

Bausteine Forschungsdatenmanagement
Empfehlungen und Erfahrungsberichte für die Praxis von
Forschungsdatenmanagerinnen und -managern

Bedarfsgesteuerte Entwicklung einer Forschungsdateninfrastruktur am Beispiel des generischen Repositoriums RADAR

Kerstin Soltauⁱ

Dorothea Streckerⁱⁱ

2019

Zitiervorschlag

Soltau, Kerstin und Dorothea Strecker. 2019. Bedarfsgesteuerte Entwicklung einer Forschungsdateninfrastruktur am Beispiel des generischen Repositoriums RADAR. *Bausteine Forschungsdatenmanagement. Empfehlungen und Erfahrungsberichte für die Praxis von Forschungsdatenmanagerinnen und -managern* Nr. 2/2019: S. 68-75. DOI: [10.17192/bfdm.2019.2.8162](https://doi.org/10.17192/bfdm.2019.2.8162).

Dieser Beitrag steht unter einer
[Creative Commons Namensnennung 4.0 International Lizenz \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/).

ⁱFIZ Karlsruhe - Leibniz-Institut für Informationsinfrastruktur. ORCID: [0000-0002-6368-1929](https://orcid.org/0000-0002-6368-1929)

ⁱⁱFIZ Karlsruhe - Leibniz-Institut für Informationsinfrastruktur. ORCID: [0000-0002-9754-3807](https://orcid.org/0000-0002-9754-3807)

1 Abstract

RADAR¹ (Research Data Repository) bietet akademischen Einrichtungen und Forschenden eine generische Infrastruktur zur langfristigen Archivierung und Publikation digitaler Forschungsdaten. Das Forschungsdatenrepositorium wird von FIZ Karlsruhe – Leibniz-Institut für Informationsinfrastruktur² betrieben.

Die DINI/nestor-AG Forschungsdaten³ lud im Januar 2019 im Rahmen der Workshop-Reihe "FDM am Standort: von der initialen Idee zum dauerhaften Service" in Kooperation mit dem Projekt UNEKE⁴ der Universitätsbibliothek Duisburg-Essen und dem IT Center der RWTH Aachen⁵ zum Workshop "Bedarfserhebungen – Grundlage für passgenaue Infrastrukturen?" ein.

Dort wurde die bedarfsgesteuerte Entwicklung von RADAR während der Projektphase und im Produktivbetrieb vorgestellt. Darüber hinaus wurde über die grundlegenden Funktionalitäten und Dienstmerkmale der Forschungsdateninfrastruktur sowie über Pläne für deren zukünftige Weiterentwicklung informiert.

2 RADAR als disziplinübergreifendes Forschungsdatenrepositorium

RADAR bietet akademischen Einrichtungen und Forschenden eine generische Infrastruktur zur sicheren Archivierung und Publikation digitaler Forschungsdaten. Es wird von FIZ Karlsruhe – Leibniz-Institut für Informationsinfrastruktur als disziplinübergreifender "All-in-One" Cloud-Dienst betrieben und zielt insbesondere auf das institutionelle Forschungsdatenmanagement im sogenannten „Long Tail“ der Forschung, also Disziplinen, die eine Vielzahl kleiner, heterogener Datensätze produzieren. Das Repositorium richtet sich derzeit vorrangig an Hochschulen und außeruniversitäre Forschungseinrichtungen, die keine eigene Forschungsdateninfrastruktur aufbauen und unterhalten oder die RADAR ergänzend zu existierenden fachspezifischen Angeboten nutzen möchten.

Die technische Infrastruktur von RADAR wird von den Rechenzentren des KIT Karlsruhe⁶ (Steinbuch Centre for Computing⁷, SCC) und der TU Dresden⁸ (Zentrum für Infor-

¹<https://www.radar-service.eu> (geprüft: 13.09.2019).

²<https://www.fiz-karlsruhe.de> (geprüft: 13.09.2019).

³<https://dini.de/ag/dininestor-ag-forschungsdaten/> (geprüft: 13.09.2019).

⁴UNEKE: Vom USB-Stick zur NFDI – Entwicklung eines kriteriengeleiteten Entscheidungsmodells für den Aufbau von Forschungsdateninfrastrukturen: <https://uneke.de> (geprüft: 13.09.2019).

⁵<http://www.itc.rwth-aachen.de> (geprüft: 13.09.2019).

⁶<https://www.kit.edu> (geprüft: 13.09.2019).

⁷<https://www.scc.kit.edu> (geprüft: 13.09.2019).

⁸<https://tu-dresden.de> (geprüft: 13.09.2019).

mationsdienste und Hochleistungsrechnen⁹, ZIH) bereitgestellt. Die Forschungsdaten werden dort in drei Kopien an geographisch getrennten Standorten gespeichert (2 Kopien am SCC, 1 Kopie am ZIH). Als deutsches DataCite-Mitglied¹⁰ registriert die TIB Hannover¹¹ DOIs für sämtliche über RADAR publizierte Forschungsdatensätze.

3 Bedarfssteuerung in der Projektphase

RADAR wurde im Rahmen eines DFG-Projekts¹² (2013 – 2016) unter dem Motto „Aus der Community für die Community“ von einem interdisziplinären Projektkonsortium entwickelt, das sowohl Infrastruktureinrichtungen als auch Vertreterinnen und Vertreter naturwissenschaftlicher Fachcommunities umfasste.

Als disziplinübergreifender Dienst steht RADAR seit Projektbeginn vor der Herausforderung, disziplinspezifische Anforderungen aus der wissenschaftlichen Gemeinschaft in einer generischen Infrastruktur zusammenzuführen. Bei der Bedarfserhebung verfolgte das RADAR-Projektteam verschiedene Ansätze:

Die Konzeption während der Projektphase wurde durch Anforderungsanalysen begleitet und sollte eine bedarfsgesteuerte, standardkonforme und dem Stand der Technik entsprechende Entwicklung ermöglichen. Fachwissenschaftliche Anforderungen an RADAR wurden sowohl über einen Vergleich mit etablierten internationalen Serviceangeboten als auch durch eine Analyse bestehender nationaler, überwiegend disziplinspezifischer Forschungsdatenrepositorien ermittelt. Dabei wurde deutlich, dass in der bestehenden Infrastrukturlandschaft insbesondere disziplinübergreifende, nicht kommerzielle Angebote unter deutschem Recht fehlten. Weiterhin sollte bereits während der Projektphase ein nachhaltiges, sich selbst tragendes Geschäftsmodell erarbeitet werden, welches den von Projektmitteln unabhängigen Betrieb des Dienstes sicherstellt.

Die enge Zusammenarbeit mit den Projektpartnern aus den Pilotdisziplinen Chemie und Biochemie stellte sicher, dass die zentralen Dienstleistungen und das Dienstleistungsspektrum von RADAR bestmöglich an die Bedarfe und Arbeitsabläufe von Forschenden an Hochschulen und an außeruniversitären Forschungseinrichtungen angepasst wurden. Das Metadatenschema und der Funktionsumfang des Dienstes wurden nach der projektinternen Evaluierung durch weitere Kooperationspartner aus anderen Fachdisziplinen (Material-, Geistes-, Sport- und Gesundheitswissenschaften) mit konkreten Anwendungsfällen auf Anwendbarkeit hin geprüft und weiter optimiert.

⁹<https://tu-dresden.de/zih> (geprüft: 13.09.2019).

¹⁰<https://datacite.org> (geprüft: 13.09.2019).

¹¹<https://www.tib.eu> (geprüft: 13.09.2019).

¹²zur Durchführung des DFG-Projekts siehe dessen Abschlussbericht vom Oktober 2016: https://www.radar-service.eu/sites/default/files/publications/Abschlussbericht_DFG-Projekt_RADAR_Vero%CC%88ffentlichung.pdf (geprüft: 13.09.2019).

Um die praxismgerechte Ausrichtung von RADAR sicherzustellen, strebte das Projektteam den kontinuierlichen Austausch mit zukünftig nutzenden Institutionen an. Während der Projektphase wurde die Fachöffentlichkeit regelmäßig über den Projektstand informiert; jährlich wurden öffentliche Workshops mit annähernd 90 Teilnehmenden veranstaltet. Im zweiten Projektjahr wurde auf Grundlage der bisherigen Arbeitsergebnisse ein Prototyp erstellt, den Forschende unterschiedlicher Fachdisziplinen sowie Verantwortliche wissenschaftlicher Bibliotheken und Rechenzentren aus über 20 Einrichtungen testeten. Deren Feedback floss in die iterative Verbesserung der Funktionalität des Prototyps ein.

Darüber hinaus wurde bereits zu Projektbeginn ein wissenschaftlicher Beirat eingerichtet, dessen Vertreterinnen und Vertreter verschiedene Infrastruktureinrichtungen repräsentieren.

RADAR wurde vom Beirat bewertet und in Hinblick auf die Anforderungen der Forschungsgemeinschaft erfolgreich evaluiert. Weiterhin sprach der Beirat Empfehlungen für die funktionale Weiterentwicklung und Verstetigung des Dienstes aus, die insbesondere das Geschäftsmodell sowie die Integration in das wissenschaftspolitische Umfeld betrafen.

Diese Maßnahmen stellten sicher, dass bei der Entwicklung von RADAR die Anforderungen verschiedener Akteure des wissenschaftlichen Umfelds berücksichtigt werden konnten.

4 RADAR Dienstleistungen

Der RADAR-Service umfasst drei zentrale Dienstleistungen¹³:

Die Datenarchivierung dient der format- und disziplinunabhängigen Verwahrung von Forschungsdaten über eine flexibel wählbare Haltefrist (z. B. entsprechend der DFG-Leitlinien zur Sicherung guter wissenschaftlicher Praxis¹⁴). Archivierte Datensätze werden standardmäßig nicht veröffentlicht, Datengeberinnen und Datengeber können jedoch selektiv einzelne Datensätze mit anderen Nutzerinnen und Nutzern teilen.

Bei der Datenpublikation wird eine Haltefrist von mindestens 25 Jahren garantiert. Jeder über RADAR veröffentlichte Datensatz erhält einen Persistent Identifier (DataCite-DOI), wird automatisch bei DataCite indexiert und über standardisierte Protokolle (OAI-PMH) zum Harvesting angeboten. Dies sorgt für maximale Verbreitung und Auffindbarkeit der Forschungsdaten. Falls die Forschungsdaten nicht sofort veröffentlicht werden sollen, kann eine Embargofrist von bis zu einem Jahr festgelegt werden. Für jeden veröffentlichten Datensatz muss eine Lizenz (z.B. Creative Commons 4.0¹⁵) ge-

¹³<https://www.radar-service.eu/de/dienstleistungen> (geprüft: 13.09.2019).

¹⁴https://www.dfg.de/foerderung/grundlagen_rahmenbedingungen/gwp/index.html (geprüft 13.09.2019).

¹⁵<https://creativecommons.org/> (geprüft 12.09.2019).

wählt und somit die Nachnutzungsbedingungen für urheberrechtlich geschützte Inhalte festgelegt werden.

Darüber hinaus unterstützt RADAR den Peer Review-Prozess von Forschungsdaten durch externe Gutachterinnen und Gutachter. Über einen sicheren Link können die einer Publikation zugrunde liegenden Forschungsdaten vorab mit Reviewerinnen und Reviewern geteilt werden.

5 RADAR Metadatenschema und Rollen-/Rechtemodell

Zu den zentralen Dienstmerkmalen von RADAR zählt das disziplinübergreifende, mit bestehenden Standards wie DataCite und DublinCore¹⁶ kompatible Metadatenschema¹⁷. Es legt 10 Pflichtfelder fest, die für die DOI-Registrierung erforderlich sind. Mit weiteren optionalen Feldern können Datensätze detaillierter beschrieben werden.

Neben dem Metadatenschema zeichnet sich RADAR durch eine offene Systemarchitektur sowie ein klar definiertes Rollen- und Rechtemodell¹⁸ aus. Administratorinnen und Administratoren verwalten dabei RADAR-Arbeitsbereiche, die als zentrale Einstiegspunkte für Forschende dienen, um die Forschungsdaten für die Archivierung oder Publikation zu organisieren. Bei den Datengebenden selbst wird zwischen 2 Rollen unterschieden:

- Kuratorinnen und Kuratoren als Datengebende mit vollen Rechten können Forschungsdaten in RADAR hinterlegen, mit Metadaten beschreiben, archivieren und publizieren.
- Subkuratorinnen und Subkuratoren als Datengebende mit eingeschränkten Rechten (in der Regel die Forschenden selbst) können lediglich Forschungsdaten in RADAR hinterlegen und mit Metadaten beschreiben.

Somit erlaubt RADAR die delegierte Administration durch die nutzende Einrichtung. Das System kann flexibel an Prozesse und Verantwortlichkeiten vor Ort angepasst werden und unterstützt dadurch die Aufbereitung, Qualitätssicherung und Kuratierung von Forschungsdaten.

6 RADAR Preis- und Geschäftsmodell

Die Nutzung von RADAR durch Institutionen setzt den Abschluss eines Dienstleistungsvertrags voraus, für den eine jährliche Grundgebühr anfällt. Die in Anspruch genommenen Dienstleistungen Datenarchivierung und -publikation werden darüber hin-

¹⁶<https://www.dublincore.org/> (geprüft: 13.09.2019).

¹⁷<https://www.radar-service.eu/de/radar-schema> (geprüft: 13.09.2019).

¹⁸<https://www.radar-service.eu/rollen-und-rechtemodell> (geprüft: 13.09.2019).

aus nutzungsbasiert je nach angefallenem Datenvolumen in Rechnung gestellt. Während für archivierte Daten innerhalb der gewählten Haltefrist jährlich Kosten berechnet werden, fällt für publizierte Daten nur im Jahr der Veröffentlichung eine Einmalzahlung an.

Grundsätzlich verfolgt FIZ Karlsruhe mit RADAR keine kommerziellen Interessen, vielmehr wird das Angebot von RADAR als Teil des öffentlichen Auftrags gemäß dem Leitmotiv "Advancing Science" betrachtet¹⁹. Ziel ist, den gesamten wissenschaftlichen Wertschöpfungsprozess zu unterstützen, z. B. über die Bereitstellung geeigneter Daten- und Informationsdienste. Daher übernimmt FIZ Karlsruhe die Hälfte der operativen Fixkosten von RADAR aus der eigenen Grundfinanzierung. Lediglich der verbleibende Fixkostenanteil und die variablen Betriebskosten werden über die Nutzungsgebühren eingenommen.

7 Bedarfssteuerung im Produktivbetrieb

Auch nach Aufnahme des Produktivbetriebs wird die Weiterentwicklung von RADAR durch Bedarfserhebungen geleitet.

Die Fachöffentlichkeit wird weiterhin stetig durch Präsentationen²⁰ auf Konferenzen und Workshops über den aktuellen Funktionsumfang informiert und in die Fortentwicklung des Dienstes eingebunden. Die jährlich stattfindenden RADAR-Workshops adressieren sowohl bestehende Nutzende als auch Einrichtungen, die Interesse an der Implementierung von RADAR haben. Um den Austausch mit den Teilnehmenden zu maximieren und deren Bedarfe zu sammeln, wird interaktiven Elementen wie Fragerunden, Teilnehmerbeiträgen und Diskussionen viel Raum eingeräumt.²¹

Zudem bietet RADAR interessierten Institutionen weiter die Möglichkeit, den Dienst über das öffentlich zugängliche Testsystem²² unverbindlich kennenzulernen, zu evaluieren und Feedback zu geben.

Daneben werden im direkten Austausch mit potentiell nutzenden Einrichtungen oder in Vertragsverhandlungen dem RADAR-Team Bedarfe und Anforderungen kommuniziert, die in Entscheidungen über die funktionale oder technische Fortentwicklung des Dienstes und dessen Betriebsszenarien einfließen.

Neben der Kernzielgruppe geben auch wissenschaftspolitische Forderungen und die dynamischen Entwicklungen in der nationalen und internationalen Forschungsdaten-

¹⁹<https://www.fiz-karlsruhe.de/de/ueber-uns/ueber-uns#unser+profil> (geprüft: 13.09.2019).

²⁰<https://www.radar-service.eu/publikationen> (geprüft: 13.09.2019).

²¹Zusammenfassung des 4. RADAR-Workshops vom Juni 2018: <https://www.radar-service.eu/sites/default/files/presentations/RADAR%20Workshop%204%20-%20Zusammenfassung.pdf> (geprüft: 13.09.2019).

²²<https://www.radar-service.eu/de/radar-testsystem> (geprüft: 13.09.2019).

landschaft (z.B. NFDI²³, EOSC²⁴, RDA²⁵, GOFAIR²⁶) zahlreiche Impulse für die Weiterentwicklung von RADAR.

Alle auf den genannten Wegen gesammelten Anregungen werden vom RADAR-Team ins Product Backlog aufgenommen, reflektiert und gegebenenfalls im Sinne agiler Softwareentwicklung für die Fortentwicklung des Dienstes aufbereitet.

Wichtige Impulse zur Identifizierung und Priorisierung zukünftiger Funktionserweiterungen oder ergänzender Unterstützungsmaßnahmen erhofft sich RADAR von einem NutzerInnenbeirat, der noch in 2019 eingerichtet werden soll. Der konstruktive Austausch mit nutzenden Einrichtungen soll auf diese Weise intensiviert und formalisiert werden und zur weiteren bedarfsgesteuerten Entwicklung des Dienstes beitragen.

8 Zukünftige Erweiterungen von RADAR

RADAR ist bestrebt, Forschende und akademische Einrichtungen dabei zu unterstützen, die Sichtbarkeit, Nachvollziehbarkeit, Reproduzierbarkeit und Transparenz von Forschungsergebnissen zu fördern. RADAR will somit den Zugang zu und die langfristige Verfügbarkeit von archivierten und veröffentlichten Forschungsdaten zum Wohle der Forschungsgemeinschaft gewährleisten.

Als Infrastrukturanbieter reagiert RADAR dabei auf neue Anforderungen der Zielgruppe und die sich dynamisch entwickelnden wissenschaftspolitischen Rahmenbedingungen im Themenfeld Forschungsdatenmanagement. Zugleich wird, basierend auf dem öffentlichen Auftrag von FIZ Karlsruhe, die langfristige Verfügbarkeit des Dienstes gewährleistet.

Die geplanten Anpassungen und Erweiterungen zielen darauf ab, die Archivierung und Publikation digitaler Forschungsdaten für Forschende, für Institutionen und für die wissenschaftliche Community, beispielsweise im Rahmen der NFDI, einfacher, attraktiver und bedarfsgerechter zu gestalten.

Im speziellen sind in der Product Roadmap unter anderem folgende Schritte vorgesehen:

²³Rat für Informationsinfrastrukturen. (2018). Zusammenarbeit als Chance - Zweiter Diskussionsimpuls zur Ausgestaltung einer Nationalen Forschungsdateninfrastruktur (NFDI) für die Wissenschaft in Deutschland. Göttingen. Abgerufen von <http://www.rfii.de/?p=2529>.

²⁴European Open Science Cloud (EOSC): <https://ec.europa.eu/research/openscience/eosc> (geprüft: 13.09.2019).

²⁵Research Data Alliance (RDA): <https://www.rd-alliance.org> (geprüft: 13.09.2019).

²⁶Die GO FAIR Initiative zielt darauf ab, die Auffindbarkeit, Zugänglichkeit, Interoperabilität und Wiederverwendbarkeit von Forschungsdaten über Länder- und Disziplingrenze zu befördern: <https://www.go-fair.org> (geprüft: 13.09.2019).

Neben regelmäßigen Funktionserweiterungen wird RADAR noch 2019 verstärkt institutionsspezifische Customizing-Optionen²⁷ ermöglichen. Diese umfassen beispielsweise Anpassungen ans Corporate Design (Logo, Farbe), Verweise auf institutionseigene Unterstützungsangebote, die Vergabe von DOIs mit einem institutionseigenen Präfix und die Möglichkeit einer institutionellen Sicht auf die eigenen publizierten Datensätze.

Dieser Ausbau der Customizing-Optionen schließt auch die Einbindung institutionseigener Speicherkapazität ein. Neben diesem Betriebsmodell, in welchem die RADAR-Software der nutzenden Einrichtungen weiterhin als gehosteter Dienst zur Verfügung gestellt wird, ist auch ein weiteres Szenario vorgesehen, in dem die nutzende Institution die RADAR-Software, unterstützt über einen Wartungsvertrag und regelmäßige Software-Updates durch FIZ Karlsruhe, auf eigener Hardware betreibt. In beiden neuen Betriebsszenarien löst sich RADAR vom Ansatz des „All-in-One“ Dienstleisters und öffnet sich einem breiteren und von der Zielgruppe nachgefragten Anwendungsspektrum.

Weiterhin wird sich RADAR zukünftig in mehreren der potenziellen NFDI-Konsortien als Infrastrukturpartner durch die Bereitstellung eines bedarfsgerechten Forschungsdatenrepositoriums einbringen. Damit RADAR als generische Infrastruktur auch für diese disziplinspezifischen Angebote genutzt werden kann, werden bestehende Möglichkeiten der formalen und inhaltlichen Beschreibung von Forschungsdaten mit Metadaten an die jeweiligen fachspezifischen Bedarfe angepasst.

Darüber hinaus lässt sich RADAR aktuell nach dem CoreTrustSeal²⁸ zertifizieren. Durch die Zertifizierung als verlässliches Datenarchiv soll die Qualität und Vertrauenswürdigkeit des Dienstes bestätigt und an die wissenschaftliche Gemeinschaft weitergegeben werden.

²⁷ siehe Präsentation bei den eScience Tagen, März 2019: https://www.radar-service.eu/sites/default/files/presentations/eScienceTage_2019_RADAR4KIT.pdf (geprüft: 13.09.2019).

²⁸ <https://www.coretrustseal.org> (geprüft: 13.09.2019).