

Bausteine Forschungsdatenmanagement
Empfehlungen und Erfahrungsberichte für die Praxis von
Forschungsdatenmanager*innen

Forschungsdaten im Wechsel der Jahreszeiten: LZVxCoscine.nrw als Wegbereiter für Langzeitverfügbarkeit von Forschungsdaten

Ilona Langⁱ Magdalene Cyra-Wolfⁱⁱ Philip Grossⁱⁱⁱ

2025

Zitiervorschlag

Lang, Ilona, Cyra-Wolf, Magdalene und Gross, Philip. 2025. Forschungsdaten im Wechsel der Jahreszeiten: LZVxCoscine.nrw als Wegbereiter für Langzeitverfügbarkeit von Forschungsdaten. *Bausteine Forschungsdatenmanagement. Empfehlungen und Erfahrungsberichte für die Praxis von Forschungsdatenmanager*innen* Nr. 3/2025: S. 1-9. DOI: [10.17192/bfdm.2025.3.8848](https://doi.org/10.17192/bfdm.2025.3.8848).

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

ⁱRWTH Aachen University. ORCID: [0000-0002-7202-5982](https://orcid.org/0000-0002-7202-5982)

ⁱⁱLandesinitiative für Forschungsdatenmanagement - fdm.nrw. ORCID: [0000-0001-7738-2703](https://orcid.org/0000-0001-7738-2703)

ⁱⁱⁱHochschulbibliothekszenrum NRW. ORCID: [0000-0002-9837-3205](https://orcid.org/0000-0002-9837-3205)

Abstract

In Forschungsprojekten werden häufig Forschungsdaten erhoben, deren Erhaltungswert über eine reine Speicherung (Bit Stream Preservation) hinausgeht. Forschende sind dabei auf einen Service angewiesen, der die Langzeitverfügbarkeit (LZV) von Forschungsdaten möglichst von Beginn und ohne einen Wechsel der Plattform entlang des Forschungsdatenlebenszyklus (FDLZ) sicherstellt, indem er eine kontinuierliche und dauerhafte Interpretierbarkeit sowie vielfältige Nachnutzbarkeit gewährleistet. Die Sicherstellung dieser langfristigen Perspektive stellt eine erhebliche technische und organisatorische Herausforderung dar, für deren Erfolg die Bündelung der Kompetenzen des FDM-Dienstes Coscine.nrw und des LZV-Dienstes LZV.nrw unerlässlich ist. Im Artikel stellen wir das avisierte Projekt LZVxCoscine.nrw vor, in dessen Rahmen der Funktionsumfang der FDM-Plattform Coscine um erforderliche LZV-Funktionen, wie Risikomanagement, Monitoring, LZV-Metadatenprofile und Kurationsprozesse, erweitert werden soll. In einer Kooperation der RWTH Aachen University und des Hochschulbibliothekszenentrums NRW soll eine zukunftsorientierte kooperative und skalierbare Informationsinfrastruktur geschaffen werden, die in effizienter Weise die LZV von erhaltenswerten Forschungsdaten sichert und dadurch eine entscheidende Grundlage für innovative und nachhaltige Wissenschaft erhält.

1 Vom Frühling zum Winter: Der Forschungsdatenlebenszyklus

Der Forschungsdatenlebenszyklus (FDLZ) beschreibt den Prozess, in dem Forschungsdaten von der Planung und Erhebung über die Speicherung und Analyse bis hin zur Veröffentlichung, Archivierung und Wiederverwendung verwaltet werden. Forschungsdaten, welche sich im FDLZ in der aktiven Projektphase befinden, werden häufig auch als *warme* Daten bezeichnet – diese Phase beginnt mit der Erhebung der Daten und läuft bis zur abgeschlossenen Analyse und Verarbeitung ([Abbildung 1](#)).

In NRW können diese *warmen* Daten vom Beginn ihres FDLZ über die Forschungsdatenmanagement (FDM)-Plattform Coscine (<https://about.coscine.de>) im Sinne der FAIR-Prinzipien (Wilkinson et al., 2016) verwaltet werden. Coscine wird über Coscine.nrw (Politze & Lang, 2022) allen Hochschulen der Digitalen Hochschule NRW (DH.NRW) zur Verfügung gestellt. Coscine verfolgt das Ziel, durch die Integration verschiedener sog. Ressourcentypen eine umfassende Abdeckung diverser generischer Systeme für das FDM zu ermöglichen (z.B. Speicher, GitLab, Linked Data). Über Speicherressourcen ermöglicht Coscine den Zugang zur DH.NRW-Forschungsdatenspeicherinfrastruktur (FDSI.nrw, <https://www.fdm.nrw/forschungsdatenspeicherinfrastruktur>): DataStorage.nrw (Objektspeicher) und DataArchive.nrw (Tape-Speicher). Die über Coscine verwalteten Forschungsdaten können im Laufe des FDLZ auch in der *kalten* Archivierungsphase ([Abbildung 1](#)) – nach



Abbildung 1: Der Forschungsdatenlebenszyklus – von warmen zu kalten Daten. (Bildnachweis: Ilona Lang)

Projektende – auf der FDSI.nrw gemäß der Guten Wissenschaftlichen Praxis (GWP) für 10 Jahre verweilen und von dort aus durch Projektmitglieder nachgenutzt werden.

Die Zuweisung der FDSI.nrw-Speicherressourcen erfolgt nach einem niederschweligen initialen Kurationsverfahren über die Plattform JARDS (<https://jards.coscine.de/>) und inkludiert die Prüfung der Bedingungen für die FDSI.nrw-Nutzung (v. a. Wissenschaftlichkeit, Hochschulzugehörigkeit, Projektbezug, Einhaltung der FAIR-Prinzipien) durch FDM-Personal der DH.NRW-Hochschulen (für eine detaillierte Prozessbeschreibung s. Lang, Nellesen & Politze, 2025). Der maximale Speicherplatz-Umfang für einen Projektantrag beträgt 125 TB. Größere Speicherplatz-Bedarfe werden nur nach einer zusätzlichen wissenschaftlichen Begutachtung provisioniert. Während der gesamten aktiven Projektphase können Forschungsdaten über Coscine verwaltet werden, dabei wird die Speicherung gemäß der FAIR-Prinzipien als FAIR Digital Objects (FDO) sichergestellt (Lang, Jansen & Politze, 2025), u. a. dank PID-Vergabe und einem umfangreichen Metadatenmanagement über sog. Metadatenprofile. Forschungsdaten, die veröffentlicht werden sollen, müssen durch Forschende selbständig in ein geeignetes Repository überführt werden, wie z. B. zukünftig DataPublication.nrw oder in

ein entsprechendes institutionelles oder fachspezifisches Repositorium. Forschungsdaten, die nicht veröffentlicht werden sollen und keinen längerfristigen Erhaltungswert aufweisen, werden 10 Jahre archiviert. Danach werden die Forschungsdaten von der FDSI.nrw gelöscht – zurzeit ohne Option für eine längere Aufbewahrung.

In Forschungsprojekten entstehen zum Teil auch Forschungsdaten, deren Erhaltungswert über 10 Jahren nach Projektende hinausgeht. Für diese Daten ist die reine Speicherung auf der FDSI.nrw und die dortige Bitstream Preservation (BSP) nicht ausreichend, da weder Format-Prüfungen noch eine semantische Präservation erfolgen. Eine semantische Präservation stellt sicher, dass neben der technischen auch die inhaltliche Interpretierbarkeit der Daten erhalten bleibt (s. auch CoreTrustSeal Standards and Certification Board, 2024). Gerade in der Beratung von Forschenden aus NRW wurde dieser Bedarf mehrfach im Bezug zu Coscine.nrw und der FDSI.nrw angegeben, sowie im Landeskonzept FDM festgehalten (DH.NRW I AG FDM, 2024).

Forschungsdaten mit hohem Erhaltungswert sind auf Prozesse angewiesen, die eine Langzeitverfügbarkeit (LZV) ermöglichen. LZV bedeutet, eine kontinuierliche und dauerhafte Interpretierbarkeit sowie vielfältige Nachnutzbarkeit der Forschungsdaten zu gewährleisten und wird in NRW durch den Dienst LZV.nrw (<https://www.lzv.nrw/>) vorangetrieben. LZV.nrw stellt produktiv laufende LZV-Services für DH.NRW-Hochschulen bereit. Der Zugang zu den Services erfolgt nicht direkt durch Forschende, sondern vermittelt durch deren Hochschulbibliotheken. Fokussiert werden sowohl Bestände der Hochschulbibliotheken als auch Forschungsdaten aus institutionellen Repositorien. Diese Materialien werden in einem kooperativen Verfahren gemeinsam mit den Hochschulen geharvestet, ggf. angereichert, validiert und in die LZV-Infrastruktur auf Basis von Rosetta eingeliefert. Die Speicherung erfolgt hierbei zentral in einem hierarchischen Speichermanagement-System, das Teil der LZV.nrw-Infrastruktur ist. Eine Anbindung externer Speichersysteme wie dem FDSI.nrw über Coscine erweitert das Servicekonzept und stellt vor allem für speicherintensive Forschungsdaten eine nachhaltige und ressourcenschonende Nutzung gegebener Infrastrukturkomponenten dar.

2 LZVxCoscine.nrw – Ein System für jede Forschungsdatenjahreszeit

Aktuell müssen Forschende in NRW, die ihre Daten langfristig verfügbar halten wollen, einen Infrastrukturwechsel von Coscine.nrw zu LZV.nrw vornehmen. Mit dem geplanten Projekt LZVxCoscine.nrw sollen Synergien der DH.NRW-Dienste genutzt werden, um Forschenden die Verwaltung ihrer Forschungsdaten während des gesamten FD-LZ ohne Plattformwechsel zu ermöglichen. Für Forschungsdaten mit einem längerfristigen Erhaltungswert soll im Rahmen von LZVxCoscine.nrw von Beginn des FD-LZ eine zusätzliche Funktion in Coscine angeboten werden: die Überführung in eine LZV-Ressource (s. [Abbildung 2](#)). Die technische Basis der LZV-Ressource bildet

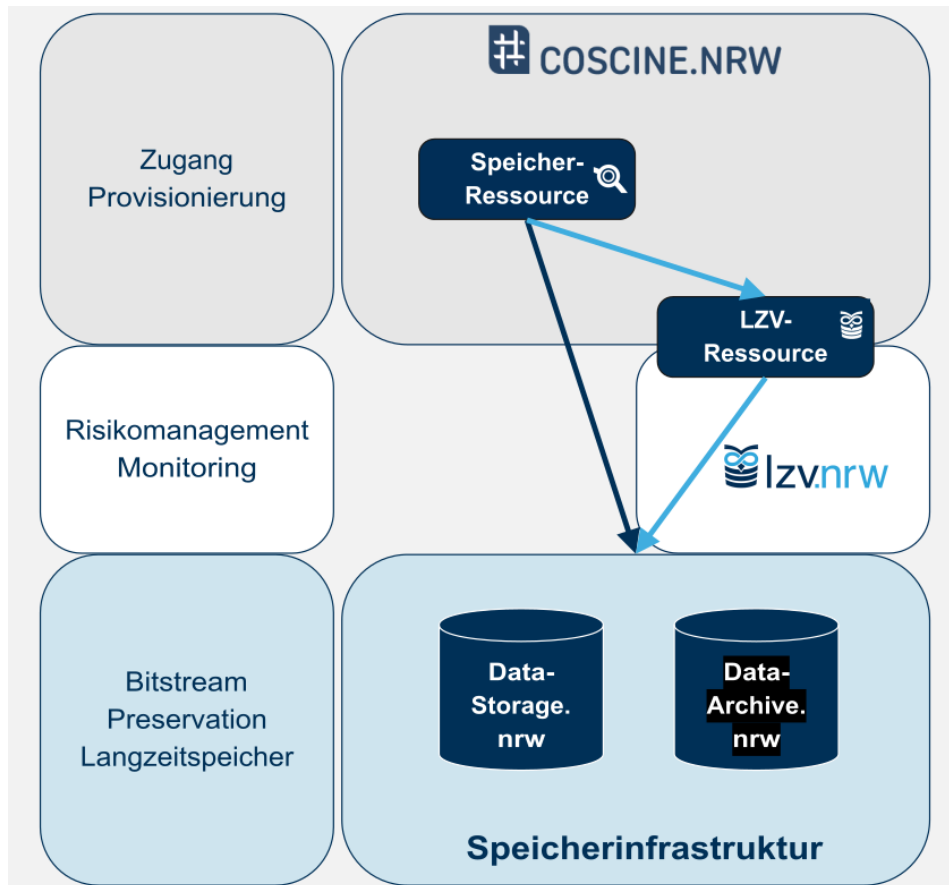


Abbildung 2: Die geplante Struktur von LZVxCoscine.nrw. Der Zugang für Forschende und die Provisionierung erfolgen über Speicher- und LZV-Ressourcen in Coscine. Das Risikomanagement und Monitoring werden von LZV.nrw u.a. durch die Software Rosetta sichergestellt. Die Speicherinfrastruktur setzt sich aus dem Objektspeicher DataStorage.nrw und dem Tape-Speicher DataArchive.nrw zusammen. Die Pfeile signalisieren den geplanten Datenfluss (Bildnachweis: Ilona Lang)

Coscine mit der darüber provisionierten FDSI.nrw. Im Vergleich zu „normalen“ Speicherressourcen soll bei LZV-Ressourcen ein Monitoring und Risikomanagement durch die im Rahmen von LZV.nrw aufgebaute Infrastruktur mit der Software Rosetta von Exlibris implementiert werden. Ziel ist es, durch geeignetes Monitoring Risiken frühzeitig zu erkennen und entsprechende Gegenmaßnahmen einzuleiten. Dies betrifft beispielsweise Prüfsummenfehler, drohende Formatobsoleszenz oder bevorstehenden Technologiewandel. Zusätzlich müssen semantische Bewahrungsmaßnahmen geplant werden, zur Erhaltung der inhaltlichen Interpretierbarkeit der (Meta)Daten (CoreTrustSeal Standards and Certification Board, 2024), damit ein „Auftauen“ der kalten Daten für eine Nachnutzung in neuen Forschungsprojekten gelingen kann. Für die benötigten Metadaten soll im Rahmen von LZVxCoscine.nrw die Entwicklung eines LZV-Metadatenprofils erfolgen, welches auf internationalen Standards basiert. Das LZV-Metadatenprofil kann im Anschluss durch Forschende als Baustein für eigene fachspezifische Metadatenprofile genutzt werden und ermöglicht so eine Eintragung aller notwendigen Metadatenfelder bereits während der warmen Projektphase.

Nicht alle Daten eines Forschungsprojekts sind für eine LZV geeignet; die Entscheidung über die kosten- und zeitintensive Aufbewahrung sollte basierend auf dem Erhaltungswert der Daten getroffen werden – wie dieser zu ermitteln ist, muss im Projekt definiert werden und ist eine der großen Herausforderungen von LZVxCoscine.nrw. Dank der Nutzung der FDSI.nrw ist in erster Instanz nicht die Größe, sondern vor allem der Erhaltungswert der Forschungsdaten entscheidend. Diverse Stakeholder sind zur Festlegung zu beteiligen, u. a. Expert*innen bspw. aus der Fachcommunity zur Einschätzung bestehender Optionen und Risiken, technische Expert*innen zur Umsetzung sowie Datenverantwortliche zur letztendlichen Entscheidung. Die Beantragung einer LZV-Ressource soll, wie bei den normalen Speicherressourcen, durch die Forschenden über JARDS erfolgen. Die Begutachtung der Anträge sollte von qualifiziertem FDM-Personal der nutzenden DH.NRW-Hochschulen durchgeführt werden – die Erarbeitung des konkreten Begutachtungsprozesses ist Teil von LZVxCoscine.nrw. Die zusätzlichen Kurationschritte für LZV-Ressourcen (z. B. Verwendung eines LZV-Metadatenprofils, Nutzungsrechte, Dateiformate) sollen als ein Meilenstein von LZVxCoscine.nrw für einen Anforderungskatalog recherchiert, analysiert und verbindlich festgelegt werden.

Zu Beginn des Projektes soll neben dem benötigten technischen Architekturmodell der Anforderungskatalog für Forschungsdaten erstellt, ein Kurationsverfahren analysiert sowie ein vertragliches Rahmenwerk vorbereitet werden. Darauf aufbauend soll dann die kooperative Informationsinfrastruktur für LZV innerhalb von Coscine entwickelt und anschließend verstetigt werden.

Zusammenfassend soll die LZV-Ressource den Forschenden die FDSI.nrw inklusive der notwendigen technischen und organisatorischen Maßnahmen für eine LZV zur Verfügung stellen. Dafür sind die Kooperation und der Wissensaustausch zwischen den Diensten Coscine.nrw und LZV.nrw essenziell. Die Kooperationspartner bündeln im Projekt ihre technische und organisatorische Expertise, um Forschenden einen

möglichst reibungslosen Übergang zwischen den Forschungsjahreszeiten – von der Daten-Erhebung bis zur LZV – zu ermöglichen. So wird Coscine um einen wertvollen Baustein bereichert: eine langfristige Aufbewahrung von Forschungsdaten mit möglichst hoher Nachnutzbarkeit und Interpretierbarkeit.

3 Vorsorgen für die kalte Jahreszeit: Gute FDM-Planung als Wegbereiter für LZV

In einem perfekten FDLZ sollten alle Informationen zu einem Forschungsprojekt unabhängig vom jeweils verwendeten FDM-Dienst übertragbar sein und möglichst keinen manuellen Prozess von Forschenden verlangen.

Die LZV richtet am FDLZ-Ende den Blick in ein „neues Jahr“ – den nächsten beginnenden Zyklus. Dabei stellt sich die systemimmanente Frage, was bereits zu Beginn des ersten FDLZ unternommen werden kann, damit Wissen – einschließlich aller Forschungsdaten – für zukünftige Forschende interpretierbar bleibt. Kern dieser Frage sind neben den gegenwärtigen FDM-Bedarfen an Forschungsdaten (Sichtbarkeit, FAIR usw.) auch die Bedürfnisse künftiger Projekte und Forschende. Durch LZVxCoscine.nrw soll Coscine um genau diese LZV-Perspektive erweitert werden – von der ersten Beantragung von Speicherplatz über die tägliche Metadatenverwaltung bis zur finalen Speicherung in einer LZV-Ressource. Dafür sollen Erkenntnisse aus den weiteren LZV-Initiativen wie EOSC-Eden (<https://eden-fidelis.eu/>), nestor (<https://www.langzeitarchivierung.de>), oder Digitale LZV im Bibliotheksverbund Bayern (<https://lzv-bayern.de>) gesammelt und in konkrete Maßnahmen überführt werden. Da Coscine oftmals als erster FDM-Dienst im FDLZ fungiert, profitieren alle anschließenden Dienste, wie z.B. Repositorien zur Veröffentlichung oder Langzeitarchive, von einer möglichst umfassenden initialen und fortlaufenden Umsetzung der FAIR-Prinzipien (Lang, Jansen & Politze, 2025), z. B. durch die frühe Annotation mit Metadaten.

Zukünftig sollen Forschende direkt beim ersten Speicherplatzantrag für Speicherressourcen in Coscine nach dem vermuteten Bedarf einer LZV befragt werden. So können benötigte LZV-Speicherbedarfe der FDSI.nrw ermittelt werden. Zudem verringert sich der Kurationsaufwand für FDM-Personal im Vergleich zur Beantragung nach Projektende: Während der aktiven Projektphase können alle Forschenden geschult, geeignete Dateiformate und Lizenzen genutzt und benötigte Metadaten inklusive dem LZV-Metadatenprofil gepflegt werden. Für normale Speicherressourcen bleibt die Verwendung des LZV-Metadatenprofils optional – erst für LZV-Ressourcen ist sie obligatorisch. Durch diese frühen LZV-Maßnahmen soll die zeitlich auseinanderfallende warme und kalte Projektphase in Coscine möglichst gut ineinandergreifen – von der initialen Speicherung bis zur LZV. Auch bei Übertragung in einen fachspezifischen LZV-Dienst außerhalb von Coscine werden die Forschungsdaten so möglichst früh gut vorbereitet.

Infrastrukturen für Forschungsdaten

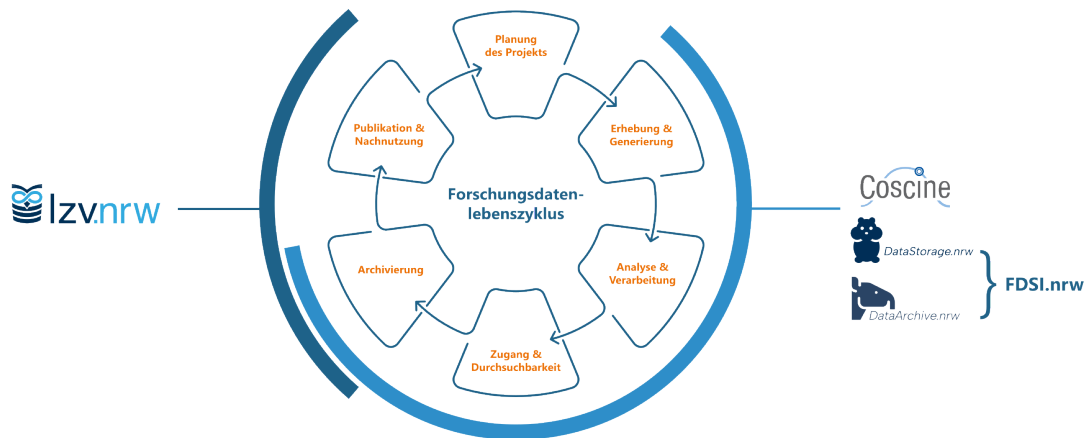


Abbildung 3: Abdeckung des Forschungsdatenlebenszyklus durch die Dienste Coscine.nrw, DataStorage.nrw, DataArchive.nrw und LZV.nrw (Bildnachweis: Landesinitiative fdm.nrw, CC-BY 4.0, <https://www.fdm.nrw/technische-fdm-landesdienste>)

4 Fazit: Immer wieder kommt ein neuer Frühling

Der FDLZ ist ein dynamischer Prozess, der bestenfalls von der Erhebung bis zur langfristigen Nachnutzbarkeit reicht und somit stetig neue Zyklen hervorbringt. In NRW entsteht das Projekt LZVxCoscine.nrw, basierend auf den Diensten Coscine.nrw, FDSI.nrw der DH.NRW und LZV.nrw. Diese Dienste haben ihre Expertise in FDM und LZV bewiesen und sich als Anbieter für Hochschulen und Forschende in NRW etabliert. Die Nutzung solcher Synergien wird in NRW hochschulpolitisch gefördert: Das im Jahr 2024 veröffentlichte Landeskoncept Forschungsdatenmanagement (DH.NRW | AG FDM, 2024) beschreibt Leitplanken für den Aufbau und Betrieb grundlegender FDM-Landesdienste, die allen Forschenden im Land zugänglich gemacht werden sollen. Diese FDM-Landesdienste sollen den gesamten FDLZ abdecken (s. [Abbildung 3, https://www.fdm.nrw/technische-fdm-landesdienste](https://www.fdm.nrw/technische-fdm-landesdienste)). Die Hochschulen in NRW haben in einem gemeinsamen Ausarbeitungsprozess die Notwendigkeit dieser Strukturen erkannt und unterstützen gemeinsam mit dem Ministerium für Kultur und Wissenschaft in NRW die fachliche und technische Zusammenarbeit zu diesem Thema.

Im geplanten Projekt LZVxCoscine.nrw sollen die FDM- und LZV-Kompetenzen zusammengeführt werden, um Forschenden in NRW eine medienbrucharmer LZV für Forschungsdaten anzubieten (s. [Abbildung 2 & 3](#)). Der Funktionsumfang der FDM-

Plattform Coscine soll um erforderliche LZV-Funktionen, wie Risikomanagement, Monitoring, LZV-Metadatenprofile und Kurationsprozesse, erweitert werden. Ein besonderes Augenmerk liegt auch auf der Schulung und Unterstützung von FDM/LZV-Personal und den Forschenden selbst, um die Übergänge zwischen den verschiedenen Phasen des FDLZ zu gewährleisten. So kann Coscine – unabhängig der jeweiligen Forschungsdatenjahreszeit – Forschende bei der Verwaltung ihrer (Meta-)Daten unterstützen. LZVxCoscine.nrw profitiert von den Erfahrungen, die in anderen Projekten und Bundesländern bereits gesammelt wurden.

Schließlich sind nachhaltige technische und organisatorische Lösungen unerlässlich. Dies erfordert nicht nur langfristige Investitionen in Hardware und Software, sondern auch in qualifiziertes Personal sowie einen Fokus auf automatisierte Prozesse. Basierend auf der deutschlandweiten Nutzung von Coscine (z. B. im Rahmen der NFDI), sollen alle Lösungen skalierbar für eine länderübergreifende Nachnutzung entwickelt werden. So sind die verwendeten technischen Elemente auch in anderen Bundesländern in Verwendung (bspw. S3-Speicher, Rosetta).

Aus unserer Perspektive ermöglichen erst die Kooperation und der Austausch zwischen FDM- und LZV-Initiativen, dass Forschungsdaten während ihres gesamten Lebenszyklus – von warm zu kalt – stets die bestmögliche Umgebung vorfinden. Gemeinsam möchten wir im Rahmen von LZVxCoscine.nrw daran arbeiten, dass wertvolle Forschungsdaten mit potenziell großem Mehrwert für künftige Generationen in Coscine langfristig verfügbar bleiben und somit zur stetigen Weiterentwicklung der Wissenschaft beitragen können.

Literaturverzeichnis

- CoreTrustSeal Standards and Certification Board. (2024). Curation & Preservation Levels: CoreTrustSeal Position Paper.
<https://doi.org/10.5281/zenodo.11476980>
- DH.NRW I AG FDM. (2024). Digitales Ökosystem DH.NRW - Landeskonzzept FDM.
<https://doi.org/10.5281/zenodo.11242683>
- Lang, I., Jansen, K., & Politze, M. (2025). Implementing the FAIR Principles with Coscine. <https://doi.org/10.5281/zenodo.14886335>
- Lang, I., Nellesen, M., & Politze, M. (2025). Die Triathlon-Staffel der Speicherplatz-Provisionierung. *E-Science Tage, Heidelberg*.
<https://doi.org/10.5281/zenodo.15111261>
- Politze, M., & Lang, I. (2022). Coscine.nrw Proposal.
<https://doi.org/10.5281/zenodo.15189277>
- Wilkinson, M., Dumontier, M., & Aalbersberg, I. e. a. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3, 160018. <https://doi.org/10.1038/sdata.2016.18>