

# GO UNITE! Empfehlungen zur Sammlung und Sichtbarkeit von Forschungsdaten der eigenen Institution

Markus Putnings

GO UNITE! Arbeitsgruppe "Sammlung und Sichtbarkeit von Forschungsdaten der eigenen Institution" (SamS), aktiv derzeit: Anna Daudrich (0000-0002-7663-5630), Nils Dille (0000-0003-4495-209X), Anastasiia Iarkaeva (0000-0002-7370-1663), Janina Kühner (0000-0001-9211-8695), Markus Putnings (0000-0002-6014-9048), Alica Wollmann (0009-0009-1605-6408)<sup>1</sup>. Kontakt:

<https://go-unite.de/index.php/ag-sammlung-und-sichtbarkeit/>.

## 1 Einführung, Ziele und Herausforderungen

### 1.1 Einführung

Innerhalb von GO Unite<sup>2</sup> (GO FAIR, n.d.) haben sich im Herbstmeeting 2021 verschiedene Arbeitsgruppen zusammengefunden, die relevante Fragen und Aufgabenstellungen im Bereich Forschungsdatenmanagement (FDM) eigenverantwortlich bearbeiten. Der Fokus der Arbeitsgruppen entstand durch die bei den beteiligten Einrichtungen bzw. Akteuren beobachteten Bedarfe und Fragestellungen für den Praxisalltag.

Die Arbeitsgruppe "Sammlung und Sichtbarkeit von Forschungsdaten der eigenen Institution" (kurz SamS) beschäftigt sich mit der zentralen Frage, wie verstreute Forschungsdaten effizient gesichtet, identifiziert und der eigenen Einrichtung zugeordnet werden können. Konkret geht es darum, Daten von Plattformen wie Zenodo, Dryad, RADAR, Figshare und fachspezifischen Repositorien (Recommended Repositories, n.d.; Registry of

---

<sup>1</sup> Zusätzlicher Input kam v. a. durch die frühere SamS-Leiterin Esther Asef von der Universitätsbibliothek der Freien Universität Berlin, von Dr. Dietmar Kammerer vom Weizenbaum-Institut sowie bspw. in den einzelnen Breakout-Sessions der GO UNITE!-General Meetings u. a. durch Andreas Christ der Universitätsbibliothek der Christian-Albrechts-Universität zu Kiel, Stefan Kellendonk vom Servicezentrum eSciences der Universität Trier, Jan Klaus vom Fraunhofer-Institut für Produktionstechnik und Automatisierung (IPA), Dr. Evamaria Krause von der Universitätsbibliothek der Universität Augsburg, Francesca Schulze von der Universitätsbibliothek der Technischen Universität Hamburg sowie Jan Taubitz vom QUEST Center des Berliner Instituts für Gesundheitsforschung in der Charité. Mit Blick auf den Call for Comments and Contributions danken wir Robert Ulrich für seine wertvollen Hinweise.

<sup>2</sup> GO UNITE! - früher auch bekannt als GO UNI - ist das deutsche GO FAIR Chapter des Implementation Network Data Stewardship Competence Centers (DSCC). Für weiterführende Informationen siehe <https://www.go-fair.org/implementation-networks/overview/dscc/#pt-item-germany-go-unite> (abgerufen am 27.09.2023). GO FAIR ist eine von Forschungsdatenmanagement-Interessengruppen getragene und selbstverwaltete Initiative, die darauf abzielt, die FAIR-Datengrundsätze umzusetzen, um Forschungsdaten entsprechend möglichst auffindbar (findable), zugänglich (accessible), interoperabel (interoperable) und wiederverwendbar (re-usable) zu machen.

Research Data Repositories (re3data), n.d.) zentralisiert zu erfassen. Ziel ist es, diese Daten entweder automatisiert in das eigene Forschungsdaten-Repository einzuspielen oder zumindest die Metadaten im Current Research Information System (CRIS) oder anderen relevanten Verzeichnissen (z. B. Hochschulbibliographie, Bibliothekskatalog) zu erfassen.

Eine Beschäftigung mit diesen Punkten kann im Rahmen der FAIR-Prinzipien u. a. zur Verbesserung der Auffindbarkeit ("F for Findability in FAIR) beitragen und zudem interne Monitoringsaufgaben erleichtern. Durch die Indexierung der Daten der eigenen Einrichtung an zentraler Stelle verbessert sich vor allem die Auffindbarkeit und Nachnutzung im Institutionskontext. Gleichzeitig kann durch die Indexierung die Auffindbarkeit der Daten in weiteren Suchmaschinen und Aggregatoren verbessert werden.

Ziel ist es also, als Einrichtung einen ganzheitlichen Überblick zu gewinnen, welche Forschungsdaten an der eigenen Einrichtung entstanden sind und veröffentlicht wurden und wo diese Daten liegen. Praktische Anlässe hierzu können z. B. (Fakultäts-)Evaluierungen sein, der Einbezug von Daten in die Leistungsorientierte Mittelvergabe (LOM) oder eine Bestandsaufnahme vor Einführung oder Aktualisierung von FDM-Policies, um den Ist- und Soll-Zustand einschätzen und entsprechende Maßnahmen festlegen zu können.

Die in dieser Arbeit verwendeten Personenbezeichnungen beziehen sich – sofern nicht anders kenntlich gemacht – auf alle Geschlechter.

## 1.2 Zielskala der Sichtung, Sammlung bzw. des Monitorings

Je nach entsprechender lokaler Situation<sup>3</sup> kann die Zielskala hierbei differieren. Hiervon abhängig variieren die unter 2 genannten Handlungsfelder und mögliche Maßnahmen:

### 1.2.1 Anlassbezogener Überblick, keine permanent-verbindliche Zielsetzung

In diesem Szenario ist ein punktueller Überblick nötig. Anlässe können intrinsischer<sup>4</sup> oder extrinsischer<sup>5</sup> Natur sein. Es empfehlen sich v. a. die Maßnahmen unter 2.1.2, 2.1.3 und 2.1.4, alternativ kann bei einem anlassbezogenen Überblick auch ein nationaler bzw. globaler Dienstleister hinzugezogen werden, der die Aufgaben des Sammelns und Aggregierens der verstreuten Forschungsdaten gegen Bezahlung übernimmt (*Consultancy*, n.d.; Kompetenznetzwerk Bibliometrie, 2022) oder eine Testphase bzw. möglichst kurze Lizenzierung von kommerziellen Datenmonitoren (siehe Kapitel 2.1.5.).

---

<sup>3</sup> Z. B. Reife- und Ausbaugrad der FDM-Infrastrukturen und -Policies, Verbindlichkeit der Policies, Einbettung von FDM in relevante Zielvereinbarungen (z. B. Innovationsbündnis), leistungsorientierte Mittelvergaben, institutionelle Führungsinformationen und Kennzahlen, universitär-strategische Fragestellungen, Monitoringsverfahren für Rankings und Ratings.

<sup>4</sup> z. B. Fakultätsevaluation, quantitative Erhebung vor der Erstellung einer FDM-Policy, um die Relevanz und Verbreitung von Forschungsdaten (FD)-Publikationen zu bewerten, Erstbefüllung eines neu implementierten CRIS, das FD-Publikationen unterstützt.

<sup>5</sup> z. B. an die Einrichtung herangetragene Umfragen, etwa von Digital Science zu „The State of Open Data“, von nationalen oder bundeslandspezifischen Ministerien oder von europäischen Einrichtungen bzw. Förderern.

### 1.2.2 Permanentes Monitoring als Ziel, jedoch geringe Verbindlichkeit

In diesem Szenario ist neben einer möglicherweise anfänglichen Ersterhebung (s. 1.2.1) ein permanentes Monitoring das Ziel, jedoch ohne forcierende Elemente wie z. B. eine Verpflichtung der Forschenden, ein bestimmtes Forschungsdatenrepositorium (FD-Repositorium) zu wählen und/oder die bibliografischen Metadaten zu den Forschungsdaten in ein vorgesehenes zentrales Verzeichnis einzutragen. Es empfehlen sich die Maßnahmen unter 2.1.1, 2.2.1, 2.1.3 und 2.1.5, eine Prüfung der Wirtschaftlichkeit bzw. Finanzierbarkeit von Datenmonitoren unter 2.1.5, eine nicht bindende Policy- und Repositorien-Empfehlung (2.2.1 und 2.2.2) und begleitende PR-Maßnahmen (v. a. 2.3.1 und 2.3.2).

### 1.2.3 Permanentes Monitoring als Ziel plus hoher Verbindlichkeitsgrad

In diesem Szenario ist neben einer möglicherweise anfänglichen Ersterhebung (s. 1.2.1) ein permanentes Monitoring das Ziel; dies wird begleitet durch forcierende Elemente.

Ein mögliches Beispiel: eine Policy, die *prioritär* das eigene FD-Repositorium empfiehlt. Weitere externe, FAIRe FD-Repositorien<sup>6</sup> sind aus fachspezifischen Gründen (z. B. Empfehlungen der DFG) ebenfalls zugelassen, dann *müssen* jedoch die Metadaten samt Link auf das externe Repositorium im CRIS eingetragen werden. Falls man ein kommerzielles, nicht-FAIRes Repositorium nutzt, *muss* zusätzlich auch in einem FAIRen Repositorium das komplette Dataset zweitveröffentlicht werden.

Der hohe Verbindlichkeitsgrad kann sich auch auf andere Akteure erweitern, z. B. in Form von Anweisungen an das CRIS-Team, FD-Publikationen automatisch zu importieren (etwa mittels Anbindung von Systemen wie dem Data Citation Index von Clarivate oder Elseviers Data Monitor).

Es empfehlen sich alle u. g. Maßnahmen und bei 2.2.1 und 2.2.2 verpflichtende Vorgaben in den FD- und Repositorienpolicies sowie eine harte Forcierung des Eintrags in das vorgesehene zentrale Verzeichnis.

### 1.2.4 Nationales, permanentes Monitoring von einer nichtkommerziellen Infrastruktureinrichtung (z. B. Bibliothek, Forschungszentrum)

Im Sinne von FAIRen Daten wäre es sinnvoll, auf lange Sicht möglichst keine Aufwände an mehreren Einrichtungen zu haben. Ebenso sollten keine kostenpflichtigen Tools wie Elseviers Data Monitor für das Auffinden, Monitoren und interoperable Nachnutzen nötig sein. Stattdessen sollte sich eine wissenschaftliche Nonprofit-Organisation um eine kostenfreie Zugänglichkeit solcher Sammel-, Aggregations- und Zuordnungsservices bemühen, insbesondere in Bezug auf bestimmte Affiliationen.

---

<sup>6</sup> Mit FAIRen Repositorien sind solche gemeint, die bei Dateneinspielungen die Einhaltung der FAIR-Prinzipien berücksichtigen bzw. die als Merkmal die CoreTrustSeal certification bzw. "FAIRification of repositories" gemäß FAIRSFAR durchlaufen haben und/oder das CoreTrustSeal + FAIRenabling Capability Maturity Model erfüllen.

Als Vorbilder könnten z. B. der Open Access Monitor des FZ Jülich (Zentralbibliothek des Forschungszentrums Jülich, n.d.) oder das Hybrid Open Access Dashboard (HOAD) (Niedersächsische Staats- und Universitätsbibliothek Göttingen, n.d.) der SUB Göttingen dienen, mit dem Publikationsanalysen möglich sind. Eventuell ergäbe auch eine nationale Datensammel- und -drehscheibe mitsamt Zuordnungs- und Zustellungsservices nach dem Vorbild von DeepGreen (Kooperativer Bibliotheksverbund Berlin-Brandenburg, n.d.) Sinn, die Uploads in Zenodo, Dryad, Figshare, RADAR etc. aufnimmt und an die jeweiligen Einrichtungen der Datenautorinnen und -autoren liefert.

Die Tätigkeiten von SamS sollen den Weg hierfür ebnen bzw. erste Vorarbeiten für entsprechende (z. B. Drittmittel-)Projekte darstellen sowie Gespräche innerhalb der Community hierüber anstoßen.

### 1.3 Herausforderungen

Im Themenfeld der Suche, Identifikation, Zuordnung, Sammlung und Sichtbarmachung etwaiger verstreut publizierter Forschungsdaten gibt es diverse Herausforderungen, die im Folgenden ansatzweise kategorisiert werden.

#### 1.3.1 Technische Herausforderungen

Hierunter subsumiert sind alle plattformspezifischen Hindernisse. Dies können z. B. fehlende erweiterte Suchmöglichkeiten bei den FD-Repositoryen sein (z. B. nach Affiliation oder Research Organization Registry (ROR)), unzureichende Exportmöglichkeiten oder fehlende Importmöglichkeiten bzw. Dokumenttypen und Metadaten im Metadatenschema des vorgesehenen zentralen Verzeichnisses.

Es empfiehlt sich hierzu, die Entwicklungsplanungen der Plattformen im Blick zu behalten. So plant Zenodo z. B. 2023 auf InvenioRDM umzusteigen, mit einer besseren Affiliation Auto-Completion via ROR. (Nielsen, 2022) Der ROR-Blog (*ROR Blog*, n.d.) liefert auch Neuigkeiten, welche FD-Repositoryen neu oder künftig ROR unterstützen.

#### 1.3.2 Monetäre Herausforderungen

Dies kann bspw. die Lösung der unter 1.3.1 skizzierten technischen Herausforderungen beinhalten, etwa indem man für die jeweilige Community oder dortige Service-Provider (etwa The Library Code (The Library Code, n.d.) für DSpace) Entwicklungsgelder bereitstellt, um Forschungsdaten der eigenen Einrichtung besser identifizieren oder z. B. ins CRIS einspielen zu können. Aber auch Mittel für Personal, um z. B. gemäß 1.2.2 und 1.2.3 ein permanentes Monitoring gewährleisten zu können oder um selbst Open-Source-Lösungen bzw. Erweiterungen zu programmieren.

#### 1.3.3 Herausforderungen im Bereich der Aufgabenorganisation und Zuständigkeiten

Unter Umständen in Abhängigkeit von 1.3.2 stellt sich die Frage, wer die Aufgaben in diesem Themenfeld konkret und idealerweise dauerhaft übernimmt.

Ein mögliches Beispiel: Teils engagiert sich die Bibliothek in den Bereichen Forschungsdatenkatalogisierung, -sammlung und -speicherung, teils das Rechenzentrum,

teils eigene FDM-Competence-Units; teils in punkto Monitoring das CRIS-Team, teils auch die Bibliothek, teils eigene Präsidialstab-Stellen für Daten für Rankings und Ratings, teils für (z. B. Fakultäts-)Evaluierungen die Forschungsreferentinnen und -referenten.

Je nach Zielskala unter 1.2 und Personalkapazitäten bzw. -mitteln gemäß 1.3.2 sind verschiedene Lösungen möglich, um Verantwortlichkeiten klar(er) festzulegen bzw. abzugrenzen. Eine Möglichkeit wäre die Orientierung am sogenannten DIAMANT 2.0-Modell zur Modellierung des FDM-Referenzprozesses inklusive Empfehlungen für die Implementierung einer institutionellen FDM-Servicelandschaft (Lemaire et al., 2023).

### 1.3.4 Herausforderungen im Bereich Standards und Regeln, u. a. bei bibliothekarischen Erschließungsregeln

Je nach Bundesland sind im Bereich bibliothekarischer Normdaten und Erschließung derzeit noch unterschiedliche Spezifika zu beachten. So gibt es bei der bibliothekarischen Katalogisierung Unterschiede bei der Feldbelegung für Forschungsdaten je nach Verbund, aber auch bibliothekskommissionspolitisch unterschiedliche Entscheidungen, ob Forschungsdaten überhaupt katalogisiert werden sollen, auf welchem Wege bzw. in welcher Granularität.

Für ein Beispiel siehe BibliotheksVerbund Bayern (BVB), Auszug aus dem Bericht der AG Metadaten für Forschungsdaten:

„Da es unter den B3Kat-Bibliotheken unterschiedliche Bedürfnisse gibt, hat man sich darauf verständigt, zwei Varianten zur freien Wahl zu stellen:

Variante 1: Die Primärererschließung findet im lokalen Repository statt, die Daten werden hieraus als Stufe 8-Sätze in den B3Kat exportiert.

Variante 2: Die Primärererschließung findet im B3Kat statt. [...] Die Metadaten zu den Forschungsdaten werden direkt im B3Kat, genauer gesagt in Aleph, mit Katalogisierungsstufe 5 erfasst. Über die OAI-PMH-Schnittstelle ‚B3Kat Open Data‘ können sie geharvestet und in lokale Services eingebunden werden. Um z. B. die Metadaten-Forschungsdaten einer bestimmten Bibliothek aus dem Datenpool herauszufiltern, können entsprechende Eintragungen in 078q vorgenommen werden.“ (Kommission für Erschließung und Metadaten, 2021).

Für ein anderes Beispiel siehe K10plus vom Bibliotheksservice-Zentrum Baden-Württemberg (BSZ) und Gemeinsamen Bibliotheksverbund (GBV): „Es wird empfohlen, bei Forschungsdaten die Felder 1131 und 4262 zu belegen. [...] Es wird empfohlen, in den Aufnahmen der zugehörigen Publikationen das Feld 4261 zu belegen.“ (Verbundzentrale des GBV (VZG), 2019), (Verbundzentrale des GBV, 2019).

Es empfiehlt sich, die jeweiligen Seiten der bundeslandspezifischen Verbunddatenbanken/-zentralen bzw. Bibliothekskommissionen zu verfolgen (vgl. Bibliotheksverbund Bayern, n.d.).

Weitere Herausforderungen umfassen Querverbindungen zu allen vorgenannten Kategorien: Notwendige Plattformentwicklungen sollten z. B. so vorangetrieben und finanziert werden, dass sie der gesamten Community zugute kommen, d.h. idealerweise

Open Source, basierend auf gängigen Standards (z. B. Kerndatensatz Forschung (Geschäftsstelle der Kommission für Forschungsinformationen in Deutschland (KFiD), n.d.)), Schnittstellen/Formaten (z. B. REST, SWORD, JSON, OAI-PMH) und unter Einbeziehung von Persistent Identifiern (PID) (ROR, n.d.), (Copyright Clearance Center, n.d.), (*ORCID iD Connecting Research and Researchers*, n.d.) und Normdaten. Dies erfordert personelle Kapazitäten sowie klare Verantwortlichkeiten und Festlegungen.

### 1.3.5 Strategisch-regulatorische Herausforderungen

Insbesondere beim Ziel eines hohen Verbindlichkeitsgrades gemäß 1.2.3 und entsprechenden harten Eingriffen in die Alltagspraxis der Forschenden stellt sich die Frage, wer die Strategie aufstellt und die diesbezüglichen Regularien erlässt und ob sich diese durchsetzen lassen (Monitoring, Sanktionen, etc.). Dies spielt u. U. auch in 1.3.2 und 1.3.3 hinein.

Zudem sind mögliche Wechselwirkungen mit anderen Policies und Verpflichtungen bzw. Rechtsgrundlagen zu berücksichtigen (z. B. Intellectual Property (IP)-Policies, Datenschutzbestimmungen), die Motive und Wünsche der für die Einrichtung relevanten Förderer (z. B. DFG, BMBF) und mögliche Tools sowie Empfehlungen für die Erstellung, Änderung oder Adaption von Strategien und FDM-Policies (Putnings, 2021, 2-13).

## 2 Handlungsfelder, in denen Einrichtungen bzw. Akteure aktiv sein müssen

Für die in 1.2 definierten Ziele sind je nach Skala bzw. Detailausprägung diverse direkte und indirekte Maßnahmen möglich. Im Folgenden ein Überblick der u. a. in den Sessions der GO UNITE!-General Meetings, #vBIB22 (Putnings & Asef, 2021) und über die Mailingliste IP-OA\_Forum (Freie Universität Berlin, n.d.) gesammelten Best-Practice-Ansätze und Ideen.

### 2.1 Technische Maßnahmen

#### 2.1.1 Sicherstellung der Eignung des vorgesehenen zentralen Verzeichnisses

Um den erzielten Überblick abbilden bzw. die Metadaten der Forschungsdaten importieren oder händisch eintragen zu können, muss sichergestellt sein, dass das Zielsystem die Datensätze abbilden kann, d. h. entsprechende Metadatenfelder existieren oder geschaffen werden, idealerweise auf Basis bestehender gängiger Standards wie dem Kerndatensatz Forschung.

Hier eine Auswahl an entsprechender CRIS-Software:

- DSpace-CRIS (<https://wiki.lyrasis.org/display/DSPACECRIS>): Die Erfassung von Forschungsdaten über eine CKAN-Schnittstelle ist seit der CKAN-Add-on Modulintegration ab Version 5.6.1 Ende 2016 möglich (ckan, n.d.), (*AVAILABLE: DSpace-CRIS 5.6 and DSpace-CKAN!*, 2016) (4Science, 2016). Zum Import bereits vorhandener DataCite-DOIs ist in Bamberg zudem ein Plugin für DSpace-CRIS 7.5 entstanden (Staudinger, 2019).

- Pure (<https://www.elsevier.com/solutions/pure>): Die Erfassung von Forschungsdaten ist als eigener Content-Type „Forschungsdaten“ möglich. Pure ermöglicht zudem die Integration von Elseviers Data Monitor. In Pure lassen sich dann entsprechende Verknüpfungen vornehmen, Berichte zum Forschungsdatenaufkommen einer Einrichtung etc. umsetzen sowie eine (Zweit-)Veröffentlichung der gefundenen Forschungsdatensätze im Pure-eigenen Open Science Portal.
- Elements (<https://www.symplectic.co.uk/theelementsplatform/>): Die Erfassung von Forschungsdaten war in Elements seit jeher als eigener Datentyp (mit dem Anzeigenamen „Dataset“) möglich. Bei der Konfiguration von Elements kann man den Datentyp zudem umbenennen (z. B. in „Forschungsdaten“) oder einen eigenen Datentyp mit einem eigenen Satz von Metadatenfeldern erstellen. Für Elements ist „Dataset“ Untertyp der Kategorie „Publikation“, die Kategorie lässt sich jedoch ebenfalls umbenennen (z. B. in „Scholar and Creative Work“, falls man die semantische Bandbreite erhöhen möchte, bspw. mit Blick auf nicht veröffentlichte Forschungsdaten wie solche mit „restricted access“). Elements enthält eine Integration mit Figshare.com, die die Aufnahme von Metadaten zur Beschreibung von Forschungsdaten unterstützt und bei der Aufnahme einen Datensatz direkt mit einem Benutzer (Person) im Elements-Objektmodell verknüpft. Ebenso kann Elements so konfiguriert werden, dass es Metadaten zur Beschreibung von Forschungsdaten aus dem institutionellen Repository einer Einrichtung aufnimmt. Derzeit unterstützt Elements Integrationen mit DSpace, EPrints, Hyrax und Figshare for Institutions. Mittels der Elements-Benutzeroberfläche können zudem Metadaten zur Beschreibung von Forschungsdaten hinzugefügt werden, oder man kann die Elements-API verwenden, um Metadaten zur Beschreibung von Forschungsdaten aus anderen strukturierten Datenquellen anzubieten, für die Zugriffsrechte bestehen.
- Converis (<https://clarivate.com/webofsciencegroup/solutions/converis/>): In der CRIS-Software sind ebenfalls schon seit jeher individuelle Anpassungen am Datenmodell und Metadatenschema möglich (aber entsprechend auch nötig); als Beispiel siehe z. B. die Converis-Instanz der Universität Kassel, Publication type „Research data publication“. Auch das CRIS-Team der FAU (Friedrich-Alexander-Universität Erlangen-Nürnberg, 2019) kann neben Kassel bei entsprechenden Anpassungen angesprochen werden, an der FAU steht eine entsprechende Erweiterung des Datenmodells kurz vor der Live-Schaltung (vgl. Anfang 2024).

Sofern diese Maßnahme noch nicht realisiert, jedoch geplant bzw. angestrebt ist, kann sie auch als Ziel in einer FDM-Policy formuliert werden, siehe z. B. die Research Data Management policy der Vrije Universiteit Amsterdam als Muster und die fortfolgenden Kapitel unter 2.2 für weitere strategische bzw. regulierende Maßnahmen:

„The VU [Vrije Universiteit] ensures that research data that are generated at the VU are Findable (F in FAIR) by including descriptions of these datasets in the

Current Research Information System (CRIS) of the VU.“ (Vrije Universiteit Amsterdam, 2020)

Forschungsdaten bzw. ihre Metadaten können zudem auch im institutionellen Repository verzeichnet werden (siehe 2.2.2). Beispiele für diese Repositoryensoftwarelösungen sind u. a.:

- MyCoRe - <https://www.mycore.de/en/>
- DSpace - <https://dspace.lyrasis.org/> (das auch eine CRIS-Version anbietet, siehe oben)
- EPrints - <https://www.eprints.org/uk/> (*EPrints for Research Data – EPrints Services*, 2015)
- Invenio - <https://inveniosoftware.org/> (als Open-Source-Framework für groß anzulegende digitale Repositorien oder spezifisch InvenioRDM für Forschungsdaten)

Des Weiteren können Forschungsdaten in der Hochschulbibliographie verzeichnet sein oder in den Bibliothekskatalog aufgenommen werden (für letzteres siehe 2.2.5). Ein Beispiel für Forschungsdaten in einer Hochschulbibliographie findet sich an der Universität Heidelberg:

„Veröffentlichte Forschungsdaten von Angehörigen der Universität werden, unabhängig davon, ob diese Daten in Heidelberg oder bei externen Anbietern liegen, über das universitäre [Forschungsdatenverzeichnis](#) nachgewiesen“<sup>7</sup>

Ähnlich hält es die Technische Universität Dortmund, dort heißt es

„Ziel der [...] Hochschulbibliographie ist es, die Forschung der Wissenschaftlerinnen und Wissenschaftler der Technischen Universität Dortmund darzustellen.“ (Universitätsbibliothek Dortmund, n.d.)

Ein weiteres Beispiel ist die Universität Bamberg, die seit Anfang 2022 Forschungsinformationen zu Forschungsdaten in ihrer Universitätsbibliographie im FIS verzeichnet und mit Forschungsinformationen zu Personen, Einrichtungen, Projekten, Publikationen und Auszeichnungen verknüpft (Franke & Ziegler, 2023, 1–13).

Alle drei Universitäten ermöglichen es, die Forschungsdaten selbst zu melden, die TU Dortmund und die Universität Bamberg unterstützen außerdem eine Synchronisation mit ORCID-Profilen.

---

<sup>7</sup> Vgl.

<https://www.umm.uni-heidelberg.de/bibliothek/literaturverwaltung-schreiben-bewerten/open-science-und-forschungsdatenmanagement/> (abgerufen am 24.11.2023).

### 2.1.2 Prüfung einer reproduzierbaren und möglichst automatisierbaren Abfragemöglichkeit der für die eigene Einrichtung relevanten FD-Repositoryn (z. B. APIs)

Je nach Anlass (s. 1.2) ist eine einmalige Abfrage oder ein kontinuierliches Prüfen auf dezentral veröffentlichte Forschungsdaten bei verschiedenen FD-Repositoryn nötig.

Hier empfiehlt es sich, zunächst die für die eigene Einrichtung relevanten FD-Repositoryn (generisch wie z. B. Zenodo, Dryad, OSF, etc. und ggf. auch die fachspezifischen, je nach vorhandenen Fakultäten bzw. Fachbereichen) zu listen und insbesondere im Falle eines permanenten Monitorings gemäß 1.2.2 und 1.2.3 die jeweiligen Abfrage- und API-Anbindungsmöglichkeiten zu prüfen und reproduzierbar auszugestalten. Falls z. B. bei einer Affiliationsabfrage viele verschiedene Schreibweisen der eigenen Einrichtung auf Deutsch, Englisch, mit Akronymen bzw. Abkürzungen wie „Univ.“ etc. genutzt werden, empfiehlt es sich, diese Liste zu speichern, damit sie bei der nächsten Abfrage wieder genutzt werden kann.

Gleiches gilt für die Methoden-Dokumentation der jeweiligen Export-, Vereinheitlichungs- bzw. Aggregationsschritte (z. B. Zugehörigkeitscheck, ob die Daten tatsächlich der Institution zugehörig sind, Dublettencheck, falls ein Datensatz sowohl im institutionellen als auch einem fachspezifischen FD-Repository gespeichert wurde, etc.), für eine leichtere Wiederholbarkeit der Prozedur bzw. eine potenzielle Einspeisung in ein vorgesehenes zentrales Verzeichnis.

Im Folgenden einige Beispiele und Ansätze:

#### 2.1.2.1 Zenodo

##### Zenodo - Search

Auf den ersten Blick erscheint die Zenodo-Suche sehr rudimentär. Unter <https://help.zenodo.org/guides/search/> findet man jedoch vielfältige mögliche Suchbefehle, die man u. a. auf das Metadatenfeld *creators.affiliation* oder auch *contributors.affiliation* anwenden kann.

Hier ein Beispiel, das in das Search-Feld kopiert werden kann, samt Erläuterung:

```
creators.affiliation:(("Friedrich-Alexander"~ AND Uni*) OR (Uni* AND
Erlangen-N*) OR ("FAU" AND Erlangen-N*) OR ("Universitätsklinik Erlangen") OR
("Universitätsklinikum Erlangen") OR ("University Hospital Erlangen"~1) OR
("University Clinic Erlangen"~1))
```

"Friedrich-Alexander"~ ist eine kombinierte Phrasen- (siehe Anführungszeichen) und Näherungssuche ("Proximity search", siehe ~). Diese Kombination ist Friedrich AND Alexander AND Uni\* vorzuziehen, da es auch Fälle geben kann, in denen diese drei Begriffe auftauchen, ohne dass die FAU enthalten bzw. gemeint ist<sup>8</sup>.

<sup>8</sup> Vgl. z. B. <https://zenodo.org/record/1140383> als Falschtreffer bei der Suche nach Friedrich AND Alexander AND Uni\*, da hier die Rheinische **Friedrich**-Wilhelms-**U**niversität und das Zoologische Forschungsmuseum **Alexander** Koenig in den Affiliationen vertreten sind.

"University Hospital Erlangen"~1 bzw. "University Clinic Erlangen"~1 nutzt ebenfalls eine kombinierte Phrasen- (siehe Anführungszeichen) und Näherungssuche (siehe ~), in der zudem innerhalb der Näherung auch 1 zusätzliches Wort erlaubt ist. So findet man z. B. gleichfalls die University Clinic **of** Erlangen oder University Hospital **of** Erlangen. Aufgrund der Näherungssuche wird auch das Erlangen University Hospital gefunden.

Mit "FAU" AND Erlangen-N\* deckt man Fälle ab, in denen das Akronym FAU genutzt wurde und gleichzeitig mittels Ortsangabe die Florida Atlantic University ausgeschlossen werden soll (man könnte alternativ auch mit NOT arbeiten).

Falls das Ziel keine ein- bzw. erstmalige Komplettabfrage gemäß 1.2.1 ist, sondern jährlich für eine Evaluierung abgefragt wird, lassen sich mittels `publication_date:[20xx-01-01 TO 20xx-01-01]` auch Abfrage-Zeitspannen festlegen bzw. von Jahr zu Jahr aktualisieren. Im Folgenden wird auf diese Einschränkung jedoch verzichtet.

Demnächst ist zudem bei Zenodo eine Migration auf eine neue Softwareversion geplant, ab hier wird dann auch ROR u. a. mit dem Search-Feld `creators.affiliations.id: <ror_value>` unterstützt.<sup>9</sup> (Nielsen, 2023)

### Zenodo - REST API

Die REST-API<sup>10</sup> kann genutzt werden, um die o. g. Suchergebnisse bequem als JSON-Antworten abzufragen und zu parsen.<sup>11</sup> Dazu kopiert man zunächst nach dem Abschicken der Befehle die Such-URL (also, die resultierende URL bei der Suche) nach „q=“ aus dem Browser

```
https://zenodo.org/search?page=1&size=20&q=creators.affiliation:(%22Friedrich-Alexander%22~%20AND%20Uni*)%20OR%20(Uni*%20AND%20Erlangen-N*)%20OR%20(%22FAU%22%20AND%20Erlangen-N*)%20OR%20(%22Universit%C3%A4tsklinik%20Erlangen%22)%20OR%20(%22Universtit%C3%A4tsklinikum%20Erlangen%22)%20OR%20(%22University%20Hospital%20Erlangen%22~1)%20OR%20(%22University%20Clinic%20Erlangen%22~1))
```

und übernimmt sie für die REST API in „`https://zenodo.org/api/records/?q=`“, konkret also

```
https://zenodo.org/api/records/?q=creators.affiliation:(%22Friedrich-Alexander%22~%20AND%20Uni*)%20OR%20(Uni*%20AND%20Erlangen-N*)%20OR%20(%22FAU%22%20AND%20Erlangen-N*)%20OR%20(%22Universtit%C3%A4tsklinik%20Erlangen%22)%20OR%20(%22Universtit%C3%A4tsklinikum%20Erlangen%22)%20OR%20(%22University%20Hospital
```

<sup>9</sup> E-Mail-Auskunft vom Zenodo Support vom 01.02.2023.

<sup>10</sup> Vgl. <https://developers.zenodo.org/#rest-api> (abgerufen am 27.09.2023).

<sup>11</sup> Bei den Abfragen der Zenodo-API muss das Rate Limit beachtet werden, das die Anzahl möglicher Abfragen pro Zeiteinheit begrenzt. Angemeldete User können dabei mehr Abfragen durchführen. Vgl. <https://developers.zenodo.org/#rate-limiting> (abgerufen am 27.09.2023).

**%20Erlangen%22~1)%20OR%20(%22University%20Clinic%20Erlangen%22~1)).**

Die API gibt bei dieser Abfrage standardmäßig 10 Ergebnisse aus. Weitere Ergebnisse werden durch den Zusatz an die URL "&page=2&size=10" ausgegeben bzw. entsprechend durch "&page=<Seitenzahl>&size=10". Man kann den Wert für <size> hochsetzen: hierfür ermittelt man am besten die Zahl der Treffer über die Angabe im JSON unter {hits:total} und wählt dann mindestens diesen Wert für <size>.

Das resultierende Ausgabeformat ist JSON. Man kann das Ergebnis in OpenRefine (OpenRefine, n.d.) einspielen (oder die Abfrage direkt aus OpenRefine starten) und in ein CSV Format exportieren, das sich wiederum leichter als JSON in bestimmte CRIS- bzw. FD-Repositoriensoftwares importieren lässt. In OpenRefine könnte man zudem Routinen bauen, um die Metadaten direkt für den Import in ein konkretes CRIS aufzubereiten.

#### *Zenodo - OAI-PMH*

Falls die Einrichtung gemäß 2.3.2.1 mit einem Funktionsaccount eine eigene kostenlose Zenodo-Community angelegt hat, in die Datensätze prioritär eingestellt werden, lassen sich die entsprechenden Metadaten über verschiedene Anfragen extrahieren, u. a. über OAI-PMH. Dazu benötigt man den Namen der jeweiligen Community, er befindet sich in der URL der Community direkt hinter „...communities/“. Bei <https://zenodo.org/communities/fau/?page=1&size=20> also „fau“. Mittels der OAI-PMH-Abfrage ListRecords kann man dann alle Metadaten z. B. in Dublin Core ausgeben lassen. Dazu muss dem Community-Namen noch „user-“ vorangestellt werden, im vorangegangenen Beispiel also dann konkret „user-fau“: [https://zenodo.org/oai2d?verb=ListRecords&set=user-fau&metadataPrefix=oai\\_dc](https://zenodo.org/oai2d?verb=ListRecords&set=user-fau&metadataPrefix=oai_dc). Für weitere Ausgabeformate siehe <https://zenodo.org/oai2d?verb=ListMetadataFormats>, möglich wäre z. B. marxml bzw. marc21 für bibliothekarische Zwecke gemäß 2.2.5.

Für besondere Auswertungen wäre das Laden der Zenodo-Community-Daten in einen Suchindex möglich (Lohmeier, n.d.); auch mit OpenRefine könnte man, wie im vorherigen Abschnitt beschrieben, die OAI-PMH-Schnittstelle abfragen und Ergebnisse aufbereiten.

#### *Zenodo - Search - Skriptabfragen (z. B. der ORCID IDs der eigenen Forschenden)*

Über die Affiliationsangaben wird man aller Wahrscheinlichkeit nach nie 100 % aller Datensätze der eigenen Einrichtung erhalten, da die Forschenden – auch in Abhängigkeit von 2.3.2 – zum Teil abweichende Affiliationsangaben eingeben (z. B. nur Nennung des Departments, nicht der Universität) oder sich schlicht vertippen. Gleiches gilt für das Befüllen der einrichtungsspezifischen Zenodo-Community gemäß 2.3.2.1 durch die Forschenden. Entsprechend können bei einem hohen Anspruch auf Vollständigkeit weitere bzw. ergänzende Abfragen sinnvoll sein.

Sofern die Einrichtung einen Überblick über die ORCID iDs der eigenen Forschenden hat (z. B. gemäß 2.3.2 und/oder einer Hinterlegung aller ORCID ins lokale CRIS bzw. Identity-Management-System), lassen sich diese gesammelt mittels Skript abfragen. Als potenzielles Softwaretool kann der Zenodo Harvester der FAU genannt werden, der

basierend auf der extensiven ORCID-Abfrage JSON für das CRIS zurückliefert (FAU, 2019). Aktuell ist der Harvester rein für die FAU ausgelegt, er müsste für andere Institutionen angepasst werden.

Einschränkend muss jedoch erwähnt werden, dass dies nur für Momentaufnahmen ginge und eher nicht für retrospektive Erhebungen, da i. d. R. sowohl im eigenen Identity Management (IDM) System die Beschäftigungszeiten aus Datenschutzgründen nicht längerfristig bzw. dauerhaft nachgewiesen werden als auch die retrospektiven Zugehörigkeitsangaben in ORCID nur bedingt durchgängig befüllt bzw. verlässlich sind.

### 2.1.2.2 Dryad

#### *Dryad - Search*

Dryad unterstützt seit 2019 ROR als eindeutigen Einrichtungs-Identifizier (Gould & Lowenberg, 2019). Entsprechend kann bei der einfachen Suche einfach die ROR der eigenen Institution, exemplarisch z. B. <https://ror.org/00f7hpc57> in das Suchfeld hineinkopiert werden.

Eine weitere Möglichkeit ist der Weg über „Explore Data“, unter „Institution“ auf „more »“ (sofern nicht direkt dort gelistet), das Durchklicken bis zur eigenen Einrichtung und dann das Bookmarken dieses Links, exemplarisch z. B.

[https://datadryad.org/search?f%5Bdryad\\_author\\_affiliation\\_name\\_sm%5D%5B%5D=University+of+Erlangen-Nuremberg](https://datadryad.org/search?f%5Bdryad_author_affiliation_name_sm%5D%5B%5D=University+of+Erlangen-Nuremberg).

Bei der Abfrage bitte beachten, dass ein logisches Oder, wie z. B.

„(<https://ror.org/00f7hpc57>) OR (<https://ror.org/0030f2a11>)“ im Gegensatz zu Zenodo nicht zu funktionieren scheint. In Dryad wird stets eine logische Und-Suche durchgeführt. Gleiches gilt für den Weg über „Explore Data“, z. B.

[https://datadryad.org/search?f%5Bdryad\\_author\\_affiliation\\_name\\_sm%5D%5B%5D=University+of+Erlangen-Nuremberg&f%5Bdryad\\_author\\_affiliation\\_name\\_sm%5D%5B%5D=Universit%C3%A4tsklinikum+Erlangen](https://datadryad.org/search?f%5Bdryad_author_affiliation_name_sm%5D%5B%5D=University+of+Erlangen-Nuremberg&f%5Bdryad_author_affiliation_name_sm%5D%5B%5D=Universit%C3%A4tsklinikum+Erlangen).

Eine Einschränkung der Abfrage auf bestimmte Publikationszeitspannen scheint aktuell gleichfalls über die Suchoberfläche nicht möglich zu sein.

Alle Daten bei Dryad sind auf Zenodo gespiegelt und können u. U. schon bei einer Suche in Zenodo gefunden werden.<sup>12</sup>

#### *Dryad - API*

Die Dryad API (Data Dryad, n.d.), (Data Dryad, n.d.) ist ebenfalls eher eingeschränkt in den Funktionalitäten. Die API-URL beginnt mit „<https://datadryad.org/api/v2/search?>“, danach werden die Suchbefehle angehängt, konkret etwa „affiliation=“.

---

<sup>12</sup> Vgl. „Dryad stores a copy of all datasets in Zenodo for enhanced preservation services.“ <https://datadryad.org/stash/faq> (abgerufen am 26.01.2023).

„affiliation -- a ROR identifier specifying an institutional affiliation that must be present in the list of dataset authors. The identifier should be in the full "https" format and should be URL-encoded, e.g., [https%3A%2F%2Fror.org%2F00f7hpc57](https://ror.org/00f7hpc57)." (Data Dryad, 2019)

Beim Beispiel der ROR <https://ror.org/00f7hpc57> wäre die API-Abfrage <https://datadryad.org/api/v2/search?affiliation=https%3A%2F%2Fror.org%2F00f7hpc57>

Gemäß der Dokumentation ließen sich Publikationszeitspannen nur über das Feld „lastModificationDate“ abfragen:

“modifiedSince -- a timestamp for limiting results. Datasets will only be returned that have been modified since the given time. The time must be specified in ISO 8601 format, and the time zone must be set to UTC, e.g., 2020-10-08T10:24:53Z.” (Data Dryad, 2019)

Eine mit „&“ kombinierte ROR- und Zeitabfrage ist demnach z. B.

<https://datadryad.org/api/v2/search?affiliation=https%3A%2F%2Fror.org%2F00f7hpc57&modifiedSince=2022-01-01T10:24:53Z>.

Die Abfragen liefern JSON zurück. Eine Aufbereitung ist mittels OpenRefine möglich.<sup>13</sup>

### 2.1.2.3 Mendeley Data

#### *Mendeley Data - Search*

Bei Mendeley Data gibt es ähnlich wie bei Zenodo unter

<https://data.mendeley.com/faq?q=advanced+search> vielfältige Suchbefehle; die für die Affiliation umfassen INSTITUTION() und INSTITUTION\_ID(), wobei INSTITUTION\_ID() derzeit nicht zu funktionieren oder unzureichend befüllt zu sein scheint, mit der testweisen Eingabe von Scopus Institution IDs erhält man keinerlei Treffer.

Die Abfrage von INSTITUTION() und entsprechenden Strings mit Trunkierung etc. ist jedoch erfolgreich. Entsprechend hier ein Abfragebeispiel:

```
INSTITUTION(("Friedrich-Alexander" AND Uni*) OR (Uni* AND Erlangen-N*) OR
("FAU" AND Erlangen-N*) OR ("Universitätsklinik Erlangen") OR
("Universitätsklinikum Erlangen") OR ("University Hospital Erlangen") OR
("University Clinic Erlangen"))
```

Wie bereits unter Zenodo erwähnt: über die Affiliationsangaben wird man aller Wahrscheinlichkeit nach nie 100% aller Datensätze der eigenen Einrichtung erhalten. Entsprechend können bei einem hohen Anspruch auf Vollständigkeit weitere bzw. ergänzende Abfragen sinnvoll sein, etwa über die ORCID iDs der Forschenden, sofern die Einrichtung einen Überblick über diese hat:

<sup>13</sup> Vgl. Hinweise dazu unter dem Unterabschnitt *Zenodo - REST API* zu [2.1.2.1 Zenodo](#).

AUTHOR\_ID(0000-0002-1163-3271 OR 0000-0002-0835-0086 OR ...)

### *Mendeley Data - API*

Die APIs (Mendeley Data, n.d.) können wiederum z. B. für JSON Ausgaben genutzt werden. Dazu ist vorab ein Antragsverfahren für einen Access Token etc. nötig. Eine Aufbereitung ist mittels OpenRefine möglich.<sup>14</sup>

#### 2.1.2.4 Figshare

Derzeit sind bei Figshare gezielte einrichtungsbezogene Abfragen nur für Figshare for Institutions möglich, siehe Figshare documentation:

„:institution: - string, this is only for institutions that have figshare and you would use the string id from the URL (for example, Loughborough University has the URL <https://lboro.figshare.com/> and you would type ":institutions: lboro" into the search bar)“ (figshare, n.d.)

Gemäß Auskunft des figshare Supports vom April 2023 wird es künftig jedoch eine ROR Unterstützung (und damit mutmaßlich perspektivisch auch eine Abfragemöglichkeit) geben: „We are adding RORs, they will be in place in the next few months.“<sup>15</sup>

#### 2.1.2.5 OSF

Ähnlich wie bei Figshare ist das primäre Monitoring augenscheinlich nur mit OSF für Institutionen möglich, siehe OSF support:

„Affiliated Institution

OSF objects affiliated with [Member Institutions](#) can be searched for using the following prefix: affiliated\_institutions:“ (OSF Support, 2023)

Tatsächlich ließe sich der Suchparameter jedoch auch für Nichtmitgliedsorganisationen verwenden, also z. B. affiliated\_institutions:”University of Erlangen-Nürnberg“. Die Ergebnislisten erscheinen jedoch eher unzuverlässig mit Blick auf die Affiliationszuordnung, d. h. mutmaßlich wird hier eine Näherungssuche durchgeführt ohne tatsächliche Berücksichtigung exakt und ausschließlich der gewünschten Affiliation.

### 2.1.3. Prüfung einer reproduzierbaren und möglichst automatisierbaren Abfragemöglichkeit für die eigene Einrichtung bei relevanten FD-Suchmaschinen und FD-Indexen

Je nach Anlass (s. 1.2) ist eine einmalige Abfrage oder eine regelmäßige Abfrage von veröffentlichten Forschungsdaten über eine FD-Suchmaschine zeitsparender als Ansatz 2.1.2, da damit potenziell gleich mehrere FD-Repositoryn durchsucht werden können, insbesondere fachspezifische Repositoryn jenseits der großen Anbieter.

<sup>14</sup> Vgl. Hinweise dazu unter dem Unterabschnitt *Zenodo - REST API* zu [2.1.2.1](#) Zenodo.

<sup>15</sup> E-Mail-Auskunft von figshare Support <info@figshare.com> vom Do 20.04.2023 10:16.

Dennoch gibt es bisher wenige Suchmaschinen, die auf FD ausgelegt bzw. spezialisiert sind. Daher ist auch hier die Vollständigkeit und zusätzlich die häufig unbekannte Grundgesamtheit (welche Repositorien und Veröffentlichungsorte werden indexiert) eine Herausforderung. Im Folgenden werden einige Beispiele vorgestellt, die sich als mögliche Ressourcen anbieten.

#### 2.1.3.1 DataCite Commons

In DataCite Commons (DataCite, n.d.) werden DOIs, ORCID iDs und ROR IDs mit ihren Metadaten in einem Graphen modelliert.

Über die öffentliche DataCite Commons Web-Suchschnittstelle kann „Organizations“ ausgewählt und eine ROR in die Suche eingegeben werden. Wählt man die entsprechende Organisation aus der Ergebnisliste aus, bekommt man eine Übersichtsseite zur Einrichtung zusammen mit einer Liste der zugeordneten Digitalen Objekte, die mit einer DOI versehen sind. In den Filtern kann diese Liste unter „Work Type“ auf „Dataset“ eingeschränkt werden. Am Beispiel der Freien Universität sieht das wie folgt aus

- <https://commons.datacite.org/ror.org/046ak2485?rorid=046ak2485&resource-type=dataset>

DataCite bietet eine REST API, eine OAI-PMH-Schnittstelle und eine GraphQL API.

#### 2.1.3.2 Scholixexplorer

Scholix steht für „Scholarly Link Exchange“ und ist ein Zusammenschluss von verschiedenen Organisationen wie z. B. Verlagen und Datenanbietern, mit dem Ziel, Verlinkungen zwischen Textpublikationen und Daten in einem Informationssystem sichtbar zu machen (Scholix, n.d.). Im vom H2020-Projekt OpenAIRE2020 und der Research Data Alliance Europe geförderten Projekt *Scholixexplorer* werden Informationen aus verschiedenen Datenservices (u. a. CrossRef, DataCite, OpenAIRE aber auch fachspezifische Anbieter wie Pangaea und Europe PubMed Central (Scholixexplorer, n.d.)) in einem Graphen zusammengeführt. Ziel des Projekts ist u. a. zur Standardisierung der Verlinkung zwischen Text- und Datenpublikation beizutragen. Daher sind im Scholixexplorer vorwiegend Daten verzeichnet, die einen zugehörigen Artikel haben.

Scholixexplorer bietet eine Suche an, die Funktionalitäten und mögliche Operatoren sind nicht weiter erklärt. Es scheint jedoch keine Suche nach ROR-ID oder Ähnlichem möglich zu sein, denn diese Informationen sind nicht in den sichtbaren Metadaten verzeichnet. Eine einfache Suche nach „Freie Universität Berlin“ führt zu 248.235 Treffern, die im Rahmen dieser Arbeit aber nicht näher untersucht wurden (Stand 30.01.2023).

Das Projekt bietet ebenfalls eine API an, hier sind Abfragen über PIDs (von digitalen Objekten), Veröffentlichungsdienste oder Datenlieferanten möglich. Es scheint daher insbesondere eine mögliche Ressource zu sein, wenn schon eine Liste an Publikationen mit PIDs vorliegt, zu denen man die zugehörigen Datensätze finden möchte.

### 2.1.3.3 BASE

BASE steht für „Bielefeld Academic Search Engine“ und ist eine der weltweit größten wissenschaftlichen Suchmaschinen. Unter den Datenlieferanten (BASE, 2019) sind aktuell auch ca. 150 FD-Repositoryn indexiert, darunter Zenodo, Dryad, OSF (unter dem Namen COS Center for Open Science) sowie verschiedene institutionelle Figshare-Instanzen.

Unter der erweiterten Suche kann man auf die Dokumentart Forschungsdaten einschränken und z. B. die ORCID der eigenen Forschenden abfragen, wenn man einen vollständigen Überblick über diese hat, z. B.

orcid:0000-0002-4726-739X OR orcid:0000-0003-2238-2049 OR  
orcid:0000-0002-5035-5825 OR orcid:0000-0002-5707-0988 OR ....

Zu beachten ist jedoch, dass im BASE-Index nur ein kleiner Prozentsatz der Autorinnen bzw. Autoren mit einer ORCID versehen ist (in Abhängigkeit von den Datenquellen und ob diese eine ORCID haben). Zudem ist die Länge des Suchstrings auf 1000 Zeichen begrenzt. Des Weiteren ist die Anzahl der Treffer auf 1000 beschränkt.<sup>16</sup>

Die Suchanfrage müsste also basierend auf diesen Begrenzungen unter Umständen je nach Zahl ORCID-IDs bzw. erzielten Treffern auf mehrere Abfragen gesplittet bzw. aufgeteilt werden.

### 2.1.3.4 OpenAIRE Explore

OpenAIRE ist ein pan-europäisches Forschungsinformationssystem zur Darstellung und Verknüpfung von (i. d. R. EU-geförderten) Forschungsergebnissen. In der Advanced Search kann man auf Research products einschränken und hier wiederum einen Filter auf Research Data setzen. Die Affiliationssuche ist leider etwas unbequem: es existiert das Searching field „Organization“ und wenn man beginnt, den Affiliationsnamen einzutippen, erhält man die verschiedenen indexierten Schreibweisen.

Mittels des +-Buttons und dem Operator OR könnte man all diese verknüpfen, jedoch ist dies ein langwieriger Aufwand mit aktuell nur sehr wenigen Forschungsdatentreffern als Ergebnis.

D. h. es bleibt einem selbst überlassen, den Aufwand zu leisten, andere der bislang beschriebenen Maßnahmen sind jedoch mit Blick auf Aufwands-/Ertragsverhältnis ergiebiger.

Auch hier wären ORCID-Abfragen möglich, falls man einen Überblick über die ORCID-IDs der eigenen Forschenden hat. Für entsprechende Vorgehensweisen und Erfahrungsberichte siehe die ORCID-DE Best Practices (Fenner, n.d.), z. B. seitens der Technischen Universität Dortmund (TUD), die die ORCID-Profile mit der TUD-Identität verknüpfen und die Publikationsdaten aus den ORCID-Profilen mit der Hochschulbibliographie synchronisieren. Beides geschieht technisch mit dem ORCID-Member-API.

---

<sup>16</sup> Auskunft von der Abteilung LibTec - Bibliothekstechnologie und Wissensmanagement der Universitätsbibliothek Bielefeld vom 01.02.2023.

### 2.1.3.5 OpenAlex

OpenAlex (OpenAlex, n.d.) präsentiert sich als ein offener Katalog für Publikationen, Autorenschaften, Journale und Institutionen. Die Abfragen der Datenbank erfolgen über eine API. Diese können z. B. mithilfe von vorgefertigter Pakete in Python (<https://pypi.org/project/pyalex/>) oder R ([https://docs.ropensci.org/openalexR/articles/A\\_Brief\\_Introduction\\_to\\_openalexR.html](https://docs.ropensci.org/openalexR/articles/A_Brief_Introduction_to_openalexR.html)) vorbereitet und ausgeführt werden. Es bietet sich an, unter dem Entity-Typ „work“ auch die Option „dataset“ zu wählen. Allerdings ist diese Datengrundlage noch nicht uneingeschränkt vertrauenswürdig, da der Typ „dataset“ oft für den Typ „journal article“ verwendet wird. Daher sollte jede Institution diese Suchmaschine auf eigene Daten testen.

Trotzdem kann diese Methode auch für die unter Punkt 2.1.4 genannte Methode zur (semi-)automatischen Extraktion von Informationen zu veröffentlichten Datensätzen aus Publikationen genutzt werden, nämlich um die Publikationen interessierter Institutionen in OpenAlex zu erkennen.

### 2.1.4 Erfassung von mit Publikationen verknüpften Datensätzen

Ein weiterer Ansatz ist die (semi-)automatische Extraktion von Informationen zu veröffentlichten Datensätzen aus Publikationen. Da an vielen Standorten bereits eine nahezu vollständige Liste von Textpublikationen (z. B. in einer Hochschulbibliographie) vorliegt, könnte diese als Basis dienen, zugehörige Datensätze zu finden.

Einen solchen Ansatz verfolgt ein Projekt im QUEST Center for Responsible Research am Berlin Institute of Health at Charité. Dort wurde eine semi-automatische Abfrage entwickelt (Iarkaeva et al., n.d.), um Hinweise auf Open Data aus Zeitschriftenartikeln zu extrahieren. Die Gruppe stellt ein genaues Methodenprotokoll und den Code (*Quest-Bih/oddpub: Algorithm for Open Data Detection in Publications (ODDPub)*, n.d.) sowie einen detaillierten Workflow-Artikel (im Moment als Preprint ) zur Verfügung (Iarkaeva et al., 2023). Die Methode ist jedoch für die Charité und mit dem Fokus auf die Medizin entwickelt worden. Dieser Weg kann daher nicht ohne Anpassungsschritte auf andere Institutionen oder Fachbereiche angewandt werden.

Lib4RI, die Forschungsinstitute der ETH Zürich, haben einen DataCite-Alert in ihrem Repositorium implementiert, der DataCite Commons nach relation-types (z. B. IsSupplementTo) sowie Verknüpfungen zwischen Publikationen und Forschungsdaten durchsucht.

Einen eher manuellen Ansatz verfolgt die Zürcher Hochschule für Angewandte Wissenschaften (ZHAW). Dort wird im Repositorium ein Feld geführt, über das Links zu publizierten Forschungsdaten erfasst werden können. Bei der Durchsicht der Einträge werden Informationen aus Data Availability Statements dann vom FD-Repositorien-Team gelesen, recherchiert, heruntergeladen und mitsamt Metadaten aufgenommen.

## 2.1.5 Datenmonitore

### 2.1.5.1 Elseviers Data Monitor

Wie die vorherigen Kapitel zeigen, geht es SamS eher um eine Substitution des Elsevier Data Monitors (Elsevier, n.d.) als um eine Bewerbung. Entsprechend im Folgenden nur kurz die Vorstellung der Produkteigenschaften, Details können auf der Homepage nachgelesen werden.

Elseviers Data Monitor harvestet Forschungsdaten aus (nach eigenen Angaben) mehr als 2.000 generischen bzw. fachspezifischen Repositorien. Die Auswahl der Repositorien richtet sich nach einer Content Policy (Elsevier, 2022). Eine genaue Auflistung ist zumindest online nicht zu finden. Die Metadaten werden nach dem Harvesting normalisiert, dedupliziert und im Bedarfsfall mit weiteren Metadaten angereichert (z. B. aus zugehörigen Publikationen und hierüber ggfs. auch Autor, Institution, etc.).

Data Monitor ist in Pure (s. 2.1.1) und Digital Commons bzw. Digital Commons Data integriert. Dort können institutionelle Datensätze importiert werden. Darüber hinaus gibt es APIs, die für die Integration in andere Systeme genutzt werden können.

### 2.1.5.2 OpenDataMonitor

Der OpenDataMonitor (*Open Data Monitor*, 2021) entstand im Rahmen eines Projekts des 7. Forschungsrahmenprogramms der EU durch ein OpenDataMonitor-Konsortium unter Federführung der SYNYO GmbH als Koordinator (*Open Data Monitor Imprint*, 2021). Er ist für die Zwecke von SamS ungeeignet, da er ausschließlich Open Government Datenportale integriert und nur Auswertungen auf Landesebene, z. B. Deutschland ermöglicht.

## 2.2 Regulierende Maßnahmen

### 2.2.1 Forcierung des Eintrags in das vorgesehene zentrale Verzeichnis

Diese Maßnahme kann durch entsprechende regulierende Leitungsvorgaben oder Stellschrauben der jeweiligen Einrichtung realisiert werden. Konkret etwa durch die Bitte, Empfehlung bzw. Forderung der Eintragung von Forschungsdaten in ein zentrales Verzeichnis innerhalb der Open Science- oder FDM-Policy, durch die Konzeption von Anreizsystemen wie eine Berücksichtigung von Forschungsdaten bei Tenure-Track-Verfahren, in der leistungsorientierten Mittelvergabe bzw. Fakultätsevaluierung, und ähnliches.

Eine verpflichtende Forderung könnte z. B. wie folgt aussehen, siehe Research Data Management Policy der Buckinghamshire New University, Version No. 1.0, Review Date: Jun-2021:

„All research data must be registered with the University whether it is hosted by the University or retained elsewhere“ (Buckinghamshire New University, 2016).

Exemplarisch könnte, als ein mögliches Anreizsystem, die Berücksichtigung von Forschungsdaten in Tenure-Track-Verfahren wie folgt vorgesehen werden, siehe Standing Rules, Institute of Arctic & Alpine Research, Last updated December 2021:

- „A. Annual appraisal procedures for Tenure Track and Research Associate Fellows  
[...] a. Research  
- Peer reviewed publications [...]  
- Open research (data sets published/open source code published).“ (Institute of Arctic & Alpine Research University of Colorado Boulder, 2021).

Das erfolgt im deutschsprachigen Raum zum Beispiel schon in der Psychologie bei der LMU München, dort gibt es Stellenausschreibungen, die um Angaben zu Open Data bei der Bewerbung bitten (Schönbrodt, 2016).

Zusätzlich können in Abhängigkeit von 2.1.1 unter Umständen auch (semi-)automatisierte Importroutinen geschaffen werden, wie etwa zwischen Pure und Elseviers Data Monitor. Das jeweils für das zentrale Verzeichnis zuständige Team kann hier entsprechend unabhängig von den Forschenden die Einträge einspielen und erfassen. Auch manuelle Such- und Erfassungsvorgänge sind – mit einem gewissen Aufwand – denkbar, etwa für einen anlassbezogenen Überblick gemäß 1.2.1.

### 2.2.2 Definition von Repositorienpolicies und -scopes („Sammlungsauftrag“)

Es existieren eine Vielzahl fachübergreifender und -spezifischer FD-Repositorien nebst etwaigen institutionellen Speichermöglichkeiten. Forschungsförderer und Akteure wie z. B. DINI definieren vermehrt Anforderungen an geeignete FD-Repositorien bzw. priorisieren existierende Ablagemöglichkeiten<sup>17</sup> (Gollwitzer et al., 2020, #).

In Orientierung an diesen kann es Sinn ergeben, eigene entsprechende Repositorienpolicies und -scopes zu definieren, v. a. falls mehrere Optionen in Frage kommen (z. B. RADAR, Zenodo-Community, empfohlene fachspezifische Repositorien etc.) und man zum einen Hilfestellung leisten möchte, in welchen Szenarien welche zu wählen sind und/oder man garantieren möchte, dass die Metadaten bzw. Datensätze auch in eigenen Verzeichnissen bzw. Repositorien nebst externen landen.

Dies bedarf entsprechend einer Berücksichtigung von

- internen FD-Repositorien, -Speichermöglichkeiten und zentralen Verzeichnissen,
- externen generischen FD-Repositorien wie z. B. Zenodo, Dryad, Figshare, OSF, Mendeley Data etc.,

<sup>17</sup> Vgl. z. B. DFG Fachkollegium „Psychologie“ zum Umgang mit Forschungsdaten: „[Es] ist ein vertrauenswürdigen öffentliches Repositorium (z. B. PsychArchives des ZPID24, datorium beiGESIS25 oder ein gut ausgebautes universitäres Repositorium) einem Zeitschriftenrepositorium vorzuziehen. Von einer Bereitstellung auf privaten oder persönlichen universitären Webseiten wird dringend abgeraten“, s. <https://doi.org/10.31234/osf.io/hcxtm>.

- externen fachspezifischen FD-Repositoryen sowie
- Auswahlkriterien hierzu (z. B. Orientierung an Vorgaben von Drittmittelgebern wie Erfüllung der FAIR-Kriterien, zertifizierte Repositoryen, Erfüllung der „Criteria for the Selection of Trustworthy Repositories“ von Science Europe und den Vorgaben von Horizon Europe (European Commission, 2023); oder ggf. auch Ausschluss von FD-Repositoryen wie z. B. kommerzielle).

Ein konkretes Beispiel aus den Guidelines for Research Data Management at ETH Zurich (RDM Guidelines) 1 July 2022 RSETHZ 414.2 (© ETH Zürich):

Diese empfehlen die Nutzung des internen Research Data Repository. Falls man ein anerkanntes FAIRes *externes* Repository wählt, wäre das gleichfalls möglich, dann sollen jedoch die dortigen Metadaten in die ETH Research Collection eingetragen werden, samt Link auf das externe Repository. Sofern ein kommerzielles, *nicht*-FAIRes Repository genutzt wird, *muss* das komplette Dataset zusätzlich auf einem FAIRen Repository zweiteröffentlicht werden:

„Repositories

a. The ETH Research Collection is available as institutional repository for Research Data publication. If a recognised ETH-external FAIR repository is used for depositing Research Data or Programming Code, it is recommended to register a metadata-only item in the ETH Research Collection linking to the dataset.

b. Commercial platforms which are well-suited for specific purposes (e.g. code development) are usually not considered as FAIR repositories. In case such platforms are used, a copy of the version of Programming Code belonging to a result publication must be archived in a FAIR repository.

c. Commercial repositories should be avoided whenever possible“ (ETH Zurich, 2022)

Der Sammlungsauftrag und -scope kann noch weitere Aspekte außer der Repositorywahl beinhalten, z. B. auf wen sich diese Richtlinien beziehen (z. B. auch Gastwissenschaftlerinnen bzw. Gastwissenschaftler), deren Verantwortlichkeiten, Qualitätskriterien von zu speichernden Daten, auf welche Weise Daten vorbereitend zur Veröffentlichung planvoll dokumentiert werden sollen (z. B. Datenmanagementplan), welche Mindestaufbewahrungsfristen für die Daten gelten sollen, ob es hiervon Ausnahmen gibt und wie diese zu dokumentieren sind etc.

Hier ein Beispiel aus der Max-Weber-Stiftung, das mehrere der genannten Aspekte beinhaltet:

„Verfügbarmachung, dauerhafte Vorhaltung und Archivierung

Die Wissenschaftlerinnen und Wissenschaftler sichern [...] Forschungsdaten bzw. Forschungsergebnisse [...] und bewahren sie für einen angemessenen Zeitraum entsprechend der FAIR-Prinzipien auffindbar, zugänglich, interoperabel und wiederverwendbar auf. Die Aufbewahrung analoger Daten erfolgt in der Regel in den Einrichtungen, an denen diese entstanden sind. Die Aufbewahrung von

digitalen Daten und Forschungssoftware ist zu Beginn der einschlägigen Forschungsarbeiten in Abstimmung mit der Geschäftsstelle festzulegen. Die entsprechenden Einrichtungen und die Geschäftsstelle stellen sicher, dass die hierfür jeweils erforderliche Infrastruktur verfügbar ist oder beraten bei der Auswahl eines standortübergreifenden Repositoriums. Die Aufbewahrungsfrist beginnt mit dem Datum der Herstellung des öffentlichen Zugangs und beträgt in der Regel zehn Jahre. Sofern verständliche Gründe dafür existieren, bestimmte Daten nicht oder über einen kürzeren Zeitraum aufzubewahren, legen die Wissenschaftlerinnen und Wissenschaftler dies gegenüber der Institutsleitung dar; diese Gründe werden in der Regel bis zum Ablauf der zehn Jahre dokumentiert.“ (Max Weber Stiftung, 2022)

### 2.2.3 Aufstellen von Regeln zur Langzeitarchivierung und -verfügbarkeit

Bei der Datenspeicherung kann zwischen sog. „hot data“ (d. h. Speicherung von Daten, auf die noch häufig zugegriffen wird, bzw. die noch geändert werden), „cold data“ (abgeschlossene, i. d. R. gespeicherte bzw. publizierte und damit weitgehend unveränderliche Datensätze) und Daten auf der Langzeitarchivierungsebene unterschieden werden.

Für letztere sind Langzeitarchivierungs- bzw. -verfügbarkeitsregeln unter Berücksichtigung etwa des möglichen „Wertes“ der Forschungsdaten für die Einrichtung unter Abwägung der Kosten der Langzeitarchivierung sinnvoll.

Siehe z. B. Guidelines for the publication and archiving of research data at the TU Braunschweig, V1.4/2021-02-26:

„Category 1: Very valuable research data

Research data that

- are unique or very difficult to restore, and
  - are of interest to a large community and/or for a wide range of subsequent uses.
- > Definitely archive / publish!

Category 2: Limited valuable research data

Research data that either

- can be reproduced with reasonable effort or
  - are only interesting for a short time, a small target group or only for the proof of research results.
- > Archive/publish if effort and cost do not argue against it.

Category 3: Research data of little value

Research data that is

- created only for testing purposes,

- not readable or not sufficiently described.  
-> Data are not suitable for archiving/publication“ (TU Braunschweig, 2021)

Zusätzlich können, sofern erwünscht, fachspezifische Festlegungen getroffen werden, die in der jeweiligen Disziplin Usus sind.

Ein Beispiel könnte, im Bereich der Klimaforschung, die Festlegung auf dortige, für die Langzeitarchivierung geeignete Dateiformate und -standards wie Network Common Data Format (NetCDF) und GRIdded Binary (GRIB) sein, siehe DKRZ Long Term Archive (LTA) Preservation and Storage Policy des Deutschen Klimarechenzentrums (Deutsches Klimarechenzentrum (DRKZ), 2022).

#### 2.2.4 Datentestamente

Es kann bei aktiven Forschungsprojekten der Fall eintreten, dass noch keine Festlegung getroffen werden kann, was nach Abschluss des Projekts langzeitarchiviert werden soll; jedoch dann plötzlich die Daten verwaisen, da das Projekt vorzeitig eingestellt wird, die ursprünglichen Dateneigentümerinnen bzw. -eigentümer das Projekt verlassen, oder ähnliches. Für solche Fälle empfiehlt es sich, bereits frühzeitig ein sogenanntes Datentestament aufzusetzen, wer dann Zugriff auf die Daten erhält und über weitere Schritte wie z. B. Löschung oder Langzeitarchivierung entscheiden darf.

Dies kann, je nach Festlegung, z. B. der jeweilige Vorgesetzte (Projektleitung, Abteilungsleitung) sein und/oder der Data Steward, sofern es diese Rolle an der Einrichtung oder dem Fachbereich gibt.

Zu Datentestamenten gibt es bislang wenige Praxisbeispiele (FAU Erlangen, n.d.), man könnte sich jedoch z. B. an solchen Formulierungen orientieren, siehe Ausschnitte aus der Research Data Policy guidance der University of Bath:

„[...] the Data Steward for the project will be responsible for the management of the data created by them, or by members of their research group [...]. This may involve the Data Steward transferring a copy of data onto resilient project storage prior to the end of the project, and ensuring the long-term preservation and publication of data via an appropriate data archive.

[..] Before leaving the University [...], researchers must first discuss management and retention of all research data for which they are responsible with their line manager, doctoral or project supervisor, or Head of Department as appropriate, to determine whether a copy of the data will be retained by the University for continued study, and whether the researcher will take a copy. These decisions must conform with requirements and restrictions arising from applicable legislation, funding, contracts, ethical approval, and policies.“ (University of Bath, n.d.)

#### 2.2.5 Festlegung von Regeln für die bibliothekarische Katalogisierung

Je nach Granularität und Zahl der FD-Publikationen der eigenen Einrichtung kann die bibliothekarische Katalogisierung immensen Aufwand bedeuten. Es stellt sich damit die

Frage, welche Forschungsdaten bedeutsam sind und katalogisiert werden sollen, unter welcher Arbeitsteilung (jede Einrichtung für sich, Patenschaften, etc.), in welcher Granularität und mit welchen Metadaten (z. B. wer ist Urheber).

Insbesondere mit Blick auf mögliche Katalogisierungspatenschaften und/oder technische Lösungen (z. B. Einspielen der Metadaten bestimmter FD-Repositoryen als Collection in Resource Discovery Systeme) sollten solche Festlegungen nicht individuell getroffen werden, sondern in Abstimmung mit dem jeweiligen Bibliotheksverbund bzw. den dort zuständigen Kommissionen (s. 1.3.4).

## 2.3 IK- und PR-Maßnahmen

### 2.3.1 Maßnahmen zur Bewerbung der FDM-Regularien und des zu nutzenden zentralen Verzeichnisses

Je nach Verbindlichkeitsgrad und Reichweite<sup>18</sup> der unter 2.2 beschriebenen Regularien zur besseren Sammlung und Sichtbarkeit von Forschungsdaten ist es wichtig bis essenziell, dass die Forschenden hinreichend darüber und etwaige Folgewirkungen bzw. Sanktionen bei Nichtbeachtung (z. B. bei Tenure-Track-Verfahren) informiert werden.

Entsprechende Kommunikationsverfahren könnten z. B. Briefings bzw. Unterlagen in Onboarding-Verfahren, Rundmails, Informations- (Helbig, 2021, 239–253) und Multiplikatorveranstaltungen (z. B. Sitzungstermine im Fakultätsrat), Aufnahme bzw. Querverweise in Policies und Satzungen u. v. m. umfassen.

Äquivalente IK- und PR-Maßnahmen sollten natürlich zudem für das vorgesehene zentrale Verzeichnis (z. B. CRIS, hier zu nutzende Metadatenfelder gemäß 2.1.1) und das institutionelle FD-Repositoryum bzw. weitere empfohlene FD-Repositoryen (gemäß 2.2.2) getroffen werden.

### 2.3.2 Maßnahmen zur Affiliations- und PID<sup>19</sup>-Angabe

Um die Sammlung von Forschungsdaten und entsprechende Abfragen zu vereinfachen, empfehlen sich Vorgaben bzw. Handreichungen zur bestmöglichen Angabe von Zuordnungs-Metadaten bei (FD-)Publikationen durch die Autorinnen und Autoren der eigenen Einrichtung.

Dies umfasst neben der korrekten Affiliationsangabe, etwa durch eine Affiliationsrichtlinie (Friedrich-Schiller-Universität Jena, n.d.), auch die Forcierung der Nutzung von persistenten Einrichtungs- und Autorenidentifiern wie ORCID und ROR (FAU Erlangen, n.d.).

---

<sup>18</sup> So betrifft 2.2.2 primär die lokale Bibliothek, kann jedoch auch weitere Auswirkungen haben, sofern z. B. der CIO auch Bibliotheks- und Verbundkataloge in Punkto Outreach von Forschungsdaten im Blick hat und als wertvoll erachtet.

<sup>19</sup> Persistente Identifier, z. B. zur eindeutigen Identifikation der Affiliation der Autorinnen bzw. Autoren.

Es können entsprechende Informationsseiten und Vorgaben formuliert werden und auch die Teilnahme im ORCID-DE Konsortium (*ORCID - Open Researcher and Contributor ID*, n.d.) ist möglich. Als Mitglied im Konsortium steht die effektive Implementierung von ORCID an der Einrichtung im Vordergrund, etwa die organisationsübergreifende Kommunikation der Vorteile von ORCID mittels Materialien des Konsortiums sowie die technische Umsetzung der Best-Practice-ORCID-Implementierungsworkflows (*Getting Started With Your Integration* -, n.d.).

### 2.3.2.1 Empfehlungen zu weiteren möglichen Zuordnungsmechanismen

Bei verschiedenen generischen und disziplinären FD-Repositoryn sind teils neben Einrichtungs- und Autorenidentifizier auch weitere eindeutige Zuordnungsfunktionalitäten möglich; auch diese können beworben bzw. empfohlen werden, um eine einfachere Sammlung und Extraktion zu ermöglichen.

Je nach FD-Repository sind z. B. eigene Communities möglich, Branding-Möglichkeiten oder zentrale Berichte zu allen Datenpublikationen der eigenen Forschenden. Im Folgenden exemplarisch drei Beispiele für entsprechende Maßnahmen:

- Zenodo:
  - Für FDM-Entscheider (z. B. Universitätsbibliothek in Abstimmung mit CIO): Anlegen einer kostenlosen Community (<https://zenodo.org/communities/>) für die eigene Einrichtung.<sup>20</sup> Die Zenodo-Community sollte in dem Fall nicht von einem persönlichen (Arbeits-)Account administriert werden, sondern durch einen Zenodo-Account, der mit einem Funktionsaccount der Forschungseinrichtung verbunden ist, z. B. aus dem Bereich Bibliothek oder Servicezentrum FDM. Andernfalls gibt es beim Ausscheiden der Privatperson ggfs. Probleme.
  - Für Forschende, die auf Zenodo publizieren: Angabe der Affiliation (möglichst standardisiert, z. B. gemäß Affiliationsrichtlinie) und ORCID beim Upload.
- Open Science Framework (OSF):
  - Für FDM-Entscheider: Prüfen des Abschlusses von OSF for institutions (<https://www.cos.io/products/osf-institutions>, Achtung Kosten!).
  - Für Forschende, die auf OSF publizieren: bei „Settings“ Angabe der Affiliation (möglichst standardisiert, z. B. gemäß Affiliationsrichtlinie) unter „Employment“ und der ORCID unter „Social“ sowie Ergänzung der Affiliation zu allen OSF Projekten, wo fehlend (OSF Support, 2023).
  - Sofern ein Vertrag zu OSF für die Institution existiert: Nutzung der Möglichkeiten bezüglich Projektaffiliation / Branding (OSF Support, 2023).

<sup>20</sup> Vgl. z. B. <https://zenodo.org/communities/fau/?page=1&size=20> (abgerufen am 27.09.2023).

- Dryad:
  - Für FDM-Entscheider: Prüfen des Abschlusses einer Dryad Membership ([https://datadryad.org/stash/join\\_us](https://datadryad.org/stash/join_us), Achtung Kosten!) und Nutzung der dortigen Administrations-/Reportingmöglichkeiten: „Institutions may view data for all affiliated researchers while publishers view all data coming through in association with their journals“ (*Join the Community*, n.d.).
  - Für Forschende, die auf Dryad publizieren: Angabe der Affiliation (möglichst standardisiert, z. B. gemäß Affiliationsrichtlinie; bei Dryad erfolgt ein Abgleich via ROR) und ORCID beim Upload.

Generell wäre für jegliche FD-Repositorynsoftware und -plattform die Einführung eines Eingabefelds für einen eindeutigen Einrichtungs-Identifizier (bevorzugt ROR, daneben ggfs. auch Ringgold) beim Upload und/oder beim Autorenprofil wünschenswert und sei hiermit den Entwicklungscommunities dringend angeraten.

Exemplarisch sei als Vorbild RADAR (<https://radar.products.fiz-karlsruhe.de/de>) genannt, hier wurde mit dem RADAR-Software-Release 1.15 das Metadatenschema aktualisiert und ROR integriert:

„Die Einbindung von ROR in RADAR erfolgt über die ROR API. Sobald Nutzerinnen und Nutzer mit dem Tippen des Institutionsnamens beginnen, wird eine Vorschlagsliste angezeigt, aus der bequem der gewünschte Eintrag ausgewählt werden kann. Der Organisationsname ist anschließend auf der RADAR-Landingpage dauerhaft mit dem persistenten Eintrag im ROR-Register verlinkt, in dem weitere aktuelle Informationen zur entsprechenden Organisation zu finden sind. Die standardisierte ROR ID wird auch in den Metadaten gespeichert.“ (FIZ Karlsruhe - Leibniz-Institut für Informationsinfrastruktur, 2021)

### **2.3.3 Forderungen an FD-Suchmaschinen, FD-Repositoryn und -Softwareentwicklungscommunities zur besseren Affiliations- und PID-Unterstützung sowie -Abfragemöglichkeit**

Wie die Abschnitte 2.1.2 und 2.1.3 sowie die dort noch fehlenden Akteure – wie z. B. Google Dataset Search, B2FIND, CORE etc. – zeigen, werden normierte Affiliations- und PID-Abfragen (z. B. ROR oder Ringgold) über entsprechende Suchfelder in vielen Fällen von FD-Anbietern bzw. -Aggregatoren (wie z. B. FD-Suchmaschinen, FD-Repositoryn) noch nicht unterstützt bzw. die hierfür nötigen Informationen nicht mit indexiert oder gematcht (ROR, n.d.), sofern überhaupt vorhanden.

Neben dem Vorhandensein und den Abfragemöglichkeiten sind teils auch die Exportfunktionalitäten nicht oder nur unzureichend vorhanden oder die exportierten

Daten müssen zum Re-Import in eigene CRIS- oder Repositoriensysteme unter Umständen noch einmal aufwendig bspw. mittels OpenRefine aufbereitet werden.

Als Folge stellt sich für die Go UNITE! Arbeitsgruppe „Sammlung und Sichtbarkeit von Forschungsdaten der eigenen Institution“ (SamS) die Frage, was das beste Vorgehen wäre und wer die nötigen Verantwortlichen aus unserer FDM-Community und Ansprechpersonen sind, um diese Desiderate für die Entwicklungsplanung der jeweiligen Plattformen zu adressieren und Verbesserungen anzustoßen, z. B.

- hinsichtlich fehlenden Abfragemöglichkeiten zu spezifischen Einrichtungen (Affiliations- und ROR-Felder)
- hinsichtlich fehlenden Metadaten-Eingabefeldern bei der Submission und dem Upload-Prozess von Forschungsdaten bei FD-Repositorien (dabei möglichst Fokus auf Normierung, d. h. ROR oder bei Affiliationen standardisierte Auswahllisten anstatt oder zusätzlich zu Freitextfeldern)
- hinsichtlich der Schaffung möglicher „Schaufenster“ (z. B. Communities) für die eigene Einrichtung
- hinsichtlich besserer Such-, Matching- und Indexingfunktionalitäten.

Die Frage soll explizit nach außen hin geöffnet werden – ist es ein Wunsch, dass sich SamS oder eine GO UNITE Folge AG darum kümmert oder hat es mehr Gewicht, wenn sich die jeweilige Community<sup>21</sup> oder ein prominenter Akteur wie RDA (z. B. Data Discovery Paradigms IG (Research Data Alliance, n.d.) darum bemüht und diese Forderungen aufstellt sowie adressiert?

Mit der Veröffentlichung dieser Empfehlungen sind diese Akteure gerne dazu aufgerufen, sich entsprechend mit uns in Verbindung zu setzen und sich abzustimmen.

## 2.4 Monetäre Maßnahmen

Die teils langsamen Fortschritte bspw. mit Blick auf die Integration von ROR und ORCID oder auch bei der Ausweitung der Metadatenschemata um solche Metadaten plus Neu- bzw. Folgeindexierung sind sicherlich, je nach Anbieter und Community der Plattform, auch ressourcenbedingt.

Es läge damit a) an den Plattformen, sich um hinreichende Mittel für nötige Fortentwicklungen und Anstrengungen zu bemühen und b) an Communities, diese Mittel zur Verfügung zu stellen. Dies kann über direkte Beziehungen geschehen (z. B. Zenodo Spendenmöglichkeit (CERN & Society Foundation, n.d.), „Ways to Give“ für das Center for Open Science als Betreiber von OSF (Center for Open Science, n.d.)), aber auch über Zwischenakteure.

Hier sei v. a. The Global Sustainability Coalition for Open Science Services (SCOSS) (SCOSS, n.d.) erwähnt, welche in der Vergangenheit bspw. DSpace und Dryad gefördert haben

---

<sup>21</sup> sofern vorhanden - z. B. die EUDAT CDI Community für B2FIND etc.

(SCOSS, n.d.), sowie die Invest in Open Infrastructure (IOI) (*Invest in Open Infrastructure*, n.d.).

Zudem sind unter Umständen, bei Erfüllen der jeweiligen Fördervoraussetzungen, auch diesbezügliche Drittmittelprojekte möglich, z. B. über das DFG-Förderprogramm „Informationsinfrastrukturen für Forschungsdaten“ (Deutsche Forschungsgemeinschaft, n.d.).

### 3 Einrichtungen bzw. Akteure, die an der Umsetzung beteiligt sind oder künftig noch beteiligt werden sollten

Die skizzierten Handlungsfelder bzw. Maßnahmen betreffen mehrere Akteure. Diese sind zur Umsetzung aufgerufen. Die verschiedenen Akteure können z. B. anhand des DIAMANT 2.0 Modells (Lemaire et al., 2023) in ein Beziehungsmodell zueinander gesetzt werden. Ausgangspunkt ist dabei der Forschungsprozess. Die Forschenden sollen dabei das Ziel der Forschungsintegrität und -ökonomie im Blick haben, demnach, v. a. in unserem Kontext, ein Bewusstsein für die Notwendigkeit der Auffindbarkeit und Nutzbarkeit der Forschungsdaten haben und alle damit zusammenhängenden lokalen bzw. internen Angebote, Informationen und Vorgaben (z. B. FDM-Regularien, zentrale Verzeichnisse wie ein CRIS, Datentestamente) sowie externen (z. B. Nutzung von persistenten Einrichtungen- und Autorenidentifiern wie ORCID, Kenntnis um Communities bei externen FD-Repositoryn) kennen und berücksichtigen.

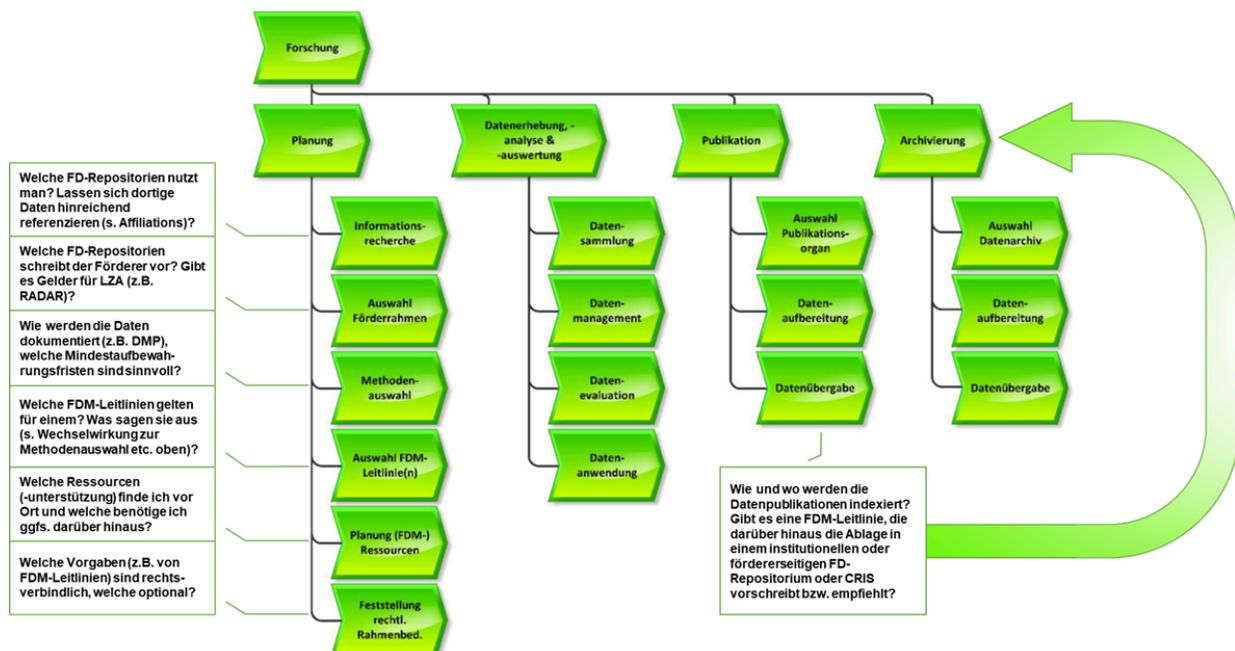


Bild: FDM-Prozesslandschaft aus LEMAIRE, Marina, et al. Das DIAMANT-Modell 2.0: Modellierung des FDM-Referenzprozesses und Empfehlungen für die Implementierung einer institutionellen FDM-Service-Landschaft. 2020.

<https://doi.org/10.25353/ubtr-xxxx-f5d2-fffb>. CC BY: Creative-Commons-Lizenz 4.0 International. Bild wurde angepasst (Ergänzung der Fragen).

Dieses Wissen, v. a. um die konkreten Abläufe, benötigt jedoch das Zusammenwirken von verschiedenen Beteiligten aus IT-Infrastruktur, Verwaltung, Bibliothek und Personalmanagement. Konkret etwa das Bekanntmachen mit den Informationen bereits bei Einstellung bzw. bei Berufung und ein stetiges Wieder-ins-Bewusstsein-rufen durch Schulungen oder praktische Unterstützungsangebote während der jeweiligen Forschungsvorhaben z. B. zu diesen Fragen (siehe links und rechts unten im o. g. Bild).

Mögliche Verantwortlichkeiten sind dabei gemäß DIAMANT 2.0 Organisationssicht der FDM-Rollen wie folgt:

- **Strategische Steuerung** (in Abstimmung mit FDM-Steuerungseinheit, Forschungsmanagement und Vertretern der Forschenden): Festlegung der Zielskala gemäß 1.2.1-1.2.3 sowie des Verbindlichkeitsgrads der Vorgaben, wie z. B. die bloße Bitte, Empfehlung bzw. verbindliche Forderung des Eintrags von Forschungsdaten in ein zentrales Verzeichnis innerhalb der Open Science- oder FDM-Policy gemäß 2.2
- **FDM-Steuerungseinheit:** Sie ist für die FDM-Services bzw. deren Steuerung verantwortlich. Die FDM-Steuerungseinheit implementiert damit zu nutzende zentrale Verzeichnisse, dezentrale Communities (z. B. bei Zenodo) und Maßnahmen zur Bewerbung der FDM-Regularien gemäß 2.3 und kümmert sich um Detailausgestaltungen sowie praktische Aspekte der FDM-Regularien (wie z. B. im Kontext Datentestamente) gemäß 2.2 und die Datensammlung je nach Zielskala gemäß 2.1.
- **Forschende:** Die Forschenden sind zum einen Adressat der FDM-Regularien als auch Nutzende der Services und des zentralen Verzeichnisses, sofern vorhanden und zur Nutzung geeignet (gemäß 2.1.1). Ihre Anforderungen sind zudem sowohl Anker- als auch Ausrichtungspunkt etwaiger Maßnahmen bzw. Infrastrukturweiterentwicklungen. So sollten langfristig gesehen jegliche Maßnahmen nicht allein auf eine händische Aufgabenerfüllung der Forschenden abzielen (z. B. darauf, ein zentrales Verzeichnis zu befüllen), sondern es müssen Arbeitserleichterungen und Automatismen gemäß 2.1.2-2.1.5 vorgesehen und realisiert werden.
- **Datenkuratorinnen bzw. -kuratoren:** Die Datenkurations-Rollen sind u. a. in jegliche Prozesse involviert, Forschungsdatensätze so gut wie möglich für die Langzeitarchivierung auszuwählen bzw. vorzubereiten. Sie müssen v. a. bei 2.2.3-2.2.4 involviert bzw. eingesetzt werden.
- **Forschungsmanagement:** Die Rollen des Forschungsmanagements umfassen in der Praxis verschiedene Unterrollen. So müssen neu eingestellte Forschende bzw. der wissenschaftliche Nachwuchs gemäß 2.3 gebrieft werden, z. B. zu lokalen Policies und Infrastrukturen. Im Kontext der Projektantragsberatung und -betreuung gibt es verschiedene Tätigkeitsaspekte rund um FDM, sowohl administrative als auch finanzielle. Im Kontext von 2.2.2 müssen z. B. bei Drittmittelprojekten die Vorgaben

und FD-Policies dieser Drittmittelgeber beachtet werden – sollten diese bspw. selbst Forschungsdaten zentralisiert monitoren bzw. eingespeist haben wollen.

- **Rechtsberatung:** Im letztgenannten Kontext kommt meist auch die Rechtsberatung für die rechtliche Einordnung und Bewertung entsprechender Policies in Kombination mit der jeweiligen Rechtsgebung (z. B. Datenschutz) und Lizenzen für Infrastrukturen und Software etc. mit ins Spiel.
- **Methodenberatung:** Die Methodenberatung kann im Lichte unserer Empfehlungen mehrere Aspekte umfassen. Zum einen Methodenberatung gegenüber den Forschenden, zum anderen jedoch auch gegenüber den anderen Rollen. Bspw. welche Zielskala die Universität gemäß 1.2.1-1.2.3 einnehmen soll oder wie man den Herausforderungen gemäß 1.3 methodisch begegnet, etwa durch vorangehende Analysen.
- **Publikations- und Archivierungsdienst:** Das DIAMANT 2.0 Modell unterscheidet zwischen Publikations- und Archivierungsdienst, teils kann es aber auch ein System für beides sein. Wichtig sind hier die Betrachtungen unter 2.1.1-2.1.2 und u. U. das Monitoring von mit Publikationen verknüpften Daten gemäß 2.1.4. Sowie im Kontext der Archivierung die Aspekte unter 2.2, vor allem 2.2.2-2.2.4.
- Weitere Rollen umfassen das **Vertragsmanagement** für die Unterstützung von Vertragsabschlüssen, etwa im Kontext von 2.1.1, 2.1.2, 2.3.2 und 2.3.2.1 (z. B. Abschluss von OSF oder Figshare for Institutions, einer Dryad oder ORCID Membership etc.) und/oder
- **Soft- und Hardware-Dienste** zur Bereitstellung der jeweils benötigten technischen Infrastrukturen unter 2.1
- sowie die jeweilige Finanzierung über das **Finanzmanagement**, siehe auch 2.4.

Insbesondere die FD-Suchmaschinen, -Repositorien und -Softwareentwicklungscommunities sind als **Soft- und Hardware-Dienste** zur besseren Affiliations- und PID-Unterstützung sowie -Abfragemöglichkeit (gemäß 2.3.3) und zum Abbau der technischen Herausforderungen gemäß 1.3.1 als akute Handlungsempfehlung aufgerufen.

## Referenzen

BASE. (2019, March 9). *Search*. Search. Retrieved November 2, 2023, from [https://www.base-search.net/about/de/about\\_sources\\_date.php?&search\\_type=Forschungsdaten](https://www.base-search.net/about/de/about_sources_date.php?&search_type=Forschungsdaten)

Bibliotheksverbund Bayern. (n.d.). *Willkommen auf den Seiten der Kommission für Erschließung und Metadaten*. Aufgaben der Kommission für Erschließung und Metadaten. Retrieved October 11, 2023, from <https://www.bib-bvb.de/web/kem/home>

- Buckinghamshire New University. (2016, 06). *Research Data Management Policy*. Research Data Management Policy Buckinghamshire New University. Retrieved November 14, 2023, from [https://www.bucks.ac.uk/sites/default/files/2021-03/research\\_data\\_management\\_policy.pdf](https://www.bucks.ac.uk/sites/default/files/2021-03/research_data_management_policy.pdf)
- Center for Open Science. (n.d.). *Support COS - Donate*. Center for Open Science. Retrieved November 14, 2023, from <https://www.cos.io/about/support-cos>
- CERN. (2021, July 23). *Version v5.0.0 - Turn-key research data management repository*. InvenioRDM. Retrieved October 11, 2023, from <https://inveniordm.docs.cern.ch/releases/versions/version-v5.0.0/#affiliations-and-ror-support>
- CERN & Society Foundation. (n.d.). *Donate now*. CERN & Society Foundation. Retrieved November 14, 2023, from [https://donate.cernandsocietyfoundation.cern/zenodo/~my-donation?\\_cv=1](https://donate.cernandsocietyfoundation.cern/zenodo/~my-donation?_cv=1)
- ckan. (n.d.). *The world's leading open source data management system*. CKAN - The open source data management system. Retrieved November 2, 2023, from <https://ckan.org>
- Consultancy*. (n.d.). Clarivate. Retrieved September 28, 2023, from <https://clarivate.com/products/scientific-and-academic-research/research-analytics-evaluation-and-management-solutions/consultancy/>
- Copyright Clearance Center. (n.d.). *Get an Organization Identifier*. Retrieved October 11, 2023, from <https://www.ringgold.com/ringgold-identifier/>
- DataCite. (n.d.). *DataCite Commons*. DataCite Commons. Retrieved November 2, 2023, from <https://commons.datacite.org/>
- Data Dryad. (n.d.). *apis*. GitHub. Retrieved November 2, 2023, from <https://github.com/CDL-Dryad/dryad-app/tree/main/documentation/apis>
- Data Dryad. (n.d.). *Stash API*. Dryad. Retrieved November 2, 2023, from <https://datadryad.org/api/v2/docs/>
- Data Dryad. (2019, March 9). *Dryad Search API*. Dryad Search API. Retrieved November 2, 2023, from <https://github.com/CDL-Dryad/dryad-app/blob/main/documentation/apis/search.md>
- Deutsche Forschungsgemeinschaft. (n.d.). *Förderprogramm "Informationsinfrastrukturen für Forschungsdaten"*. Deutsche Forschungsgemeinschaft. Retrieved November 14, 2023, from [https://www.dfg.de/foerderung/programme/infrastruktur/lis/lis\\_foerderangebote/forschungsdaten/index.html](https://www.dfg.de/foerderung/programme/infrastruktur/lis/lis_foerderangebote/forschungsdaten/index.html)
- Deutsches Klimarechenzentrum (DRKZ). (2022, December 13). *DKRZ Long Term Archive (LTA) – Preservation and Storage Policy*. World Data Center for Climate. Retrieved November 14, 2023, from <https://www.wdc-climate.de/docs/DKRZ-LTA-PreservationAndStoragePolicy.pdf>

Elsevier. (n.d.). *Data Monitor - Track and analyse your institutional research data*.

Elsevier. Retrieved November 3, 2023, from

<https://www.elsevier.com/solutions/data-monitor>

Elsevier. (2022, March 3). *Data Monitor Content and Data Policy*. Elsevier.

Retrieved November 3, 2023, from

[https://www.elsevier.com/\\_data/assets/pdf\\_file/0018/1243341/ACADRLLIBDataMonitorContent-Policy.pdf](https://www.elsevier.com/_data/assets/pdf_file/0018/1243341/ACADRLLIBDataMonitorContent-Policy.pdf)

*EPrints for Research Data – EPrints Services*. (2015, July 9). EPrints. Retrieved

October 30, 2023, from

<https://www.eprints.org/uk/index.php/flavours/eprints-for-research-data-1/>

ETH Zurich. (2022, July 1). *Guidelines for Research Data Management at ETH*

*Zurich (RDM Guidelines)*. Rechtssammlung der ETH Zürich. Retrieved November

14, 2023, from <https://rechtssammlung.sp.ethz.ch/Dokumente/414.2en.pdf>

European Commission. (2023, April 1). *EU Grants AGA - Annotated Grant*

*Agreement EU Funding Programmes 2021-2027*. EU Grants AGA - Annotated Grant

Agreement EU Funding Programmes 2021-2027. Retrieved November 14, 2023,

from

[https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/common/guidance/aga\\_en.pdf](https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/common/guidance/aga_en.pdf)

*FAIR Principles*. (n.d.). GO FAIR. Retrieved September 28, 2023, from

<http://www.go-fair.org/fair-principles/>

FAU. (2019, March 9). *Zenodo harvester (for CRIS data quality)*. Zenodo harvester

(for CRIS data quality). Retrieved November 2, 2023, from

<https://gitlab.rrze.fau.de/cdi/labs/system/data-quality/cris-zenodo-harvester/-/tree/main/>

FAU Erlangen. (n.d.). *Autoren- und Einrichtungsidentifikatoren –*

*Universitätsbibliothek*. Autoren- und Einrichtungsidentifikatoren –

Universitätsbibliothek. Retrieved November 14, 2023, from

<https://ub.fau.de/forschen/autoren-und-einrichtungsidentifikatoren/>

FAU Erlangen. (n.d.). *FAUDataCloud - IT-Services für Forschungsdaten | CDI*.

FAUDataCloud - IT-Services für Forschungsdaten | CDI. Retrieved November 14,

2023, from

<https://www.cdi.fau.de/services/faudatacloud-it-services-fuer-forschungsdaten/>

Fenner, M. (n.d.). *Best Practices - ORCID in Deutschland*. ORCID DE. Retrieved

November 2, 2023, from <https://www.orcid-de.org/support/best-practices>

figshare. (n.d.). *Figshare documentation*. Figshare documentation. Retrieved

November 2, 2023, from <https://docs.figshare.com/>

FIZ Karlsruhe - Leibniz-Institut für Informationsinfrastruktur. (2021, September

22). *RADAR Software Release (Version 1.15)*. RADAR Software Release (Version

1.15). Retrieved November 14, 2023, from

<https://radar.products.fiz-karlsruhe.de/de/nachricht/radar-software-release-version-115>

- 4Science. (2016, November 30). *AVAILABLE: DSpace-CRIS 5.6 and DSpace-CKAN!* 4Science. Retrieved November 2, 2023, from <https://www.4science.com/2016/11/30/available-dspace-cris-5-6-and-dspace-ckan/>
- Franke, F., & Ziegler, B. (2023). 6 in 1: Publikationen, Forschungsdaten, Projekte, Forschende, Einrichtungen, Auszeichnungen. *O-Bib. Das Offene Bibliotheksjournal*, 10, 1–13. <https://doi.org/10.5282/o-bib/5868>
- Freie Universität Berlin. (n.d.). *ipoa-forum Infoseite*. ipoa-forum - Expert\*innenforum für die Informationsplattform Open Access. Retrieved October 30, 2023, from <https://lists.fu-berlin.de/listinfo/ipoa-forum>
- Friedrich-Alexander-Universität Erlangen-Nürnberg. (2019, March 9). *Willkommen im CRIS, dem Forschungsinformationssystem der Friedrich-Alexander-Universität Erlangen-Nürnberg*. CRIS Home. Retrieved October 30, 2023, from [https://cris.fau.de/converis/portal?lang=de\\_DE](https://cris.fau.de/converis/portal?lang=de_DE)
- Friedrich-Schiller-Universität Jena. (n.d.). *Affiliationsrichtlinie*. Affiliationsrichtlinie. Retrieved November 14, 2023, from <https://www.uni-jena.de/affiliation>
- Geschäftsstelle der Kommission für Forschungsinformationen in Deutschland (KFiD). (n.d.). *Der KDSF – Standard für Forschungsinformationen in Deutschland*. Kerndatensatz Forschung: KDSF. Retrieved October 11, 2023, from <https://www.kerndatensatz-forschung.de/>
- Getting started with your integration* -. (n.d.). ORCID. Retrieved November 14, 2023, from <https://info.orcid.org/documentation/integration-guide/getting-started-with-your-orcid-integration/>
- GO FAIR. (n.d.). *Data Stewardship Competence Centers (DSCC)*. GO FAIR. Retrieved September 28, 2023, from <https://www.go-fair.org/implementation-networks/overview/dscc/#pt-item-germany-go-unite>
- Gollwitzer, M., Abele-Brehm, A., Fiebach, C., Ramthun, R., Scheel, A., Schönbrodt, F., & Steinberg, U. (2020). Management und Bereitstellung von Forschungsdaten in der Psychologie: Überarbeitung der DGPs-Empfehlungen. *PsyArXiv Preprints*. <https://doi.org/10.31234/osf.io/hcxtm>
- Gould, M., & Lowenberg, D. (2019, July 10). *ROR-ing Together: Implementing Organization IDs in Dryad*. Research Organization Registry (ROR). Retrieved November 2, 2023, from <https://ror.org/blog/2019-07-10-ror-ing-together-with-dryad/>
- Helbig, K. (2021). 3.3 Schulungs- und Beratungskonzepte. In *Praxishandbuch Forschungsdatenmanagement* (pp. 239-253). De Gruyter.
- Iarkaeva, A., Bobrov, E., Taubitz, J., Carlisle, B. G., & Riedel, N. (n.d.). Semi-automated extraction of information on open datasets mentioned in articles. *protocols.io*. <https://dx.doi.org/10.17504/protocols.io.q26g74p39gwz/v1>

Iarkaeva, A., Nachev, V., & Bobrov, E. (2023). Workflow for detecting biomedical articles with underlying open and restricted-access datasets. *MetaArXiv Preprints*. <https://doi.org/10.31222/osf.io/z4bkf>

Institute of Arctic & Alpine Research University of Colorado Boulder. (2021, September 22). *Standing Rules*. Standing Rules. Retrieved November 14, 2023, from <https://www.colorado.edu/instaar/sites/default/files/attached-files/INSTAAR-Standing-Rules-2023-02-16.pdf>

*Invest in Open Infrastructure*. (n.d.). Invest in Open Infrastructure. Retrieved November 14, 2023, from <https://investinopen.org>

*Join the community*. (n.d.). Dryad. Retrieved November 14, 2023, from [https://datadryad.org/stash/join\\_us](https://datadryad.org/stash/join_us)

Kommission für Erschließung und Metadaten. (2021, October 5). *9. Sitzung der Kommission für Erschließung und Metadaten Sitzungsperiode 2019/2021 - Protokoll*. Kommission für. Retrieved October 11, 2023, from <https://www.bib-bvb.de/documents/11108/10890485/9.Sitzung/95e65dd5-8234-4e4f-a41e-43d465b45737>

Kompetenznetzwerk Bibliometrie. (2022, August 30). *Kompetenznetzwerk Bibliometrie*. Startseite - Kompetenznetzwerk Bibliometrie. Retrieved September 28, 2023, from <https://bibliometrie.info/>

Kooperativer Bibliotheksverbund Berlin-Brandenburg. (n.d.). *DeepGreen*. DeepGreen. Retrieved October 11, 2023, from <https://info.oa-deepgreen.de/>

Lemaire, M., Gerhards, L., Kellendonk, S., Blask, K., & Förster, A. (2023, 06 29). *Das DIAMANT-Modell 2.0: Modellierung des FDM-Referenzprozesses und Empfehlungen für die Implementierung einer institutionellen FDM-Servicelandschaft*. <https://doi.org/10.25353/ubtr-xxxx-f5d2-ffff>

The Library Code. (n.d.). The Library Code DSpace Service Providerr. Retrieved October 11, 2023, from <https://www.the-library-code.de/index.de.html>

Lohmeier, F. (n.d.). *5.1 Zusätzliche Daten über OAI-Schnittstelle laden und mit OpenRefine verarbeiten · Seminar Praxis der Digitalen Bibliothek*. 5.1 Zusätzliche Daten über OAI-Schnittstelle. Retrieved November 2, 2023, from <https://felixlohmeier.gitbooks.io/seminar-praxis-der-digitalen-bibliothek/content/5-1-zusaetzliche-daten-ueber-oai-schnittstelle.html>

Max Weber Stiftung. (2022, May 13). *MWS-Regeln zur Sicherung guter wissenschaftlicher Praxis Beschlossen vom Stiftungsrat am 13. Mai 2022*. *INHALTSVERZEICHNIS Entsp*. Max Weber Stiftung. Retrieved November 14, 2023, from [https://www.maxweberstiftung.de/fileadmin/user\\_upload/MWS-Regeln\\_GWP.pdf](https://www.maxweberstiftung.de/fileadmin/user_upload/MWS-Regeln_GWP.pdf)

Mendeley Data. (n.d.). *API Docs*. Digital Commons Data API. Retrieved November 2, 2023, from <https://data.mendeley.com/api/docs/>

Niedersächsische Staats- und Universitätsbibliothek Göttingen. (n.d.). *HOAD - About Hybrid Open Access Dashboards (HOAD)*. Retrieved October 11, 2023, from <https://subugoe.github.io/hoaddash/about.html>

Nielsen, L. H. (2022, December 7). *Zenodo's next generation platform - InvenioRDM*. Zenodo. Retrieved October 11, 2023, from <https://blog.zenodo.org/2022/12/07/2022-12-07-zenodo-on-inveniordm/>

Nielsen, L. H. (2023, October 13). *Zenodo launched on next generation platform - InvenioRDM*. Zenodo. Retrieved October 30, 2023, from <https://blog.zenodo.org/2023/10/13/2023-10-13-zenodo-rdm/>

OpenAlex. (n.d.). *A free index for the world's research*. OpenAlex: The open catalog to the global research system. Retrieved November 2, 2023, from <https://openalex.org/>

*Open Data Monitor*. (2021, September 22). Open Data Monitor. Retrieved November 14, 2023, from <https://opendatamonitor.eu/frontend/web/index.php>

*Open Data Monitor Imprint*. (2021, September 22). Open Data Monitor Imprint. Retrieved November 14, 2023, from <https://opendatamonitor.eu/frontend/web/index.php?r=site%2Fimprint>

OpenRefine. (n.d.). *OpenRefine*. OpenRefine. Retrieved November 2, 2023, from <https://openrefine.org>

*ORCID iD Connecting research and researchers*. (n.d.). ORCID iD. Retrieved October 11, 2023, from <https://orcid.org/>

*ORCID - Open Researcher and Contributor iD*. (n.d.). Home - ORCID in Deutschland. Retrieved November 14, 2023, from <https://www.orcid-de.org/home>

OSF Support. (2023, March 21). *Getting Started with OSF Search*. Getting Started with OSF Search. Retrieved November 2, 2023, from <https://help.osf.io/article/574-advanced-search-on-osf#inst>

OSF Support. (2023, May 5). *Affiliate an OSF Project with an Institution*. OSF Support. Retrieved November 14, 2023, from <https://help.osf.io/article/192-affiliate-an-osf-project-with-an-institution>

Putnings, M. (2021). Förderpolitische Maßgaben (Preprint Langfassung). In *Praxishandbuch Forschungsdatenmanagement* (pp. 53-88). De Gruyter. <https://dx.doi.org/10.22000/289>

*quest-bih/oddpub: Algorithm for Open Data Detection in Publications (ODDPub)*. (n.d.). GitHub. Retrieved November 3, 2023, from <https://github.com/quest-bih/oddpub>

*Recommended Repositories*. (n.d.). PLOS. Retrieved September 28, 2023, from <https://journals.plos.org/plosone/s/recommended-repositories>

Registry of Research Data Repositories (re3data). (n.d.). Home | re3data.org. Retrieved September 28, 2023, from <https://doi.org/10.17616/R3D>

- Research Data Alliance. (n.d.). *Data Discovery Paradigms IG / RDA*. Research Data Alliance. Retrieved November 14, 2023, from <https://www.rd-alliance.org/groups/data-discovery-paradigms-ig>
- Research Data Alliance. (n.d.). *Data Repository Attributes WG / RDA*. Research Data Alliance. Retrieved November 14, 2023, from <https://www.rd-alliance.org/groups/data-repository-attributes-wg>
- ROR. (n.d.). *Match organization names to ROR IDs*. ROR Documentation. Retrieved November 14, 2023, from <https://ror.readme.io/docs/match-organization-names-to-ror-ids>
- ROR. (n.d.). *ROR*. Research Organization Registry (ROR) | Home. Retrieved October 11, 2023, from <https://ror.org/>
- ROR Blog*. (n.d.). Research Organization Registry (ROR). Retrieved October 11, 2023, from <https://ror.org/blog>
- Schoexplorer. (n.d.). *Statistics*. ScholeXplorer - The Data Literature Interlinking Service. Retrieved November 2, 2023, from <https://schoexplorer.openaire.eu/#/statistics>
- Scholix. (n.d.). *About Scholix*. SCHOLIX. Retrieved November 2, 2023, from <http://www.scholix.org/about>
- Schönbrodt, F. (2016, January 6). *Changing hiring practices towards research transparency: The first open science statement in a professorship advertisement – nicebread.de*. Changing hiring practices towards research transparency: The first open science statement in a professorship advertisement – nicebread.de. Retrieved November 14, 2023, from <http://www.nicebread.de/open-science-hiring-practices/>
- SCOSS. (n.d.). SCOSS – The Global Sustainability Coalition for Open Science Services – Facilitating funding to help ensure the long-term sustainability of the world's Open Science infrastructure. Retrieved November 14, 2023, from <https://scoss.org>
- SCOSS. (n.d.). *Current funding calls – SCOSS – The Global Sustainability Coalition for Open Science Services*. SCOSS. Retrieved November 14, 2023, from <https://scoss.org/how-it-works/current-funding-calls/>
- Servicezentrum eSciences der Universität Trier & Leibniz-Zentrum für Psychologische Information und Dokumentation (ZPID). (n.d.). *Das DIAMANT-Modell*. FDM Uni Trier - FDM-Referenzmodell "DIAMANT". Retrieved October 11, 2023, from <https://fdm.uni-trier.de/>
- Staudinger, J. (2019, March 9). *DataCite plugin for import*. DataCite plugin for import. Retrieved November 2, 2023, from <https://github.com/DSpace/DSpace/pull/8593>
- TU Braunschweig. (2021, 02 26). *Guidelines for the publication and archiving of research data at the TU Braunschweig*. Guidelines for the publication and archiving of research data at the TU Braunschweig. Retrieved November 14, 2023, from

[https://www.tu-braunschweig.de/fileadmin/Redaktionsgruppen/Einrichtungen/UB/Guideline\\_Research\\_data\\_publication\\_TUBS\\_V1.4.pdf](https://www.tu-braunschweig.de/fileadmin/Redaktionsgruppen/Einrichtungen/UB/Guideline_Research_data_publication_TUBS_V1.4.pdf)

Universitätsbibliothek Dortmund. (n.d.). *Leitlinien - UB - TU Dortmund*.

Universitätsbibliothek Dortmund. Retrieved October 30, 2023, from

<https://ub.tu-dortmund.de/forschen-publizieren/hochschulbibliographie/leitlinien/>

University of Bath. (n.d.). *Research Data Policy guidance*. Research Data Policy guidance. Retrieved November 14, 2023, from

<https://www.bath.ac.uk/guides/research-data-policy-guidance/>

Verbundzentrale des GBV. (2019, March 9). *AG K10plus Katalogisierung*. AG

K10plus Katalogisierung. Retrieved October 11, 2023, from

<https://wiki.k10plus.de/display/K10PLUS/AG+K10plus+Katalogisierung>

Verbundzentrale des GBV (VZG). (2019, March 9). *Forschungsdaten*.

Forschungsdaten - K10plus - Startseite. Retrieved October 11, 2023, from

<https://wiki.k10plus.de/display/K10PLUS/Forschungsdaten>

Vrije Universiteit Amsterdam. (2020, Februar 17). *Research Data Management policy*. Vrije Universiteit Amsterdam. Retrieved October 30, 2023, from

<https://assets.vu.nl/d8b6f1f5-816c-005b-1dc1-e363dd7ce9a5/29d6b549-653f-4ec9-9d07-1598b2437c64/RDM-policy-VU-2020-EN-v2.0.pdf>

Zentralbibliothek des Forschungszentrums Jülich. (n.d.). Open Access Monitor.

Retrieved October 11, 2023, from <https://open-access-monitor.de/>