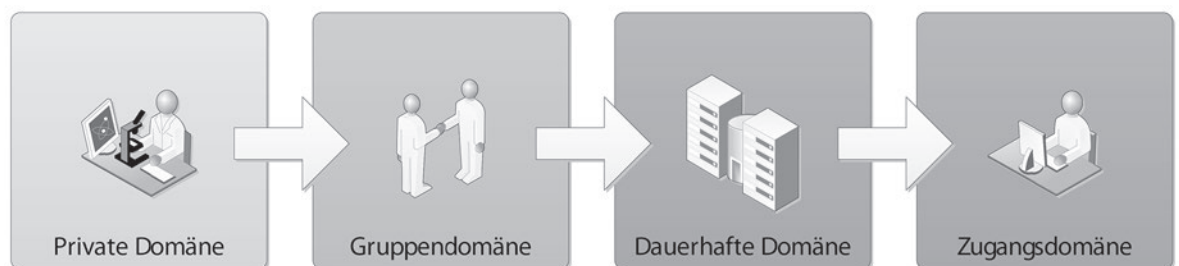


Abb. 1: Erweitertes Domänenmodell

gen wahrzunehmen, an denen die Überlegungen für die Entwicklung der Forschungsdateninfrastruktur ansetzen sollten.

Im Kontext der laufenden Debatte um die adäquate Förderung der Forschungsdateninfrastruktur in Deutschland ist es nun interessant, sich die Verteilung der Projekte, die in diesem Bereich gefördert wurden und werden, anzusehen. Im *Radieschen*-Projekt wurde dies für die Projekte der Ausschreibung »Infrastrukturen für Forschungsdaten« durchgeführt. Dafür wurden innerhalb einer Bestandsaufnahme der verschiedenen Aktivitäten im Bereich Forschungsdaten die besagten Projekte in den vier Domänen verortet. Das Resultat ist in Abbildung 2 dargestellt.

Es zeigt sich, dass in diesem Förderungsprogramm verstärkt Projekte gefördert wurden, die ihren Schwerpunkt in den Bereich der Langzeitarchivierung und der Veröffentlichung von Forschungsdaten gelegt haben. Beispiele hierfür sind die Erarbeitung von Archivkonzepten, Entwicklung der entsprechenden Softwarelösungen oder die Bereitstellung von Webportalen. Im Bereich der Infrastrukturen für die Gruppendomäne, für das kooperative Datenmanagement, sind deutlich weniger Projekte präsent. Für Infrastrukturen in der privaten Domäne werden noch kaum Projekte gefördert (bzw. beantragt).

VON DER PRIVATEN ZUR GRUPPENDOMÄNE

Wie eingangs beschrieben, sind es aber besonders die wissenschaftlichen Kooperationen der Gruppendomäne und, bezüglich der inhomogenen Sekundärdaten, der Übergang zwischen Gruppen- und privater Domäne, welche das Datenmanagement vor besondere Herausforderungen stellen. Der wichtigste (und einzige) Förderansatz besteht hier in den bereits angesprochenen INF-Projekten in den SFB/Transregio der DFG. Seit 2007 kann im Rahmen eines SFB/Transregio ein solches Projekt mitbeantragt werden. Das Projekt soll dazu genutzt werden, eine Dateninfrastruktur bzw. eine virtuelle Forschungsumgebung für den SFB/Transregio aufzubauen und insbesondere das »systematische und auf langfristige Nutzung angelegte Management der im Kontext des Sonderforschungsbereichs erhobenen relevanten Daten«⁵ zu gewährleisten. Seit 2009 werden die ersten INF-Projekte gefördert. Im Januar 2013 waren es bereits 27. Insgesamt förderte die DFG zum gleichen Zeitpunkt 232 SFB/Transregio. Um einen Austausch zwischen den verschiedenen INF-Projekten zu fördern, veranstaltete das Projekt *Radieschen* am 11. April 2013 an der SUB Göttingen einen »Gemeinsamen Workshop der SFB-INF-Projekte«⁶. Hierbei standen die organisatorischen und die technischen Schwierigkei-

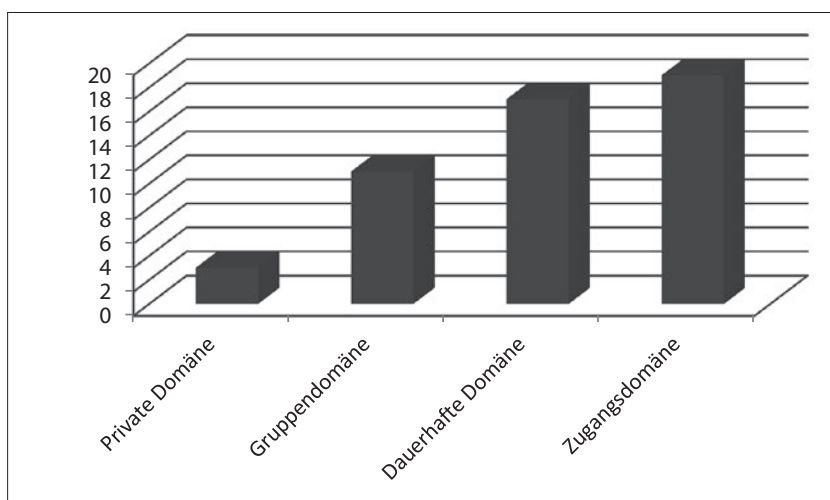


Abb. 2: Verteilung der Projekte des DFG-Calls »Infrastrukturen für Forschungsdaten« über die vier Domänen des Forschungsdatenmanagements. Die Gesamtzahl der betrachteten Projekte beträgt 31. Dadurch, dass verschiedene Projekte in mehr als einer Domäne verortet sind, ist die Addition der Anzahl in den einzelnen Domänen höher als diese Zahl

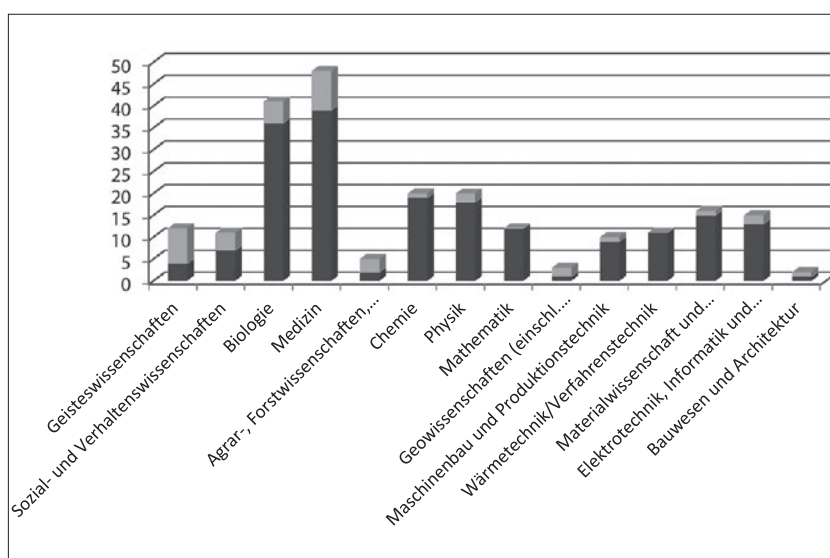


Abb. 3: Verteilung der Sonderforschungsbereiche mit INF-Projekt (hellgrau) und ohne (dunkelgrau) über die verschiedenen Fachgebiete. Aufgrund der Vergleichbarkeit werden hierbei nur die SFB bzw. TRR betrachtet, deren Förderung im Jahr 2009 oder später begonnen hat

ten in diesen Projekten im Vordergrund. Die Veranstaltung zeigte einerseits, wie wichtig die INF-Projekte für die Forschungsdateninfrastruktur in Deutschland sind, und andererseits, wie groß der Bedarf an Diskussion zu diesem Programm ist.

Im *Radieschen*-Projekt wurde auch untersucht, wie die INF-Projekte über die verschiedenen Fachgebiete verteilt sind. Als Schema wurde hierbei die DFG-eigene Fachsystematik verwendet.⁷ Die ermittelte Verteilung ist in Abbildung 3 dargestellt. Es ist ersichtlich, dass in den Geistes- und Sozialwissenschaften sowie in den Lebenswissenschaften bereits ein signifikanter Anteil der SFB/Transregio über ein INF-Projekt verfügt. Die Naturwissenschaften (mit Ausnahme der Geowissen-

INF-Projekte in den SFB/Transregio der DFG

schaften) und die Ingenieurwissenschaften sind, was dieses Förderungsinstrument angeht, schwächer vertreten.

Diese Verteilung resultiert nicht unbedingt aus einem Ungleichgewicht bei der Vergabe der Förderung, sondern kann auch darauf zurückzuführen sein, dass die Wissenschaftlerinnen und Wissenschaftler in diesen Disziplinen das spezielle Förderungsinstrument INF weniger in Anspruch nehmen. Interessanterweise existieren gerade in den Forschungsdisziplinen, die im Allgemeinen nicht mit großen Datenmengen (Stichwort: Big Data) assoziiert werden, mehr INF-Projekte. Dies kann zum einen als Zeichen gewertet werden, dass auch in vermeintlich weniger computeraffine Disziplinen ein digitaler Umgang mit Forschungsdaten Einzug hält. Zum anderen waren in diesen Disziplinen auch vor der Einführung von modernen IT-Systemen der Aufbau und die Nutzung von Archiven mit Forschungsdaten fester Bestandteil des wissenschaftlichen Arbeitsprozesses. Bei den Geisteswissenschaften zeigt sich dies in einer traditionellen Nähe zu Bibliotheken, die schon immer sowohl Lieferant als auch Abnehmer der Forschungsobjekte, in der Regel Literatur, waren. Sowohl in den Sozial- und Wirtschaftswissenschaften als auch in bestimmten Lebenswissenschaften und den Geo- und Klimawissenschaften haben Forschungsdaten einen dezidiert episodischen Charakter und lassen sich in der Regel nicht beliebig reproduzieren. Beispiele hierfür reichen von statistischen Erfassungen der Bevölkerung eines Landes bis zu Bohrkernen, die zur Bestimmung der Umweltbedingungen zu bestimmten erdgeschichtlichen Zeiten entnommen werden. In der Vergangenheit sind daher große Sammlungen und Archive entstanden, um diese Daten und Objekte Forscherinnen und Forschern zur Verfügung zu stellen, die nicht direkt an der Produktion der Daten beteiligt waren.

Im Gegensatz dazu sind die übrigen Naturwissenschaften und die Ingenieurwissenschaften in der Vergangenheit stark durch die technologische Entwicklung geprägt gewesen. Als Folge werden ältere Forschungsdaten mit dem Aufkommen neuerer Technologien, Messmethoden und -geräte schnell obsolet, weil verbesserte experimentelle Daten vorliegen. Erst in neuerer Zeit, mit IT-basierten Analysemethoden und Data-Mining, werden auch ältere Datensammlungen wieder interessant. Sofern der mit der Digitalisierung verbundene Aufwand vertreten werden kann, gilt dies auch für analog vorliegende Forschungsdaten, wie beispielsweise Fotoplatten aus der Astronomie.⁸ Die Möglichkeit, computergestützt neue Schlussfolgerungen aus alten Datenbeständen zu gewinnen, hat in jüngerer Zeit zur Entwicklung von Datenarchi-

ven geführt, die jedoch weniger an den klassischen Gedächtnisorganisationen angesiedelt sind, sondern, auch aufgrund der benötigten IT-Ressourcen, eher die Nähe zu Rechenzentren suchen oder zum Aufbau von entsprechenden Kapazitäten an Forschungseinrichtungen führen.

WEITERENTWICKLUNG DER INF-PROJEKTE

Die INF-Projekte der DFG stellen zurzeit einen vielversprechenden Ansatz dar, die oben diskutierten Defizite in der Forschungsdateninfrastruktur im Übergang von privater Domäne zur Gruppendomäne und in der Gruppendomäne selbst abzubauen. Fast fünf Jahre nach Einführung stellen sich nun die Fragen, welche Erfahrungen aus den bisherigen INF-Projekten gewonnen werden können und wie sich das Programmelement weiterentwickeln könnte.

Zunächst ist festzuhalten, dass jeder SFB/Transregio über ein solches Infrastrukturprojekt verfügen sollte. Nur so ist es möglich, die in den einzelnen Projekten erzeugten Forschungsdaten fachgerecht zu archivieren und, natürlich unter Berücksichtigung von rechtlichen Einschränkungen und anderen sachlichen Hemmnissen, zukünftigen Forschergenerationen zugänglich zu erhalten. Weiter oben ist dargestellt, dass nur ca. 10 % der SFB/Transregio-Projekte das Angebot, das die INF-Projekte darstellen, wahrnehmen. Sicher mag es auch Bereiche geben, in denen so wenige Forschungsdaten anfallen, dass ein solches Projekt nicht zweckmäßig erscheint. Dies wird jedoch nicht für die Mehrzahl der SFB/Transregio gelten. Es ist daher in jedem Fall seitens des Förderers darauf zu bestehen, dass innerhalb des Beantragungsprozesses belastbare Datenmanagementpläne formuliert und dann innerhalb eines INF-Projektes umgesetzt werden.

Bei der Bewilligung der INF-Projekte, aber auch von anderen Infrastruktur-Projekten, sollte nicht entscheidend sein, ob die in dem Projekt verwendeten Infrastrukturkomponenten besonders neuartig oder einzigartig sind (Kriterien, die für die Bewertung eines Forschungsvorhabens entscheidend sind), vielmehr sollte darauf Wert gelegt werden, dass die in der jeweiligen Community erarbeiteten Standards und Formate verwendet werden und das INF-Projekt optimal in einen bereits existierenden infrastrukturellen Kontext eingebunden ist.

Die DFG erwartet von den Universitäten, deren Verwaltungen der SFB/Transregio-Beantragung zustimmen müssen, dass diese nach Ende des SFB/Transregio die entstandenen Forschungsdaten in geeignete Strukturen übernehmen. Dieser Anforderung stehen jedoch meist nur unzureichende Infrastrukturen auf

Seiten der Universität entgegen. Um diese Lücke zu schließen und entsprechende Kapazitäten an den der Universität zugehörigen Bibliotheken bzw. an den IT-Abteilungen oder universitären Rechenzentren aufzubauen, existieren verschiedene Ansätze. Teilweise wird sich hierbei darauf gestützt, dass die Mittel aus INF-Projekten gleichzeitig dem Aufbau solcher universitätsweiten Strukturen dienen sollen. Ein solcher Ansatz wird aber im Ergebnis keine Verbesserung bringen, da die Mittel des INF-Projektes für ein solches Unterfangen zu gering sind. Dadurch, dass der Fokus von der zielgerichteten Unterstützung der Forscherinnen und Forscher bei der Überführung von Forschungsdaten von der privaten Domäne in die Gruppendomäne verloren geht, wird allerdings die eigentliche Stärke des Programmelementes der INF-Förderung verwässert.

Das Nichtvorhandensein von adäquaten Forschungsdateninfrastrukturen an den Universitäten, in den Fachgemeinschaften oder allgemeiner in der Gruppendomäne führt noch zu weiteren Phänomenen. INF-Projekte eignen sich kaum dafür, von dem recht eingeschränkten Fokus eines Projektes zu einer Einrichtung für eine ganze Forschungsdisziplin oder gar mehrere erweitert zu werden. Es ist jedoch bei einer Reihe von INF-Projekten der Anspruch zu erkennen, insbesondere wenn Überlegungen zur Nachhaltigkeit des Projektes formuliert werden, genau dies zu leisten. Es ist sicherlich so, dass von den INF-Projekten ein gewisser normativer Einfluss auf die Community ausgeht, da die SFB/Transregio eine substantielle Konzentration von Forschern zu einem bestimmten Thema darstellen. Im Zuge der Projekte können Werkzeuge entstehen, die auch außerhalb SFB/Transregio ein valides Einsatzszenario haben. Nichtsdestotrotz entsteht in der Regel jedoch keine übergreifende Infrastruktur, die die vorhandenen Lücken im Forschungsmanagement in der jeweiligen Community adäquat zu schließen vermag.

Aus diesem Grund, und gerade jetzt, da die ersten SFB/Transregio mit INF-Projekt ihrem Ende zugehen, ist es notwendig, sich mit der Frage nach dem Weiterbestehen der INF-Projekte über das Projektende hinaus zu beschäftigen. Da im Idealfall die Infrastruktur des INF-Projektes auf die Unterstützung des Forschungsprozesses im SFB/Transregio optimiert ist, ist diese Frage in der Regel zu verneinen. Dies bedeutet jedoch nicht, dass alle Arbeitsergebnisse des Projektes wertlos sind. Es ist im Falle einer Abwicklung in jedem Fall darauf zu achten, dass die Datenbestände des Projektes sauber in Archive einer bereits vorhandenen Institution überführt werden. Dies ist nicht nur aus guter wissenschaftlicher Praxis geboten, sondern

ist auch der einzige Weg, um eine Nachnutzung zu ermöglichen. Entwickelte Software sollte möglichst im Quelltext veröffentlicht werden und so potentiell weiter benutzt- und wartbar sein. Unter Umständen können geschaffene organisatorische Strukturen mit anderen, bereits etablierten Institutionen verschmelzen. Die Frage, welche Strukturen bereit stehen, um die Arbeitsergebnisse aufzunehmen, ist hierbei noch zu klären.

FORSCHUNGSDATEN IN DEN COMMUNITIES

Die überwiegende Anzahl von SFB/Transregio-Projekten ist lokal, also an einen bestimmten Standort gebunden. Daraus resultiert, dass die mit ihnen verbundenen INF-Projekte kein geeignetes Instrument für die Unterstützung von nationalen oder internationalen Vorhaben sind, die nicht um ein lokales Zentrum herum entstehen. Eine komplementäre Förderungsmöglichkeit – wiederum von Seiten der DFG – war hier die schon diskutierte Ausschreibung »Infrastrukturen für Forschungsdaten«. Wie bereits gezeigt, sind in dieser Ausschreibung jedoch Projekte, die direkt Probleme der Forschungsdateninfrastruktur in der Gruppendomäne adressieren, unterrepräsentiert.

Die Gruppendomäne wird in vieler Hinsicht stark durch institutionelle Strukturen geprägt. Diese sind jedoch sehr Community-spezifisch und diffizil. Einerseits sind, gerade in den Sozialwissenschaften, einzelne Institute damit beschäftigt, Daten systematisch aufzunehmen, aufzubereiten und dann externen Wissenschaftlerinnen und Wissenschaftlern zur Verfügung zu stellen. Wie schon einleitend erwähnt, stehen diese Forscherinnen und Forscher dann vor dem Problem, keine Infrastruktur für die Organisation ihrer eigenen, von den Daten des Providers abgeleiteten Daten zu haben. Andererseits existieren aber auch, beispielsweise in der Astrophysik, Kooperationen, die nur lose an mehrere Institutionen gebunden sind, jedoch trotzdem große Mengen an Daten verarbeiten bzw. stark datenbasiert arbeiten. Diese Gruppen, die oft einen eher informellen Start haben und grenzüberschreitend arbeiten, sind mit dem Problem konfrontiert, keine organisatorische Struktur und kein Zentrum zu besitzen, um eine Infrastruktur aufzubauen, die in die bestehenden Förderungskonzepte passt. Diese Komplexität beeinflusst den Umgang mit Forschungsdaten in der Gruppendomäne auf verschiedenste Weise. Die unzureichende Unterstützung der IT-basierten kollaborativen Arbeitsweise in der Gruppendomäne war so auch einer der Hauptantriebe für die Entstehung von Virtuellen Forschungsumgebungen.

Nachnutzung

Einrichtung für eine ganze Forschungsdisziplin

Weiterbestehen über das Projektende hinaus

Es ist festzustellen, dass die Ergebnisse der diversen Projekte zum Datenmanagement in den einzelnen Communities oft noch nicht adäquat angenommen werden. Inzwischen existieren in den meisten Disziplinen vielfältige Vorarbeiten zu Themen wie Formate, Standards, Metadaten, Identifikatoren etc. Die Nutzung dieser Arbeiten in neuen Projekten ist jedoch noch verbesserungswürdig. Seitens der Förderer sollte daher noch stärker darauf geachtet werden, dass die in den Communities erarbeiteten Standards auch verwendet werden. In diesem Zusammenhang ist es auch wünschenswert, dass andere Projekte die Kompetenz der verschiedenen Gremien der Communities und der Wissenschaftsorganisationen nutzen. Hier ist eine stärkere Kommunikation und Community-Arbeit seitens dieser Einrichtungen notwendig.

In der Vergangenheit hat es sich gezeigt, dass sich Infrastrukturprojekte häufig zu Werkzeugen mit vielen Features entwickeln, die in der Folge jedoch nur unzureichend durch die Forscherinnen und Forscher verwendet werden. In der Regel scheiterte dies an einer mangelnden Integration der neuen Werkzeuge in die bereits existierenden wissenschaftlichen Arbeitsabläufe. Von Vorteil wäre es daher, wenn es den einzelnen (Sub-)Communities gelänge, aufgrund von Analysen der wissenschaftlichen Arbeit ihre grundlegenden Workflows herauszuarbeiten, sie zu dokumentieren und innerhalb der Community zu diskutieren. Durch die Konzentration auf diese Workflows wäre es dann möglich, neue Werkzeuge zu schaffen (oder alte weiterzuentwickeln), die in den Communities auch angenommen werden.

Im Bereich der Metadaten stellt sich inzwischen heraus, dass sich nur Systeme durchsetzen, die einerseits von den fachlichen Notwendigkeiten des jeweiligen Fachgebiets getragen werden und andererseits nicht durch ein zu spezielles Vokabular überkomplex werden. Bei der Neu- und Weiterentwicklung von Metadaten sollte daher der Schwerpunkt nur auf der Auffindbarkeit der Daten und der Nachvollziehbarkeit der Entstehung (oder *Provenance*) liegen, und nicht auf einer möglichst kompletten Abbildung aller zur Nachnutzung notwendigen Kontextinformationen. Auch immer komplexere Metadaten-Schemata, die mit entsprechendem Aufwand erzeugt und gepflegt werden müssen, können nicht das fachliche Kontextwissen, das zur Nachnutzung der Daten nötig ist und durch persönliche Kontakte oder Literaturstudien erworben werden kann, ersetzen. Auch die Innovationsrate des Forschungsprozesses spiegelt sich in den Datensätzen und lässt die tiefe semantische Erschließung durch Metadaten, die notwendigerweise post festum erfolgt, zu rasch veralten.

DIE ROLLE VON VIRTUELLEN FORSCHUNGsumGEBUNGEN

Weiter oben ist bereits bemerkt worden, dass der Hauptantrieb für die Entstehung von Virtuellen Forschungsumgebungen (VRE = Virtual Research Environment) in der unzureichenden Unterstützung von IT-basierten, kollaborativen Arbeitsumgebungen durch die vorhandenen IT-Infrastrukturen zu suchen ist. Unter der Bezeichnung VRE sind sehr verschiedene Projekte durchgeführt worden. Manche Initiativen haben sich auch erst im Rückblick als solche herausgestellt. Ohne auf die Diskussion einzugehen, ob ein solches »Label VRE« für das jeweilige Projekt sinnvoll ist oder nicht, kann sehr wohl die Frage gestellt werden, welches der Projekte über die Förderung hinaus eine funktionierende Infrastruktur hinterlassen hat oder zumindest Werkzeuge geschaffen hat, die immer noch genutzt werden. Wie auf einer gemeinsamen Veranstaltung der Arbeitsgemeinschaft VRE der Allianz der deutschen Wissenschaftsorganisationen und der Arbeitsgemeinschaft VRE der Deutschen Initiative für Netzwerkinformation e.V. (DINI) im Frühjahr 2013 festgestellt werden musste, sieht diese Nachhaltigkeitsbilanz nicht sehr erfreulich aus. Nur solche Projekte, die in einen schon bestehenden Kontext eingebettet waren bzw. die einen bestimmten, existierenden Workflow effizient implementieren konnten und so eine Nutzerbasis gewonnen haben, haben eine realistische Chance zu überdauern. Aus den bisherigen Erfahrungen lassen sich auch weitere Kriterien für den Erfolg oder den Misserfolg einer VRE ableiten. So hat es sich erwiesen, dass das Bestreben, sehr viele Komponenten bzw. Workflows in eine VRE zu integrieren, eher hinderlich für ihren Erfolg ist. Ein Grund hierfür ist die aus ihrer Komplexität resultierende steile Lernkurve für die Nutzung der VRE.⁹ Für Projekte in Deutschland ist außerdem wichtig, dass bei der Übertragung von Konzepten aus den anglo-amerikanischen Ländern die doch sehr verschiedenen Strukturen der deutschen Forschungslandschaft besser beachtet werden.

Die unter der Bezeichnung VRE erfolgten Arbeiten sind primär aufgrund von Lücken der vorhandenen Infrastruktur an den Universitäten und den anderen Forschungseinrichtungen durchgeführt worden. Diese Lücken sind jedoch nach wie vor vorhanden. Auch wenn das Instrument VRE, wie beschrieben, Schwächen aufweist und überarbeitet werden muss, kann es immer noch einen wichtigen Baustein bei der Umstrukturierung der derzeitigen Forschungsdateninfrastruktur darstellen. Insbesondere für die Unterstützung kooperativer, überinstitutioneller Arbeitsstrukturen sind sie hervorragend geeignet. Die nur mittelfristige Dauer (einige Jahre) von VRE ist hierbei keine Einschränkung.

DISZIPLINÜBERGREIFENDE STRUKTUREN

Wann immer über Forschungsdateninfrastruktur debattiert wird, stehen insbesondere Lösungen im Vordergrund, die sich an mehr als nur eine Fach-Community wenden. Auch die Förderer legen einen Schwerpunkt auf diese Interdisziplinarität. Neben der wünschenswerten Vermeidung von Doppelstrukturen und der Förderung neuer Ideen durch den Austausch mit fachfremden Wissenschaftlerinnen und Wissenschaftlern birgt dies jedoch auch Gefahren. Bei kooperativen Projekten, die eng an interdisziplinäre Forschungsvorhaben geknüpft sind, sind natürlich Forschungsdateninfrastrukturen notwendig, die disziplinübergreifend fungieren können. Disziplinübergreifende Strukturen sind auch dann sinnvoll und fruchtbringend, wenn sie vorhandene disziplinäre Strukturen integrieren. Wird jedoch bei einem disziplinspezifischen Projekt die Interdisziplinarität stark in den Vordergrund gestellt, so ist dies ein fast sicherer Indikator für eine unzureichende Nutzerbasis innerhalb der Community. Im Resultat werden dann Strukturen geschaffen, die zwar fachübergreifend sind, aber nicht wahrgenommen werden. Insbesondere in den Disziplinen, in denen das Forschungsdatenmanagement noch in den Kinderschuhen steckt, sollten sich Projekte zunächst auf die Schaffung von Basisstrukturen konzentrieren, bevor Strukturen auf angrenzende Felder übertragen werden.

Es gibt jedoch Themen, in denen disziplinübergreifenden Lösungen anzustreben sind. Insbesondere im organisatorischen Bereich, in der Kommunikation der Ergebnisse der diversen Forschungsdatenprojekte sowie für fachferne Querschnittsthemen wie Nachhaltigkeit liegt dies nahe. Hier können interdisziplinäre Kompetenznetzwerke und Expertengremien hilfreich sein. Hierfür kommen im Datenmanagement erfahrene Personen aus den Communities infrage, aber natürlich auch aus Rechenzentren oder aus den Bibliotheken. Es ist jedoch wichtig, dass diese Gruppen nicht abgehoben von dem Bedarf in den Communities agieren, sondern gezielt aus den Communities beauftragt werden. Den Förderern oder den Wissenschaftsorganisationen könnte hier eine Vermittlerrolle zufallen. Die Expertinnen und Experten müssen für ihren Einsatz selbstverständlich adäquat gefördert werden, jedoch sollte dies, um den so wichtigen Bezug zu den eigentlichen Projekten nicht zu verlieren, hauptsächlich über selbige erfolgen. Vereinfacht gesagt, ist es entscheidend, dass die Initiative für eine Zusammenarbeit von den Projekten (bzw. mittelbar durch die jeweilige Förderinstitution) ausgeht, und es nicht die Aufgabe des Netzwerkes ist, neue »Kunden« zu akquirieren.

FAZIT

Die gewachsenen Möglichkeiten, mit digitalisierten Datensammlungen auch neue Methoden für die wissenschaftliche Arbeit einzusetzen, kombiniert mit der Möglichkeit, auch über Entfernungen hinweg gemeinsam wissenschaftlich zusammenzuarbeiten, machen es nötig, die vorhandene Forschungsdateninfrastruktur gezielt weiterzuentwickeln. In den letzten Jahren hat sich daher auch eine Vielzahl verschiedener Initiativen und Projekte in den Communities, an den Bibliotheken und Rechenzentren und bei den Förderern und den Wissenschaftsorganisationen gebildet. Bei der Bewertung dieser Vorhaben hat sich das im Artikel diskutierte Domänenmodell als ein hervorragendes analytisches Hilfsmittel erwiesen. In gewisser Weise ermöglicht es, die Projektideen zu »erden«. So wird beispielsweise offensichtlich, dass ein Vorhaben, das sich über mehr als zwei Domänen erstreckt, nicht mit zwei bis drei Mitarbeitern realisiert werden kann. Es ermöglicht aber auch, die beabsichtigten Änderungen an bestehenden Workflows gezielt zu verorten, die Arbeiten in ihrer Gesamtstruktur auf adäquate Zielsetzungen und Methoden zu hinterfragen und die Rollen der am Forschungsprojekt Beteiligten einzuordnen.

Im Bereich der SFB/Transregio der DFG stellen die INF-Projekte eine geeignete Maßnahme dar, das Datenmanagement in der Gruppendomäne zu verbessern. In Zukunft sollte daher noch stärker darauf hingewirkt werden, dass jeder SFB/Transregio eine solche Komponente besitzt. Die INF-Projekte sollten sich aber nur auf den direkten Kontext des SFB/Transregio beschränken und nicht für den Aufbau fehlender Strukturen in den Hochschulen oder den Forschungsdisziplinen zweckentfremdet werden.

Außerhalb der SFB/Transregio steht die Förderung vor dem Problem, dass die Gruppendomäne stark durch komplexe und teils informelle Strukturen geprägt ist, die die Grenzen von Institutionen (und auch von Nationen und Kontinenten) überschreiten. Virtuelle Forschungsumgebungen, die nicht auf eine ganze Forschungsdisziplin gerichtet sind, sondern spezifische auf die jeweilige Kollaboration zugeschnittene Dienste bereitstellen, können hier hilfreich sein.

Bei den immer wieder betonten disziplinübergreifenden Maßnahmen ist kritisch zu hinterfragen, ob das Vorhaben, in der Regel eine Erweiterung eines bestimmten Dienstes oder eines Werkzeuges auf eine angrenzende Disziplin, wirklich schon genug Akzeptanz in der ursprünglichen Zielgruppe besitzt. Für Querschnittsthemen, wie beispielsweise Nachhaltigkeit, können interdisziplinäre Kompetenzteams einen Wissenstransfer und somit eine Entlastung der in den

Domänenmodell als analytisches Hilfsmittel

interdisziplinäre Kompetenznetzwerke und Expertengremien

Infrastrukturprojekten arbeitenden Wissenschaftlerinnen und Wissenschaftler darstellen.

Abschließend ist festzustellen, dass fortdauernde Entwicklung im Bereich der Forschungsdaten es weiterhin nötig macht, die Rollen von Bibliotheken, Hochschulrechenzentren und anderen Institutionen neu zu überdenken. Es ist absehbar, dass die traditionelle wissenschaftliche Arbeitsweise mehr und mehr durch eine datenbasierte Forschung ergänzt wird. Hierdurch verschiebt sich jedoch der Fokus vom Zugang und der Publikation von Artikeln, Büchern usw. auf den Zugang und die Publikation von Datensammlungen. Die technischen und wissenschaftlich-kontextuellen Anforderungen, die hieraus erwachsen, werfen die Frage auf, inwieweit diese noch sinnvoll in den Bibliotheken vertortet werden können. Die universitären IT-Strukturen ihrerseits sind – dem jetzigen Anspruch nach – gehalten, alle Dienste bereitzustellen, die die Forscherinnen und Forscher an der jeweiligen Institution benötigen. Eine solche Aufgabe ist jedoch ohne eine Zusammenarbeit über die Grenzen der einzelnen Hochschulen hinaus realistisch nicht zu bewältigen. Hier könnten Einrichtungen wie regionale oder nationale Rechenzentren Abhilfe schaffen. In der Regel sind diese jedoch in ihrer Ausrichtung und Konzeption auf die Bereitstellung von Rechenkapazitäten für das Scientific Computing beschränkt. Wie sich die Aufgaben des Forschungsdatenmanagements in der Zukunft über die verschiedenen Akteure verteilen werden, ist jedoch zum gegenwärtigen Stand der Diskussion nur in Ansätzen erkennbar.¹⁰

Dieser Artikel beschränkt sich bewusst darauf, die Wirkung der Förderprogramme und der Projekte in Bezug auf die Gruppendomäne zu diskutieren. Daher wird weder auf europäische noch internationale Initiativen bezüglich der Forschungsdateninfrastruktur eingegangen, obwohl auch diese die Gesamtsituation mit beeinflussen. Eine kritische Bestandsaufnahme liegt jedoch dafür nicht vor und würde auch den Rahmen des Artikels sprengen. Es ist hier vielmehr versucht worden, für die Fortentwicklung der Forschungsdateninfrastruktur Bereiche zu durchleuchten, in denen der größte Handlungsbedarf besteht und wo, im Unterschied zu den Bereichen der dauerhaften Domäne und der Zugangsdomäne, viele grundlegende Strukturen, Workflows und Werkzeuge noch zu erarbeiten sind.

LITERATUR

DFG, *Grafik zur Systematik der Fächer und Fachkollegien der DFG für die Amtsperiode 2012–2015*, www.dfg.de/download/pdf/dfg_im_profil/gremien/fachkollegien/amtsperiode_2012_2015/fachsystematik_2012_2015_de_grafik.pdf (abgerufen 22.7.2013).

DFG, *Merkblatt Sonderforschungsbereich*, DFG-Vordruck 50.06 – 06/12. Bonn, 2012, S. 8 f., www.dfg.de/formulare/50_06/50_06_de.pdf (abgerufen 28.7.2013).

Enke, H., Wambsganß, J. (2012) in: Neuroth, H. et al. (Hrsg.) (2012) *Langzeitarchivierung von Forschungsdaten – Eine Bestandsaufnahme*, Universitätsverlag Göttingen S. 275 ff.

LOFAR, *Systembeschreibung*, www.lofar.org/about-lofar/system/introduction (abgerufen 26.7.2013).

Treloar, A. und Harboe-Ree, C. (2008) *Data management and the curation continuum: how the Monash experience is informing repository relationships*, Proceedings of VALA 2008, Melbourne, February.

Radieschen (2013a), *Bericht Organisation: Empfehlungen zur Einbettung des Forschungsdaten-Managements in die Forschungsinfrastruktur*, http://dx.doi.org/10.2312/RADIESCHEN_005 (abgerufen 28.7.2013).

Radieschen (2013b), *Bericht Synthese: Eine Zusammenfassung der gewonnenen Erkenntnisse, Handlungsempfehlungen und Ausblick auf die Zukunft*, http://dx.doi.org/10.2312/RADIESCHEN_007 (abgerufen 28.7.2013).

Radieschen (2013c), *DFG-Projekt Radieschen: Gemeinsamer Workshop der SFB-INF-Projekte*, <http://gfzpublic.gfz-potsdam.de/pubman/faces/viewItemFullPage.jspx?itemId=escidoc:253073> (abgerufen 28.7.2013).

¹ Ob die Unterscheidung zwischen Primär- und Sekundärdaten im Allgemeinen eine sinnvolle Hilfe ist, wird immer wieder diskutiert. Im Zusammenhang mit Großprojekten, wie sie hier beschrieben werden, erachten wir sie jedoch als gerechtfertigt.

² Im Folgenden wird Fachgemeinschaft und Community oft synonym verwendet. Community kann jedoch auch als Bezeichnung für interdisziplinäre Gemeinschaftsvorhaben verwendet werden und wird daher bevorzugt, wenn es um die potentielle Nutzergemeinschaft einer Infrastruktur zur Unterstützung des Forschungsprozesses geht.

³ Vgl. *Radieschen* (2013a).

⁴ Vgl. Treloar, A. und Harboe-Ree, C. (2008).

⁵ Vgl. DFG: *Merkblatt Sonderforschungsbereiche*, www.dfg.de/formulare/50_06/50_06_de.pdf (abgerufen 28.7.2013).

⁶ Vgl. *Radieschen* (2013c).

⁷ Vgl. Grafik zur Systematik der Fächer und Fachkollegien der DFG für die Amtsperiode 2012–2015 auf der Webseite der DFG.

⁸ Vgl. Enke (2012).

⁹ Eine von der DFG finanzierte Studie zur Nachhaltigkeit von VRE wird derzeit durchgeführt.

¹⁰ Vgl. Zukunftsszenarien in *Radieschen* (2013b).

DIE VERFASSER

Dr. Jochen Klar, Leibniz-Institut für Astrophysik Potsdam (AIP) Supercomputing und E-Science, An der Sternwarte 16, 14482 Potsdam, Tel.: 0331 – 7499 536, E-Mail: jklar@aip.de

Dr. Harry Enke, Leibniz-Institut für Astrophysik Potsdam (AIP) Supercomputing und E-Science, An der Sternwarte 16, 14482 Potsdam, Tel.: 0331 – 7499 433, E-Mail: henke@aip.de