

Jochen Schirrwagen

Repositorien und Forschungsinformationssysteme bilden keine Dichotomie

Zusammenfassung: Dieser Beitrag zeichnet gemeinsame Entwicklungslinien von Forschungsinformationssystemen und Repositorien nach. Beide Systemwelten haben sich zunächst unabhängig voneinander entwickelt, was ihrer Verankerung in unterschiedlichen Verwaltungseinheiten wissenschaftlicher Einrichtungen geschuldet ist. Im Zuge von Open Access und Open Science, sowie steigenden Anforderungen an die Forschungsberichterstattung haben sich mehr und mehr Schnittmengen ergeben. In dessen Folge treiben Initiativen auf europäischer und nationaler Ebene die Konvergenz und Integration über Grundsatzvereinbarungen, Roadmaps, Leitlinien und Implementierungen beider Systemwelten voran.

Schlüsselwörter: Forschungsinformationssystem, Open-Access-Repository, Interoperabilität, Publikationsmanagement, Integration

Repositories and Current Research Information Systems do not form a dichotomy

Abstract: This article traces common lines of development of research information systems and repositories. Both system worlds initially developed independently of each other, which is due to their anchoring in different administrative units of academic institutions. In the course of Open Access and Open Science, as well as increasing demands on research reporting, more and more intersections have emerged. As a result, initiatives at the European and national levels are driving convergence and integration through fundamental agreements, roadmaps, guidelines and implementations of both system worlds.

Keywords: Research information system, open access repository, interoperability, publication management, integration

1 Einleitung

Dieser Beitrag zeichnet die gemeinsamen Entwicklungslinien von Forschungsinformationssystemen und Repositorien nach. Beide Systemwelten haben sich in unterschiedlichen Fachcommunities in den 1990er Jahren zunächst unabhängig voneinander entwickelt. Im Zuge der Einführung von Open Access Policies, Resolutionen zum Forschungsdatenmanagement, steigenden Anforderungen an die Forschungsberichterstattung haben sich mehr und mehr Schnittmengen ergeben. Maßgebliche Initiativen auf europäischer Ebene wie euroCRIS, COAR und OpenAIRE, sowie auf nationaler Ebene beteiligen sich an gemeinsamen Erfahrungsaustausch, koordinieren

abgestimmte Roadmaps und Leitlinien, die schliesslich in technische Verknüpfungen und Integrationen beider Systemwelten münden.

2 Entwicklungslinien von Forschungsinformationssystemen und Repositorien

Ausgehend davon, dass Forschungsinformationssysteme (FIS) und Open Access Repositorien (OAR) in unterschiedlichen administrativen Einheiten wissenschaftlicher Einrichtungen beheimatet sind, haben sich diese Systeme seit den 1990er Jahren zunächst eigenständig mit klar voneinander abgegrenzten Funktionalitäten, Datenmodellen, unterschiedlichen Zielsetzungen, Bedarfen und Zielgruppen entwickelt. Dadurch konnte zunächst aus organisatorischer wie technischer Perspektive der Eindruck entstehen, es handle sich um eine „angebliche Dichotomie“ zwischen beiden Systemwelten.¹

Dieser Eindruck könnte sich verfestigen, wenn man sich lediglich auf die Definitionen für FIS und OAR festlegen würde.

Demnach vereinigt ein OAR als System eine Reihe von Diensten zur Aufnahme, Speicherung, Verwaltung, und Anzeige von digitalen Objekten.² Darüberhinaus ermöglicht es die Abfrage und Verbreitung bibliographischer Metadaten über diese digitalen Objekte. Zu diesen gehören alle Formen wissenschaftlicher Publikationen, aber auch Forschungsdaten und Forschungssoftware. Dabei sind die Volltexte, Forschungsprimärdaten und Softwarecode in der Regel frei und unverzüglich zugänglich.

Sehr häufig sind OAR an wissenschaftlichen Bibliotheken verankert und in der Regel verbunden mit weiteren Support- und Beratungsdienstleistungen zum Open-Access-Publizieren.

Unter einem FIS wird ein Informationssystem, etwa in Form einer Datenbank oder einem Decision Support System, verstanden, in welchem Daten über Forschungsaktivitäten gespeichert und verwaltet werden. Es ist ein wesentliches Werkzeug für die Forschungsberichterstattung, Forschungsadministration und Forschungsevaluation.³ FIS verwalten aggregierte Informationen, die aus anderen IT-Verwaltungssystemen über Personal, Finanzen oder Campusmanagementsystemen stammen. Sie sind in der Regel an Forschungsadministrationen wissenschaftlicher Einrichtungen verankert.

Für die europäischen FIS hat sich seit den 90er Jahren das Common European Research Information Format (CERIF)-Informationsmodell⁴ etabliert. Es wird von der euroCRIS-Initiative als Standard gepflegt und weiterentwickelt. Es umfaßt als Kern Entitäten zur Person, zur Organisationseinheit und zum Projekt. Es wird ergänzt um Entitäten über Forschendenprofile, Forschungsinfrastrukturen, Aktivitäten, Statements, Dokumente

1 De Castro, Shearer und Summann (2014).

2 Pinfield (2009).

3 Herwig und Schlattmann (2016).

4 <https://eurocris.org/services/main-features-cerif> .

und Forschungsoutput, worunter auch Publikationen fallen. In Deutschland wurde mit dem Kerndatensatz Forschung – Standard für Forschungsinformationen in Deutschland (KDSF)⁵ eine Spezifikation in Anlehnung an den CERIF-Standard geschaffen, der für das deutsche Wissenschaftssystem eine standardisierte Erfassung und Vorhaltung von Forschungsaktivitäten festlegt.

Demgegenüber ist die grundlegende Entität bei OAR die „Publikation“, die über weitere Attribute zur inhaltlichen Beschreibung ergänzt wird. OARs exponieren ihre Metadaten über die OAI-PMH-Schnittstelle⁶ meist in Form flacher Datenformate mit Dublin Core⁷ als Minimalstandard.

Während OARs demnach ihren Fokus auf die Inhalte von Forschungsergebnissen, wie Nachweise und Volltexte von Open-Access-Publikationen und deren Sichtbarkeit und Verbreitung legen, liegt der Fokus bei FIS auf der kontextualisierten Repräsentation und Sichtbarkeit von institutionellen Forschungsaktivitäten und -ergebnissen, häufig verbunden mit Funktionen für das Monitoring und Analysen für die Forschungsevaluation. Der Bereich „Publikationen“ ist dabei die entscheidende, wenn auch nicht alleinige Schnittmenge zwischen beiden Systemwelten.

3 Impulse, Standards und Richtlinien

Die Erkenntnis, dass zwischen den beiden Systemwelten vielfältige Gemeinsamkeiten als Schnittmengen existieren, wurde durch eine Reihe von Impulsen aus den Communities der OAR und FIS befördert, die in Standards und Richtlinien mündeten und sich mit Schnittstellen, Integrationen, Metadatenmappings und Austauschformaten auseinandergesetzt haben.

So wurden auf europäischer Ebene im Rahmen der Knowledge Exchange Initiative die Herausforderungen und Vorteile der Interoperabilität zwischen FIS und OAR herausgearbeitet.

Wesentliche Vorteile bestehen zum Beispiel darin, wenn die Dateneingabe mit allen dafür erforderlichen Prozessen in nur einem System erfolgt und dann von anderen Systemen unter Vermeidung von Redundanzen nachgenutzt und weiterverarbeitet werden kann. Aus Sicht von Nutzer:innen ergeben sich Effizienzgewinne und eine konsistente Datenqualität.⁸

Die auf Initiative von euroCRIS 2011 verabschiedete Rome Declaration on CRIS and OAR⁹ ist ein Bekenntnis der FIS und OAR-Communities zur Relevanz von Verfügbarkeit, Zugang und nahtloser Integration qualitativ hochwertiger Forschungsinformationen im Interesse aller an Forschungsaktivitäten beteiligten Zielgruppen.

5 <https://www.kerndatensatz-forschung.de/>.

6 <https://www.openarchives.org/pmh/>.

7 <https://www.dublincore.org/specifications/dublin-core/>.

8 Razum, M.; Simons, E.; Horstmann, W. (2007).

9 <https://www.openaire.eu/rome-declaration-on-cris-and-oar/view-document>.

Zum Erreichen dieser Ziele ist eine abgestimmte Koordination und die Anwendung offener Standards unabdingbar.

Mit der Roadmap for Future Directions for Repository Interoperability¹⁰ der Confederation of Open Access Repositories (COAR) wurden wichtige Trends für Repository Services vor dem Hintergrund des grundlegenden Wandels des wissenschaftlichen Kommunikationssystems und einer zunehmend komplexeren Infrastrukturlandschaft identifiziert. Darin wird ein Plan für kurz-, mittel- und langfristige Maßnahmen skizziert, der alle am wissenschaftlichen Kommunikationsprozess partizipierenden Zielgruppen und Services adressiert, darunter Forschungsinformationssysteme. Es wird auf die Notwendigkeit und das Potential hingewiesen, das OARs mit anderen lokalen und externen Systemen bzgl. des Austauschs von Metadaten und der Verlinkung digitaler Objekte profitieren können.

Der Weg hin zu mehr Interoperabilität, Konvergenz oder auch Verschmelzung beider Systemwelten ist evolutionär und kann nur schrittweise erfolgen.¹¹ Gründe dafür sind u.a. die vielfältige Landschaft an unterschiedlichen Plattformen und Entwicklungsstufen bzgl. der Datenmodelle und Schnittstellen, was für OAR und FIS gleichermaßen zutrifft.¹²

Ein wesentlicher Treiber für Interoperabilität von OAR und FIS mit aggregierenden Infrastrukturen sind die OpenAIRE-Guidelines¹³ seit 2010, die inhaltlich auf den DRIVER Guidelines¹⁴ basieren und ein Application Profile auf Basis des Dublin Core Metadatenstandards empfehlen. Sie wurden in entscheidenden Punkten weiterentwickelt, indem sie heute verschiedene Typen von Datenprovidern und Metadatenstandards zur Beschreibung von Forschungsinformationen bzw. Forschungsprodukten unterstützen. So wurden für OARs semantisch-syntaktische Metadatenstrukturen geschaffen, um Forschungspublikationen mit Projektinformationen europäischer Forschungsrahmenprogramme in Beziehung setzen zu können.¹⁵ Mit Veröffentlichung der „OpenAIRE Guidelines for CRIS Managers 1.0“¹⁶ wurde 2015 erstmals ein Grundstein gelegt, der die Registrierung und Aggregation von FIS in der OpenAIRE-Infrastruktur im CERIF-XML Format ermöglicht hat. Folgeversionen dieser Guidelines nehmen praktische Erfahrungen verschiedener FIS-Plattformen bzgl. der Integration mit OpenAIRE auf und berücksichtigen die aufgetretenen Schwierigkeiten bei Aggregation und Transformation von CERIF in den OpenAIRE-Index.¹⁷

10 Summann; Shearer (2015).

11 De Castro; Shearer; Summann (2014).

12 Scholze; Summann(2009).

13 <https://guidelines.openaire.eu/en/latest/>.

14 Vanderfeesten; Summann; Slabbertje (2008).

15 Diese Art der Verknüpfung über das „info:eu-repo“-Vokabular ist gewissermaßen ein Vorläufer der heute gängigen Praxis der Verknüpfung über persistente Identifier, z.B. FundRef-DOIs.

16 Houssos, N.; Joerg, B.; Dvořák, J. (2015).

17 De Castro et.al. (2017).

Im nationalen Kontext kommt den DINI AGs E-Pub und FIS¹⁸ in Zusammenarbeit mit dem KDSF-Helpdesk¹⁹ bzgl. Handlungsempfehlungen und Standardisierung eine besondere Rolle zu.

Hervorzuheben sind insbesondere das in regelmäßigen Abständen aktualisierte DINI-Zertifikat als qualitatives Gütesiegel für Open-Access-Publikationsdienste.²⁰

Um den Weg für den interoperablen Metadaten austausch zwischen OAR und KDSF-konformen FIS zu ebnen, erfolgte eine Überarbeitung des Gemeinsamen Vokabulars für Publikations- und Dokumenttypen.²¹ Es berücksichtigt neue Publikationsformate und internationale Vokabulare wie z.B. von COAR²² und schlägt erstmals ein Mapping von Publikationstypen in OAR zu denen im KDSF für FIS vor.

4 Technisch-Organisatorische Szenarien für Interoperabilität

In einer 2016 durchgeführten Umfrage der European University Information Systems (EUNIS) und euroCRIS wurde festgestellt, dass 62% der an der Umfrage teilnehmenden Einrichtungen sowohl ein OAR als auch ein FIS betreiben und 18% dieselbe Softwareanwendung einsetzen.²³

Bei den am meisten integrierten Standards handelt es sich um OAI-PMH als Protokoll, CERIF als Datenmodell und ORCID als Autor:innenidentifizier.

Wie vielfältige Praxisbeispiele, präsentiert und diskutiert zum Beispiel auf dem Workshop Repository/CRIS Interoperability der Open Repositories Conference 2019²⁴, zeigen, liegen die Schwerpunkte für Interoperabilität bzw. Integration auf Aspekten, wie:

- interoperable Workflows zum Management von Publikationen und Daten
- plattformübergreifender Austausch von Förderinformationen über Drittmittelprojekte
- die Nutzung persistenter Identifikatoren für die eindeutige und dauerhafte Adressierung von Entitäten und deren Verlinkung bzw. semantische Verknüpfung.

Es ergeben sich verschiedene Interaktions- und Integrationsszenarien:

- Kopplungen FIS-OAR; FIS mit Repositoryfunktionen, Repository erweitert um FIS-Funktionen.

18 <https://dini.de/ag/> .

19 Der KDSF-Helpdesk wird mit der Einführung der Kommission für Forschungsinformationen in Deutschland (KfiD) abgelöst, <https://www.kerndatensatz-forschung.de/index.php?id=helpdesk> .

20 <https://dini.de/dienste-projekte/dini-zertifikat/> .

21 DINI AG EPUB; DINI AG FIS (2022).

22 <https://vocabularies.coar-repositories.org/> .

23 Ribeiro, De Castro und Mennielli (2016).

24 OR19 Workshop on Repository/CRIS Interoperability (2019).

- IR als Datenquelle über wissenschaftliche Publikationen für das FIS
 - Open Access Status, Publikationsstatus, Verlag, aber auch damit verknüpfte Personen und deren Einrichtungen, Projekte und Förderer
- FIS als konsistente Publikationsdatenbank, die vor dem Hintergrund von Normdaten und kontrollierter Vokabulare die Qualität von Metadaten verbessern kann.

Als integrierte Systeme werden FIS um Funktionen eines Repositories erweitert (u.a. durch Adaption einer OAI-PMH-Schnittstelle) oder Repositorien erweitern ihre Datenmodelle um Entitäten von Forschungsinformationen, hierbei mit einem Schwerpunkt auf der Integration von Projekt-, Förderer- und Einrichtungsinformationen.

Dabei zu berücksichtigen sind

- Rollen und Akteure
- die Organisation von Geschäftsprozessen
- die Integration von persistenten Identifikatoren, wie URN, DOI, ORCID, ROR um Objekte zu verlinken, Metadaten anzureichern, um auf Informationen zu verweisen
- neben der manuellen Erfassung von Publikationen und Hochladen von Volltexten der automatisierte Import aus Zitationsdatenbanken und weiteren Archiven über REST-APIs und das Simple Web-service Offering Repository Deposit (SWORD)

Modernes Publikationsmanagement kann als Teil des Managements von Forschungsinformation verstanden werden, denn es berücksichtigt:

- Anforderungen bzgl. Leitlinien oder Mandaten zu Open Access und Forschungsdaten
- Anforderungen, Forschungsaktivitäten und -ergebnisse umfassend darzustellen (neben Publikationen auch Projekte, Patente, Forschungsinfrastrukturen, wissenschaftl. Preise)
- Aufbau und Pflege von Hochschul- bzw. Forschungsbibliographien wissenschaftlicher Einrichtungen.

5 **Ausblick**

Die gemeinsamen Entwicklungslinien von FIS und OAR verdeutlichen einen evolutionären Prozess, der beide Systemwelten zunehmend zusammenführt. Dieser Trend bzgl. des interoperablen oder alternativ integrativen Zusammenwirkens von FIS und OAR sowohl auf technischer wie auch organisatorischer Ebene wird insbesondere bei der Einführung von FIS bzw. bei Migrationen bestehender FIS und OAR-Plattformen mehr und

mehr berücksichtigt. So listet das Directory of Research Information Systems (DRIS)²⁵ für Deutschland zur Zeit 39 Einträge für installierte oder sich in der Einführung befindliche FIS. Demgegenüber stehen 284 Einträge für Repositorien im Verzeichnisdienst OpenDOAR²⁶.

Neben wissenschaftspolitischen Aspekten, wie der Implementierung von Open-Access-Mandaten oder der Umsetzung des KDSF, ergeben sich auch rein praktische Effizienz- und Optimierungspotentiale bei Betrieb und praktischer Nutzung dieser Systeme, sei es auf organisatorischer Ebene hinsichtlich abgestimmter Workflows, gemeinsamen Betriebskonzepten, auf inhaltlicher Ebene hinsichtlich Vollständigkeit und Qualität bereitgestellter Forschungsinformation oder bezüglich des Komforts für die Nutzenden.

Weiterhin entfalten FIS und OAR ihre Potentiale bei der Unterstützung der Open-Access-Transformation und bilden wichtige Infrastrukturen für eine offene Wissenschaft. Unter anderem zeigt sich dies an den ursprünglich im Kontext von Forschungsdaten aufgestellten FAIR-Prinzipien. Diese können im Grundsatz sowohl auf andere Forschungsergebnisse, wie Forschungsliteratur, angewandt werden, wie aktuelle Entwicklungen in OpenAIRE²⁷ oder die im Projekt FAIRIO erfolgte Prüfung ihrer Anwendbarkeit auf Forschungsinformation²⁸ zeigen.

Literaturverzeichnis

Czerniak, A. (2022): FAIR Prinzipien aus Sicht des OpenScience Aggregators - OpenAIRE. RDA Deutschland Tagung 2022 (RDA-DE 2022). Verfügbar unter doi: 10.5281/zenodo.6277720

De Castro, P.; Shearer, K.; Summann, F. (2014): The gradual merging of repository and CRIS solutions to meet institutional research information management requirements. In: *Procedia Computer Science*, 33, 39-46. doi:10.1016/j.procs.2014.06.007

De Castro, P.; Schirrwagen; J., Karaiskos, D.; Dvořák, J.; Bollini, A.; Bonis, V.; Gasparis, N.; et al. (2017). Progress in the Implementation of the OpenAIRE Guidelines for CRIS Managers. *Procedia Computer Science*, 106, 104-111. Verfügbar unter doi:10.1016/j.procs.2017.03.042

DINI AG EPUB; DINI AG FIS (2022): Gemeinsames Vokabular für Publikations- und Dokumenttypen. Humboldt-Universität zu Berlin. Verfügbar unter doi: 10.18452/24147

European Commission. Directorate General for Research and Innovation. (2021): Towards a reform of the research assessment system: scoping report. Publications Office. Verfügbar unter doi: 10.2777/707440

25 <https://eurocris.org/services/dris>

26 <https://v2.sherpa.ac.uk/opensoar/>

27 Czerniak (2022)

28 Hauschke et.al. (2021)

Hauschke, C.; Nazarovets, S.; Altemeier, F.; Kaliuzhna, N. (2021): Roadmap to FAIR Research Information in Open Infrastructures. In: *Journal of Library Metadata*, 21:1-2, 45-61. Verfügbar unter doi:

10.1080/19386389.2021.1999156

Herwig, S., & Schlattmann, S. (2016): Eine wirtschaftsinformatische Standortbestimmung von Forschungsinformationssystemen. In: Mayr Heinrich C., Pinzger Martin (Hrsg.), *Informatik 2016 – Informatik von Menschen für Menschen*. Köllen, Bonn, S. 901–914

Houssos, N.; Joerg, B.; Dvořák, J. (2015): OpenAIRE Guidelines for CRIS Managers 1.0. Zenodo. Verfügbar unter doi: 10.5281/zenodo.17065

Knowledge Exchange (2010): CRIS/OAR Interoperability Project. Verfügbar unter <https://www.knowledge-exchange.info/event/cris-oar-interoperability>, zugegriffen am 03.04.2022

OR19 Workshop on Repository/CRIS Interoperability (2019). Open Repositories 2019 Conference Workshop. Verfügbar unter <https://dspacecris.eurocris.org/cris/events/events01208>, zugegriffen am 04.04.2022

Pinfield, S (2009): Journals and repositories: an evolving relationship? In: *Learned Publishing* 22 (3), 165–175. Verfügbar unter doi:10.1087/2009302

Razum, M.; Simons, E.; Horstmann, W. (2007). Exchanging Research Information (Expert Report for the Knowledge Exchange Workshop on Institutional Repositories). Knowledge Exchange. Verfügbar unter <http://nbn-resolving.de/urn:nbn:de:0070-pub-17595188>, zugegriffen am 04.04.2022

Ribeiro, L.; De Castro, P.; Mennielli, M. (2016): CRIS/IR Survey report. Eunis. Verfügbar unter <https://www.eunis.org/blog/2016/03/01/crisir-survey-report/>, zugegriffen am 03.04.2022

Scholze, F.; Summann, F.(2009): Forschungsinformationen und Open Access Repository-Systeme. In: *Wissenschaftsmanagement*, 15 (2009) 3, 41-42

Summann, F. (2014): Interoperability – the IR (COAR) perspective. euroCRIS Strategic Membership Meeting Autumn 2014, Amsterdam. Verfügbar unter <http://dspacecris.eurocris.org/handle/11366/351>, zugegriffen am 03.04.2022

Summann, F.; Shearer, K. (2015): COAR Roadmap Future Directions for Repository Interoperability. Göttingen: COAR Confederation of Open Access Repositories. Verfügbar unter https://www.coar-repositories.org/files/Roadmap_final_formatted_20150203.pdf, zugegriffen am 03.04.2022

Vanderfeesten M.; Summann F.; Slabbertje M. (Hrsg.) (2008): DRIVER Guidelines 2.0 : Guidelines for content providers - Exposing textual resources with OAI-PMH (November 2008). Verfügbar unter: <http://nbn-resolving.de/urn:nbn:de:0070-pub-24916109>