
Storage for Science – Aktueller Stand und anstehende Entwicklungen eines verteilten FDM-Systems

Dirk von Suchodoletz ^{*}, Ulrich Hahn , Jonathan Bauer , Kolja Glogowski  und Mark Seifert 

^{*}RZ Universität Freiburg

Der Speicherbedarf über die verschiedenen wissenschaftlichen Fach-Communities hinweg ist in den letzten Dekaden erheblich gestiegen. Neben der reinen Verarbeitung und Speicherung kommen zunehmend Anforderungen an Such-, Auffind- und Verfügbarkeit der Daten hinzu, die sich aus der „Guten wissenschaftlichen Praxis“ oder den FAIR-Prinzipien ergeben [1, 9]. Da sich große Speichersysteme mit Laufzeiten von mindestens zehn Jahren nicht mehr durch kleinere Arbeitsgruppen realisieren lassen, wurde für das Speichersystem bwSFS (Storage-for-Science) ein föderativer Ansatz gewählt und über einen gemeinsamen Antrag der Universitäten Freiburg und Tübingen umgesetzt. Die fachliche Zuordnung der wissenschaftlichen Arbeitsgruppen orientiert sich an den bwHPC-Communities und folgt den Konzepten aus bwDATA Phase III. Das Speichersystem bwSFS realisiert eine georedundant verteilte technische Plattform für Basis-Speicherdienste mit darauf aufsetzendem Forschungsdatenmanagement und erlaubt das Teilen wissenschaftlicher Daten über deren gesamten Datenlebenszyklus hinweg. Es ist damit sowohl ein zentraler Baustein für das Data Intensive Computing der BinAC- und NEMO-HPC-Communities [3, 4] und stellt gleichzeitig Kapazitäten und Dienste für Forschende ohne HPC-Bezug der beteiligten Universitäten, des Science Data Centers BioDATEN und des NFDI-Konsortiums DataPLANT bereit.

1 bwSFS – Hardwaregrundlage und Basisdienste

Die Zielgruppen und die zu integrierenden Forschungsinfrastrukturen umfassen primär die Arbeitsgruppen aus den Anträgen, sowie zusätzliche AGs aus den abgedeckten Fachbereichen und lokale AGs zur Grundversorgung mit Langzeitspeicherdiensten [5]. Dabei strebt bwSFS eine effiziente und langfristig gesicherte Ablage von Forschungsdaten an, die ergänzend und integrierend zu bereits bestehenden Repositorien der Fachwissenschaften operiert. Mit bwSFS werden die infrastrukturellen Ressourcen einzelner Fachwissenschaften für das Forschungsdatenmanagement gebündelt, um eine bessere Unterstützung in

Das hier beschriebene Poster ist in der Open Access-Plattform der Universität Heidelberg heiDOK unter der DOI: <https://doi.org/10.11588/heidok.00029643> veröffentlicht.

der Umsetzung spezifischer FDM-Anforderungen zu erreichen. Die notwendige umfassende Beratung der Forschenden sollte über parallel laufende Aktivitäten im FDM der jeweiligen Einrichtung oder durch FDM-orientierte Projekte gewährleistet werden. Zur Umsetzung der FAIR-Anforderungen und OpenAccess-Prinzipien wird initial auf Datenmanagementpläne gesetzt, die durch Vorgaben der jeweiligen Fach-Communities mit Richtlinien für Metadatenmanagement, Archivierung und Lizenzmodelle unterstützt werden. Organisatorische Fragen auf Seiten der verschiedenen institutionellen Ebenen werden durch den von den Science Data Centern und dem Arbeitskreis FDM erstellten „Leitfragen zum verantwortungsvollen Umgang mit Forschungsdaten“ vorangetrieben. Im Folgenden wird der aktuelle Stand der aus [6], [7] und [8] weiterentwickelten Überlegungen sowie die noch anstehende Entwicklung eines verteilten FDM-Systems vorgestellt.

Die zentralen Speicherinstallationen befinden sich an den Standorten Tübingen und Freiburg, zusätzlich kommen Cache-Systeme an den Universitäten Konstanz und Stuttgart zum Einsatz. bwSFS stellt insgesamt knapp 20 Petabyte nutzbare Speicherkapazität auf Basis von NetApp-Komponenten der FAS- und StorageGrid Produktlinien bereit. Das System arbeitet mit Speicherplatzoptimierung durch Kompression und Deduplizierung, so dass insbesondere für noch in Verwendung befindliche, ungepackte Daten virtuell zusätzliche Kapazität zur Verfügung steht. Die Basis-Speicherdienste sind als Netzwerkdateisysteme NFS und SMB sowie als Objektspeicher (S3) ausgeführt. Die Filesysteme werden primär lokal an den Hauptstandorten oder via Caching-Komponente zusätzlich für Arbeitsgruppen in Stuttgart und Konstanz transparent lokal verfügbar gemacht. Ein Teil des Objektspeichers wird weltweit verfügbar sein, um insbesondere in verteilten Workflows und Kooperationen eingesetzt werden zu können. Eine Anbindung an die bwHPC-Systeme erfolgt mittels SFTP und S3. Zudem stehen verschiedene Überlegungen zum effizienten Datenaustausch innerhalb der Baden-Württembergischen Datenföderation im Rahmen von bwHPC-S5 an. Ein weiteres Ziel besteht im Angebot automatisierter Workflows für die Speicherverwaltung und Anbindung an Dienste von Fachwissenschaften. Das System verfügt über eine solide Hardwarebasis mit moderner Überwachung und verschiedenen, teilweise über die Standortgrenzen hinweg reichenden Redundanzen in Form einer kompletten Spiegelung des Filesystem-Bereichs und Erasure Coding für den Objektspeicher. Die Installation ist auf Erweiterbarkeit an beiden Standorten angelegt; zusätzliche Komponenten wurden bereits für die de.NBI-Cloud am Standort Freiburg und das QBIC in Tübingen hinzugefügt.

2 Dienste für das Forschungsdatenmanagement

bwSFS wird eine Reihe von FDM-Diensten für einzelne Fachwissenschaften der Antragsteller, der beteiligten Projekte sowie Universitäten offerieren, die im Backend auf die Basisdienste aufsetzen. Zur Unterstützung von Datenpublikationen wird innerhalb von bwSFS InvenioRDM verwendet, welches ein komfortables Userinterface und die OAI-PMH-Schnittstelle bereits beinhaltet. In diese Entscheidung wurden frühzeitig alle am FDM-Prozess beteiligten zentralen Einrichtungen und Projekte einbezogen. In Tübingen

sind das die Universitätsbibliothek, die Core-Facility eScience-Center und das SDC Bio-DATEN. In Freiburg erfolgt die Koordination mit zentralen Einrichtungen und die Communities durch die Research Data Management Group. Für die DOI-Vergabe in Invenio wird auf etablierte Dienste der beteiligten Universitätsbibliotheken zurückgegriffen. Als Speicher-Backend wird von InvenioRDM das S3-Protokoll nativ unterstützt und erlaubt so die Anbindung an die Object-Storage-Infrastruktur von bwSFS. In der Implementierung macht sich Invenio die vollen Vorteile von S3 zu nutze und dient dabei als Broker. Das System händigt bei Datenübertragungen pre-signed URLs aus, um eine direkte Verbindung zwischen Clients und Objektspeicher zu erlauben. Auf diese Weise wird die hohe Verfügbarkeit und Performanz der zugrundeliegende bwSFS-Infrastruktur direkt genutzt. Die umfangreiche REST-API von InvenioRDM bietet weiterhin viele Möglichkeiten zur Integration in Drittsysteme. Einige Workflows konnten bereits evaluiert werden, etwa das automatische Erstellen von Publikationsentwürfen aus unterschiedlichen Systemen heraus. Dazu zählen Code-Versionierungsplattformen wie GitLab oder GitHub bei neuen Code-Releases, andere Compute-Umgebungen wie HPC oder Galaxy nach Beendigung von Jobs oder auch in SDC-Gateways über Veröffentlichungstemplates. Nutzer könnten diese Entwürfe dann auf der InvenioRDM-Webseite fertigstellen und, falls erwünscht, veröffentlichen. InvenioRDM etabliert keinen Ersatz bestehender Daten-Repositorys, sondern komplementiert fehlende Angebote und bietet Ergänzungen zu existierenden Systemen, um beispielsweise eine gesicherte, lokale Zweitkopie eines Datensatzes zu hinterlegen.

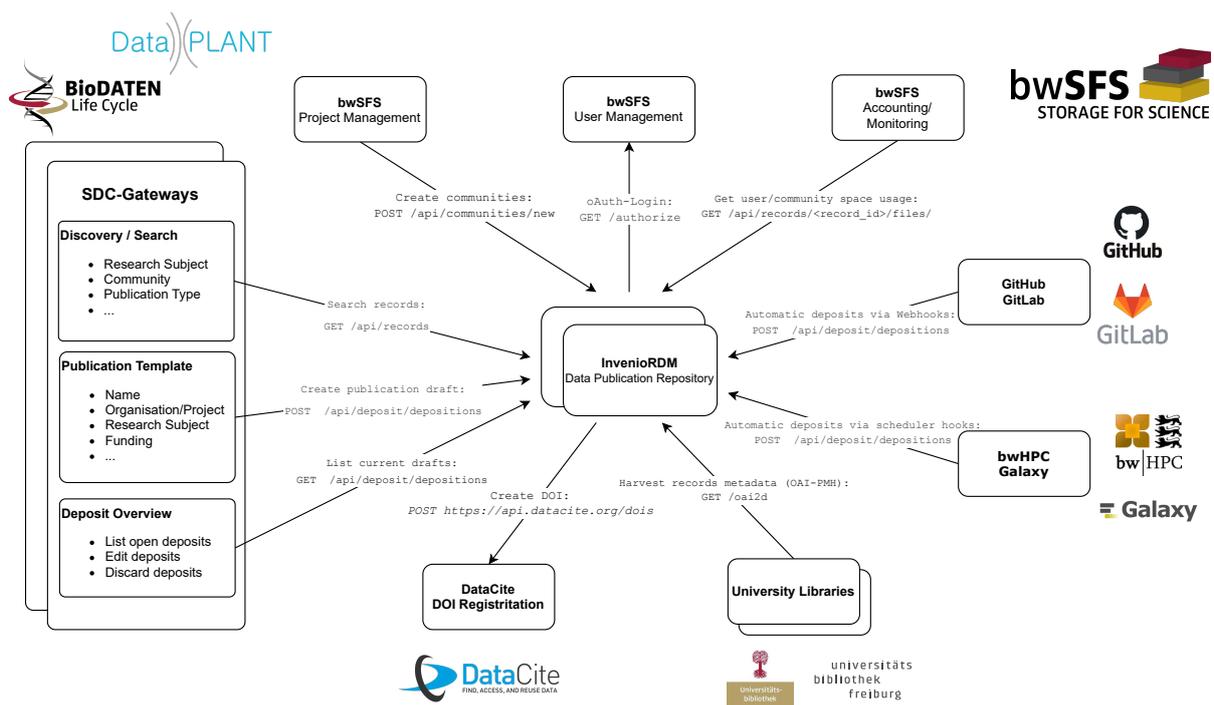


Abbildung 1: InvenioRDM als zentraler Baustein für das Publizieren und Teilen von Forschungsdaten.

In Freiburg wird eine GitLab-Instanz für Versionierung, Kollaboration und zum Teilen von Daten und Code laufender Projekte zum Einsatz kommen, ein Konzept welches bereits seit längerem durch die Research Data Alliance für viele Fachwissenschaften vorgeschlagen wird. Um dabei effizient mit großen Datenmengen umzugehen und diese möglichst direkt im Objektspeicher von bwSFS ablegen zu können, werden aktuell im Rahmen von Bio-DATEN und DataPLANT verschiedene Ansätze untersucht. Das Bilddatenmanagement-System OMERO bietet nutzerfreundliche Schnittstellen für den Zugriff, die Darstellung und die Arbeit mit Bilddaten aus der Mikroskopie. Je nach Bedarf können OMERO Instanzen als Gruppen-eigenes Repositorium z.B. auch für sensible Daten, als Kollaborationsplattform oder zur Bereitstellung öffentlicher Daten für Websites und Publikationen eingesetzt werden. OMERO legt einen starken Fokus auf den Erhalt und die Erweiterung der Metadaten. Mit OMERO erhalten Forschende der beteiligten Universitäten die Möglichkeit die stetig wachsende Menge an Mikroskopie-Bilddaten aufsetzend auf der bwSFS-Infrastruktur zu verwalten. Dazu wäre eine technische Basisinfrastruktur aus sicherer Speicherung im Filesystem und gehosteter OMERO-Instanz denkbar, die über automatisierte Deployments beispielsweise mit Kubernetes erzeugt werden kann. Das erlaubt mehrere parallel laufende Instanzen mit unterschiedlicher Konfiguration, die durch die jeweilige Arbeitsgruppe an ihre Bedürfnisse angepasst und durch diese selbst verwaltet wird. Dieses entlastet die Forschenden ebenso wie im Fall InvenioRDM und GitLab vom Betrieb eigener Basisinfrastrukturen und erlaubt ihnen den Fokus auf ihre fachlichen Belange zu richten.

Da nicht für alle Bedarfe der Antragstellenden auf bereits existierende Software zurückgreifen kann, sollen zudem eigene Dienste¹ der Fachwissenschaften gehostet werden können. Weiterhin ist eine sichere Langzeitspeicherung, abseits etablierter Repositorien, für eine DFG-konforme Ablage nicht veröffentlichter Daten mit S3-Backend vorgesehen. Die technische Grundlage der FDM-Dienste basiert auf einer "Hyper Converged Infrastructure"(HCI), die mit Kubernetes und Rancher orchestriert wird und die verschiedenen Micro-Services des Gesamtsystems betreibt. Für die Bereitstellung umfangreicherer Dienste wird je nach fachwissenschaftlicher Zuordnung auf Ressourcen der bwCloud oder der de.NBI-Cloud zurückgegriffen.

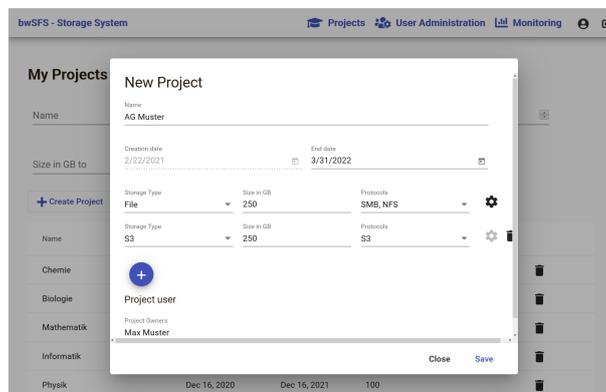
3 Nutzer- und Projektmanagement

Schon in der Implementierungsphase der Software und Dienste, die die Fachwissenschaften einbezieht, zeichnet sich ab, dass die klassischen Methoden des Identitätsmanagements nicht genügen. Im Vergleich zu HPC-Diensten erfordern Speicherdienste wegen ihrer vieltaligen Nutzerschaft und langen Haltefristen von Forschungsdaten eine wesentlich tiefere Integration bestehender Infrastrukturen und ein flexibleres Nutzermanagement. Um die vorgesehene Nutzerbasis des Systems von verschiedenen Standorten und aus den unterstützten Fachwissenschaften verwalten zu können und zukünftig eine nahtlose Inte-

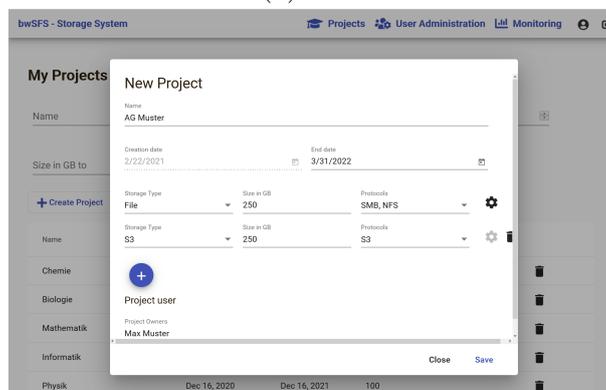
¹Zum Beispiel die Oberflächendatenbank, <https://contact.engineering/>, gefördert im Rahmen des ERC StG-757343, <https://cordis.europa.eu/project/id/757343> und livMatS

gration in die Datenföderation zu erreichen, ist ein föderiertes Management der Projekt-, User- und Gruppendaten notwendig. Hierbei wird einerseits auf etablierte Strukturen aus dem HPC-Umfeld wie bwIDM aufgesetzt, in denen ein Teil der Nutzerbasis bereits beheimatet ist. Andererseits sind weitere Quellen für Benutzer-Identitäten beziehungsweise -informationen, wie Elixir-AAI oder zusätzliche Dienste wie ORCID vorgesehen. Hierzu erfolgt eine enge Kooperation mit dem geplanten bwIDM2-Projekt, welches die Weiterentwicklung der badenwürttembergischen Identitäts Föderation um Aspekte des FDM, wie beispielsweise Fragen zur langfristigen Nutzeridentifizierung, berücksichtigt.

Unter Einbeziehung der Anforderungen der Speichersysteme und FDM-High-Level-Dienste soll ein einrichtungsübergreifender persistenter Identifikator zum bwIDM Datensatz hinzugefügt werden. Hier folgt bwSFS der Empfehlung des AK FDM in Baden-Württemberg und setzt auf die langzeitstabile ORCID ID [9], welche eine Infrastruktur für die Wiedererkennung derselben Person unabhängig von ihrer aktuellen Heimateinrichtung erlaubt.



(a) UI 1



(b) UI 2

Abbildung 2: Web-GUI der eigens entwickeltes Nutzerinterface zur Projektverwaltung.

Zur Anlage, Verwaltung und Konfiguration von Speicherprojekten wird parallel zur Inbetriebnahme der Hardware eine Projekt-Management-Software für bwSFS entwickelt. Nach der Beantragung und Bewilligung eines Speicherprojekts kann hier ein berechtigter Administrator das Projekt anlegen, projektverantwortliche Personen und technische Projektadministratoren festlegen und die dazugehörigen Speicherressourcen erstellen. Die

Software umfasst drei wesentliche Komponenten, die in einem Kubernetes-Cluster bereitgestellt werden. Als Schnittstelle zum NetApp-ONTAP-System werden Ansible-Playbooks zur Provisionierung der Speicherressourcen, derzeit NFS-/CIFS-Shares, später ebenfalls S3-Tenants/Buckets, geschrieben und durch die Integration in einer AWX-Instanz über REST-API angesteuert. Die Projekt- und Nutzerverwaltung werden als eigenständige Micro-Services mit dem Spring-Framework für die Java-Plattform realisiert; die anfallenden Daten werden in einer PostgreSQL Datenbank abgelegt. Das in Angular2 entwickelte Web-Frontend dient schließlich als graphische Schnittstelle zu den Micro-Services und ist momentan nur für berechtigte Administratoren des Systems zugänglich. Um die Basis-Funktionalität des Web-Frontends erweitern zu können, wird ein Plugin-Mechanismus entwickelt. Durch die Entwicklung eigener Plugins können weitere Speicherressourcen, wie beispielsweise der Zugriff auf FDM-Dienste, in der Web-Oberfläche angeboten werden und dabei auf die Schnittstellen der Projekt- und Nutzerverwaltung zurückgreifen. Darüber hinaus werden Monitoring und Accounting Informationen über eine im Kubernetes-Cluster betriebenen Grafana-Instanz bereit gestellt, um unter anderem eine graphische Übersicht über die Belegung aller Speicherressourcen innerhalb eines Speicherprojektes zu erhalten.

4 Fazit und Ausblick

Die Entwicklung der Konzepte für bwSFS, die Definition des Funktionsumfangs und die Ausschreibung des Systems werden nun mit der Inbetriebnahme der ersten Dienste auf einen produktiven Stand gebracht. Die Etablierung dieses Gesamtsystems über zwei Hauptstandorte und zwei Cache-Standorte und für verschiedene Fachwissenschaften stellt eine große Herausforderung dar. Dies gilt insbesondere weil es keine vorgefertigten kommerziellen Lösungen gibt, sondern eine Kombination aus einzelnen Komponenten auf Hard- und Softwareebene geschaffen werden musste. Alle im Rahmen von bwSFS entwickelten Softwarekomponenten werden unter eine Open Source Lizenz gestellt, um sowohl die Nachnutzung, als auch Erweiterung und Anpassung offen und transparent zu gestalten. Während die Basisdienste bereits verfügbar sind und von ersten Antragstellern genutzt werden, sind viele High-Level-Dienste noch in Entwicklung oder Erprobung. Auch die breite Unterstützung des Forschungsdatenmanagement stellt ein erhebliches Unterfangen dar. Die Kooperation aller beteiligter Einrichtungen und Forschenden und der übergreifende Austausch von Daten und Erfahrungen müssen angestrebt werden, da der Umfang der zu berücksichtigenden Aspekte von den einzelnen Einrichtung dauerhaft nicht zu stemmen ist. Im Erfolgsfall wird eine solche Infrastruktur die Basis, oder zumindest eine Blaupause, für größere und umfassendere Aktivitäten, wie die NFDI bilden können. Jedoch erschöpfen sich die Kosten des FDMs nicht in der Hardware [10]. Deshalb ist ein Begleitprogramm auf organisatorischer und technischer Ebene wie beispielsweise durch BioDATEN und DataPLANT notwendig, um die personellen Ressourcen für die notwendige dauerhafte Betreuung der Forschenden vorzuhalten.

Danksagungen

Wir danken der Deutsche Forschungsgemeinschaft DFG für die Unterstützung der Projekte bwSFS und DataPLANT. bwSFS wird durch die Deutsche Forschungsgemeinschaft DFG gefördert: GZ: INST 37/1046-1 FUGG, GZ: INST 37/1047-1 LAGG, GZ: INST 39/1099-1 FUGG, GZ: INST 39/1098-1 LAGG DataPLANT wird durch die Deutsche Forschungsgemeinschaft DFG gefördert: GZ: 670407 (NFDI 7/1) auf Basis der Bund-Länder-Vereinbarung zum Aufbau einer nationalen Forschungsdateninfrastruktur (NFDI) vom 26. November 2018 finanziert. Wir Danken dem Land Baden-Württemberg für die Unterstützung des Science Data Centers BioDATEN und bwSFS-Infrastruktur.

ORCID IDs

- Dirk von Suchodoletz  <https://orcid.org/0000-0002-4382-5104>
- Ulrich Hahn  <https://orcid.org/0000-0003-4471-9263>
- Jonathan Bauer  <https://orcid.org/0000-0002-5624-2055>
- Kolja Glogowski  <https://orcid.org/0000-0002-1361-5712>
- Mark Seifert  <https://orcid.org/0000-0002-1042-6107>

Literaturverzeichnis

- [1] Deutsche Forschungsgemeinschaft. *Sicherung guter wissenschaftlicher Praxis/Safeguarding Good Scientific Practice. Denkschrift/Memorandum*. Wiley Online Library, 2013. URL: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/9783527679188>, <https://doi.org/10.1002/9783527679188>.
- [2] Mark D Wilkinson, Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, Jan-Willem Boiten, Luiz Bonino da Silva Santos, Philip E Bourne, et al. The fair guiding principles for scientific data management and stewardship. *Scientific data*, 3(1):1–9, 2016.
- [3] Jens Krüger, Volker Lutz, Felix Bartusch, Werner Dilling, Anna Gorska, Christoph Schäfer, and Thomas Walter. Bioinformatics and Astrophysics Cluster (BinAC). In *Proceedings of the 3rd bwHPC-Symposium*, pages 91–95, 2017. URL: <https://books.ub.uni-heidelberg.de/heibooks/reader/download/308/308-4-79226-1-10-20171002.pdf>, <https://doi.org/10.11588/heibooks,308.418>.
- [4] Michael Janczyk, Dirk von Suchodoletz, and Bernd Wiebelt. bwforcluster nemo. In Michael Janczyk, Dirk von Suchodoletz, and Bernd Wiebelt, editors, *Proceedings of*

- the 5th bwHPC Symposium*, pages 29–50. TLP, Tübingen, 2019. URL: <http://hdl.handle.net/10900/87655>, <https://doi.org/10.15496/publikation-29041>.
- [5] Rahmenkonzept der Hochschulen des Landes Baden-Württemberg für datenintensive Dienste – bwDATA Phase III (2020-2024), 2021. URL: https://ub01.uni-tuebingen.de/xmlui/bitstream/handle/10900/114548/bwDATA_III__1_001.pdf, <https://doi.org/10.15496/publikation-55923>.
- [6] Dennis Wehrle, Bernd Wiebelt, and Dirk von Suchodoletz. Design eines FDM-fähigen Speichersystems. In *10. DFN-Forum Kommunikationstechnologien, 30.-31. Mai 2017, Berlin, Gesellschaft für Informatik eV (GI)*, pages 145–154, 2017. URL: <https://dl.gi.de/bitstream/handle/20.500.12116/470/paper10.pdf>.
- [7] Dirk von Suchodoletz, Ulrich Hahn, Bernd Wiebelt, Kolja Glogowski, and Mark Seifert. Storage infrastructures to support advanced scientific workflows. In Michael Janczyk, Dirk von Suchodoletz, and Bernd Wiebelt, editors, *Proceedings of the 5th bwHPC Symposium*, pages 263–279. TLP, Tübingen, 2019. URL: <http://hdl.handle.net/10900/87672>, <https://doi.org/10.15496/publikation-29058>.
- [8] Felix Bartusch, Kolja Glogowski, Ulrich Hahn, Michael Janczyk, Steve Kaminski, Jens Krüger, Volker Lutz, Gerhard Schneider, Mark Seifert, Dirk von Suchodoletz, Thomas Walter, and Bernd Wiebelt. Defining the future scientific data flow for multi-disciplinary research data. In *E-Science-Tage 2019: Data to Knowledge*, pages 110–127. heiBOOKS, March 2020. URL: <https://books.ub.uni-heidelberg.de/heibooks/reader/download/598/598-4-88224-1-10-20200325.pdf>, <https://doi.org/10.11588/heibooks.598>.
- [9] Dirk von Suchodoletz, Elisabeth Böker, Peter Brettschneider, and Franziska Rapp. Entwicklung in Baden-Württemberg: ORCID und ROR IDs als Standard für langfristige Personen- und Institutionen-Identifizierer. *Bausteine Forschungsdatenmanagement*, (2):80–88, 2020.
- [10] Jan Leendertse and Dirk von Suchodoletz. Kosten und Aufwände von Forschungsdatenmanagement. *Bausteine Forschungsdatenmanagement*, (1):1–7, April 2020. URL: <https://bausteine-fdm.de/article/view/8246>, <https://doi.org/10.17192/bfdm.2020.1.8246>.