



Research Data Repository

Abschlussbericht

an die Deutsche Forschungsgemeinschaft (DFG)

Oktober 2016

DFG-Geschäftszeichen:	BE 1042/7-1 HE 5985/4-1 NE 1352/2-1 RO 2273/4-1 WE 1467/14-1
AOBJ:	602266, 602267, 602268, 602269, 602270
Kennwort:	RADAR
Internetadressen:	https://www.radar-projekt.org https://www.radar-service.eu
Berichtszeitraum:	01.09.2013 – 31.08.2016
Projektlaufzeit:	01.09.2013 – 31.08.2016
Antragsteller:	Prof. Dr. Thomas Bein Dr. Leni Helmes Prof. Dr. Bernhard Neumair Uwe Rosemann Prof. Dr. Ludger Wessjohann

DFG Deutsche
Forschungsgemeinschaft

1.3 Liste der wichtigsten Publikationen

a) Arbeiten aus Publikationsorganen mit wissenschaftlicher Qualitätssicherung (chronologisch)

- **Kraft A., Razum M., Potthoff J., Porzel A., Engel T., Lange F., van den Broek K.** (2016) *Archivierung und Publikation von Forschungsdaten: Die Rolle von digitalen Repositorien am Beispiel des RADAR Projekts*. Bibliotheksdienst, Band 50, Heft 7, S. 623–635. doi: 10.1515/bd-2016-0077
- **Kraft A., Razum M., Potthoff J., Porzel A., Engel T., Lange F., van den Broek K., Furtado F.** (2016) *The RADAR Project—A Service for Research Data Archival and Publication*. ISPRS Int. J. Geo-Inf. 5, 28. doi:10.3390/ijgi5030028
- **Potthoff J., van Wezel J., Razum M., Walk M.** (2014) Anforderungen eines nachhaltigen, disziplinübergreifenden Forschungsdaten-Repositoriums. In: Müller P., Neumair B., Reiser H., Rodosek G.D. (Hrsg.), Proceedings 7. DFN-Forum Kommunikationstechnologien, Lecture Notes in Informatics, GI, S. 11-20.
- **Razum M., Neumann J.** (2014) *Das RADAR Projekt: Datenarchivierung und -publikation als Dienstleistung - disziplinübergreifend, nachhaltig, kostendeckend*. o-bib - das offene Bibliotheksjournal, Verein Deutscher Bibliothekare (VDB), 1(1): 30-44. doi: 10.5282/o-bib/2014H1S30-44
- **Razum M., Neumann J., Hahn M.** (2014) *RADAR – Ein Forschungsdaten-Repositorium als Dienstleistung für die Wissenschaft* Zeitschrift für Bibliothekswesen und Bibliographie, 61(1):18-27. doi: 10.3196/186429501461150

b) Andere Veröffentlichungen (chronologisch)

- **Potthoff J., Razum M., Kraft A.** (2015) *RADAR-Team stellt Testsystem auf zweitem Projekt-Workshop in Frankfurt vor* SCC-News 2015, S. 25-26. PDF: https://www.scc.kit.edu/downloads/oko/SCC-news01_2015.pdf
- **Pawlak A.** (2015) *Sag mir, wo die Daten sind...* Physik-Journal Nr. 11, S. 24-27. Online: http://www.pro-physik.de/details/physikjournalArticle/8483981/Sag_mir_wo_die_Daten_sind.html
- **Kraft A.** (2015) *RADAR - A repository for long tail data*. Proceedings of the IATUL Conferences, Paper 1. PDF: <http://docs.lib.purdue.edu/iatul/2015/mrd/1>
- **Hackl T.** (2014) *Forschungsdatenmanagement - Im Sinne der Nachhaltigkeit* In: CU Chemie UniHamburg, Die Mitarbeiterzeitung, Ausgabe 14, Dezember 2014, S. 18-19. PDF: http://www.chemie.uni-hamburg.de/zeitung/Mitarbeiterzeitung_17__Dez_2014.pdf
- **Ockenfeld M.** (2014) *FIZ Karlsruhe setzt auf Forschungsdatenverwaltung* In: Ball R (Ed.) b.i.t. online Heft 6/2014, b.i.t.online - Zeitschrift für Bibliothek, Information und Technologie, S. 586. PDF: <http://www.b-i-t-online.de/heft/2014-06-reportage-ockenfeld.pdf>
- **Gotzner P.** (2014) *Digital Amnesie | Wie lassen sich geschäftliche Daten oder private Fotos über Jahrzehnte speichern? Erstaunlich: Keiner weiß es – bisher* Wirtschaftswoche, Handelsblatt GmbH, Nr. 12, S 70.
- **Neumann J., Kraft A.** (2014) *Das RADAR Projekt – Research Data Repository*. Online: <http://www.forschungsdaten.org/index.php/RADAR>
- **Engel T., van den Broek K.** (2014) *Forschungsdaten sichern* GDCh Nachrichten aus der Chemie, Zeitschrift der Gesellschaft Deutscher Chemiker, 62:47. Online: <https://www.gdch.de/publikationen/nachrichten-aus-der-chemie.html>

2. Arbeits- und Ergebnisbericht

2.1 Ausgangsfragen und Zielsetzung von RADAR

Mit der Durchführung des DFG-geförderten RADAR - Research Data Repository - Projekts wurde eine generische Forschungsdateninfrastruktur geschaffen. RADAR ist ein Repository zur Archivierung und Publikation von Forschungsdaten. Zielgruppen von RADAR sind in erster Linie wissenschaftlich arbeitende Institutionen. Langfristiges Ziel ist es, mit RADAR einen Baustein für eine nationale Forschungsdateninfrastruktur zu etablieren. Das disziplinübergreifende Projektteam besteht aus folgenden Kooperationspartnern: FIZ Karlsruhe – Leibniz-Institut für Informationsinfrastruktur, das Steinbuch Centre for Computing (SCC) des Karlsruher Instituts für Technologie (KIT), die Ludwig-Maximilians-Universität München (LMU) mit der Fakultät für Chemie und Pharmazie, das Leibniz-Institut für Pflanzenbiochemie (IPB) in Halle und die Technische Informationsbibliothek (TIB) in Hannover. Der Betreiber des aus dem Projekt resultierenden Dienstes ist das FIZ Karlsruhe. Das SCC stellt die technische Basisinfrastruktur für den Betrieb und die Datenarchivierung bereit. Die Technische Informationsbibliothek ist über ein Memorandum of Understanding in den Betrieb eingebunden und bringt sich in den Bereichen Akquise, Beratung und strategische Weiterentwicklung ein.

2.2 Etablierung generischer RADAR-Dienstleistungen

Von Beginn an wurde RADAR als disziplinübergreifendes Repository konzipiert. Der Dienst bietet eine zentrale Anlaufstelle für die gesicherte Archivierung und Publikation vielfältiger Daten und Dateiformate mit einem zweistufigen Forschungsdaten-Dienstleistungsmodell: Die Archivierung der Daten mit und ohne Datenpublikation.

- Bei dem **Archivierungsangebot** werden die Daten in Form von paketierten Zusammenstellungen für eine vom Kunden gewählte Haltefrist (5-15 Jahre) langzeitarchiviert und erhalten einen eindeutigen Identifier. Nach Ablauf der Haltefrist kann der Kunde zwischen einer Verlängerung der Haltefrist und der Löschung der Daten wählen.
- Beim Angebot der **Datenpublikation** wird von RADAR eine unbegrenzte Haltefrist angestrebt, d.h. die Daten werden dauerhaft archiviert und nicht gelöscht. Faktisch wird eine Haltefrist von mindestens 25 Jahren kalkuliert. Die Auffindbarkeit der in RADAR publizierten Datenpakete wird über eine auf dem DataCite-Metadatenchema aufbauende Metadatenbeschreibung, die automatisierte Anbindung des DOI-Systems und die zur Verfügungsstellung der (Meta-)Daten via OAI-PMH und im national und international vernetzten Discovery-System des DataCite-Metadatenstore¹ sichergestellt.

Darüber hinaus bietet RADAR:

- Ein transparentes, auf jährlichen und Einmalzahlungen beruhendes Gebührenmodell.
- Die Möglichkeit des Peer-Reviews eines Datenpakets, z.B. im Rahmen eines Begutachtungsverfahrens, bei dem ein Manuskript bei einem wissenschaftlichen Journal eingereicht wird. Der Kunde erhält die Möglichkeit, die zugehörigen Daten über einen gesicherte Review-URL dem Journal-Editor und so auch potentiellen Gutachtern zugänglich zu machen. Nach Annahme des Artikels kann eine Verknüpfung von Textpublikation und zugrundeliegenden Forschungsdaten z.B. über den DOI von Artikel und zugehörigen Forschungsdaten erfolgen.
- Eine moderne und rollenbasierte Benutzungsoberfläche.
- Eine Programmierschnittstelle (API) zur Einbindung der Dienstleistungen aus RADAR in eigene Systeme.
- Die Integration der ORCID-Identifikation² in Metadaten und den Registrierungsprozess, um Autoren und anderen Mitwirkende eindeutig zu identifizieren und so die Verbindung von Forschern mit ihren Forschungsergebnissen herzustellen.

¹ <https://mds.datacite.org/>

² <http://orcid.org/>

- Die Integration der CrossRef FundRef-Funktion³, welche eine mensch- und maschinenlesbare Zuordnung von Fördergebern zu archivierten und publizierten Datenpaketen ermöglicht.
- Ein generisches Metadatenschema, das zu 100 % kompatibel zu DataCite⁴ und DublinCore⁵ Schemas ist.

Zur Sicherstellung eines zuverlässigen Betriebs wurden bei der Implementierung von RADAR wichtige Konzepte und Standards berücksichtigt. RADAR entspricht den Anforderungen eines OAIS⁶-konformen Langzeitarchivierungssystems. Entscheidend für die Langzeitarchivierung ist vor allem das Format der Archival Information Packages (AIP nach OAIS). Diese liegen als BagIt-Strukturen⁷ in ZIP-Containern vor, d.h. die Daten werden in wohldokumentierten Containern in einer u.a. von der Library of Congress entwickelten Struktur gespeichert. Jedes archivierte Datenpaket besteht aus einer einzelnen Datei, die in sich selbstbeschreibend ist (also Daten und beschreibende Metadaten enthält). Für die Interpretation der BagIt-Strukturen ist keine spezielle Software notwendig, die Inhalte sind sowohl maschinenlesbar wie auch durch Menschen interpretierbar. Alle Aktionen, die im Verlauf der Datenübernahme ins Archiv erfolgen, werden dokumentiert und in Form von PREMIS-Records abgespeichert.

Forschungsdatenmanagement, -archivierung und -publikation berühren eine Vielzahl komplexer Fragestellungen organisatorischer, technischer und juristischer Art. Um hier den Forschenden erste Hilfestellungen und Orientierung zu bieten, erstellte das Projektteam eine zweisprachige Projekt⁸- und Produktwebsite.⁹ Die dort verfügbaren Inhalte wie Glossar, Formatempfehlungen und Metadatendokumentationen wurden unter der Creative Commons Lizenz Namensnennung 4.0 International (CC BY 4.0) veröffentlicht. Viele Themen lassen sich nicht von einem einzelnen Projekt, sondern nur übergreifend angehen. Deshalb wurden Kooperationen mit verwandten nationalen und internationalen Initiativen aufgebaut, u.a. mit der Research Data Alliance¹⁰, EUDAT¹¹, SoWiDataNet¹² sowie INFO-MatWerk.¹³

2.3 Definition der Zielgruppen für RADAR und Entwicklung eines Geschäftsmodells

RADAR richtet sich an die Forschung im nationalen und zukünftig auch europäischen Rahmen. Dabei spricht RADAR vier Zielgruppen an: Projekte/Forschende, Institutionen, wissenschaftliche Verlage und Gedächtniseinrichtungen. In der Einführungsphase fokussiert RADAR zuerst auf die institutionellen Nutzer (Hochschulbibliotheken und außeruniversitäre Forschungseinrichtungen). Zu einem späteren Zeitpunkt folgen dann Projekte bzw. Forschende, die unabhängig von ihrer Einrichtung RADAR nutzen wollen. Mit wissenschaftlichen Verlagen wurden während der Projektlaufzeit Gespräche geführt und Workflows eruiert, jedoch nicht zuletzt auf Empfehlung des wissenschaftlichen Beirats (siehe Abschnitt 2.5) zugunsten übergreifender Lösungen (z.B. im Kontext der Research Data Alliance) zurückgestellt. Die Idee weitergehender Angebote für Gedächtniseinrichtungen wie Museen und Archive zur Archivierung und Bereitstellung von Digitalisaten kamen während der Projektlaufzeit auf, konnten jedoch noch nicht vertieft und ausspezifiziert werden. Die Abgrenzung zu anderen, bereits existierenden Angeboten zur Archivierung von Forschungsdaten erfolgt anhand zweier Kriterien. Zum einen ist RADAR im Unterschied zu fachspezifischen Forschungsdatenzentren disziplinübergreifend ausgerichtet.

³ <http://www.crossref.org/fundingdata/>

⁴ <https://www.datacite.org/>

⁵ <http://dublincore.org/>

⁶ http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=57284

⁷ <http://www.digitalpreservation.gov/documents/bagitspec.pdf>

⁸ <http://www.radar-projekt.org/>

⁹ Ab März 2016: <http://www.radar-service.eu>

¹⁰ <https://rd-alliance.org/node>

¹¹ <http://eudat.eu/>

¹² <https://sowidatanet.de/>

¹³ INFOMatWerk ist ein DFG-Projektantrag der Deutschen Gesellschaft für Materialkunde, DGM e.V und wird von RADAR unterstützt. Eine Nutzung der RADAR-Infrastruktur durch INFOMatWerk ist Bestandteil des INFOMatWerk-Antrags.

Zum anderen grenzt sich RADAR von vergleichbaren Initiativen wie ZENODO¹⁴, Dryad¹⁵ und figshare¹⁶ durch eine zentrale Datenspeicherung innerhalb Deutschlands und den hier gültigen rechtlichen Rahmen ab. Dadurch kann RADAR seinen Nutzern neben der Publikation auch eine gesicherte Datenarchivierung ohne öffentlichen Zugriff anbieten. Zudem unterstützt RADAR eine Vielzahl von Lizenzen für veröffentlichte Forschungsdaten und verlangt keine vollständige Freigabe aller Nutzungsformen über z.B. eine Creative-Commons-0-Lizenz.

Ein zentrales Ziel des Projekts war die Entwicklung eines nachhaltigen und **sich selbst tragenden Geschäftsmodells** für RADAR, das nach Aufnahme des Produktivbetriebs innerhalb von fünf Jahren Kostendeckung erreicht und einen dauerhaften Betrieb von RADAR ohne weitere Projektförderung sicherstellt. Das Geschäftsmodell sieht zwei unterschiedliche Produkte (Datenarchivierung mit und ohne Publikation) mit zwei Preismodellen (jährliche Zahlungen und Einmalzahlungen) vor. Die jährliche Zahlungsweise richtet sich an institutionelle Nutzer und funktioniert insbesondere für reine Datenarchivierung. Für die Publikation von Daten und (zu einem späteren Zeitpunkt) für nicht-institutionelle Nutzer empfiehlt sich die Einmalzahlung.

2.4 Disziplinübergreifende Evaluierung und Validierung der Dienstleistungen

Die Evaluation und Anpassung der generischen Funktionalität des Dienstes erfolgte durch in Projektverlauf neu gewonnene Kooperationspartner, v.a. außerhalb der Bereiche Chemie und Biochemie. Die in 2.2 genannten Dienstleistungen wurden durch Datenmodelle von Institutionen aus verschiedenen Fachdisziplinen geprüft. Nachfolgend sind exemplarisch drei Beispiele aufgeführt. Darüber hinaus wird das Testsystem von 23 Hochschulen und außeruniversitären Forschungseinrichtungen aus Deutschland und der Schweiz genutzt, um das System zu evaluieren und Feedback zu geben.

- Ingenieurwissenschaften: Deutsche Gesellschaft für Materialkunde e.V. (DGM)
Die DGM hat es sich mit dem DFG-Projektantrag „Sichern, Archivieren und Bereitstellen von Forschungsdaten im interdisziplinären Fachgebiet der Materialwissenschaft und Werkstofftechnik -INFO-MatWerk-“ zur Aufgabe gemacht, durch die Entwicklung von Workflows für disziplinspezifische Publikationsprozesse, Qualitäts- und Zitierstandards, wissenschaftlicher Leitfäden und fachspezifischen Metadatenprofilen geeignete Strukturen für die Archivierung und Publikation von Forschungsdaten innerhalb der Disziplin der Materialwissenschaft und Werkstofftechnik zu etablieren. Als technische Grundlage wurden vor Antragsstellung die in RADAR entwickelten, forschungsnahen Metadatenprofile und Verfahren zur systematischen, zitierfähigen Datenerschließung geprüft. Die DGM hat zum Ziel, ausgehend von den RADAR-Systemkomponenten, eine für die Materialwissenschaften optimierte Lösung zum Forschungsdatenmanagement zu entwickeln.
- Geisteswissenschaften: Caspar-David-Friedrich Institut (CDFI)
Das CDFI hat zur Evaluierung und Anpassung der RADAR-Datenstrukturen und des Metadatenschemas für geisteswissenschaftliche Fachrichtungen exemplarische Forschungsdaten aus dem Projekt „Historisches Raum-Zeitmodell der Stadt Nürnberg“ zur Verfügung gestellt.
- Sport- und Gesundheitswissenschaften: Fakultät für Sport- und Gesundheitswissenschaften der Technischen Universität München
Anhand von biomechanischen Forschungsdaten, die mehrfach verzweigte, anonymisierte Bewegungs- und Fallanalysen aus dem Leistungssport enthielten, konnte eine Erweiterung der RADAR-Datenstrukturen insbesondere im Hinblick auf die Einrichtung von Ordnerverzeichnissen und relevanten Metadatenbeschreibungen geschaffen werden.

¹⁴ <http://zenodo.org/>

¹⁵ <http://datadryad.org/>

¹⁶ <http://figshare.com/>

Eine weitere wichtige Quelle für die Validierung und für Feedback waren die drei öffentlichen RADAR-Workshops mit jeweils über 30 vertretenen Einrichtungen aus dem deutschsprachigen Raum.

2.5 Empfehlungen des wissenschaftlichen Beirats

Auf Anregung des Ausschusses für wissenschaftliche Bibliotheken und Informationssysteme (AWBI) wurde 2015 ein **wissenschaftlicher Beirat**¹⁷ eingeführt. Dieser wird durch Vertreter folgender Infrastruktureinrichtungen repräsentiert:

- Geoforschungszentrum Potsdam, Roland Bertelmann
- Helmholtz Open Science Koordinationsbüro, Dr. Heinz Pampel
- Verein zur Förderung eines Deutschen Forschungsnetzes (DFN) e.V., Ulrich Kähler
- GESIS - Leibniz-Institut für Sozialwissenschaften, Dr. Jonas Recker
- Deutsches Klimarechenzentrum GmbH, Dr. Michael Lautenschlager
- Universität Siegen, Volker Hess

Im Rahmen von zwei Beiratssitzungen wurde 2015 der Projektstand und 2016 das fertig entwickelte System evaluiert und bewertet. Zudem unterstützt der Beirat RADAR in rechtlichen Fragestellungen, z.B. bzgl. der Haftungsverpflichtung eines Datenrepositoriums, das einen kostenpflichtigen Dienst anbietet (vergl. 2.6.6 AP 6: Geschäftsmodell und Rahmenbedingung). Alle Mitglieder des wissenschaftlichen Beirats haben sich bereit erklärt, diese Funktion auch nach Projektende wahrzunehmen.

Nach zwei Beiratssitzungen wurde vom Beirat festgestellt, dass die zusätzlich umgesetzten Funktionen (z.B. Shibboleth-Integration über DFN-AAI, ORCID und FundRef-Anbindung, Peer-Review Funktion für Forschungsdaten, Einbindung der deskriptiven Metadaten als RDF/a auf Landing Pages der Datenpakete) den Anforderungen aus der Forschungsdatencommunity Rechnung tragen und an vielen Stellen über die im Projektantrag dargestellten Funktionen hinaus gehen.

In der zweiten Beiratssitzung kurz dem Ende der Projektlaufzeit am 10. August 2016 hat der Beirat u.a. folgende Kernempfehlungen zur weiteren Verstetigung RADARs gegeben, die auch nach Projektende von Projektpartnern eingebunden wurden bzw. werden:

- Die umfangreichen Funktionalitäten des Gesamtsystems RADAR sollen vor der den Zielgruppen, der Forschungsdatencommunity und den Drittmittelgebern noch stärker beworben werden.
- Zunächst sollte sich RADAR bei der Kundengewinnung auf Deutschland konzentrieren, im nächsten Schritt dann auf den deutschsprachigen Raum. Sofern deutscher Gerichtsstand und deutsche Rechtsprechung anwendbar sind, kann in einem dritten Schritt eine Öffnung von RADAR für europäische Kunden erfolgen.
- Es wird eine Positionierung und Zusammenarbeit mit den 2016 gestarteten bzw. geplanten Vorhaben GeRDI und LEIBNIZ DATA empfohlen. RADAR soll zu beiden Projekten frühzeitig informative Gespräche anstreben.
- Der Beirat empfiehlt, dass RADAR möglichst schnell online gehen sollte. Die Einbeziehung eines zweiten Rechenzentrums neben dem SCC könnte die Haftungsrisiken, die in Abschnitt 2.6.6 AP 6: Geschäftsmodell und Rahmenbedingung dargestellt werden und einen derzeitigen Onlinegang noch verhindern, minimieren. Die Mitglieder des GeRDI-Konsortiums bieten sich hierfür als primäre Absprechpartner an. Allerdings bedingt die Einbindung eines weiteren Rechenzentrums zusätzliche Arbeiten, die den Start verzögern und die Kosten steigen lassen.
- Die Gründung einer eigenständigen Rechtsform für den Betrieb von RADAR ist problematisch und langwierig; ebenso ist eine Übernahme des Dienstes durch DFN

¹⁷ <http://www.radar-projekt.org/display/RD/Kontakt#Kontakt-Beirat>

schwierig, da hier nur das Haftungsproblem verschoben wird. Alternative Ansätze, z.B. von PANGAEA mit einer freiwilligen Zahlung, oder ein kostenloses Angebot würden eine komplette Überarbeitung des RADAR-Geschäftsmodells bedeuten, und wären voraussichtlich nicht zielführend. Letztlich erscheint dem Beirat die Idee, die Haftung durch entsprechende Vertragsgestaltung auf den Kunden zu begrenzen und die Datenkuratoren als Erfüllungsgehilfen zu betrachten, die im Auftrag des Kunden Datenpakete einstellen, als am erfolgsversprechendsten. Damit würde das Haftungsrisiko deutlich reduziert werden.

- Das zukünftige Kommunikations- und Marketingmodell RADARs soll zwei regelmäßige Komponenten enthalten: Zum einen eine RADAR-Tagung für Nutzer und interessierte Personen, z.B. einmal jährlich, und zum anderen die Bildung eines Kundenbeirats, bei dem RADAR bewusst Kunden auswählt, von denen man sich konstruktives Feedback erwartet.

2.6 Durchgeführte Arbeitsschritte und Ergebnisse im Berichtszeitraum

2.6.1 AP 1: Projektmanagement

Die Zusammenarbeit zwischen den RADAR-Projektpartnern fand im Rahmen wöchentlicher Webkonferenzen, über ein projektinternes Wiki als Kommunikations- und Arbeitsplattform sowie in halbjährlichen Projekttreffen statt. RADAR wurde im Projektzeitraum auf **21 nationalen Workshops und Konferenzen** vorgestellt und mit den dort anwesenden wissenschaftlichen und informationstechnischen Experten diskutiert. Alle drei geplanten Workshops wurden im Projektzeitraum durchgeführt. Die Workshops führten zu direktem Feedback der RADAR-Anwender und wurden in der Systementwicklung aufgegriffen. Darüber hinaus konnten **23 Institutionen als Testnutzer** gewonnen werden. 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.

Die Dokumentation der Workshops ist im Rahmen der Projekthomepage verfügbar¹⁸. Zur weiteren Bekanntmachung des Projekts wurden 13 Print- und Online-Artikel (vergl. Abschnitt 1.3) durch fachwissenschaftliche, informationstechnische und wirtschaftliche Medien kommuniziert, um einen möglichst breiten Kunden- und Interessentenkreis zu erreichen. Die Projektwebsite stellt Informationen in übersichtlicher Form zur Verfügung. Sie wird auch nach Projektende weiter aktualisiert und langfristig durch das Online Portal des Dienstes (www.radar-service.eu) abgelöst.

2.6.2 AP 2: Anforderungsanalyse

Die Analyse der **fachwissenschaftlichen Anforderungen** an eine Dienstleistung zur Datenarchivierung und -publikation wurde zu Projektbeginn durchgeführt. Sie umfasste sowohl einen Vergleich mit etablierten internationalen Dienstleistern als auch eine Analyse bestehender nationaler, überwiegend disziplinspezifischer Forschungsdatenrepositorien.

Während sich disziplinspezifische Datenbanken meist nur auf einen sehr eingegrenzten Fachbereich (z.B.: PDB¹⁹ für Proteinstrukturen; Birmingham Metabolite Library²⁰ für Metabolomics) mit einem spezifischen Nutzerkreis fokussieren, bieten generische Forschungsdatenzentren wie ZENODO, figshare und Dryad eine fachübergreifende Datenspeicherung mit Publikation und Digital Object Identifier (DOI)-Vergabe²¹ an. ZENODO und figshare bieten dabei kostenlose Angebote für private Nutzer mit limitierten Volumina an. Diese Repositorien werden im Rahmen des EU-Projekts OpenAire²² bzw. durch das Unternehmen

¹⁸ <https://www.radar-projekt.org/display/RD/News>

¹⁹ <http://www.pdb.org>, Protein Data Bank

²⁰ <http://www.bml-nmr.org/>

²¹ Digital Object Identifier (DOI). Siehe auch <http://www.doi.org/>

²² OpenAire Projekt (EU FP7). Vgl. auch <https://www.openaire.eu/>

Digital Science²³ finanziert. Digital Science bietet zudem kostenpflichtige, institutionelle Services an, z.B. für Verlage. Dryad bietet eine Unterstützung für die Veröffentlichung von zu einer Publikation gehörenden Daten mit „peer-review“ an, sowie die Implementierung von Standards für die Datenarchivierung mit einer Qualitätskontrolle über Kuratoren. Dryad ist dabei für den Nutzer i.d.R. kostenpflichtig. Im Bereich der Sozialwissenschaften werden Forschungsdaten bereits erfolgreich durch die GESIS - Leibniz-Institut für Sozialwissenschaften e.V. - archiviert und publiziert²⁴. Darüber hinaus wurden fachspezifische Anforderungsanalysen aus dem Bereich Chemie und Biochemie von den Partnern des IPB und der LMU durchgeführt. Im Vergleich zu den bestehenden Alternativen ergab die Anforderungsanalyse die Forderungen nach den folgenden, zentralen Eigenschaften an RADAR. Diese Eigenschaften unterscheiden RADAR, vor allem in ihrer Gesamtheit, von den zuvor genannten Angeboten. Gleichzeitig wird das System somit für den Einsatz im primärem Kundenkreis der Hochschulen und außeruniversitären Forschungseinrichtungen in Deutschland optimiert:

- Die **Wahl eines zweistufigen Dienstleistungsmodells** mit dem Angebot Archivierung und Publikation (s. 2.2) und dem entsprechenden Rollen- und Rechtemanagement trägt einem flexiblen Management der anfallenden Forschungsdaten Rechnung. So können bei der Archivierung Daten aus dem laufenden Forschungsprozess wie auch Daten abgeschlossener Projekte, Master- und Promotionsarbeiten u.ä. sicher, langfristig und rückholbar gespeichert werden, ohne sofort veröffentlicht zu werden. Ein klassischer Anwendungsfall für das Archivierungsangebot sind z.B. nicht qualitätsgesicherte Rohdaten. Das Angebot Publikation zielt auf alle Daten, die eigenständig oder in Zusammenhang mit einer Textpublikation öffentlich zugänglich gemacht und zitierbar sowie für andere nachnutzbar sein sollen. Dies umfasst sowohl Daten, die bereits zu wissenschaftlichen Erkenntnissen geführt haben als auch noch nicht interpretierte Daten, die der Wissenschaftler aber der Allgemeinheit zugänglich machen will, insbesondere in Hinblick auf eine mögliche spätere Verwendung für neue Fragestellungen oder neue methodische Ansätze.
- Die **Datenspeicherung nach deutschem Recht**, mit allen zugehörigen Ansprüchen für Kunden (z.B. deutscher Gerichtsstand, Urheberrecht und Haftung und Absicherung im Verlustfall).
- Die **Speicherung von Forschungsdaten in Form von paketierte Zusammenstellungen** einer oder mehrerer Dateien („Datenpaketen“) nach internationalen Standards (*BagIt*, s. Abschnitt 2.2).
- Durch ein **abgestuftes Rechtemanagement** soll ausgewählten Nutzern, wie z.B. Kooperationspartnern oder Gutachtern, Zugriff auf die Daten gewährt werden. Damit ist dieses Angebot besonders geeignet für die Speicherung von Originaldaten entsprechend den Regeln der guten wissenschaftlichen Praxis, für interne Daten, für Daten aus institutionsübergreifenden Kooperationen, für Positiv- und Negativ-Daten aus abgeschlossenen Projekten und für Daten, die noch für Publikationen oder Patente verwendet werden sollen.
- Da eine Qualitätskontrolle der hinterlegten Daten aufgrund ihrer Heterogenität und des generischen Charakters von RADAR nur durch die jeweilige Fach-Community erfolgen kann, soll eine **Peer-Review Funktion für Forschungsdaten** geschaffen werden. So kann, z.B. im Rahmen eines Einreichungsprozesses, bei dem ein Manuskript bei einem wissenschaftlichen Journal eingereicht wird, auch das zugehörige Datenpaket durch einen Dritten anonym begutachtet werden.

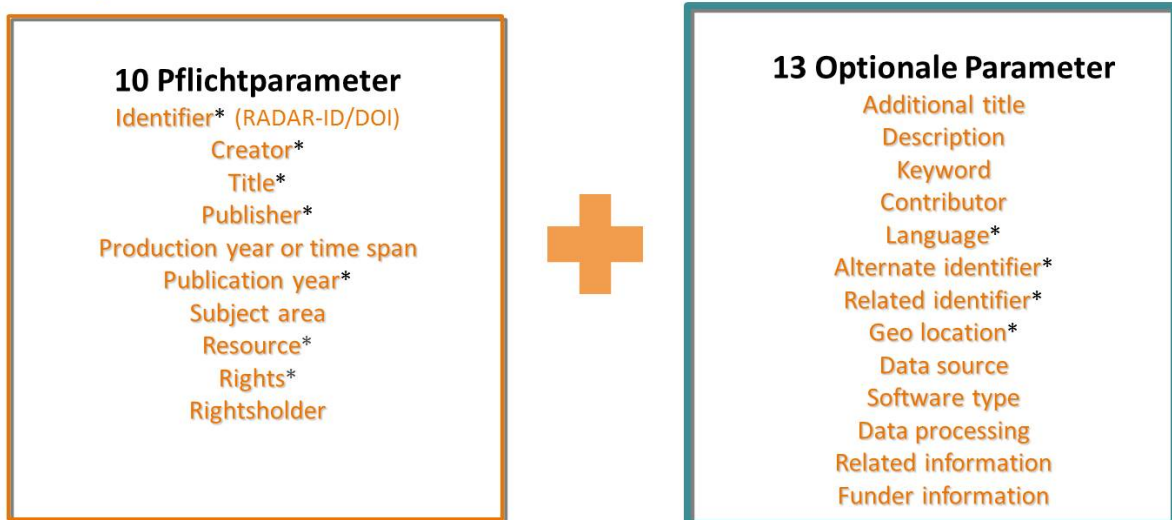
²³ <https://www.digital-science.com/>

²⁴ <http://www.gesis.org>

2.6.3 AP 3: Metadatenprofile

Die zentrale Aufgabe des AP 3 war die Erstellung eines **Metadatenprofils, das sowohl einen interdisziplinären, zentralen Nachweis der in RADAR archivierten und publizierten Forschungsdaten erlaubt als auch die fachspezifischen Anforderungen zur Suche und zur Nachnutzung dieser Daten erfüllt**. Dazu wurden im Anschluss an die Anforderungsanalyse in Anlehnung an das DataCite-Metadatenchema v4.0 und an geeignete fachspezifische NMR und 2D-DIGE-Daten Metadatenparameter für die geplanten RADAR-Dienstleistungsangebote identifiziert und ausgewählt. Diese Metadaten schemata umfassen die Definition von 10 Pflichtfeldern, welche zusammen den allgemeinen, deskriptiven Teil des Metadatenprofils bilden, sowie die Definition, Abstimmung und Anpassung von 13 optionalen Feldern, welche die fachspezifischen Beschreibungen der Datensätze abbilden (Abb. 1). Die Pflichtfelder des entwickelten Metadatenprofils enthalten die Grundanforderungen für eine DOI-Registrierung und stellen so die Integration in den Workflow der Datenpublikation im Rahmen des AP 5 sicher. Um die heterogenen Ansprüche verschiedener wissenschaftlicher Fachgebiete und die generische Ausrichtung von RADAR zusammenzubringen, wurde bei der Definition der 13 optionalen Metadatenfelder mit einer Kombination von kontrollierten Listen und Freitextfeldern gearbeitet. Bei der Definition der kontrollierten Listen wurde auf weltweit anerkannte, verständliche Standards (z.B. ISO-Normen für die Sprache und das Entstehungsland der Forschungsdaten) zurückgegriffen.

Adaptives RADAR-Metadatenchema



Ziel: Metadaten, die das Datenpaket hinreichend beschreiben

* basiert auf  DataCite
 Metadata Kernel v4.0

Abb. 1: Pflichtparameter und optionale Felder des entwickelten RADAR-Metadatenchemas. Das Schema bietet eine volle Kompatibilität zu etablierten Standards wie DataCite und DublinCore. Die Metadaten der bei RADAR publizierten Daten stehen unter einer CC0-Lizenz und sind u.a. über den DataCite-Metadatenstore abrufbar. Des Weiteren werden die Metadaten zum Harvesting via OAI-PMH zur Verfügung gestellt.

Das RADAR-Metadatenchema bildet zusammen mit dem RADAR-Glossar und dem FAQ-Bereich die Grundlage, um Wissenschaftlern als Anwender des Online-Portals bei der nachhaltigen Beschreibung eines detaillierten Datenpakets zu unterstützen. Nach der Fertigstellung des entwickelten Metadatenprofils wurde dieses in eine XML Schema Definition (XSD) mit Anwendungsbeispielen umgesetzt. Darauf basierend wurde im Online-Portal eine Eingabemaske implementiert. Die weiteren Arbeitsschritte in diesem AP umfassten die Evaluierung des entwickelten Metadatenprofils im laufenden Forschungsbetrieb (zusammen mit AP 7) durch die wissenschaftlichen Partner und durch weitere RADAR-Stakeholder aus anderen

Fachdisziplinen Materialwissenschaften (DGM e.V.), den Sport- und Gesundheitswissenschaften (TU München) und den Geisteswissenschaften (CDFI) (vergl. 2.4). Das RADAR-Metadatenchema²⁵ und die zugehörige Dokumentation²⁶ sind online verfügbar, die Dokumentation findet sich zusätzlich im Anhang.

Das Metadatenchema wurde zusätzlich im letzten Halbjahr der Projektlaufzeit auf das im September 2016 gestartete DataCite-Metadatenchema v4.0²⁷ angepasst. Hiermit wird die zukünftige Anschlussfähigkeit des RADAR-Systems an die für Forschungsdaten immer wichtiger werdende DOI-Registrierung gesichert. Die umfassenden Anpassungen beinhalteten u.a. die Erweiterung der Pflichtfelder und optionalen Felder mit der Ergänzung der **ORCID-Funktionalität** (vergl. AP 5) und der Angabe von Förderinstitutionen und Referenzierenden. Die Metadaten der publizierten Daten stehen unter einer CC0-Lizenz und sind u.a. über den DataCite-Metadatenstore abrufbar. Des Weiteren werden die Metadaten zum Harvesting via OAI-PMH zur Verfügung gestellt.

2.6.4 AP 4: Datenmanagement

Das AP 4 umfasste die technische Anforderungs- und Designanalyse sowie die Erstellung einer Spezifikation für die Datenmanagementschicht. In der Ausarbeitung der Anforderungen an RADAR wurden ca. 50 Use Cases identifiziert. Diese Anwendungsfälle beschreiben die geplanten Nutzungsabläufe der jeweiligen Zielgruppen für das Produktivsystem. Dabei wurde von Beginn an eine auf die **OAIS konforme Implementierung des Langzeitarchivierungssystems** geachtet. Ausgehend von der Use Case-Analyse wurden Wireframes entworfen und in einem iterativen Prozess mit den Projektpartnern abgeglichen. Neben der Use Case-Analyse wurde ein Datenmodell auf Basis der Erkenntnisse aus dem EU-Projekt SCAPE²⁸ (SCAPE Digital Object Model) evaluiert und für RADAR angepasst, da dieses Datenmodell speziell im Bereich der Langzeitarchivierung großer Repositorien entwickelt wurde. Es basiert auf den beiden Standards PREMIS²⁹ und METS³⁰.

Eine der wichtigsten Faktoren bei der Archivierung von Daten ist die Verlässlichkeit der Technologie, die zum Einsatz kommt. Um die Nachhaltigkeit zu gewährleisten ist das System modular aufgebaut (Abb. 2).³¹ Es besteht aus einer Managementsoftware mit Programmierschnittstellen (API) zur Benutzeroberfläche und zur Speicherschicht. Für die Implementierung wurde überwiegend auf Open Source-Komponenten zurückgegriffen. Die Managementsoftware setzt auf Komponenten wie Cassandra³² und Elasticsearch³³ auf. Hier werden Forschungsdaten als Datensätze erfasst, auf ihre technische Integrität überprüft und mit Metadaten versehen. Die Benutzungsoberfläche wurde in Groovy on Grails³⁴ implementiert.

²⁵ <https://www.radar-projekt.org/download/attachments/753673/radar-schema-v09.zip>

²⁶ https://www.radar-projekt.org/download/attachments/753673/AP3_RADAR_Metadata_Documentation_v05.pdf

²⁷ <http://schema.datacite.org/>

²⁸ <http://www.scape-project.eu/>

²⁹ <http://www.loc.gov/standards/premis/>

³⁰ <http://www.loc.gov/standards/mets/>

³¹ Potthoff J, van Wezel J, Razum M, Walk M (2014) Anforderungen eines nachhaltigen, disziplinübergreifenden Forschungsdaten-Repositoriums. In: Müller P, Neumair B, Reiser H, Rodosek GD (Hrsg.), Proceedings 7. DFN-Forum Kommunikationstechnologien, Lecture Notes in Informatics, GI, S. 11-20

³² <http://cassandra.apache.org/>

³³ <https://www.elastic.co/de/products/elasticsearch>

³⁴ <https://grails.org/>

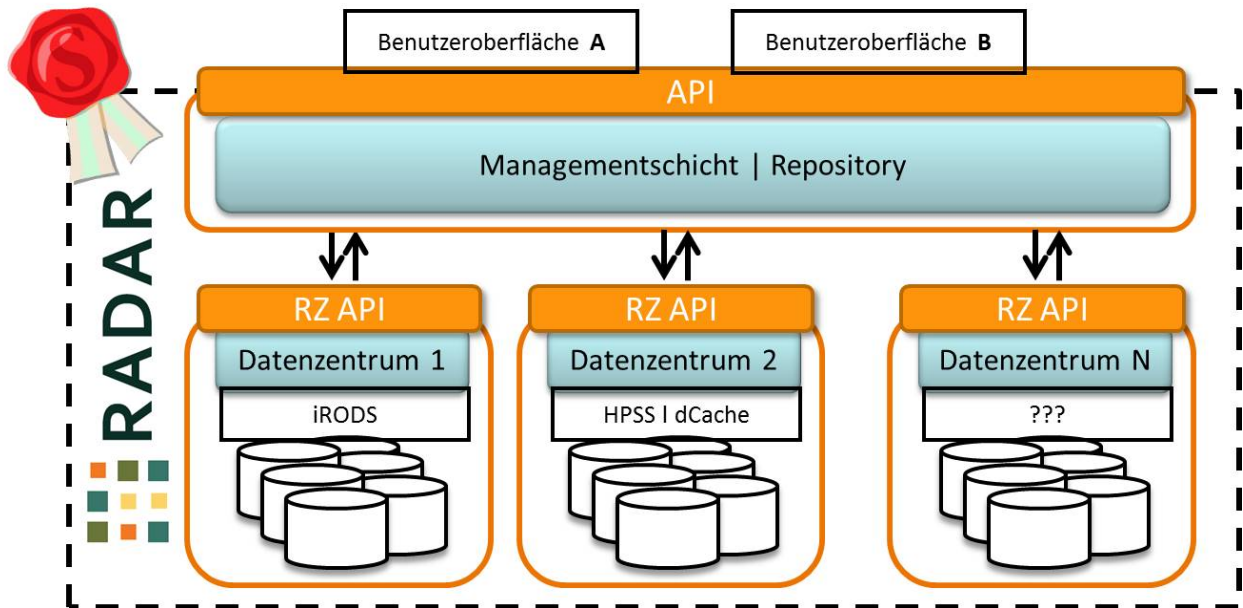


Abb. 2: Schematische Darstellung der RADAR-Systemarchitektur mit Daten-Ingest, externen Schnittstellen zu Benutzern und Systemen und der internen Schnittstelle zwischen Management- und Speicherschicht.

Die Authentifizierung ist sowohl über eine lokale Datenbank als auch über delegierte Verfahren wie z.B. Shibboleth möglich. Ein entsprechender Vertrag zur Einbindung von RADAR in die DFN-AAI-Föderation wurde mit dem DFN-Verein abgeschlossen. Die Datenübertragung zur Speicherschicht erfolgt über standardisierte Protokolle (*SFTP*, *GridFTP*), während die Datenhaltung selbst auf modernen Platten- und Bandspeicher-Systemen erfolgt. Die Basis der RADAR-Architektur bildet eine erweiterbare Schnittstelle für die Archiv Storage Services (Programmierschnittstelle zum Rechenzentrum; in Abb. 2: „RZ API“). Diese Schnittstelle soll die Nutzung verschiedener Storage Services vor dem Frontend verbergen, so dass die Nutzung eines speziellen Speicherverwaltungssystems nicht verpflichtend ist. Damit ist es möglich, auch weitere Rechenzentren in die RADAR-Infrastruktur einzubinden.

Auf der Ebene der Speicherschicht (Abb. 3) sorgt ein *High-Performance Storage System* (HPSS) für einen effizienten Daten-Zugriff und erlaubt eine kostengünstige Langzeitarchivierung. Die Zuverlässigkeit des Systems wird durch Mehrfachkopien gesichert, die an unterschiedlichen, geographisch verteilten Standorten abgelegt werden. Beim Schreiben der Daten und mit jedem Lesezugriff werden die Daten auf Integrität überprüft. Zusätzlich wird die Integrität von der Übergabe der Daten an den Archivspeicher bis zur Speicherung auf Band überwacht. 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. Über einen temporären Speicher (Festplatte), der wiederum über die genannten Sicherheitsmechanismen abgesichert ist, ist eine synchrone Bearbeitung der Daten und Metadaten durch den Nutzer gewährleistet. Die Daten selbst liegen als *BagIt*-Strukturen³⁵ 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.

³⁵ http://www.digitalpreservation.gov/news/2008/20080602news_article_bagit.html Library Develops Specification for Transferring Digital Content.

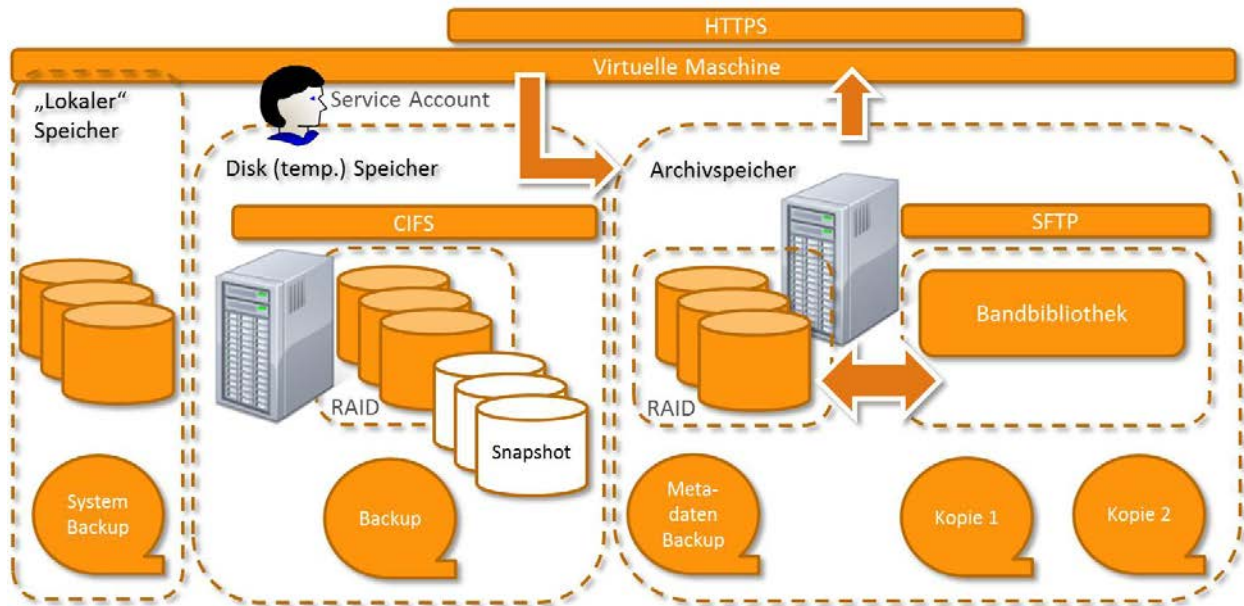


Abb. 3: Schematische Darstellung der RADAR-Speicherinfrastruktur mit der Darstellung des Datentransfers zwischen Archivdiensten „Lokaler Speicher“, „Disk bzw. temporärer Speicher“ und des „Archivspeichers“ als Endpoint.

Um die Einbindung von RADAR-Funktionen in bestehende Applikationen und die Implementierung von alternativen Benutzungsoberflächen zu ermöglichen, wurde eine RADAR Programmierschnittstelle (**Application Programming Interface, API**) geschaffen. Sie ist **REST-basiert** und enthält einen vollständigen Funktionsumfang. Der Zugriff kann über HTTP mit JSON als Austauschformat erfolgen.

Auch der gemischte Einsatz von API und Benutzungsoberfläche ist möglich. Beispielsweise kann die effiziente und automatisierte Übertragung von Daten über das API erfolgen, während die für die Verwaltung und Anreicherung der Daten weiterhin die Benutzungsoberfläche zum Einsatz kommt. Im Rahmen des Arbeitspakets wurde eine detaillierte API Dokumentation und Demonstration erstellt, die verfügbar ist.³⁶

2.6.5 AP 5: Datenpublikation

Die (traditionelle) Publikation im Kontext der wissenschaftlichen Forschung umfasst meist zwei zentrale Eigenschaften: Zum einen die öffentliche Verfügbarkeit des publizierten Mediums für einen unbegrenzten Zeitraum; zum anderen impliziert Publikation, dass das Medium zitierfähig ist, d.h. durch die Angabe einer Referenz eindeutig auf das Originalwerk verwiesen werden kann. Während für wissenschaftliche Aufsätze diese Eigenschaften meist durch die Veröffentlichung in einer Fachzeitschrift gewährleistet sind, folgt der Publikationsprozess von Forschungsdaten - soweit dieser innerhalb eines Forschungsprojekts überhaupt vorgesehen ist- meist individuellen Richtlinien einzelner Fachbereiche, Institutionen oder Arbeitsgruppen. In AP 5 wurde daher ein Workflow erstellt, der forschenden Institutionen die Möglichkeit gibt, eine nachhaltige, öffentlich zugängliche, zitierfähige und optional auch einschlägig qualitätsgeprüfte (peer-review) Datenpublikation über RADAR durchzuführen. Zur Implementation des Publikationsworkflows konnte auf das in AP 4 entwickelte **Rollen- und Rechtemanagementsystem** (Tab. 1) zurückgegriffen werden.

³⁶ <https://www.radar-projekt.org/display/RE/RADAR+API+Documentation>.

Tab. 1: Rollen- und Datenmanagementkonzept für die Kunden 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) Datenpaket	Eine Zusammenstellung digitaler Daten, die zusammengehörig einen Forschungsprozess bzw. dessen Ergebnis beschreiben. Ein Datenpaket 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 Datenpakete	
1.) In Bearbeitung: <i>Pending</i>	Initialer Status; (Sub-)Kuratoren können das Datenpaket bearbeiten (Modifizierung, Update, Löschung). Der Pending Status ist auf 6 Monate begrenzt.
2.) In Begutachtung: <i>Review</i>	Datenpaket wurde durch Kurator zur Begutachtung (Peer Review) bei einem Verlag/Editor eingereicht und ist temporär gesperrt; es erhält eine sichere, temporäre 'Review-URL'.
3.) Archiviert: <i>Archived</i> (= Service: Archivierung)	Datenpaket ist archiviert und mittels einer RADAR-ID eindeutig identifizierbar. In diesem Status ist keine weitere Bearbeitung möglich. Kuratoren dürfen ausgewählten Nutzern Leserechte erteilen.
4.) Publiziert: <i>Published</i> (= Service: Publikation)	Datenpaket ist publiziert und mittels eines DOI eindeutig identifizierbar und referenzierbar. Über eine öffentlich zugängliche Landingpage, die deskriptive und technische Metadaten anzeigt, kann das Datenpaket heruntergeladen werden. In diesem Status ist keine weitere Bearbeitung möglich.

Nach dem Rollen- und Rechtemanagement entscheiden die vom Kunden eingesetzten **Kuratoren**, welche Daten publiziert werden. Im Publikationsangebot bietet RADAR einen kombinierten Service aus Datenpublikation und Archivierung an. RADAR führt dabei eine funktionale Qualitätskontrolle durch, in dem die angegebenen Metadatenpflichtfelder (vergl. AP 3, Abb. 1) auf Vollständigkeit geprüft werden. Des Weiteren wird der Kurator bei der inhaltlichen Datenbeschreibung durch Hilfetexte und Beispieldateien, die über das RADAR-Frontend

verfügbar sind, als auch über umfangreiche Anleitungen aus dem RADAR-Glossar und den FAQ unterstützt.

Durch die Verwendung von *Digital Object Identifier* (DOI) als persistente Identifikatoren wird sichergestellt, dass jeder publizierte Datensatz eigenständig zitierfähig ist. Nach der Veröffentlichung sind die RADAR-Publikationen Open Access, mit der Möglichkeit eines optionalen Embargo-Zeitraums von bis zu einem Jahr. Die zugehörigen Metadaten sind über Schnittstellen, z.B. OAI-PMH harvestbar und neben dem RADAR-Schema auch im Format des DataCite-Schemas und des DublinCore-Schemas abrufbar.

Um die Datenpublikationen zusammen mit Manuskriptveröffentlichungen im Kontext von RADAR für die entsprechenden Fachbereiche attraktiver zu machen wurde Rücksprache mit mehreren Verlagen und dem Beirat gehalten – diese haben im Projektverlauf von der im Erstantrag beschriebenen Lösung einer direkten Plug-In Funktion für wissenschaftliche Manuskripteinreichungssysteme abgeraten. Es sollte keine Insellösung für ein einzelnes Repository wie RADAR geschaffen werden. Stattdessen wurde eine allgemein verfügbare **Peer-Review Funktion** für die in RADAR hochgeladenen Datenpakete implementiert. Bei diesem Modell wird dem Kurator eine gesicherte URL zur Verfügung gestellt, die an Editoren und Gutachter kommuniziert werden kann. Über diesen URL-Zugang können Editoren und Gutachter auf die Entwurfsversion des Datenpaketes ohne eine weitere Authentifizierung zugreifen. Nach Abschluss des Review-Prozesses kann das Datenpaket ggf. nochmals bearbeitet oder direkt publiziert und mit dem zugehörigen wissenschaftlichen Artikel verlinkt und durchsuchbar gemacht werden.

Des Weiteren wurde der Datenpublikationsworkflow um die **ORCID-Identifikation** erweitert: Im Rahmen der Kompatibilitätsanpassung mit dem im Herbst 2016 startenden DataCite-Metadatenchema ist es auch im Frontend und über die API möglich, Autoren und anderen Mitwirkende bei der Erstellung eines Datenpakets eindeutig zu identifizieren und so die Verbindung von Forschern mit ihren Forschungsergebnissen herzustellen. Darüber hinaus wird den Kuratoren ermöglicht, bei einer Datenpublikation den zugehörigen Förderer bzw. Drittmittelgeber eindeutig zuzuordnen (Integration der **CrossRef-Funktion FundRef**). Dies ermöglicht eine menschen- und maschinenlesbare Zuordnung von Fördergebern zu archivierten und publizierten Datenpaketen. Weitere in AP 5 durchgeführte Arbeitsschritte beinhalteten den Aufbau von Beratungsfunktionen wie Formatempfehlungen, Lizenz-Erläuterungen und -Empfehlungen sowie die Bereitstellung eines umfangreichen Glossars und FAQ.

2.6.6 AP 6: Geschäftsmodell, rechtliche und organisatorische Rahmenbedingung

In der Ausschreibung zum Call „Überregionale Informationsservices“ legte die DFG ausdrücklichen Wert auf den Aufbau eines nachhaltigen Dienstes. Für RADAR war deshalb von Anfang an die Entwicklung eines von Projektförderung unabhängigen Geschäftsmodells ein zentrales Ziel. Neben der Erarbeitung eines solchen Geschäftsmodells standen noch die Ausarbeitung der rechtlichen Rahmenbedingungen (etwa Verträge und Nutzungshinweise), Untersuchungen zur bilanziellen Betrachtung von Einmalzahlungen, zur EU Trennungsrechnung sowie zur Haftung auf der Agenda.

Zielgruppen

RADAR sieht primär drei Zielgruppen für den Dienst vor: institutionelle Nutzer (etwa Hochschulen und außeruniversitäre Forschungseinrichtungen), Forschende und Projektgruppen sowie wissenschaftliche Verlage. Mittelfristig können noch Gedächtniseinrichtungen wie Museen und Archive hinzukommen, die eine Archivierungs- und Zugriffsplattform für Digitalisate benötigen. Das Projektteam diskutierte diese Zielgruppen und ihre jeweiligen Bedürfnisse im Rahmen der drei durchgeführten Workshops mit potenziellen Nutzern, mit dem wissenschaftlichen Beirat sowie auch mit Verlagsvertretern.

Die Gespräche mit Verlagsvertretern offenbarten, dass eine nicht projekt- bzw. datenzentrumsübergreifende Herangehensweise zu einer asymmetrischen Verhandlungsposition und so gegebenenfalls zu wissenschaftspolitisch unerwünschten Ergebnissen (etwa einer Monetarisierung von Forschungsdaten ohne adäquate Beteiligung an den entstehenden Kosten für die notwendige Infrastruktur) führen könnte. Gleichzeitig sind wissenschaftliche Verlage überwiegend als gewinnorientierte Unternehmen zu betrachten, die aufgrund der Maßgaben der EU Trennungsrechnung zu einer anderen Preiskalkulation zwängen. Deshalb führten auch pragmatische Gründe zu einer Zurückstellung dieser Zielgruppe bei der Entwicklung des initialen Geschäftsmodells. Bei den Forschenden und Projektgruppen sieht das Projektteam zwar einen Bedarf an einer Dienstleistung, wie RADAR sie anbietet. Aber auch hier existieren pragmatische Gründe, die vorerst zu einer Zurückstellung der Zielgruppe geführt haben:

- ein hoher Akquiseaufwand, da bei der Zielgruppe RADAR weitgehend unbekannt ist;
- ein erhöhtes Haftungsrisiko, da dieses mit der Anzahl der abgeschlossenen Nutzungsverträge steigt;
- ein erhöhter Beratungsbedarf, da die Bibliotheken der Einrichtungen als Mittler wegfallen.

Damit fokussiert RADAR im initialen Geschäftsmodell auf die Zielgruppe der institutionellen Nutzer. Mittelfristig ist angedacht, zumindest die Forschenden und Projektgruppen als weitere Zielgruppe aktiv zu bedienen.

Laut destatis gibt es in Deutschland zurzeit 955 außeruniversitäre Forschungseinrichtungen und 426 Hochschulen, von denen geschätzt 250 aktiv forschen.³⁷ Damit umfasst die primäre Zielgruppe national ca. 1.200 potenzielle Nutzer. Viele dieser potenziellen Nutzer arbeiten an eigenen Lösungen, beteiligen sich an konsortialen Ansätzen oder nutzen freie Angebote wie ZENODO oder EUDAT. Damit ist die tatsächliche Menge potenzieller Nutzer geringer. Auch deswegen ist eine Ausdehnung auf das europäische Ausland zu einem späteren Zeitpunkt angedacht. Grundsätzlich denkbar ist langfristig auch ein Angebot für die forschende Industrie; hierzu gab es erste Interessensbekundungen, die aber zum jetzigen Zeitpunkt nicht weiter verfolgt wurden.

Betriebskosten

Die Betriebskosten teilen sich in fixe und variable Kosten auf. Fixkosten fallen unabhängig von der Nutzung von RADAR an, variable Kosten entstehen ausschließlich bei der Speicherung von Daten durch die Nutzer. Die gesamte IT-Infrastruktur stellt das SCC des KIT bereit. Einen Teil dieser Infrastruktur, etwa die Applikationsserver, sind den Fixkosten zuzuordnen, der größere Teil der Dienstleistung umfasst aber die eigentliche Speicherung der Daten in den Bandarchiven. Die dafür anfallenden Kosten sind vollständig den variablen Kosten zuzurechnen.

Für den Betrieb von RADAR als disziplinübergreifendes Repository sind zwei Vollzeitstellen vorgesehen: eine Produktmanagerin für Marketing, Akquise, Nutzerbetreuung und Produktfortschreibung sowie ein Software-Entwickler für Systemadministration, Softwarepflege und Weiterentwicklung. Beide Positionen sind in EG 13 eingruppiert. Hinzu kommen Sachmittel für Reisekosten, für das Dienstangebot benötigte IT-Ressourcen, Informationsmaterialien und Lizenzen. Die folgende Tabelle 2 führt die anfallenden Fixkosten auf (gerundete Beträge). Bei den Personalkosten handelt es sich um Vollkosten, d.h. inklusive Gemeinkostenzuschlägen.

³⁷ <https://www.destatis.de/DE/ZahlenFakten/GesellschaftStaat/BildungForschungKultur/Hochschulen/Tabellen/HochschulenHochschularten.html>

Tab. 2: Darstellung der Fixkosten

Position	Kosten pro Jahr
Produktmanagerin	132.000,00 €
Software-Entwickler	132.000,00 €
Lizenzen und Wartungsverträge	1.000,00 €
Reisekosten	10.000,00 €
Informationsmaterialien	5.000,00 €
Virtuelle Maschinen für Betrieb der Applikation	10.500,00 €
Temporärer Speicher für Betrieb der Applikation	1.500,00 €
Sonstiges	5.000,00 €
Summe Fixkosten	297.000,00 €

FIZ Karlsruhe als Betreiber von RADAR sieht die Archivierung und Publikation von Forschungsdaten als Teil seines Auftrags. Deswegen erklärt es sich bereit, die Stelle des Software-Entwicklers aus seiner Grundfinanzierung zu übernehmen. Damit reduzieren sich die für den Betrieb von RADAR notwendigen Kosten um 132.000,00 € auf **165.000,00 € pro Jahr**.

Die variablen Kosten umfassen lediglich die Kosten für die Speicherung der Forschungsdaten in den Bandarchiven des SCC. Hierfür berechnet das SCC aktuell 0,15 € pro Gigabyte/Jahr bei einer Speicherung von zwei Kopien an räumlich getrennten Standorten. Dieser Preis enthält bereits alle Kosten des SCC (Personal, Energie, Geräte, Bänder, Netzwerk, Zuschläge, etc.). Da FIZ Karlsruhe diese Leistung beim SCC extern einkauft, sind auf diesen Preis noch 21% Verwaltungsgemeinkosten aufzuschlagen. Somit ergibt sich ein variabler Kostenanteil von **0,18 € pro Gigabyte/Jahr**.

Preisfindung

Ziel ist es, in den ersten 5 Jahren des RADAR-Betriebs insgesamt 30 Verträge mit Hochschulen und außeruniversitären Forschungseinrichtungen abzuschließen und dabei ein Gesamtvolumen von einem Petabyte an archivierten bzw. publizierten Daten zu speichern. Entsprechend werden die Fixkosten von 165.000,00 € auf 1 Petabyte Datenvolumen³⁸ umgelegt, um zu diesem Zeitpunkt Kostendeckung zu erreichen. Damit ergibt sich die in Tabelle 3 dargestellte Preiskalkulation.

Tab. 3: Preiskalkulation für die Speicherung von Forschungsdaten in RADAR

Position	Kosten pro Gigabyte/Jahr
Fixkosten / Datenvolumen Break-even	0,17 €
Variable Kosten (Speicherung)	0,18 €
Risikozuschlag (25 % auf Fixkosten)	0,04 €
Umsatzsteuer	0,08 €
Preis pro Gigabyte/Jahr	0,47 €

Grundsätzlich strebt RADAR eine möglichst gleichmäßige und für die nutzenden Einrichtungen planbare Kostenentwicklung an. Da die Archivierung von Forschungsdaten langfristig angelegt ist, geht das Geschäftsmodell von einer jährlichen Preissteigerung von 2 % aus. Eine Ausnahme bildet die eigentliche Speicherung der Daten, bei der es von einer jährlichen Preissenkung von 5 % ausgeht. Selbstverständlich handelt es sich dabei um Projektionen, die sich auf die rückwirkende Betrachtung von Preisentwicklungen und Literaturrecherche stützt. Um unerwartete Kostensteigerungen abfedern zu können, enthält die Kalkulation daher noch einen Risikozuschlag. Weiterhin ist FIZ Karlsruhe verpflichtet, auf seine Leistungen die gesetzliche Mehrwertsteuer in Höhe von zurzeit 19 % aufzuschlagen.

³⁸ Der Einfachheit halber rechnen sowohl SCC wie FIZ Karlsruhe bei allen Angaben mit 1.000.000 Gigabyte für ein Petabyte

Im dritten RADAR-Workshop Anfang August 2016 hat das RADAR-Team dieses Preismodell vorgestellt und mit den Teilnehmern diskutiert. Grundlage des Modells ist immer ein laufender Vertrag mit der nutzenden Einrichtung. Während dies für archivierte Daten als adäquat betrachtet wurde, kam sehr deutliche Kritik hinsichtlich publizierter Daten auf. Die Teilnehmer forderten einer langfristigen Garantie für die Verfügbarkeit der Daten durch den Betreiber – auch über das Ende der Laufzeit eines Vertrags hinaus. Das ist aber mit einer jährlichen Zahlungsweise nicht zu leisten. Entsprechend hat das RADAR-Team mit Hochdruck an einem zweiten Preismodell für Einmalzahlungen gearbeitet. Dabei konnte es auf bereits geleistete Vorarbeiten zurückgreifen, da solche Einmalzahlungen für die Zielgruppe Forschende/ Projektgruppen ohnehin angedacht waren, wenn auch zu einem deutlich späteren Zeitpunkt.

Bei der Preiskalkulation für Einmalpreise ging das Projektteam von folgenden Annahmen aus: eine Haltefrist für publizierte Daten von mindestens 25 Jahre, eine jährliche allgemeine Preissteigerung von 2 % und eine jährliche Preissenkung für die Speicherung von 5 %. Auf dieser Basis entstand eine Fortschreibung der Kosten über 25 Jahre. Darüber hinaus wurde der Personalanteil der Fixkosten (Produktmanagerin) nur für die ersten fünf Jahre berücksichtigt, da nur in der Anfangszeit Beratungs- noch Akquisekosten anfallen. Aufgrund der sehr langen Haltefrist und der damit verbundenen großen Unsicherheit bei der Prognose der Preisentwicklung kommt ein sehr hoher Risikoaufschlag hinzu. Tabelle 4 zeigt die Preiskalkulation für die Einmalzahlung.

Tab. 4: Preiskalkulation für Einmalzahlungen bei einer mindestens 25-jährigen Haltefrist

Position	Kosten pro Gigabyte
Fixkosten (Personalkosten nur in den ersten 5 Jahren)	1,13 €
Variable Kosten (Speicherung)	2,97 €
Risikozuschlag (200 % auf Fixkosten)	2,27 €
Umsatzsteuer	1,21 €
Preis pro Gigabyte für 25+ Jahre	7,58 €

Das Angebot von Einmalzahlungen steht allerdings noch vor dem Vorbehalt einer Genehmigung durch den Aufsichtsrat von FIZ Karlsruhe. Insbesondere die langfristige Verpflichtung wirft komplexe Fragen auf. Für die Unterhaltsträger von FIZ Karlsruhe dürfen keine Risiken oder Verpflichtungen entstehen, sollte FIZ Karlsruhe nicht mehr in der Lage sein, die langfristigen Verpflichtungen zu übernehmen. Auch dürfen die Vorauszahlungen nicht dazu führen, dass die Grundfinanzierung von FIZ Karlsruhe (Fehlbedarfsfinanzierung) durch die Vorauszahlungen vermindert wird. Zu diesen Fragen beauftragte FIZ Karlsruhe seinen Wirtschaftsprüfer Rödl & Partner mit einem Gutachten, das das Vorhaben untersuchte und Stellung zur bilanziellen Behandlung und zur Auswirkung auf die Fehlbedarfsfinanzierung umfasste. Darüber hinaus wurden auch noch Fragen des Beihilferechts/der EU Trennungsrechnung und des Vergaberechts untersucht, die sich auf das gesamte Angebot inklusive jährlicher Zahlungsweise beziehen.

Laut des Gutachtens lassen sich die Vorauszahlungen (in Form der Einmalzahlungen) als passiver Rechnungsabgrenzungsposten verbuchen. Dies hat auch keine Auswirkungen auf die Fehlbedarfsfinanzierung von FIZ Karlsruhe. Damit einhergeht, dass die Vorauszahlungen ausschließlich zum Zweck der vorgesehenen Leistung verwendet werden können. Sollte FIZ Karlsruhe nicht mehr in der Lage sein, die versprochenen Leistungen in Zukunft zu erbringen, kann diese durch einen anderen Dienstleister erbracht und dessen Leistung aus den verbleibenden Vorauszahlungen bezahlt werden. Diese Vorsorge ist ein essentieller Teil der Exit-Strategie. Ein weiterer Baustein ist die Absprache mit andern Datenzentren, die im Fall eines Exits bereit wären, über die Übernahme der Daten und die Fortführung der versprochenen Leistung zu verhandeln. Hierzu fanden erste Gespräche statt, die kurzfristig in Form eines Letters of Intent formalisiert werden sollen.

In Vorbereitung der Entscheidung des Aufsichtsrats hat sich der wissenschaftliche Beirat von FIZ Karlsruhe inzwischen für eine Aufnahme des Betriebs von RADAR mit jährlichen Zahlungen und

Einmalzahlungen ausgesprochen und auf die Wichtigkeit eines solchen Angebots für die Wissenschaft hingewiesen. Neben den Preisen für volumenabhängige Datenarchivierung bzw. -publikation sieht das Preismodell noch eine Grundgebühr pro Nutzungsvertrag in Höhe von 500,00 € vor. Ein vorhandener Vertrag ist Voraussetzung sowohl für die Datenarchivierung wie für die -publikation. Die Grundgebühr deckt die Nutzung des temporären Speichers (200 GB), das Vertragsmanagement, die Nutzerverwaltung, die Rechnungsstellung sowie einen Anteil an einer Versicherungspolice zur Absicherung von Haftungsrisiken ab.

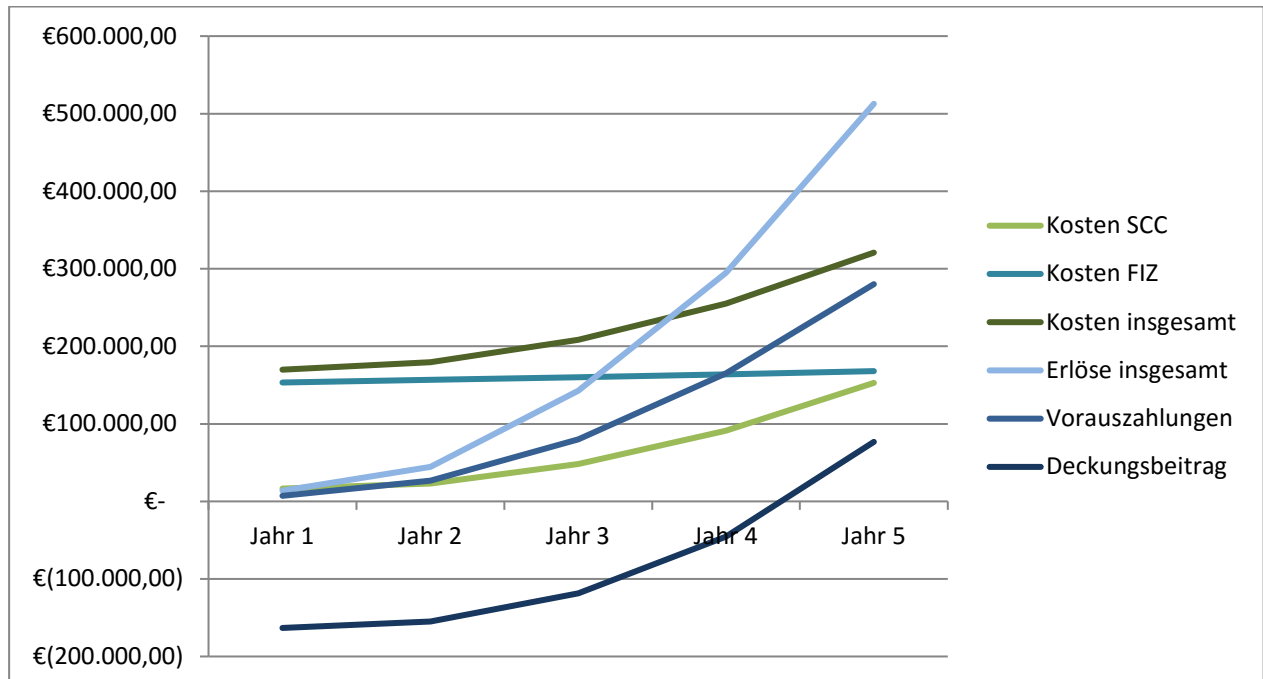


Abb. 4: Prognostizierte Kosten und Erlöse des Geschäftsmodells sowie Deckungsbeitrag mit Break-even nach knapp fünf Jahren (unterste Kurve).

Haftung

Innerhalb der deutschen Forschungsdatenlandschaft hat die Etablierung eines nachhaltigen, auf Basis von Nutzungsentgelten sich selbst tragenden Repositoriums für primär institutionelle Nutzer in Deutschland auch 2016 einen ausgeprägten Pilotcharakter. So waren bis zum Abschluss des RADAR Projekts im August 2016 keine vergleichbaren Dienstleistungen und respektive Geschäftsmodelle verfügbar. Mit Stand vom Oktober 2016 arbeiten die RADAR-Projektpartner mit Nachdruck an der auch juristisch abgesicherten, nachhaltigen Vertragsgestaltung eines solchen Repositoriums. Dabei hat sich im Projektverlauf vor allem die in Deutschland oftmals unsichere Rechtslage hinsichtlich des Kundenanspruchs auf Datenverfügbarkeit und Haftung im Schadensfall als zentrale Herausforderung erwiesen.

Die Ausgangslage stellt sich wie folgt dar: Bei einem kostenpflichtigen Angebot lässt sich nach „AGB-Recht“ kein Haftungsausschluss für sogenannte Kardinalpflichten erreichen. Zu den Kardinalpflichten gehören bei einem Datenarchiv die sichere und dauerhafte Speicherung der anvertrauten Daten. Im Falle eines Datenverlusts wäre also das Repositorium schadensersatzpflichtig. Es gibt keinerlei Anhaltspunkte, wie in einem solchen Fall der Wert der Forschungsdaten und damit die Schadenshöhe zu bemessen wäre. Zugleich wären sowohl die nutzende Institution wie auch der Datengeber schadensersatzberechtigt. Abbildung 5 zeigt die bei der Nutzung von RADAR relevanten Rollen und deren Vertrags- und Nutzungsverhältnisse.

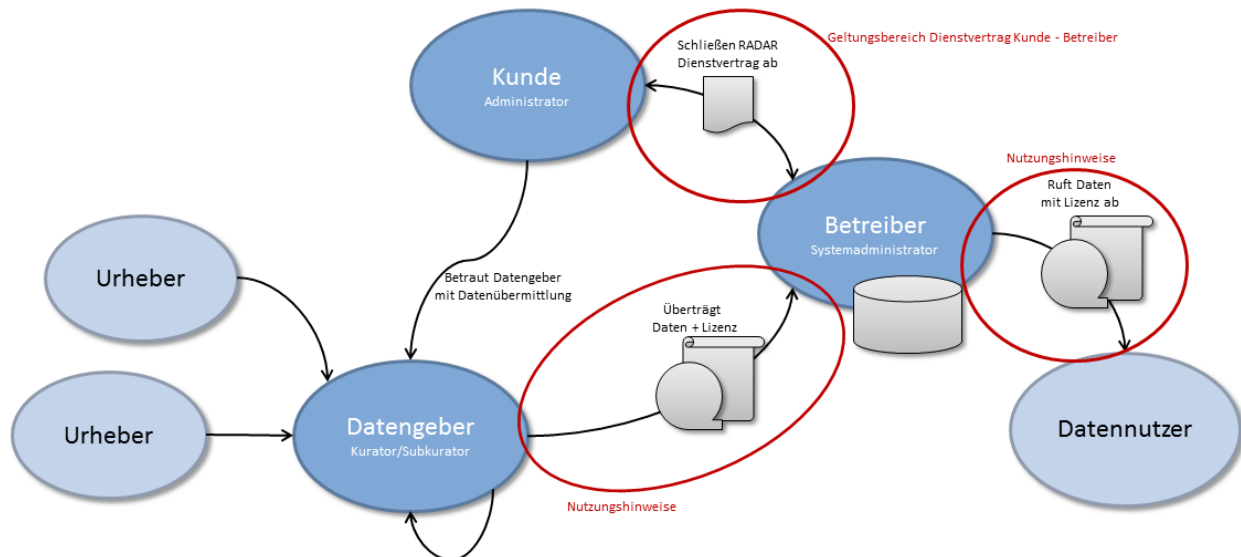


Abb. 5: Rollen, Vertrags- und Nutzungsverhältnisse bei RADAR

Durch eine entsprechende rechtliche Gestaltung kann ein Vertragsverhältnis zwischen Datengeber und Betreiber vermieden werden. Es entsteht nur ein Vertragsverhältnis zwischen Betreiber und Kunde (nutzende Institution). Diese wiederum ermöglicht ihre Forschenden oder Mitarbeiter, Daten in RADAR einzustellen. Der Betreiber haftet in diesem Konstrukt nur gegenüber der nutzenden Institution, nicht aber gegenüber den Datengebern. Die Schadenshöhe ist aber immer noch unbestimmt. Eine Begrenzung könnte über eine freie Vertragsverhandlung erfolgen, die aber in der Praxis nahezu unmöglich gerichtsfest nachzuweisen ist. Praktikabler erscheint der Ansatz, mehrere Tarife zur Auswahl anzubieten, bei denen sich die Höhe der Haftung im Preis niederschlägt. Der Kunde kann bei der Vertragsgestaltung zwischen verschiedenen Haftungssummen wählen (z.B. 50.000 €, 100.000 € oder 500.000 €) und zahlt dafür eine entsprechend angepasste Jahresgrundgebühr. Dieses Modell wird gerade rechtlich geprüft. Gleichzeitig erfolgt eine Prüfung, inwieweit sich FIZ Karlsruhe als Betreiber gegen Haftungsrisiken versichern kann und darf. Ungelöst ist bisher das Problem, dass das SCC als eigentlicher Leistungserbringer von einer Versicherung im Schadensfall in Regress genommen würde. Hierzu laufen aktuell Verhandlungen mit mehreren Versicherungen und dem KIT.

Entsprechend dieser mit Stand dieses Abschlussberichtes noch nicht abschließend geklärten Rechtsfragen und der noch ausstehenden Genehmigung des Geschäftsmodells durch den Aufsichtsrat von FIZ Karlsruhe ist der Dienst noch nicht online verfügbar. Im Anhang findet sich der letzte Stand des Vertrags zwischen nutzender Institution und FIZ Karlsruhe als Betreiber sowie die Nutzungshinweise für Datengeber und Datennutzer.

Zertifizierung:

Eine Zertifizierung ist für die Beurteilung der Vertrauenswürdigkeit eines Repositoriums essentiell. Im Rahmen der Projektlaufzeit hat sich für Datenrepositorien das Data Seal of Approval als bewährte Zertifizierung etabliert³⁹. Während im Juni 2014 schon 31 wissenschaftliche Datenzentren durch ein DSA-Siegel zertifiziert waren, sind es im September 2016 bereits 62.⁴⁰ Aufgrund der in Juni 2016 gestarteten Harmonisierung des DSA-Zertifizierungsverfahrens mit den Richtlinien des World Data System (WDS) und dem einhergehenden Aufnahmestopp von Anträgen war die Durchführung des Prozesses innerhalb der RADAR Projektlaufzeit nicht möglich. Sobald die Einreichung von Anträgen seitens des DSA Konsortiums wieder freigegeben ist (voraussichtlich gegen Ende 2016), wird RADAR den Zertifizierungsprozess starten.

³⁹ <http://www.datasealofapproval.org/en/>

⁴⁰ <http://www.datasealofapproval.org/en/assessment/>

2.6.7 AP 7: Evaluierung

Die Evaluierung der entwickelten Dienstleistungen fand projektbegleitend seit Januar 2014 auf mehreren Ebenen statt:

- Durch neu gewonnene Kooperationspartner, u.a. der Deutschen Gesellschaft für Materialkunde e.V. (DGM), des Caspar-David-Friedrich Instituts (CDFI) und der Fakultät für Sport- und Gesundheitswissenschaften der Technischen Universität München (vergl. Kapitel 2.4).
- Durch Nutzer des RADAR-Testsystems seit seiner Vorstellung im Rahmen des 2. RADAR Workshops im Juni 2015. Mittlerweile wird das RADAR Testsystem von 23 Hochschulen und außeruniversitären Forschungseinrichtungen aus Deutschland und der Schweiz genutzt.
- Durch drei Stakeholder-Workshops in Karlsruhe und Frankfurt/Main (vergl. AP 1).
- Durch interne Produkttests, die auf Seite des RADAR-Frontends regelmäßig durch alle Projektpartner, auf Ebene der RADAR-API durch das IPB und bzgl. der RADAR-Speicherschicht v.a. durch das SCC durchgeführt wurden.

Mit Stand dieses Abschlussberichtes wurden im projektinternen Ticketsystem 1226 Tickets angelegt und davon 1167 durch die Softwareentwickler des FIZ Karlsruhe bearbeitet. Als Resultat der Evaluationen wurden die Software des Management-Systems und die Schnittstelle zur Speicherinfrastruktur des SCC kontinuierlich verbessert und weiterentwickelt. Mit Abschluss der Projektlaufzeit wurden insgesamt sechs Software-Iterationen durchgeführt und als Testsystem dem potentiellen Kundenkreis zur Verfügung gestellt.

2.7 Erfahrungen über angewandte Methoden und Nachnutzungsmöglichkeiten

Die im Projektverlauf erarbeiteten Ergebnisse sind von vornherein auf Nachhaltigkeit und damit auch auf Nachnutzbarkeit angelegt. Neben dem Angebot des RADAR-Repositoryums für die wissenschaftliche Community sind folgende Ergebnisse bereits während der Projektlaufzeit auf großes Interesse gestoßen und wurden für Dritte bereitgestellt bzw. mit diesen diskutiert:

- Metadaten-Schema
- Geschäftsmodell
- Rechtliche Rahmenbedingungen und Verträge (Arbeitsstände)
- Umgang mit Haftung

Es bleibt festzuhalten, dass insbesondere die rechtlichen Fragen sehr viel komplexer waren und sind, als dies das Projektteam bei der Antragstellung annahm. Entsprechend konnte es die Arbeiten zu diesem Thema bis zum Ende des Projekts nicht abschließend klären. Das Projektteam und der zukünftige Betreiber FIZ Karlsruhe sind aber zuversichtlich, diese Fragen bis Ende 2016 so weit zu klären, dass ab 2017 RADAR an den Start gehen kann. Entsprechend wurden beide vorgesehenen Stellen zwischenzeitlich besetzt.

2.8 Eigenleistung

Die Arbeitsplätze des RADAR-Kernteam und die zur Entwicklung des RADAR-Test- und Produktivsystems notwendigen Speicher- und Rechenkapazitäten wurden durch die Projektpartner als Eigenleistung zur Verfügung gestellt. Für die Projektvorstellung auf Konferenzen wurde seitens der TIB der Druck von 250 Flyern finanziert. Die Software-Entwicklung bei FIZ Karlsruhe wurde zum überwiegenden Teil in Eigenleistung erbracht. Bis Projektende flossen hier etwas mehr als 60 Personenmonate als Eigenanteil ein. Bei der Gestaltung und rechtlichen Prüfung der Verträge und Nutzungshinweise wurde der Justiziar von FIZ Karlsruhe, Mitarbeiter des Bereichs Immaterialgüterrechte in verteilten Informationsinfrastrukturen sowie eine externe Rechtsanwaltskanzlei als Eigenleistung einbezogen. Darüber hinaus erstellte der Bereich Immaterialgüterrechte in verteilten Informationsinfrastrukturen ein Rechtsgutachten zu

Fragestellungen. Die Prüfung vergaberechtlicher Fragen, der bilanziellen Betrachtung der Vorauszahlungen durch Einmalzahlungen und von Aspekten der EU Trennungsrechnung erfolgten durch ein Gutachten der Wirtschaftsprüfer von FIZ Karlsruhe, das ebenfalls in Eigenleistung beauftragt wurde.

2.9 Öffentlichkeitsarbeit

Vorträge und Poster:

- Präsentation: Kraft A (2016) *RADAR - Ablage und Nachnutzung von Forschungsdaten*. 15. SciFinder hbz Treffen, 10.03.2016, Hamburg
- Präsentation: Razum M (2016) *Archivierung und Publikation von Forschungsdaten mit RADAR*. DPG Jahrestagung, Arbeitsgruppe Information, 07.03.2016, Regensburg
- Präsentation: Kraft A (2015) *RADAR - Research Data Repository*. Treffen der niedersächsischen Forschungsreferent/-innen zum Thema „Forschungsdatenmanagement“, 11.11.2015, Hannover
- Posterbeitrag: van den Broek K, Furtado F, Engel T (2015) *RADAR – A Research Data Repository for the “Long-Tail of Science”*. GCC 2015 - 11th German Conference on Chemoinformatics, 08-10.11.2015, Fulda
- Präsentation: Razum M, Engel T, Kraft A (2015) *RADAR – DIGITAL DATA REPOSITORY FOR THE ‚LONG TAIL‘* Wiley-VCH - RADAR Projekt Informationstreffen, 28.10.2015, Weinheim
- Präsentation: Kraft A, Engel T (2015) *RADAR – Research Data Repository: FDM-Dienstleistungen, Organisation & technische Infrastruktur*. Innovationskonferenz „eResearch und Forschungsdaten“ der Kommission Virtuelle Bibliothek (KVB) im BibliotheksVerbund Bayern (BVB), 16.10.2015, München
- Präsentation: Kraft A, Pothoff J (2015) *RADAR - Ein Datenrepositorium für den ‚Long Tail of Research Data Data‘*. Open-Access-Tage, 7-8.9.2015, Zürich. Weitere Informationen: <http://open-access.net/community/open-access-tage/open-access-tage-2015-zuerich/>
- Präsentation: Kraft A (2015) *RADAR - A repository for long tail data*. 36th Annual IATUL Conference, 5-9.7.2015, Hannover. Weitere Informationen: <http://www.iatulconference2015.org/>
- Präsentation: Razum M (2015) *RADAR. Wo steht das Projekt?* Deutschen Bibliotheksverband - Sektion 4: Fühjahrssitzung 2015, 28.-29.4.2015, Darmstadt. Weitere Informationen: <http://www.bibliotheksverband.de/fachgruppen/sektionen/sektion-4/aktivitaeten/sitzungen.html>
- Präsentation: Razum M (2014) *RADAR Project Data Archival and Publication as a Service*. LizaRDs Workshop, SWITCH, Zürich. Weitere Informationen: <https://www.switch.ch/uni/projects/lizard/>
- Posterbeitrag: Kraft A, Hahn M, Potthoff J (2014) *Securing data for the future – the RADAR project. 2nd Data Management Workshop*, 28.-29.11.2014, Köln. Weitere Informationen: <http://www.tr32db.uni-koeln.de/workshops/overview.php?wsID=4>
- Präsentation: Kraft A (2014) *RADAR: Interdisziplinäres Forschungsdatenarchiv*. Research Data Alliance Deutschland-Treffen, 20.-21.11.2014, Potsdam. Weitere Informationen: <https://europe.rd-alliance.org/Content/Events.aspx?id=222>
- Posterbeitrag: Furtado F, Porzel A, Wessjohann L (2014) *RADAR: Research Data Repository*. GDCh-NMR-Diskussionstagung, 29.09-02.10.2014, Berlin
- Präsentation: Kraft A (2014) *RADAR project: Data preservation and publication*. DataCite Annual Conference 2014, 25.-26.08.2014, Nancy, Frankreich. Weitere Informationen: <http://datacite.inist.fr/?Programme>
- Präsentation: Hahn M (2014) *RADAR Project: Data Preservation and Publication*. Research Data Management - Organizational, Legal and Technical Challenges, 15.09.2014, Karlsruhe

- Präsentation: Potthoff J (2014) *Anforderungen eines nachhaltigen, disziplinübergreifenden Forschungsdaten-Repositories*. 7. DFN-Forum Kommunikationstechnologien, 16.-17.06.2014, Fulda
- Präsentation: Neumann J, Razum M (2014) *Das RADAR-Projekt: Datenarchivierung und -publikation als Dienstleistung - disziplinübergreifend, nachhaltig, kostendeckend*. 103. Bibliothekartag: ‚Bibliotheken: Wir öffnen Welten‘, 03.-06.06.2014, Bremen
Weitere Informationen: <http://www.bibliothekartag2014.de>
- Präsentation: Razum M (2014) *RADAR*. Deutschen Bibliotheksverband - Sektion 4: Frühjahrssitzung 2014, 29.-30.04.2014, Berlin
- Präsentation: Furtado F, Porzel A, Wessjohann L (2014) *RADAR - Research Data Repository: an overview*. GDCh-NMR-Diskussionstagung, Praktische Probleme der Kernresonanzspektroskopie, 13.-14.01.2014, Erlangen
- Präsentation: Hauser R, Razum M (2013) *Bitstream Preservation: Zentrale Dienstleistungen und Kostenmodelle*. FIZ Karlsruhe, 22.10.2013, Karlsruhe
- Präsentation: Kraft A (2013) *Datenarchivierung & -publikation als Dienstleistung: Aufbau eines Research Data Repositoriums (RADAR)*. 8. Sitzung der Arbeitsgruppe „Biodiversitätsdaten“ der DFG-Senatskommission für Biodiversitätsforschung, 05.-06.12.2013, Bremen

Online-Material:

- Projektwebsite: www.radar-projekt.org
- Produkt: www.radar-service.eu
- Übersicht Workshop-Berichte und Bekanntmachungen: <https://www.radar-projekt.org/display/RD/News>
- 1. Stakeholder-Workshop: 16.09.2014 am Zentrum für Kunst und Medientechnologie (ZKM), Karlsruhe, 29 Teilnehmer.
- 2. Stakeholder-Workshop: 23.06.2015 bei der Gesellschaft für Chemische Technik und Biotechnologie e.V. (DECHEMA), Frankfurt am Main, 24 Teilnehmer.
- 3. Stakeholder-Workshop: 06-07.06.2016 bei der Gesellschaft für Chemische Technik und Biotechnologie e.V. (DECHEMA), Frankfurt am Main, 35 Teilnehmer.
- Dokumentation des deskriptiven RADAR Metadatenschemas, Version 0.5: https://www.radar-projekt.org/download/attachments/753673/AP3_RADAR_Metadata_Dokumentation_v05.pdf
- Dokumentation der RADAR REST-API: <https://www.radar-projekt.org/display/RE/RADAR+API+Documentation>

3. Zusammenfassung

Dieser Abschlussbericht stellt die während der dreijährigen Projektlaufzeit durchgeführten Prozessschritte des Projekts und die daraus resultierenden Funktionen des Dienstes RADAR (Research Data Repository) vor. In der Forschungsdatencommunity vereint RADAR einen disziplin- und institutsübergreifenden Ansatz bei der Etablierung einer nachhaltigen Infrastruktur für Forschungsdaten. RADAR hat zum Ziel, das in vielen wissenschaftlichen Disziplinen noch fehlende Forschungsdatenmanagement zu fördern und einen wichtigen Beitrag zu einer besseren Verfügbarkeit, nachhaltigen Bewahrung und (eigenständigen) Publikationsfähigkeit von Forschungsdaten zu leisten. Als Forschungsdaten werden in diesem Kontext digitale Daten verstanden, die je nach Fachkontext Gegenstand eines Forschungsprozesses sind, während eines solchen Prozesses entstehen oder sein Ergebnis sind.

Als nationales, generisches End-Point-Repository soll die institutionelle Nutzung von RADAR vor allem dann erfolgen, wenn kein fachspezifisches Datenarchiv in dem betreffenden Fachbereich existiert oder eine Archivierung bzw. Veröffentlichung der Daten innerhalb einer generischen Infrastruktur als sinnvoll angesehen wird. Die sich daraus ergebenden RADAR-Zielgruppen umfassen primär interdisziplinär arbeitende Institutionen aus dem ‚Long Tail‘-Bereich wissenschaftlicher Forschung. Dazu gehören bereits fachübergreifend arbeitende Institutionen wie z.B. (Universitäts-)Bibliotheken und Vereine, aber auch wissenschaftliche Verlage und Kultureinrichtungen.

RADAR wird nach Projektende von FIZ Karlsruhe – Leibniz-Institut für Informationsinfrastruktur angeboten („Betreiber“). Der Dienst richtet sich in der Startphase erst einmal an Hochschulen und außeruniversitäre Forschungseinrichtungen („Kunden“), die ihren Forschenden („Datengebern“) die Möglichkeit zur Datenarchivierung und -publikation anbieten wollen. In RADAR gespeicherte Daten können von Dritten („Datennutzern“) durchsucht und heruntergeladen werden, sofern die Datengeber dem zugestimmt haben.

RADAR speichert Forschungsdaten grundsätzlich in Form von paketierte Zusammenstellungen einer oder mehrerer Dateien. Diese Zusammenstellung enthält die eigentlichen Forschungsdaten sowie beschreibende Daten (Metadaten) und wird als „Datenpaket“ bezeichnet. RADAR ermöglicht die dauerhafte und unverfälschte Aufbewahrung von Datenpaketen über eine definierte Zeitspanne. Die Aufbewahrungsfrist wird durch den Datengeber bestimmt. Während der Aufbewahrungsfrist verändert RADAR die gespeicherten Datenpakete nicht mehr, sondern sichert ausschließlich deren physischen Erhalt („bitstream preservation“). Dabei kann RADAR über eine webbasierte Benutzungsoberfläche mit aktuellen Webbrowsern oder über eine REST-basierte Programmierschnittstelle (API) genutzt werden.

Weitere Informationen und die Dienstleistung sind online unter www.radar-projekt.org und www.radar-service.eu verfügbar.