

Angelina Kraft, Matthias Razum, Jan Potthoff, Andrea Porzel,
Thomas Engel, Frank Lange, Karina van den Broek

Archivierung und Publikation von Forschungsdaten: Die Rolle von digitalen Repositorien am Beispiel des RADAR- Projekts

DOI 10.1515/bd-2016-0077

Zusammenfassung: Disziplinübergreifendes Forschungsdatenmanagement für Hochschulbibliotheken und Projekte zu vereinfachen und zu etablieren – das ist das Ziel von RADAR. Im Sommer 2016 geht mit ‚RADAR – Research Data Repository‘ ein Service an den Start, der Forschenden, Institutionen verschiedener Fachdisziplinen und Verlagen eine generische Infrastruktur für die Archivierung und Publikation von Forschungsdaten anbietet. Zu den Dienstleistungen gehören u. a. die Langzeitverfügbarkeit der Daten mit Handle oder Digital Object Identifier (DOI), ein anpassbares Rollen- und Zugriffsrechtmanagement, eine optionale Peer-Review-Funktion und Zugriffsstatistiken. Das Geschäftsmodell ermutigt Forschende, die anfallenden Nutzungsgebühren des Repositoriums in Drittmittelanträge und Datenmanagementpläne zu integrieren. Publierte Daten stehen als Open Data zur Nachnutzung wie etwa Data Mining, Metadaten-Harvesting und Verknüpfung mit Suchportalen zur Verfügung. Diese Vernetzung ermöglicht ein nachhaltiges Forschungsdatenmanagement und die Etablierung von Dateninfrastrukturen wie RADAR.

Schlüsselwörter: Forschungsdaten, Forschungsdatenmanagement, digitale Infrastruktur

Dr. Angelina Kraft: Angelina.Kraft@tib.eu

Matthias Razum: Matthias.Razum@fiz-karlsruhe.de

Jan Potthoff: Jan.Potthoff@kit.edu

Dr. Andrea Porzel: Andrea.Porzelt@ipb-halle.de

Dr. Thomas Engel: Thomas.Engel@lmu.de

Frank Lange: Frank.Lange@ipb-halle.de

Karina van den Broek: Karina.vandenBroek@cup.uni-muenchen.de

Archiving and publishing of research data: The role of digital repositories, taking the example of the RADAR project

Abstract: The goal of RADAR is to simplify and establish inter-disciplinary research data management for university libraries and projects. In summer 2016, ‘RADAR – Research Data Repository‘ starts as a service that offers researchers, institutions of different disciplines and publishers a generic infrastructure for archiving and publishing their research data. Among others, services are long-term data availability with Handle or Digital Object Identifier (DOI), an adaptable role and access rights management, an optional peer review function and access statistics. The business model encourages researchers to integrate arising charges for using the repository into applications for Third-Party funding and data management plans. Published data are available as Open Data to be used by Data Mining, metadata harvesting or linking with search portals. This interlinking enables a sustainable research data management and the establishment of data infrastructures like RADAR.

Keywords: research data, research data management, digital infrastructure

1 Big Data, Long Tail Data und Repositorien

Forschungsdaten entstehen im wissenschaftlichen Alltag: Studenten, Doktoranden, Post-Docs und etablierte Wissenschaftler¹ erzeugen oder sammeln Daten im Rahmen ihrer wissenschaftlichen Arbeit. Dabei sind Art und Umfang der Daten äußerst divers: Sie reichen von wenigen Kilobyte großen Excel-Tabellen bis hin zu mehrere 100 Terabyte umfassenden Sammlungen von Satellitenbildern. Abhängig von der Forschungsdisziplin dienen die Daten als Grundlage für die Beantwortung von Forschungsfragen oder das Testen von Hypothesen. Darüber hinaus stellen zitierfähige Forschungsdaten die Integrität und bestenfalls auch die Überprüfbarkeit wissenschaftlicher Erkenntnisse sicher. Doch im Gegensatz zu Textpublikationen, deren Formate und Struktur schon seit vielen Jahrzehnten etablierten Strukturen, Peer-Review-Mechanismen und wohl verteilten Zuständigkeiten zwischen Wissenschaftlern, Gutachtern und Herausgebern folgen,

¹ Die männliche Sprachform ist alleine der besseren Lesbarkeit geschuldet und schließt ausdrücklich auch jeweils die weibliche Sprachform mit ein.

stellt der Umgang mit den vielfältig strukturierten Forschungsdaten umfangreiche Anforderungen an Hochschulen, außeruniversitären Forschungseinrichtungen, Bibliotheken, Verlage und natürlich den Wissenschaftler selbst.

Vor allem in Disziplinen außerhalb der sogenannten *Big Data* Communities und in Hochschulbibliotheken fehlen oftmals geeignete Infrastrukturen für die nachhaltige Archivierung und Publikation von Forschungsdaten. Dabei machen gerade diese ‚kleinen‘ Datensätze einen großen Anteil der erhobenen Forschungsdaten aus: Eine in Science 2011 veröffentlichte Umfrage² unter 1700 Wissenschaftlern aus unterschiedlichen Disziplinen zeigte, dass 48,3 % der Forscher mit Datensätzen von unter 1 GB Volumen arbeiten. Abhängig von institutionellen, projektbezogenen oder durch den Geldgeber festgelegten Vorgaben werden Forschungsdaten nach erfolgter Veröffentlichung der Resultate in wissenschaftlichen Publikationen auf unterschiedliche Art und Weise gespeichert, z. B. auf persönlichen Festplatten, institutionellen Netzwerklaufwerken oder mit Hilfe von Cloud-Diensten. Diese Speicherformen garantieren Forschern jedoch keine nachhaltige Datensicherheit und bieten oftmals nur eine eingeschränkte Nachnutzbarkeit der Daten. Dabei ist gerade die Verfügbarkeit, Reproduzierbarkeit und Nachvollziehbarkeit digitaler Daten essentiell für die weitere Entwicklung der Wissenschaft.

Die Gründe für die Zurückhaltung, Forschungsdaten nicht oder nur eingeschränkt verfügbar zu machen, sind so vielfältig wie die wissenschaftlichen Disziplinen selbst. Drei zu nennende Hauptfaktoren sind Zeit, Vertrauen und Anerkennung³. Geldgeber und Forscher fordern deshalb zu Recht ein Publikationssystem, das die erhobenen wissenschaftlichen Daten

- sicher speichert,
- eindeutig den wissenschaftlichen Autoren und durchführenden Institutionen zuordnet,
- zitierbar macht,
- eine langfristige Nachnutzung ermöglicht und
- eine missbräuchliche Verwendung verhindert.

² Science: Challenges and Opportunities. In: Science 11, Vol. 331 no. 6018, S. 692–693, Februar 2011. DOI: 10.1126/science.331.6018.692.

³ Siehe hierzu auch die Umfragen zum Umgang mit Forschungsdaten an Hochschulen und Forschungseinrichtungen. Eine Übersicht bietet das Portal [forschungsdaten.org](http://www.forschungsdaten.org/index.php/Umfragen_zum_Umgang_mit_Forschungsdaten_an_wissenschaftlichen_Institutionen): http://www.forschungsdaten.org/index.php/Umfragen_zum_Umgang_mit_Forschungsdaten_an_wissenschaftlichen_Institutionen [Zugriff: 23.03.2016].

Die Berücksichtigung dieser Forderungen setzt neben einer vertrauenswürdigen Speicherung eine hinreichende Beschreibung der Forschungsdaten voraus (diese werden auch als deskriptive Metadaten bezeichnet). Dazu muss eine Infrastruktur existieren, die eine solche Daten- und Metadatenverwaltung unterstützt.

Um den wachsenden Anforderungen an ein wissenschaftlich fundiertes digitales Forschungsdatenmanagement zu entsprechen, kommen vermehrt Datenrepositorien zum Einsatz. In einigen wissenschaftlichen Schwerpunktbereichen wie z. B. den Umwelt- und Klimawissenschaften ist ein langjähriger Erfahrungsschatz für die Erhebung, Speicherung und Nachnutzung wissenschaftlicher Daten vorhanden. Datenrepositorien wie PANGAEA⁴ (Publishing Network for Geoscientific and Environmental Data), das vom Alfred-Wegener-Institut für Polar- und Meeresforschung (AWI) und vom Zentrum für Marine Umweltwissenschaften der Universität Bremen (MARUM) betrieben wird, macht wissenschaftliche Umweltdaten im Sinne des Open Access nachnutzbar. Große Datenzentren wie PANGAEA verweisen in den Metadaten des publizierten Forschungsdatensatzes oftmals auf die zugehörigen Zeitschriftenaufsätze. Diese langjährig etablierten, disziplinären Datenzentren sorgen somit für eine dauerhafte und zuverlässige Speicherung und den Zugriff auf die Daten. Im Gegensatz zu den gerade genannten Big Data Gemeinschaften bildet die strukturierte, nachvollziehbare Datenarchivierung und Publikation in vielen interdisziplinär arbeitenden Forschungsbereichen und anderen wissenschaftlichen Disziplinen, der auch der ‚*Long Tail of Science*‘⁵ genannt wird, die Ausnahme. Dabei bietet sich gerade in diesen Bereichen eine Chance, wissenschaftliche Forschung auch über disziplinäre Grenzen hinweg nachnutzbar zu machen. Das ist eine der Aufgaben von multidisziplinär ausgerichteten (=generischen) Repositorien. Sie ermöglichen neben der Archivierung und Verfügbarkeit digitaler Objekte auch eine Sichtbarkeit von Forschungsleistungen über einzelne Institutionen und disziplinspezifische Grenzen hinaus.

Mit dem Fokus auf den *Long Tail*-Disziplinen strebt das RADAR – Research Data Repository – Projekt die Schaffung eines generischen Forschungsdatenrepositoriums an, das eine strukturierte, einfach bedienbare Lösung für eine nachhaltige Datenarchivierung und optionale Publikation anbietet. Entwickelt wird RADAR als DFG-gefördertes Verbundprojekt von FIZ Karlsruhe – Leibniz-Institut für Informationsinfrastruktur, der Technischen Informationsbibliothek (TIB) in Hannover und dem Karlsruher Institut für Technologie (KIT) zusammen mit der

⁴ <https://pangaea.de/PANGAEA> ist ein disziplinäres Repositorium für Geowissenschaft und Umweltforschung. [Zugriff: 23.03.2016].

⁵ Vgl. Borgman, Christine L: The Conundrum of Sharing Research Data. In: Journal of the American Society for Information Science and Technology 63, 6 (2012), S. 1059–1078.

Ludwig-Maximilians-Universität München (LMU) und dem Leibniz-Institut für Pflanzenbiochemie (IPB). Nach dem Projektstart im September 2013 wurde regelmäßig über die Planung und Fortschritte des Projekts berichtet^{6,7}. Nachfolgend wird nun das RADAR-Produktivsystem und die verfügbaren Dienstleistungen vorgestellt, die ab Sommer 2016 zur Verfügung stehen.

2 RADAR – Research Data Repository: Zwei Dienstleistungen für wissenschaftliche Institutionen

Die Dienstleistung des RADAR-Repositoryums (Abb. 1) besteht aus einem Einstiegsangebot zur Archivierung wissenschaftlicher Daten und einem erweiterten Datenangebot, um Daten zu publizieren und gleichzeitig zu archivieren:

- Im Basisangebot „Archivierung“ bietet RADAR eine formatunabhängige Datenarchivierung an. Datengebern ist die Möglichkeit gegeben, ihre Daten über langfristige Archivierungsperioden (z. B. 10 Jahre, gemäß den Empfehlungen der DFG) zu speichern. Dieser Service gewährleistet, dass wissenschaftliche Daten (ohne Publikation) sicher gespeichert werden können. Standardmäßig sind weder Daten noch Metadaten öffentlich zugänglich, der Datengeber kann jedoch abweichende Zugriffsregeln festlegen.
- Im erweiterten Angebot „Datenpublikation mit integrierter Archivierung“ bietet RADAR einen kombinierten Service aus Datenpublikation und Archivierung an. Durch die Verwendung von *Digital Object Identifier (DOI)* als persistente Identifikatoren wird sichergestellt, dass jeder publizierte Datensatz eigenständig zitierfähig ist. RADAR Publikationen sind Open Access (mit der Möglichkeit eines optionalen Embargo-Zeitraums). Die zugehörigen Metadaten sind über Schnittstellen harvestbar.

6 Razum, Matthias; Neumann, Janna; Hahn, Matthias: RADAR – Ein Forschungsdaten-Repository als Dienstleistung für die Wissenschaft. *Zeitschrift für Bibliothekswesen und Bibliographie* (2014), 61(1): 18–27. DOI: <http://dx.doi.org/10.3196/186429501461150> [Zugriff: 22.4.2016].

7 Razum, Matthias; Neumann, Janna: Das RADAR Projekt: Datenarchivierung und -publikation als Dienstleistung – disziplinübergreifend, nachhaltig, kostendeckend. *o-bib – das offene Bibliotheksjournal* (2014), Verein Deutscher Bibliothekare (VDB), 1(1): 30–44. DOI: <http://dx.doi.org/10.5282/o-bib/2014H1S30-44> [Zugriff: 22.4.2016].

Herausgeber wissenschaftlicher Literatur können ebenfalls vom RADAR-Service profitieren, indem die wissenschaftlichen Daten aus publizierten Manuskripten archiviert oder publiziert werden. Dies steigert sowohl die Sichtbarkeit, als auch die Transparenz wissenschaftlicher Arbeiten. Dabei ist es ein Ziel von RADAR, Forschungsdaten in den Peer-Review-Prozess von Manuskripten zu integrieren. Hierzu bietet RADAR eine Peer-Review Funktion für Datensätze an. In diesem Status wird dem Datengeber eine gesicherte URL zur Verfügung gestellt, die an Editoren und Gutachter kommuniziert werden kann.

RADAR Suche Metadaten+ Hochladen Preise Kontakt Registrieren Anmelden

Forschungsdaten sichern und zugänglich machen

RADAR ist ein wissenschaftliches Datenrepositorium für die Archivierung und Publikation von Forschungsdaten.

Durchsuchen Sie das interdisziplinäre Datenrepositorium.
Sie können gezielt in RADAR suchen um Daten für ihre wissenschaftlichen Arbeiten zu finden.

Archivierung und Publikation von Forschungsdaten.
RADAR bietet Ihnen vielfältige Funktionen zum Speichern, Organisieren und Publizieren ihrer Forschungsdaten.

Über RADAR | FAQ | Glossar | Datenschutzerklärung | Geschäftsbedingungen Englisch / Deutsch

Abb. 1: Homepage des RADAR – Research Data Repository – Service zur Archivierung und Publikation von Forschungsdaten. Das Repository (www.radar-projekt.org) geht im Sommer 2016 an den Start.

3 Finanzierung und Nachhaltigkeit: Offene Forschungsdaten und die Wahl zwischen einer jährlichen oder einer einmaligen Nutzungsgebühr

Das Geschäftsmodell von RADAR sieht zwei Zahlungsverfahren vor: Jährliche Zahlungen (Zielgruppe: institutionelle Kunden) und Einmalzahlungen für die gesamte Aufbewahrungsdauer (Zielgruppe: Forschungsprojekte). Zum Start wird RADAR zunächst jährliche Zahlungen anbieten. In einem zweiten Schritt sollen dann Einmalzahlungen folgen.

Zugang und Verfügbarkeit der Daten: RADAR bietet freien Zugriff auf alle publizierten Daten an. Im Publikationsangebot ist durch die Publikation eine dauerhafte Speicherung vorgesehen. Diese umfasst eine Datenverfügbarkeit für mindestens 25 Jahre. Bei archivierten Daten kann der Kunde zwischen einer Archivierungsdauer von 5, 10 und 15 Jahren wählen. Darüber hinaus bietet der RADAR-Service Nutzungsstatistiken an. Dort kann die Anzahl der monatlichen Aufrufe der Metadateninformationen und die Anzahl der Downloads eines Datensatzes eingesehen werden.

Die Kosten für die Archivierungs- und Publikationsservices von RADAR werden durch die Vertragsadministratoren (Vertragspartner) getragen. Eine Preisliste⁸ für Hochschulen und Forschungseinrichtungen als institutionelle Kunden wurde im Januar 2016 veröffentlicht und ist über die Projekthomepage (<http://radar-projekt.org>) verfügbar. RADAR bietet die Option einer jährlichen Bezahlung an. Die Kosten sind abhängig vom benötigten Datenvolumen. Mit Einführung der Einmalzahlungen strebt RADAR zukünftig auch ein Zahlungsverfahren an, dass zu den üblichen Projektstrukturen in der Forschung kompatibel ist. RADAR ermutigt Wissenschaftler die Kostenschätzung in Förder- bzw. Drittmittelanträge zu integrieren, um die Finanzierung für die Datenarchivierung und -publikation über das RADAR-Repository zu sichern. Dazu wird RADAR dann eine einfache Möglichkeit zur Erstellung von Kostenvoranschlägen anbieten. Mit seinem Geschäftsmodell zielt RADAR auf einen nachhaltigen und sich selbst tragenden Betrieb ab, der dauerhaft ohne Förderung durch Förderer wie z. B. die DFG auskommt. Diese decken allerdings nur einen Teil der entstehenden Kosten ab. Der Rest wird von den am Betrieb beteiligten Einrichtungen übernommen.

⁸ <https://www.radar-projekt.org/pages/viewpage.action?pageId=3997880> Preise für RADAR Archivierungs- und Publikationsservice verfügbar. [Zugriff: 17.03.2016].

Dieses Geschäftsmodell kann deshalb nur einen eingeschränkten Service, etwa bei der Beratung von Einrichtungen und Wissenschaftlern bezüglich geeigneter Formate, Datenmanagementpläne und Lizenzen, unterstützen. Darüber hinausgehende Angebote müssten entweder als bezahlte Dienstleistung erbracht oder aber eine dauerhafte institutionelle Förderung ermöglicht werden.

4 Modulare Systemarchitektur: Frontend – Repository-Management – Backend

Das RADAR-System ist modular aufgebaut. Es besteht aus einer Managementsoftware mit Programmierschnittstellen (API) zur Benutzeroberfläche und zur Speicherschicht. Die Managementsoftware basiert auf der Repository-Software „eSciDoc Next Generation“, kurz eSciDoc NG⁹, einer vollständig überarbeiteten Version von eSciDoc. Hier werden Forschungsdaten als Datensätze erfasst, auf ihre technische Integrität überprüft und mit Metadaten versehen. Die Datenübertragung zur Speicherschicht erfolgt über standardisierte Protokolle (*SFTP*, *GridFTP*), während die Datenhaltung selbst über moderne Platten- und Band-speicher-Systeme erfolgt. Ein sogenanntes *High-Performance Storage System* (HPSS) sorgt für einen effizienten Daten-Zugriff und erlaubt eine kostengünstige Langzeitspeicherung. Die Zuverlässigkeit des Systems wird durch Mehrfachkopien gesichert, die an unterschiedlichen, geographisch verteilten Standorten abgelegt werden. Die Daten werden einer regelmäßigen Kontrolle unterzogen und im Falle eines Fehlers wieder hergestellt. Ein regelmäßiges Backup sowie der Einsatz von *Redundant Array of Independent Disks* (RAID) und eine Systemüberwachung sorgen für die Datensicherheit der administrativen Komponenten, wie zum Beispiel der Datenbanken, im System. Die Daten selbst liegen als *BagIt*-Strukturen¹⁰ in ZIP-Containern vor, d. h. die Daten werden in dokumentierten Containern vorgehalten. Jeder archivierte Datensatz besteht aus einer Datei, die in sich selbstbeschreibend ist (also Daten und beschreibende Metadaten enthält). Für die Interpretation der *BagIt*-Strukturen ist keine spezielle Software notwendig. Die RADAR-Software wird zum Start des Produktivsystems als Open Source Software und unter einer *ASL 2.0 Lizenz*¹¹ verfügbar gemacht.

⁹ <https://github.com/escidoc-ng/escidoc-ng> [Zugriff: 23.03.2016].

¹⁰ http://www.digitalpreservation.gov/news/2008/20080602news_article_bagit.html Library Develops Specification for Transferring Digital Content. [Zugriff: 23.03.2016].

¹¹ <http://www.apache.org/licenses/LICENSE-2.0.html> Apache Licence. [Zugriff: 23.03.2016].

5 Datenmanagement in RADAR

Die Ansätze und Ziele von RADAR wurden seit Projektstart 2013 in interaktiven Workshops mit potenziellen Nutzern aus Hochschulen und Forschungseinrichtungen als auch mit Vertretern von wissenschaftlichen Verlagen diskutiert. Die Ergebnisse flossen in die Entwicklung des RADAR-Dienstes ein. Von besonderem Interesse war dabei die Möglichkeit der Integration des Systems in den Forschungsprozess mit z. B. einer Funktion für die (automatisierte) Erfassung von Forschungsdaten sowie deren beschreibenden und technischen Metadaten. Im aktuellen Produktivsystem stellt RADAR ein Metadatenprofil mit zehn fachspezifischen Pflichtfeldern und zwölf optionalen Feldern zur Verfügung, die eine erweiterte, fachspezifische Beschreibung der Forschungsdaten ermöglichen. Zur Sicherstellung der Metadatenqualität verwendet RADAR z. B. kontrolliertes Vokabular und überprüft die Vollständigkeit der Metadaten auf Datensatzebene. Darüber hinaus wird die technische Integrität des Datensatzes beim Upload und im Rahmen der *Bitstream Preservation*¹² in regelmäßigen Abständen von RADAR kontrolliert und dokumentiert. Eine weitergehende, inhaltliche Überprüfung der Datensätze liegt in der Verantwortung der Datengeber (Kuratoren) selbst (Tab. 1). Das RADAR-Metadatenchema bietet dabei eine volle Kompatibilität zu etablierten Standards wie *DataCite*¹³ und *DublinCore*¹⁴. Die Metadaten der bei RADAR publizierten Daten stehen unter einer CC0-Lizenz und sind u. a. über den *DataCite Metadatenstore*¹⁵ abrufbar. Des Weiteren werden die Metadaten zum Harvesting via *OAI-PMH*¹⁶ zur Verfügung gestellt.

12 Bitstream Preservation bezeichnet die physische Erhaltung eines gespeicherten Datenobjekts auf einem Speichermedium. Dagmar Ulrich (2009): Kapitel 8.2 „Bitstream Preservation“ (Version 2.0): urn:nbn:de:0008-20090811392. Aus: Heike Neuroth et al. (Hrsg.) Handbuch: Eine kleine Enzyklopädie der digitalen Langzeitarchivierung.

13 <https://www.datacite.org/>DataCite ist ein 2009 gegründetes internationales Konsortium, das zum Ziel hat einen weltweit einheitlichen Zugang zu Forschungsdaten zu ermöglichen. [Zugriff: 23.03.2016].

14 <http://dublincore.org/>Dublin Core Metadata Initiative. [Zugriff: 23.03.2016].

15 <http://search.datacite.org/>DataCite Metadata Search. [Zugriff: 23.03.2016].

16 <https://www.openarchives.org/OAI/openarchivesprotocol.html> The Open Archives Initiative Protocol for Metadata Harvesting. [Zugriff 23.03.2016].

Tab. 1: Rollen- und Datenmanagementkonzept von RADAR.

Datenmanagement	
A) Vertrag	Ein Kunde (z. B. Institution, Verlag, Forschungsprojekt) schließt einen Dienstleistungsvertrag mit RADAR ab
B) Arbeitsbereich	Eine organisatorische Einheit (engl. Workspace) zum Upload und Strukturierung der Forschungsdaten und zugehöriger Metadaten. Ebenfalls können dort Kontaktinformationen verwaltet und Zugriffsrechte gesteuert werden
C) Datensatz	Eine Zusammenstellung digitaler Daten, die zusammengehörig einen Forschungsprozess bzw. dessen Ergebnis beschreiben. Ein Datensatz kann z. B. aus Rohdaten, intermediären Arbeitsdaten, Sekundärdaten, Ergebnisdaten und/oder Negativdaten bestehen
Rollen und Aufgabenverteilung	
A) RADAR Betreiber	Überprüfung von Verträgen und Bereitstellung von Vertragsadministrator-Zugängen für die Kunden (interner Systemadministrator)
B) Vertragsadministrator (Kunde)	Ansprechpartner der Institution, eines Projekts oder Fachbereichs (Vertragspartner). Besitzt vollumfassende Lese- und Schreibrechte, generiert Arbeitsbereiche und ordnet Kuratoren zu
C) Kurator(en) (Datengeber)	Verfügt in einem Arbeitsbereich über volle Lese- und Schreibrechte, kann Daten hinzufügen, editieren, archivieren und publizieren sowie weitere (Sub-)Kuratoren zu ihrem Arbeitsbereich hinzufügen (z. B. Datenmanager einer Institution/Fakultät)
Status für Datensätze	
1.) In Bearbeitung: <i>Pending</i>	Initialer Status; (Sub-)Kuratoren können den Datensatz bearbeiten (Modifizierung, Update, Löschung)
2.) In Begutachtung: <i>Review</i>	Datensatz wurde durch Kurator zur Begutachtung (Peer Review) bei einem Verlag/Editor eingereicht und ist temporär gesperrt; er erhält eine sichere, temporäre 'Review-URL'
3.) Archiviert: <i>Archived</i> (= Service: Archivierung)	Datensatz ist archiviert und mittels eines Handle-Identifiers eindeutig identifizierbar. In diesem Status ist keine weitere Bearbeitung möglich
4.) Publiziert: <i>Published</i> (= Service: Publikation)	Datensatz ist publiziert und mittels eines DOI eindeutig identifizierbar und referenzierbar. Über eine öffentlich zugängliche Landingpage, die deskriptive und technische Metadaten anzeigt, kann der Datensatz heruntergeladen werden. In diesem Status ist keine weitere Bearbeitung möglich

6 Zielgruppen: Akademische Institutionen, Verlage, Kultureinrichtungen und Forschende

RADAR ermöglicht die verlässliche Archivierung von Forschungsdaten für Zielgruppen, die bisher noch keine eigene disziplinrelevante Forschungsdateninfrastruktur besitzen. Dazu gehören fachübergreifend arbeitende Institutionen wie z. B. Bibliotheken und Medienzentren kleinerer und mittlerer (forschender) Hochschulen, aber auch außeruniversitäre Forschungseinrichtungen, wissenschaftliche Verlage und Kultureinrichtungen. Diese institutionellen Kunden erhalten die Möglichkeit, RADAR als im Hintergrund agierende Dateninfrastruktur zu nutzen und an die eigenen Portale anzubinden. Somit können sie gegenüber ihrem Kundenstamm als kompetenter Datenmanagement-Dienstleister und Ansprechpartner wahrgenommen werden. Auch kann eine Institution auf diese Weise die Funktionalität des Portals an die individuellen Gegebenheiten anpassen. Akademische Institutionen sind meist an einer jährlichen Zahlungsweise interessiert, um flexibel agieren zu können und sich nicht langfristig an einen bestimmten Dienstleister zu binden. Dazu wird RADAR ein jährliches Bezahlmodell anbieten. Dieses Modell umfasst eine Kontingent-Lösung, die eine jährliche Zahlung bei speichervolumenabhängiger Preisstaffelung und die Möglichkeit des Kostenvoranschlags beinhaltet. Gleichzeitig wird bei einem Vertragsabschluss mit RADAR keine hohe Einmalzahlung erforderlich, da sich die Kosten auf viele Jahre verteilen. Das kann insbesondere für Einrichtungen interessant sein, die bereits umfangreiche Datenbestände vorhalten.

Eine weitere Zielgruppe stellen Verlage dar. Wissenschaftliche Verlage reichern zunehmend Publikationen mit Referenzen zu Forschungsdaten an. Sie steigern damit den Nutzen der Publikationen und unterstützen gleichzeitig die Forderung von Fördereinrichtungen und Wissenschaftsorganisationen, den Zugang zu Forschungsdaten zu ermöglichen. Auf der anderen Seite sind Verlage oftmals zurückhaltend, die dafür notwendige Infrastruktur bereitzustellen. RADAR als zertifiziertes Datenzentrum kann hier eine Dienstleistung für Verlage erbringen. Diese ist sowohl in Form einer Integration der Manuskript-Einreichungssysteme in das RADAR System als auch mit der zur Verfügungsstellung von Speicherressourcen in RADAR für ihre wissenschaftliche Community möglich. Darüber hinaus bietet RADAR einen Peer-Review Workflow an, um Forschungsdaten zeitgleich mit dem Review des zugehörigen Manuskripts durchführen zu können. Als weitere Zielgruppe können Kultureinrichtungen wie Archive oder Museen RADAR-Dienste nutzen, um z. B. Digitalisate in Form von Masterdateien dauerhaft vorzuhalten und gleichzeitig online Zugriffskopien anzubieten, ohne die dafür notwendige Infrastruktur selbst betreiben zu müssen.

Mit der Fokussierung auf die genannten Zielgruppen bilden Forschungsdatenrepositorien wie RADAR komplementäre Informationsinfrastrukturen zu den bestehenden, disziplinspezifischen Repositorien. Die Vernetzung beider Lösungen durch offene Schnittstellen zum Austausch von Metadaten, sowie deren Anbindung an weitere Informationsdienstleistungssysteme wie Repositorien-Verzeichnisse (z. B. re3data.org¹⁷ oder RISources.dfg.de¹⁸), Autorenverzeichnisse (z. B. ORCID¹⁹) und Forschungsinformationssysteme tragen darüber hinaus zur Reputationssteigerung der Forschenden und ihrer Institutionen bei. So soll ein Umfeld geschaffen werden, in dem sich Forschende zukünftig in gleichem Maße auf Daten- als auch auf Textpublikationen berufen können, um ihre wissenschaftliche Qualität zu untermauern und Fördermittel zu beantragen oder zu verlängern.



Dr. Angelina Kraft

Technische Informationsbibliothek (TIB)
Abteilung Forschung und Entwicklung
Welfengarten 1 B
30167 Hannover
E-Mail: Angelina.Kraft@tib.eu



Matthias Razum

FIZ Karlsruhe – Leibniz-Institut für Informationsinfrastruktur
Teilbereichsleiter e-Science
Hermann-von-Helmholtz-Platz 1
76344 Eggenstein-Leopoldshafen
E-Mail: Matthias.Razum@fiz-karlsruhe.de



Jan Potthoff

Karlsruher Institut für Technologie (KIT)
Research Group Management of Complex IT-Systems
Steinbuch Centre for Computing
Hermann-von-Helmholtz-Platz 1
76344 Eggenstein-Leopoldshafen
E-Mail: Jan.Potthoff@kit.edu

¹⁷ <http://www.re3data.org/Registry> of Research Data Repositories. [Zugriff: 24.03.2016].

¹⁸ <http://risources.dfg.de/Willkommen> beim Portal für Forschungsinfrastrukturen: RISources. [Zugriff: 24.03.2016].

¹⁹ <http://orcid.org/Connecting> Research and Researchers. [Zugriff: 23.03.2016].

**Dr. Andrea Porzel**

Leibniz-Institut für Pflanzenbiochemie (IPB)
Abteilung Natur- und Wirkstoffchemie
Weinberg 3
06120 Halle (Saale)
E-Mail: Andrea.Porzel@ipb-halle.de

**Dr. Thomas Engel**

Department Chemie
Ludwig-Maximilians-Universität München
Butenandtstr. 5–13, Haus F
81377 München
E-Mail: Thomas.Engel@lmu.de

**Frank Lange**

Leibniz-Institut für Pflanzenbiochemie (IPB)
Abteilung Natur- und Wirkstoffchemie
Weinberg 3
06120 Halle (Saale)
E-Mail: Frank.Lange@ipb-halle.de

**Karina van den Broek**

Department Chemie
Ludwig-Maximilians-Universität München
Butenandtstr. 5–13, Haus F
81377 München
E-Mail: Karina.vandenBroek@cup.uni-muenchen.de