

Konservierungsrichtlinien

Präambel

ARCHE bietet ein stabiles Hosting von digitalen Forschungsdaten für die österreichische geisteswissenschaftliche Forschungsgemeinschaft. Eine Kernfunktion von ARCHE ist die digitale Konservierung, die für die Unterstützung aktueller und zukünftiger Forschungsaktivitäten unerlässlich ist. Dieses Dokument beschreibt, wie digitale und digitalisierte Daten innerhalb von ARCHE nachhaltig bewahrt werden. Durch die Identifizierung von Schlüsselaktivitäten zur Erhaltung und Aufwertung der gehosteten Daten soll die ARCHE-Konservierungsrichtlinie einen umfassenden Entscheidungsrahmen für die zukünftige Entwicklung von Verfahren und Arbeitsabläufen bieten und eine wesentliche Dokumentation über Archivierungsmethoden und -prozesse (z. B. Dateiformate, Speicherverfahren, Sicherungsstrategien und Migrationsverfahren) bereitstellen.

Anwendungsbereich der Konservierung

Der Geltungsbereich dieser Richtlinie ist auf die digitalen Datensammlungen in ARCHE beschränkt. Sie gilt für alle Daten in allen Formaten, die in ARCHE verwaltet und zur langfristigen Aufbewahrung gespeichert werden und für die ARCHE der primäre Verwahrer ist. Interne Verwaltungsdokumente und Korrespondenz fallen nicht in den Geltungsbereich dieser Richtlinie.

1

Konservierungsgrundsätze

ARCHE wurde auf der Grundlage des [OAIS-Referenzmodells \(ISO 14721\)](#) implementiert, welches ein offenes archivisches Informationssystem (OAIS) und die damit verbundenen Strukturen und Aktivitäten beschreibt. Ein OAIS ist ein Archiv, das die Verantwortung übernommen hat, Daten zu bewahren und sie anderen zur Verfügung zu stellen. Als solches erfüllt ARCHE alle im OAIS-Referenzmodell beschriebenen obligatorischen Aufgaben, die in den folgenden Unterabschnitten beschrieben werden.

Die primäre Konservierungsstrategie von ARCHE besteht in der Migration von Formaten und nicht in der Bereitstellung von Softwareemulationen. ARCHE zielt darauf ab, seine digitalen Datensammlungen für die Wiederverwendung zu bewahren und dabei die Authentizität und Lesbarkeit zu erhalten, indem Risiken wie Verschlechterung, Beschädigung, Verlust, Korruption oder Veralterung der Datenspeicherung, der Dateiformate oder der Verbreitungswege gemindert werden.

ARCHE ist sich seiner Verantwortung bewusst, alle Veränderungen zu überwachen, die sich auf die Nachhaltigkeit und langfristige Zugänglichkeit seiner Daten auswirken könnten. Daher überwachen wir kontinuierlich die technologischen und wissenschaftlichen Entwicklungen und deren Umsetzung durch unsere technischen und organisatorischen Strategien.

Verhandlungen über angemessene Informationen

«Negotiate for and accept appropriate information from information Producers.» [Obligatorische OAIS-Aufgaben]

Die Verhandlungen über die entsprechenden Informationen finden vor und während des Datenübergabeprozesses zwischen den Datengebenden und ARCHE statt. Sie stützen sich auf die Sammlungsstrategie und den Datenübergabevertrag, die die Anforderungen an die Hinterlegung von Forschungsdaten detailliert festlegen.

Die wichtigsten Anforderungen sind:

- Die Daten müssen in den Anwendungsbereich der Sammlungsstrategie fallen (siehe [Sammlungsstrategie](#))
- Alle eingereichten Formate sollten für eine langfristige Aufbewahrung geeignet sein (siehe [Formate für die Langzeitarchivierung](#))
- Ausreichende Metadaten müssen bereitgestellt werden (siehe [ARCHE-Schema](#))
- Alle rechtlichen Fragen sind geklärt (siehe [Datenübergabevertrag](#))

Um den Datengebenden bei der Erfüllung der Anforderungen zu helfen, bietet ARCHE Informationen und Beratung vor der Dateneinreichung an. Dies trägt dazu bei, die Qualität der Daten zu gewährleisten, die Qualität der Metadaten zu verbessern, etwaige rechtliche Fragen zu klären und die Kosten im Zusammenhang mit der Aufnahme und Pflege der Daten zu senken.

Informationskontrolle für die Langzeitarchivierung

«Obtain sufficient control of the information provided to the level needed to ensure Long-Term Preservation.»
[Obligatorische OAIS-Aufgaben]

Der [Datenübergabevertrag](#) stellt sicher, dass ARCHE ausreichende Kontrolle über die Informationen für die Langzeitarchivierung erhält. In der Vereinbarung werden die Rechte und Pflichten der Datengebenden und des Repositoriums in Bezug auf die Hinterlegung, Kuratierung, Konservierung, Pflege und Verbreitung klar dargelegt.

Die:Der Datengebende muss die Vereinbarung unterzeichnen und erkennt damit an, dass sie:er das Recht hat, die Daten zu hinterlegen, dass sie:er ARCHE erlaubt, die Daten gemäß den gewählten Zugangsmodalitäten zu verbreiten, und dass sie:er alle rechtlichen bzw. ethischen Fragen berücksichtigt und geklärt hat.

Bestimmung des Zielpublikums

«Determine, either by itself or in conjunction with other parties, which communities should become the Designated Community and, therefore, should be able to understand the information provided, thereby defining its Knowledge Base.» [Obligatorische OAIS-Aufgaben]

Die designierte Zielgruppe ist in der [Sammlungsstrategie](#) festgelegt.

Sicherstellung der unabhängigen Verständlichkeit

«Ensure that the information to be preserved is Independently Understandable to the Designated Community. In particular, the Designated Community should be able to understand the information without needing special resources such as the assistance of the experts who produced the information.» [Obligatorische OAIS-Aufgaben]

Die unabhängige Verständlichkeit der Daten wird durch maßgeschneiderte Metadaten sichergestellt. Das ARCHE-Metadatenschema, das im OWL-Format formalisiert ist, macht die Semantik einzelner Eigenschaften

explizit und liefert eine Beschreibung der Daten auf Projekt-, Sammlungs- und individueller Ressourcenebene. Das Schema wurde so konzipiert, dass es generisch und flexibel genug ist, um die Dokumentation heterogener Ressourcen aus einer Vielzahl von Disziplinen zu gewährleisten. Dabei wurden internationale Standards wie die Dublin Core Metadata Initiative, DataCite und Component Metadata von CLARIN berücksichtigt. Der Metadatensatz enthält auch Eigenschaften, um Beziehungen zu anderen Publikationen, Datenquellen und verwandten Daten herzustellen.

Metadateneigenschaften können obligatorisch, empfohlen oder optional sein. Die Informationen auf der obersten Ebene der Sammlung müssen jedoch ausreichend sein, damit die Sammlung selbst verständlich ist.

Mit Hilfe von automatisch generierten Berichten führen die verantwortlichen Datenkurator:innen manuelle Qualitätsprüfungen durch, insbesondere hinsichtlich der Konsistenz und Richtigkeit der Daten. Da die von ARCHE designierte Gemeinschaft multidisziplinär ist, ist eine enge Zusammenarbeit mit den Datengebenden erforderlich. Ein:e Datenkurator:in begutachtet die bereitgestellten Metadaten, um die Konsistenz, Verständlichkeit und Vollständigkeit der Daten zu überprüfen. Wenn Probleme auftreten, kontaktiert der:die Datenkurator:in den:die Datengebende für weitere Maßnahmen. Soweit möglich, wird ARCHE die Kuratierung der hinterlegten Daten an Mitarbeiter:innen mit entsprechendem fachlichen Hintergrund übertragen.

Dokumentation und Befolgung von Grundsätzen und Abläufen

«Follow documented policies and procedures which ensure that the information is preserved against all reasonable contingencies, including the demise of the Archive, ensuring that it is never deleted unless allowed as part of an approved strategy. There should be no ad-hoc deletions.» [Obligatorische OAIS-Aufgaben]

3

Alle Maßnahmen, die für die Vorbereitung und Umsetzung einer nachhaltigen Konservierung relevant sind, werden in unseren Unterlagen und Richtlinien beschrieben. Alle wichtigen Schritte werden im ARCHE [Einreichungsprozess](#) erklärt.

Die Sammlungsstrategie hilft Datengebenden bei der Entscheidung, ob ARCHE der richtige Ort für ihre Daten ist. Die relevanten Schritte für die Vorbereitung der Daten durch die Datengebenden sind im Abschnitt „Depositionsprozess – Vor der Einreichung“ dokumentiert. Die Einreichung wird durch die Unterzeichnung des Datenübergabevertrags abgeschlossen.

Nach der Einreichung werden automatisierte Prüfungen mit speziellen Softwaretools durchgeführt und etwaige Fehler an die Datengebenden zurückgemeldet, mit der Bitte um Maßnahmen bzw. Bestätigung der vorgeschlagenen Maßnahmen. Erst nach erfolgreichem Abschluss der automatisierten Prüfungen kann die manuelle Kuratierung erfolgen. Ein allgemeiner Überblick über die Kuratierungsaktivitäten ist im Abschnitt „Depositionsprozess – Kuratierung“ bereitgestellt. Die von den ARCHE-Kurator:innen angewandten Kuratierungsverfahren werden intern dokumentiert.

Der Kuratierungsworkflow respektiert das OAIS-Referenzmodell, indem er das von der:dem Datengebenden eingereichte Submission Information Package (SIP) bei der Erstellung des Archival Information Package (AIP) unberührt lässt. Der Arbeitsablauf wurde unter Berücksichtigung der Praxis und den Empfehlungen des Data Curation Network (DCN), der Micro-Services von Archivematica, PAIMAS und IANUS entwickelt.

Der gesamte Arbeitsablauf der Kuratierung ist in sieben aufeinander folgende Phasen unterteilt:

1. **Information und Kontakt:** Beginnt mit der ersten Kontaktaufnahme zwischen der:dem Datengebenden und dem Archiv und führt nicht immer zu einer tatsächlichen Hinterlegung von Dateien. Beinhaltet im Wesentlichen Beratung.
2. **Einreichung:** Die:Der Datengebende reicht die zu archivierende Sammlung ein. Die abgelieferten Informationen stellen das SIP dar.
3. **Beitritt:** Beginnt nach Abschluss der Einreichungsphase. Diese Phase umfasst mehrere Aufgaben und ist abgeschlossen, wenn der Inhalt vollständig unter der Kontrolle von ARCHE ist.
4. **Kuratierung (von Daten und insbesondere Metadaten):** Der:Die zuständige Kurator:in analysiert die Verzeichnisstrukturen, die Dateibeziehungen und die verwendeten Namenskonventionen und prüft die Daten sowie die bereitgestellte Dokumentation und die Metadaten auf Qualitäts- und Nutzbarkeitsprobleme. Ungeklärte rechtliche Fragen, insbesondere in Bezug auf Urheberrechtsverletzungen und sensible Daten, sollten aufgedeckt werden. Der:Die designierte Kurator:in informiert die Datengebende bzw. den Datengebenden über alle gefundenen Probleme. Die Korrekturen werden in enger Zusammenarbeit durchgeführt.
5. **Transformation des SIP in AIP:** Nachdem das SIP überprüft und kuratiert wurde, kann es in ein AIP umgewandelt werden. Die Transformation kann das Umbenennen von Dateien, das Neuordnen der Ordnerstruktur, das Löschen und die Konvertierung von Dateiformaten umfassen. Alle vorgenommenen Änderungen werden im Kuratierungsprotokoll dokumentiert.
6. **AIP-Prüfung:** Das AIP wird in eine Kuratierungsinstanz des ARCHE-Systems eingespeist, wo es von der:dem Datengebenden und einem:einer zweiten Kurator:in auf etwaige Inkonsistenzen bzw. undokumentierte Transformationen überprüft wird. Diese Phase wird wiederholt, bis keine Probleme mehr gefunden werden.
7. **Einspeisung ins System:** Das AIP wird in die Produktionsinstanz des ARCHE-Systems eingespeist, und es werden automatische Integritätsprüfungen durchgeführt.

Um Datenverluste zu vermeiden, gibt es dokumentierte Verfahren zur Sicherung und Wiederherstellung. Die physische Sicherheit der aufbewahrten Daten wird vom ARZ, dem Rechenzentrum der Österreichischen Akademie der Wissenschaften (ÖAW), gewährleistet. Jede potenzielle Beschädigung von Daten wird durch regelmäßige Beständigkeitsprüfungen überwacht.

Im Falle der Auflösung des Archivs können Daten und zugehörige Metadaten in ein anderes geeignetes Archiv übertragen werden. Verbreitungsdienste zum Herunterladen von Sammlungen und entsprechenden Metadaten im ttl-Format ermöglichen den Export des Archivinhalts. Derzeit werden Aufzeichnungen über Formatmigrationen und andere Änderungen am ursprünglichen SIP außerhalb von ARCHE aufbewahrt. Die Bereitstellung solcher Provenienzinformationen in ARCHE ist derzeit in Entwicklung.

Verbreitung der bewahrten Informationen

«Make the preserved information available to the Designated Community and enable the information to be disseminated as copies of, or as traceable to, the original submitted Data Objects with evidence supporting its Authenticity.»
[Obligatorische OAIS-Aufgaben]

Alle in ARCHE aufgenommenen Daten werden durch begleitende Metadaten beschrieben. Die Metadaten können über eine spezielle Weboberfläche durchsucht und über eine Reihe von maschinenlesbaren Endpunkten, den Verbreitungsservices, abgerufen werden. Dieser wachsende Satz von Diensten liefert die Daten und Metadaten in benutzerdefinierten DIPs, je nach Datentyp und Format.

Für Sammlungen und einzelne Ressourcen stehen drei Zugriffsmodi zur Verfügung:

- Öffentlicher Inhalt (PUB), dass der Zugang für die Allgemeinheit frei und uneingeschränkt ist. Die Einstufung einer Ressource als öffentlicher Inhalt bedeutet nicht, dass sie für jeden Zweck genutzt werden darf. Die zulässigen Nutzungsarten werden mittels der Lizenz beschrieben, die jeder Ressource beiliegt.
- Akademische Inhalte erfordern eine Anmeldung mittels akademischer Anmeldedaten. Dies erfolgt durch Authentifizierung beim Identitätsanbieter der Heimatorganisation mit Hilfe der Identity Federation.
- Eingeschränkte Inhalte sind nur über spezielle Autorisierungsregeln verfügbar, die im zugehörigen Metadatensatz detailliert beschrieben werden. Normalerweise ist für den Zugriff auf Daten ein dediziertes Benutzerkonto erforderlich.

Der Zugriffsmodus ist im Datenübergabevertrag vereinbart und wird den Nutzer:innen für jede Ressource deutlich angezeigt.

Neben einem Zugriffsmodus wird den Daten auch eine Lizenz beigelegt, aus der klar hervorgeht, wie die Ressourcen weiterverwendet werden können. Wann immer möglich, wird eine offene Lizenz bevorzugt. Die verfügbaren Lizenzen sind in einem [eigenen Vokabular](#) angeführt.

Die Metadaten in ARCHE sind öffentlich und unter CC0 lizenziert. Dies ermöglicht eine weite Verbreitung, zum Beispiel durch externe Aggregatoren, wofür ein OAI-PMH-Endpunkt erforderlich ist. Der Endpunkt liefert Metadaten in einer Vielzahl von Formaten und gruppiert Metadaten auch in thematische Sets.

Formate für die Langzeitarchivierung

Da die primäre Konservierungsstrategie von ARCHE in der Migration von Formaten besteht, werden nur bestimmte Datenformate für die Archivierung akzeptiert. Die Datenformate müssen für die Langzeitarchivierung geeignet sein, was einige wichtige Voraussetzungen mit sich bringt. Geeignete Formate sollten weit verbreitet sein, (wenn möglich) offenen und nicht-proprietären Standards entsprechen und gegebenenfalls von den jeweiligen Forschungsgemeinschaften anerkannt sein. Wann immer eine Wahl der Kodierung möglich ist, sollte UTF-8 ohne Byte Order Mark (BOM) gewählt werden.

Nachstehend finden Sie eine Liste geeigneter Formate, gruppiert nach Dateitypen. Sie basiert auf den „IT-Empfehlungen“ von IANUS, die ihrerseits verschiedene internationale Standards und Empfehlungen berücksichtigen, wie z. B. die Guides to Good Practice des Archaeology Data Service oder die Sustainability of Digital Formats Website der Library of Congress. Diese Liste von Formaten kann sich im Laufe der Zeit ändern, da neue Formate entwickelt werden und andere veraltet sind.

Die Liste hilft den Datengebern und Kurator:innen bei der Entscheidung, welche Formate sie hinterlegen und akzeptieren sollen. Bevorzugte Formate sind für die Langzeitarchivierung geeignet. Akzeptierte Formate erfordern eine Konvertierung, die von ARCHE auf Anfrage angeboten werden kann. Wenn

Dateikonvertierungen notwendig werden, sollte der potenzielle Verlust von Informationen minimiert werden. Wenn eine verlustfreie Konvertierung in ein offenes oder empfohlenes Format nicht möglich ist, werden die Originaldateien zusammen mit den konvertierten Versionen aufbewahrt.

Außerdem sollten die Dateien nicht passwortgeschützt, verschlüsselt oder verlustbehaftet komprimiert sein. Wenn Dateien von Verweisen auf andere Dateien, Schriftarten oder andere externe Daten abhängen, sollten diese Objekte ebenfalls hinterlegt oder zumindest beschrieben werden, z. B. in einer README-Textdatei.

Ausnahmen sind möglich, müssen aber sorgfältig geprüft und von Fall zu Fall entschieden werden. Eine solche Ausnahme können Rohdaten sein, wie z. B. Dateien, die von einem Messgerät erzeugt werden.

Siehe Liste der [Formate für die Langzeitarchivierung](#).

Rollen und Verantwortlichkeiten

Die Konservierungsaktivitäten der ARCHE werden von einem speziellen Team durchgeführt, das in die größere Organisationsstruktur des ACDH an der Österreichischen Akademie der Wissenschaften (ÖAW) eingebunden ist. Die ÖAW stellt Mittel für Personal und Ressourcen zur Verfügung (siehe [Finanzierung und Nachhaltigkeit](#) unten) und ihre Leitung gibt die allgemeine strategische Richtung vor. Die Strategie wird von der Leitung des ACDH weiter verfeinert, die auch die ARCHE-Richtlinien beeinflusst. Die Richtlinienentwicklung und -umsetzung wird von den Arbeitsgruppenleitern innerhalb der Einheit [DH Forschung & Infrastruktur](#) des ACDH gesteuert.

Die Teammitglieder von ARCHE lassen sich grob in kuratorisches und technisches Personal unterteilen, wobei ein koordinierendes Mitglied die Kommunikation und die Sitzungen des Teams leitet und die Aktivitäten beider Parteien im Hinblick auf ihre Kompatibilität überwacht. Das Team stellt gemeinsam Dokumentationen und Benutzerinformationen über ARCHE bereit.

Die Kurator:innen sind die Ansprechpersonen für die Datenebenen. Die Kurator:innen nehmen Dateneinreichungen entgegen und prüfen die Daten auf Konsistenz, Genauigkeit und Eignung für die Archivierung und Wiederverwendung. Sie helfen auch bei der Anreicherung von Metadaten, um die Auffindbarkeit der Daten zu verbessern, und konvertieren das SIP in das AIP, um es in das System aufzunehmen. Auf einer abstrakteren Ebene ist das Kuratierungsteam auch an der Entwicklung der Ontologie und der erforderlichen Vokabulare sowie an der Formulierung der Richtlinien und Rechtstexte beteiligt. Bei rechtlichen Fragen kann die Rechtsabteilung der ÖAW konsultiert werden.

Das technische Personal ist für die Entwicklung und den Betrieb der ARCHE zugrunde liegenden Technologie verantwortlich. Für diese Aufgabe kann es auf die Bereitstellung der technischen Infrastruktur durch das Rechenzentrum der ÖAW (ARZ) zurückgreifen, das Kapazitäten für die Back- und Frontend-Entwicklung freihält. Das technische Personal stellt nicht nur die Kernkomponenten des Systems zur Verfügung, sondern entwickelt auch zusätzliche Funktionen wie z. B. einen File-Checker oder einen Metadaten-Editor.

Finanzierung und Nachhaltigkeit

Die Finanzierung des Stammpersonals erfolgt aus dem Globalbudget der Österreichischen Akademie der Wissenschaften (ÖAW). Die Wartung und Bereitstellung der Server-Hardware erfolgt durch das

Rechenzentrum der ÖAW. Beides sichert den Basisbetrieb und das Monitoring der zugrundeliegenden Systeme sowie allgemeine Managementaufgaben wie z. B. die Nachhaltigkeit für den Betrieb von ARCHE.

Kuratierung und Aufnahme von Daten institutioneller Datengebenden werden durch die Einhebung von Gebühren für den erforderlichen Aufwand ermöglicht. Datenhinterlegungen einzelner Datengebenden sowie die technische Weiterentwicklung der ARCHE-Systeme werden direkt aus Mitteln des ACDH, das aus dem Globalbudget der Akademie finanziert wird, und aus einem Mix von nationalen und internationalen drittmittelfinanzierten Projekten bestritten, die entweder direkt vom Institut oder von einer Reihe externer Kooperationspartner:innen durchgeführt werden.

Umgang mit Risiken für digitale Kontinuität

ARCHE arbeitet aktiv daran, die digitale Kontinuität der archivierten Daten, d. h. ihre Nutzbarkeit im Laufe der Zeit, sicherzustellen. Jegliche Änderungen der Organisation, der Verwaltungsprozesse oder der Technologie im Laufe der Zeit werden überwacht. Erforderliche Maßnahmen werden zeitnah geplant und umgesetzt.

Die digitale Konservierung in ARCHE konzentriert sich besonders auf die Abschwächung von Risiken, die mit der rasanten technologischen Entwicklung und der wachsenden Menge an digitalen Daten verbunden sind. Die relevanten Risiken und Maßnahmen zu ihrer Beherrschung sind nachstehend angeführt. Die Nummer der Risikopriorität bestimmt den Grad des Risikos auf der Grundlage seiner Wahrscheinlichkeit und der möglichen Auswirkungen, die es haben kann. Die Wahrscheinlichkeit und die Auswirkung werden mit einer Punktzahl von 1 (niedrigste) bis 5 (höchste) angegeben, deren Produkt die Risikoprioritätszahl darstellt.

7

Veralterung von Dateiformaten

Heute gebräuchliche Dateiformate können aufgrund von Änderungen der Hardware und insbesondere der Software in der Zukunft überholt sein und nicht mehr unterstützt werden. In ARCHE soll dieses Risiko vermieden werden, indem nur eine bestimmte Anzahl von Datenformaten für die Archivierung akzeptiert wird. Diese Liste wird sorgfältig kuratiert (siehe [Formate für die Langzeitarchivierung](#)) und bei Bedarf aktualisiert. Eine Technologiebeobachtung überwacht insbesondere die Entwicklung von Formaten mit einem hohen Risiko (siehe [NARA Digital Preservation Risk Matrix](#)), veraltet zu sein. Sollte ein Format veraltet sein, wird ein maßgeschneiderter Migrationsplan entwickelt und ausgeführt.

Risikoprioritätszahl - 10 | Wahrscheinlichkeit - 2 | Auswirkungen - 5

Komplexe digitale Objekte

Sehr neue, hochspezialisierte und komplexe Dateiformate mit zahlreichen Abhängigkeiten lassen nicht immer genaue Entscheidungen über die Konservierung zu. Künftige Probleme könnten plötzlich auftreten, wenn diese Arten von digitalen Objekten nicht allgemein angenommen werden oder wenn es ihnen an einer einheitlichen Spezifikation mangelt. Dieses Risiko lässt sich nicht vermeiden, aber es kann gemindert werden, indem sowohl das Original als auch eine migrierte Datei des Objekts zusammen mit einer ausreichenden technischen Dokumentation aufbewahrt werden.

Risikoprioritätszahl - 6 | Wahrscheinlichkeit - 2 | Auswirkungen - 3

Dateibeschädigung

Bei der Übertragung von Daten von einem Speichermedium auf ein anderes kann der Bitstrom beschädigt werden. Ein Systemausfall kann zu demselben Problem führen. Wird die Dateibeschädigung zu einem sehr späten Zeitpunkt entdeckt, ist eine Wiederherstellung möglicherweise nicht mehr möglich. Dieses Risiko wird durch die Verwendung eines angemessenen Serverspeichers, eine spezielle Sicherungsstrategie und regelmäßige Integritätsprüfungen auf der Grundlage von Prüfsummen gemindert.

Risikoprioritätszahl - 9 | Wahrscheinlichkeit - 3 | Auswirkungen - 3

Ausfall von Speichermedien

Das einzige Speichermedium, auf das ARCHE zurückgreift, ist ein spezieller institutioneller Netzwerkspeicher (in einer redundanten RAID-6-Konfiguration). Diese werden genau überwacht und im Falle eines Ausfalls ersetzt. Die Backups werden auf einem separaten Speichersystem gesichert, das an zwei physischen Standorten repliziert wird, so dass die Wiederherstellung auch im Katastrophenfall gewährleistet ist.

Risikoprioritätszahl - 5 | Wahrscheinlichkeit - 1 | Auswirkungen - 5

Menschliches Versagen (versehentliches Löschen oder Ändern)

Menschliche Fehler kommen vor und können dazu führen, dass Daten versehentlich geändert oder sogar gelöscht werden. Dieses Risiko wird dadurch gemindert, dass der Schreibzugriff auf das Archivierungssystem nur gut geschulten Datenkurator:innen gewährt wird. Außerdem fungiert eine Kuratierungsinstanz des Systems als Sicherheitsnetz, bevor die Daten in die Produktionsinstanz eingespeist werden. Vor der Übernahme in die Produktionsinstanz wird das „Vier-Augen-Prinzip“ angewendet – keine Daten werden in die Produktion übernommen, bevor sie nicht von einem:einer anderen Kurator:in freigegeben wurden. Die vollständige Historie der Metadatenänderungen wird aufbewahrt, und alle Daten werden regelmäßig gesichert, so dass alle fehlerhaften Änderungen an den Metadaten oder Löschungen des binären Inhalts rückgängig gemacht werden können.

Risikoprioritätszahl - 9 | Wahrscheinlichkeit - 3 | Auswirkungen - 3

Nichteinhaltung

Die Nichteinhaltung von rechtlichen Verpflichtungen wie Gesetzen, internen oder externen Richtlinien und Standards kann zu einem Verlust der Vertrauenswürdigkeit führen. Die manuelle Kuratierung durch geschultes Personal hilft, dieses Risiko zu vermeiden.

Risikoprioritätszahl - 4 | Wahrscheinlichkeit - 2 | Auswirkungen - 2

Institutionelles Risiko

Eine Änderung der Prioritäten der Österreichischen Akademie der Wissenschaften (ÖAW), ein Mangel an finanziellen Mitteln oder eine generelle Änderung des politischen Klimas könnten zum Ende des Archivs führen. In jedem dieser Fälle werden die Daten und Metadaten von ARCHE an ein anderes Repository übertragen oder sogar auf andere geeignete Repositorien aufgeteilt.

Risikoprioritätszahl - 3 | Wahrscheinlichkeit - 1 | Auswirkungen - 3

Relevant Standards and Frameworks

ARCHE befolgt eine Vielzahl externer Standards, Rahmenwerke und bewährter Verfahren für die Datenverwaltung im Allgemeinen und für die digitale Konservierung im Besonderen, die alle dazu beitragen, die Qualität der Daten und die Interoperabilität des Repositoriumsdiens zu maximieren. Diese Liste wird bei Bedarf überprüft und aktualisiert. Der Abschnitt [Technischer Aufbau](#) gibt Einblick in die technische Implementierung der ARCHE-Systeme.

- [Component Metadata of CLARIN \(CMDI\)](#)
- [CoreTrustSeal](#)
- [Dublin Core Metadata Initiative \(DCMI\)](#)
- [FAIR Principles](#)
- [Handle System](#) (via [ePIC](#))
- [Linked Data Platform](#)
- [OAI-PMH](#)
- [OAIS Reference Model](#)
- [OWL](#)
- [W3C Standards](#)

Quellen

Diese Richtlinie basiert auf mehrere unten angeführten Quellen, für die wir dankbar sind. Alle angegebenen Links wurden zuletzt am 11.11.2020 aufgerufen.

9

Konservierungsrichtlinien

- Archaeology Data Service (ADS): Collections Policy.
<https://archaeologydataservice.ac.uk/about/policies/collections-policy> (Version 9.0)
- British Library: Digital Preservation Strategy.
<https://web.archive.org/web/20200812100157/https://www.bl.uk/digital-preservation/strategy>
- DANS: Preservation Policy, Version 1.2. <https://dans.knaw.nl/nl/over/diensten/easy/toelichting-data-deponeren/DANSpreservationpolicyUK.pdf> [Nicht mehr verfügbar.]
- DANS: Preservation Plan, Version 1.
<https://web.archive.org/web/20190505125409/https://dans.knaw.nl/en/about/organisation-and-policy/policy-and-strategy/preservation-plan-data-archiving-and-networked-services-dans-1>
- Data Preservation Coalition (DPC): Data Preservation Handbook.
<https://www.dpconline.org/handbook>
- Open Preservation Foundation (2019-20): Digital Preservation Community Survey. Findings report. https://openpreservation.org/wp-content/uploads/public/resources/surveys/2020/OPFCommunitySurvey_Report_v03_200907.pdf
- RDA: Reagan Moore, Rainer Stotzka, Claudia Cacciari, Petr Benedikt (2015): Practical Policy. DOI: [10.15497/83E1B3F9-7E17-484A-A466-B3E5775121CC](https://doi.org/10.15497/83E1B3F9-7E17-484A-A466-B3E5775121CC)
- Scalable Preservation Elements (SCAPE): Barbara Sierman, Catherine Jones, Gry Elstrøm (2014) Catalogue of preservation policy elements. https://scafe-project.eu/wp-content/uploads/2014/02/SCAPE_D13.2_KB_V1.0.pdf
- The National Archives: Preservation Policy
2018. <http://www.nationalarchives.gov.uk/documents/preservation-policy-june-2018.pdf>

- TIB: Franziska Schwab (2019) Export und Exit-Szenario.
<https://wiki.tib.eu/confluence/display/lza/Export+und+Exit-Szenario>
- TIB: Preservation Policy der drei Zentralen Fachbibliotheken (Version 3, 2019)
<https://www.tib.eu/de/die-tib/policies/preservation-policy-der-drei-zentralen-fachbibliotheken>
- TIB: Preservation Policy of the Technische Informationsbibliothek (TIB) – German National Library of Science and Technology (Version 1.4 2019). <https://www.tib.eu/en/tib/policies/preservation-policy>
- UK Data Archive: Preservation Policy 2019. <https://dam.data-archive.ac.uk/controlled/cd062-preservationpolicy.pdf>
- University of Edinburgh: Developing a Digital Preservation Policy.
http://www.dpconline.org/component/docman/doc_download/1321-making-progress-hsbc-nov-2014-lee

Kuratierungsworkflow

- Archivematica: Micro-services.
<https://wiki.archivematica.org/Micro-services>
- Consultative Committee for Space Data Systems: Producer-Archive Interface Methodology Abstract Standard (PAIMAS). Magenta Book (2004).
<https://public.ccsds.org/Pubs/651x0m1.pdf>
- Data Curation Network (DCN): Lisa R. Johnston: Curating Research Data. Volume Two: A Handbook of Current Practice.
http://www.ala.org/acrl/sites/ala.org/acrl/files/content/publications/booksanddigitalresources/digital/9780838988633_crd_v2_OA.pdf
- IANUS: Archivierung bei IANUS.
<https://ianus-fdz.de/langzeitarchivierung>
- UC Curation Center. <https://escholarship.org/uc/item/5313h6k9>

11

Dateiformate

- Archaeology Data Service (ADS): Guides to Good Practice
<https://guides.archaeologydataservice.ac.uk/g2gp/Main>
- IANUS: IT-Empfehlungen für den nachhaltigen Umgang mit digitalen Daten in den Altertumswissenschaften. DOI: [10.13149/000.111000-a](https://doi.org/10.13149/000.111000-a)
- CLARIN: Standards recommendation.
<https://www.clarin.eu/sites/default/files/Standards%20for%20LRT-v6.pdf>
- Library of Congress: Digital Preservation at the Library of Congress.
<https://www.loc.gov/preservation/digital/>
- U.S. National Archives and Records Administration Digital Preservation Framework:
<https://github.com/usnationalarchives/digital-preservation>
- U.S. National Archives and Records Administration Digital Preservation Framework: The NARA Risk and Prioritization Matrix.
https://github.com/usnationalarchives/digital-preservation/tree/e746d3a6a6806f209b9c9272cc5c43d940cb8830/Digital_Preservation_Risk_Matrix

Risk Management

- Bodleian Libraries: Introduction to Digital Preservation: Risks to digital assets
<https://web.archive.org/web/20211009015508/https://libguides.bodleian.ox.ac.uk/digitalpreservation/risks>
- Data Preservation Coalition (DPC): Digital Preservation Handbook. Preservation issues.
<https://www.dpconline.org/handbook/digital-preservation/preservation-issues>
- The National Archives: Risk Assessment Handbook
<https://www.nationalarchives.gov.uk/documents/information-management/Risk-Assessment-Handbook.pdf>
- ZBW: Risk Management and Preservation Planning in the Digital Archive of the ZBW.
<https://www.zbw.eu/en/about-us/key-activities/digital-preservation/risk-management>