DLZA[1] Repository des Kleinen Museums (KM) mit einer Cloud-Lösung

Für den Workshop an der Memoriav-Fachtagung vom 19.3.2019 vorbereitet durch das DLZA Team des Archivs für Zeitgeschichte (Sonja Vogelsang, Jonas Arnold, Almir Ajradini) und finalisiert durch Fabian Würtz (Schweizerisches Sozialarchiv)

Annahmen für den Workshop
Das (fiktive) Kleine Museum (KM) möchte einen Bestand digitaler Videos zur Stadtgeschichte in seine Sammlung übernehmen und nutzt diese Gelegenheit, sich ein System zur digitalen Langzeitarchivierung zu beschaffen. Die ursprünglich analogen Videos wurden fachgerecht für die Langzeiterhaltung digitalisiert, aber noch nicht erschlossen. Im vorliegenden Papier wird davon ausgegangen, dass das KM eine Cloud-Lösung für sein Repository wählt. Die Variante mit stationärer Lösung wird im Dokument Memoriav-Fachtagung_2019_DLZA-Policy-Workshop_stationaereLoesung.pdf durchgespielt.

Ausgangslage
In Zukunft ist einen Anstieg digitaler Ablieferungen, sowohl in Bezug auf deren Häufigkeit als auch hinsichtlich der durchschnittlichen Datenmenge einer Ablieferung zu erwarten. Die Komplexität und Vielfalt des Materials werden ebenfalls zunehmen. Im Hinblick auf diese Veränderungen sind technische Infrastrukturen aufzubauen und Prozesse zu etablieren, die für die qualitätsgesicherte Übernahme, Erschliessung und Speicherung digitaler Daten geeignet sind.

Adressaten Policy
Die vorliegende Policy erläutert die vom KM gewählten Strategien im Umgang mit digitalen Daten und schafft dadurch Transparenz. Die Policy folgt den 16 Kriterien des "CoreTrustSeal"[2]. Die Policy richtet sich in erster Linie an die Mitarbeiterinnen und Mitarbeiter des KM. Für bestehende und potentielle Deponenten, mit dem KM verbundene Privatpersonen und Institutionen, sowie für die Nutzerinnen und Nutzer, welche mit den Beständen des KM arbeiten, wird gegebenenfalls eine gekürzte Version öffentlich zugänglich gemacht.

Digitale Daten: Definition
Digitale Daten sind in binärem Code vorliegende Daten, welche zur Nutzung mit technologischen Hilfsmitteln decodiert werden müssen. Als Erläuterung dienen die folgenden Definitionen:

- "digitale Repräsentation von Information in einer formalisierten Art, die die Interpretation, Verarbeitung bzw. den Austausch erlaubt" (DIN 31644 Kriterien für vertrauenswürdige digitale Langzeitarchive)
- "Digitale Unterlagen sind Unterlagen, die mit Hilfe von Informations- und Kommunikationstechnologien erstellt oder empfangen wurden, in digitalen Formaten vorliegen und als abgegrenzte Menge digitaler Daten Inhalte und Informationen transportieren." (Policy Digitale Archivierung, Schweizerisches Bundesarchiv, 2009)

[1] DLZA = Digitale Langzeitarchivierung
[2] CoreTrustSeal ist eine non Profit Organisation zur Förderung von nachhaltigen und vertrauenswürdigen Daten-Infrastrukturen. www.coretrustseal.org

Abgrenzung zwischen digitalisierten und digital generierten Daten
Zu unterscheiden ist zwischen digitalisierten Daten, die durch die Anlage einer digitalen Kopie von analogen Unterlagen entstanden sind, und digital generierten Daten, welche von Anfang an in digitaler Form erstellt werden (digital born).
Die Anforderungen an die Langzeitarchivierung sind für digitalisierte und digital generierte Daten praktisch identisch, die Vorgehensweisen jedoch unterschiedlich. Bei der Digitalisierung wird von katalogisiertem Museumsgut eine digitale Kopie in einem langzeittauglichen Format erstellt und anschliessend ins digitale Langzeitmagazin importiert (Ingest). Digital generiertes Museumsgut erhält das KM als digitale Originale in unterschiedlichen Ordnungszuständen und in möglicherweise grosser Formatvielfalt. Diese Unterlagen müssen analysiert, bewertet, erschlossen und für die Langzeitarchivierung sowie für die Nutzung aufbereitet werden, bevor sie in einem Langzeitmagazin archiviert und für die Nutzung zugänglich gemacht werden können.

R0: Background Information

Guidance: To assess a repository, reviewers need some information about the repository to set it in context. Please select from among the options and provide details for the items that appear in the Context requirement.

R0	**Repository Type** Die ARCHIV-CLOUD wird durch einen kommerziellen Anbieter Betrieben, der sich auf archivische Leistungen spezialisiert hat. Sie ist Teil einer OAIS-Infrastruktur, die er für mehrere Archive betreibt.
R0	**Repository's Designated Community** Das KM ist verpflichtet, der gesamten interessierten Öffentlichkeit im In- und Ausland Zugang zu seinem Museumsgut zu ermöglichen und diese bei der Nutzung zu unterstützen. Der Zugriff externer BenutzerInnen auf Museumsgut erfolgt via Museumskatalogsystem des KM auf die ARCHIV-CLOUD. Im Zusammenspiel mit dem Museumskatalog deckt die ARCHIV-CLOUD alle Zugangs- und Nutzungsbedürfnisse ab und ist für alle Interessierten lesend zugänglich.
R0	**Level of Curation** Das KM kuratiert sein digital(isiert)es Museumsgut in hohem Masse und generiert dadurch zusätzliche Wertschöpfung und Kontextwissen. Dazu gehört die Bewertung der Daten auf ihre Überlieferungswürdigkeit, die Erstellung einer logischen Struktur, die Erfassung von deskriptiven, strukturellen, administrativen und technischen Metadaten sowie die Migration in Erhaltungsmaster und Nutzungskopien. Data Curation Level C Das KM will, dass sein digital(isiert)es Museumsgut langfristig konsultierbar und interpretierbar bleibt. Da das KM ein Museum ist, ist nicht nur der Inhalt (Content), sondern nach Möglichkeit auch die ursprünglich übernommene, freilich nur in seltenen Fällen lang haltbare originale Überlieferungsform der Nachwelt zu erhalten. Das **analoge** Museumsgut wird daher

- Im Original so lang wie möglich erhalten und
- so hochwertig digitalisiert, dass im Digitalisat bei Verlust des analogen Originals wenigstens die als signifikant definierten Eigenschaften des Originals erhalten bleiben.

Das **digitale** Museumsgut wird
- im Original so lang wie möglich erhalten (Original Master) und gleichzeitig
- so hochwertig konvertiert, dass in den entstehenden Preservation Masters die als signifikant definierten Eigenschaften des Original Masters erhalten bleiben.

Insgesamt werden langfristig nur Inhalt und Überlieferungszusammenhang des Museumsguts und nicht dessen ursprüngliche Überlieferungsform überliefert werden.
Die ARCHIV-CLOUD dient als Speicher für die Erhaltungsmaster und Nutzungskopien des digitalisierten Museumsguts sowie für die digitalen Originale, Erhaltungsmaster und Nutzungskopien des digital generierten Museumsguts.

Original Masters
Das Einspeisen der Daten in die ARCHIV-CLOUD trennt diese von der technischen Umgebung, in welcher sie ursprünglich erstellt, bearbeitet und ausgelesen wurden. In manchen Fällen (z.B. interaktiven Museumspräsentationen) wird damit auch ihre Erscheinungsform (Repräsentation) verändert werden. Das KM wird die ursprüngliche technische Umgebung (z.B. ein interaktives digitales Ausstellungsstück) zwar aufbewahren, aber nicht auf lange Sicht voll funktionsfähig erhalten können.
Für den eventuellen Rückgriff auf die Originaldaten werden diese aber ebenfalls langfristig in der ARCHIV-CLOUD aufbewahrt.

Metadaten
Anlässlich der Trennung der Daten von ihrer ursprünglichen technischen Umgebung und allfälliger Normalisierungs- oder Migrationsmassnahmen muss sichergestellt werden, dass die in den Original Masters gespeicherten Metadaten nicht in die neuen Dateien transferiert werden können. Für die Katalogisierung und um BenutzerInnen bei Bedarf Zugriff auf möglichst alle ursprünglich vorhandenen Metadaten zu ermöglichen, werden diese zum frühestmöglichen Zeitpunkt extrahiert, separat gesichert und den bearbeiteten Daten beigelegt.

Preservation Masters
Das KM verfolgt langfristig eine Normalisierungs- und Migrationsstrategie: Die übernommenen Originaldaten (Original Masters) werden zwar aufbewahrt, aber in wenige ausgewählte Formate (Preservation Masters) konvertiert, welche als langfristig stabil und (aus)lesbar gelten. Das KM prüft in Zusammenarbeit mit ihren Partnern (…) sowie der Abteilung Informatik und Logistik regelmässig die Eignung der gewählten Formate für die Langzeitarchivierung und beschliesst bei Bedarf die Migration in andere, besser geeignete Langzeitformate. Als Quelle für eine Migration ist - wenn möglich - das noch vorhandene Original Master gegenüber dem Preservation Master zu priorisieren.

Nutzungskopien

	Für die Kommunikation im Internet werden von den Original Masters oder Preservation Masters in der Regel komprimierte Nutzungskopien (Access Masters) angelegt. Diese sollen hinsichtlich audiovisueller Qualität gegenüber den Originalen bzw. Erhaltungsmastern möglichst gleichwertig sein (visual losless compression ohne Verlust der als signifikant definierten Eigenschaften), so dass ein Zugriff auf Original Masters oder Preservation Masters zwar möglich, aber in vielen Nutzungsfällen nicht nötig sein wird. **Emulation** Emulation kann nötig werden, • damit das KM überhaupt lesenden Zugriff auf Daten erhält (alte oder hochproprietäre Systeme, Fachanwendungen) • um signifikante Eigenschaften (z.B. einer interaktiven Präsentation) zu erhalten und gegebenenfalls beim Zugang wieder zu gewährleisten. Im Falle nicht realistischer Perspektiven der fallweisen Emulation, wird es darum gehen, den überlieferungswürdigen Inhalt zu extrahieren und in den normalen Normalisierungs- und Migrationszyklus zu integrieren sowie die ursprüngliche Systemumgebung und die ursprüngliche Repräsentation der Daten zu dokumentieren. **Zeitpunkt für Normalisierungs- und Migrationsmassnahmen** Das KM erwartet, dass die Original Masters unter Umständen sehr schnell nach Übernahme und langfristig alle Fälle nicht oder nur mit unrealistischem Aufwand decodiert werden können und dass die Preservation Masters sowie die Access Masters ebenfalls periodisch migriert werden müssen. Das KM wird Normalisierungen und oder Migrationen zum frühestmöglichen Zeitpunkt durchführen, insbesondere bei • Original Masterformaten, deren Obsoleszenz bei der Akzession absehbar ist (Beispiel: .doc) • Original Masterformaten, welche ohne (als signifikant erachtete) Verluste in Preservation Master Formate übertragen werden können (Beispiel: avi uncompressed -> mkv) • Access Masterformaten, die angesichts des schnellen technologischen obsolet werden (Beispiel: .flv)
R0	**Outsource Partners** Dienstleister Digitalisierung • Diverse Dienstleister Repository • Cloud-Anbieter Weitere: • Zu definieren. Dokumente:

	• Vereinbarung über die Übernahme von Daten in die Cloud vom dd.mm.jjjj. • Zertifizierungen
R0	**Other Relevant Information** Keine Bemerkungen

Zentraler Anforderungskatalog für die DLZA des KM (Core Trustworthy Data Repositories Requirements)

Vorbemerkung

Im Folgenden wird die Strategie des KM zur vertrauenswürdigen Bearbeitung seines digital(isiert)en Museumsguts anhand der 16 Kriterien des "CoreTrustSeal" zur Vertrauenswürdigkeit von Langzeitarchiven[3] erläutert. Unter "Background Information" wurde erwähnt, dass die Betreuung des KM in Arbeitsteilung mit einem ARCHIV-CLOUD-Anbieter erfolgt. Die Aufgaben sind wie folgt auf das KM und dem Cloud-Anbieter verteilt.

Aufgabe (geordnet nach Ablauf der Arbeitsprozesse)	**Zuständigkeit**
Bereitstellung des Respoitories	Cloud-Anbieter
Vergabe der Gruppen-Zugriffsrechte	KM
Sichtung, Auswahl, Übernahme der Daten	KM
Transfer und die erste Sicherung der Daten	KM
Analyse und technische Metadatierung vor Ingest	KM
Vorbewertung und Löschung von Daten	KM
Strukturierung der Daten, deskriptive und administrative Metadatierung (Katalogisierung) der Daten	KM
Aufbereitung für den Ingest	KM
Technische Metadatierung anlässlich des Ingest	Cloud-Anbieter
Sichere Speicherung der Daten auf zeitgemässen Speichermedien	Cloud-Anbieter
Überwachung der Daten hinsichtlich Integrität	Cloud-Anbieter
Überwachung der Daten hinsichtlich Authentizität	Cloud-Anbieter
Überwachung der Daten hinsichtlich Les- und maschineller Interpretierbarkeit	Cloud-Anbieter
Anzeige von Erhaltungsmassnahmen	Cloud-Anbieter
Planung und Qualitätskontrolle von Erhaltungsmassnahmen	KM

[3] Das CoreTrustSeal ist der Nachfolger des Data Seal of Approval (DAS), welches aus einer Kooperation von niederländischen Forschungsinstitutionen entstand und 2007 zum ersten Mal veröffentlicht wurde. Der Hauptfokus lag ursprünglich auf der Archivierung von Forschungsdaten, die 16 Kriterien stellen jedoch ein grundlegendes Set an Minimalanforderungen für die Langzeitarchivierung dar. 2010 wurde das DSA durch das "Memorandum of Understanding to create a European Framework for Audit and Certification for Digital Repositories" Teil des Zertifizierungsverfahren für die vertrauenswürdige Langzeitarchive nach ISO 16363 und DIN 31644.

Durchführung von Erhaltungsmassnahmen	Cloud-Anbieter
Steuerung des externen Zugriffs auf Daten	Cloud-Anbieter / KM

1. Organizational Infrastructure

1.1. Mission / Scope

R1	Das KM ist Teil der Museumslandschaft der Stadt. Es sammelt analoges und digitales Museumsgut aus öffentlichem wie privatem Besitz, sichert diese und macht sie öffentlich zugänglich und nutzbar. **Fragen:** 1. Welche Dokumente müssen zur Untermauerung der Mission des KM im digitalen Bereich beigelegt (und gegebenenfalls mit Blick auf das digitale Videoarchiv wie angereichert) werden? **Antwort:** • Strategie der Stadt / Leistungsauftrag / bzw. -vereinbarung der Stadt mit dem KM • Sammlungsprofil und Auftrag des KM • Leitbild des KM, in dem die Sicherung digital generierten Museumsguts erwähnt ist Liste der Stiftungen / Donatoren, welche das KM unterstützen

1.2. Licenses

R2	Sämtliches Museumsgut des KM ist frei zugänglich (Datenschutz) und frei nutzbar (Datenschutz und Urheberrecht). **Fragen:** 1. Welche Dokumente müssen zur Untermauerung der freien Zugänglichkeit und Nutzbarkeit beigelegt (und gegebenenfalls mit Blick auf das digitale Videoarchiv wie angereichert) werden? 2. Wenn die Daten nicht frei nutzbar wären könnte trotzdem eine Cloud-Lösung verwendet werden? 3. Wie sieht das im digitalen Videoarchiv von Yvan Kohler aus? **Antwort:** **1. Dokumente zu Besitz- und Urheberrechten** • Mustervertrag mit Donatoren einschliesslich Übertragung der Urheber- bzw. Verwertungsrechte • Museumsordnung, CC-Lizenzregelungen des KM • URG • Gültige Regelungen von oder mit Verwertungsgesellschaften

	2. Nicht freie Dokumente in der Cloud Wenn die Cloud Zugangsregeln unterstützt, der Ablieferungsvertrag keine externe Aufbewahrung untersagt und die gesetzlichen Pflichten (z.B. im Bezug auf Datenschutz) erfüllt werden sollte dies möglich sein. **3. Videoarchiv Yvan Kohler:** Liegenschaften haben kein Persönlichkeitsrecht. Allenfalls könnten erkennbare Personen ein Datenschutzproblem darstellen. Die Urheberrechte sind noch nicht geklärt und müssen dem KM übertragen werden.

1.3. Continuity of Access

R3	Das KM ist verpflichtet, das von ihm kuratierten Daten des KM unbefristet aufzubewahren und zugänglich zu machen. **Frage:** Was soll hier hinsichtlich der Eventualität eingetragen werden, dass das KM oder die Cloud-Anbieter aufgelöst wird? **Antwort:** Im unwahrscheinlichen Falle der Auflösung des KM würde das Museumsgut des KM anderen Gedächtnisinstitutionen der Stadt , im Kanton Bern, in der Schweiz sowie im Ausland zur Übernahme angeboten. Das NBM sollte sich vertraglich zusichern lassen, dass im Fall eines Konkurses des Cloud-Anbieters Zugriff auf die Daten bzw. die Hardware erhält. Ebenfalls ist wichtig kontinuierlich mit dem Anbieter in Verdingung zu stehen und stets zu wissen wo die Daten physisch liegen. Ebenfalls sollte bei der Anbieterwahl mögliche Auflösung-Szenerien und Risiken abgewogen werden.

1.4. Confidentiality / Ethics

R4	Für das KM gilt Folgendes: • Das KM ist dem internationalen Code of Ethics for Museums des International Council of Museums (ICOM) verpflichtet (ICOM Code of Ethics vom 8.10.2004). • Das KM wendet bei der Katalogisierung und Dokumentation seines Museumsguts ein LIDO-kompatibles Regel- und Datenset an. • Das KM stellt sicher, dass sämtliche Datentransfers über Verbindungen erfolgen, die den aktuellen Sicherheitsstandards entsprechen.

	• Grundsätzlich unterliegt das Museumsgut des KM keinen datenschutzrechtlichen Auflagen. Das KM bearbeitet Daten gegebenenfalls unter Berücksichtigung der Erfordernisse des Persönlichkeits- und Datenschutzes und trägt diesem bei der Bereitstellung von Recherchewerkzeugen und der Vermittlung Rechnung. • Der online-Zugang zum Museumsgut des KM ist ohne Auflagen möglich, soweit keine persönlichkeitsrechtlichen Schranken tangiert werden. • Interessierte Nutzerinnen des Museumsguts werden über die Museumsordnung sowie über Disclaimer in den online-Angeboten über ihre Rechte und Pflichten informiert. **Fragen:** 1. Können diese Punkte so auch für das digitale Videoarchiv umgesetzt werden? 2. Können Sie einzelne Punkte noch spezifizieren (z.B. was sollte in einem solchen Disclaimer stehen)? 3. Welche Vorkehrungen sollten für den Fall getroffen werden, dass sich bspw. in Videoaufnahmen trotzdem datenschutzrechtlich relevante Inhalte finden? 4. Für den Fall es befänden sich gesperrte Informationen im Bestand. Wäre es ethisch vertretbar diese in der Cloud zu hosten und wenn ja zu welchen Bedingungen? **Antworten:** 1. Ja, sofern die Urheberrechte geklärt werden und keine Personen erkennbar sind. 2. Im Disclaimer könnte festgehalten werden, dass die Videos genutzt, aber nicht verändert werden dürfen, und das die Gedächtnisinstitution und der Urheber genannt werden sollten, also eine CC-BY-ND Lizenz gilt. Siehe auch Antwort 3. 3. Vorschlag: Im Disclaimer vermerken, dass sich Personen, die nachweislich auf den Videos erkennbar sind, melden können, falls dies ein Problem ist. 4. Dies hängt vom konkreten Fall/Kontext ab. Sicher gilt es abzuklären wo die Daten gelagert werden, welche Sicherheitsmassnahmen getroffen werden und wer sich physisch Zugriff verschaffen kann.

1.5. Organizational Infrastructure

R5	**Institution:** Das KM besitzt langjährige Erfahrung im Kuratieren von Museumssammlungen und ist eine anerkannte Institution in der Museumslandschaft. Mit der Unterstützung durch langjährige Gönner und / oder Stiftungen gewährleistet es eine längerfristig stabile Organisationsstruktur. Der Cloud-Anbieter existiert seit 20 Jahren und bietet diverse digitale und nicht digitale Archivdienstleistungen an. **Finanzierung:** Das KM ist eine feste Grösse in der Finanzstrategie 2019-2023 der Stadt . Der Cloud-Anbieter finanziert sich durch Aufträge von diversen Auftragsgebern. **Qualifikation:** Innerhalb des KM ist ein spezialisiertes Team für die Betreuung und Bearbeitung digital(isiert)en Museumsguts zuständig. Das Team besteht aus Museumsspezialisten, die eine Ausbildung in digitaler Langzeitarchivierung absolviert haben. Das Team bildet sich regelmässig weiter und ist für die Vernetzung innerhalb der schweizerischen (wo sinnvoll auch europäischen) Museumslandschaft besorgt. Der Cloud-Anbieter hat eine Reihe von Projekte mit namhaften Archiven durchgeführt und bildet seine Mitarbeiter*innen fortlaufend weiter. **Fragen:** Wieviel Kompetenz muss ein Archiv in der digitalen Archivierung aufweisen. Reicht ein grundlegendes Wissen, aufgrund von dem die Entscheidungen mit dem Cloud-Anbieter getroffen werden oder muss das Archiv ein vollkommen ausgebildetes Team haben. Mit Blick auf das digitale Videoarchiv: Welche Kenntnisse gehören zu einer «Aus- und Weiterbildung in digitaler Langzeitarchivierung» und was nicht, weil es in die Kompetenz des Cloud-Anbieters gehört? **Antwort:** Je mehr Wissen ein Archiv hat umso besser ist es. Ohne Grundverständnis kann man das Angebot des Cloud-Anbieter nicht einschätzen. Wenn man selber wenig Wissen hat, kann es eine Option sein, einen Cloud-Anbieter zu nehmen, der bereits grössere Archive als Kunde hat, da hier die Hoffnung besteht, dass die anderen Archive mit mehr Ressourcen eine gewisse Kontrollfunktion übernehmen.

Expert Guidance

R6	**Fragen:**

	1. Mit welchen Institutionen, Dienstleistern und Arbeitsgruppen sollte das KM und der Cloud-Anbieter in regelmässigem Austausch stehen, um die Herausforderung der digitalen Langzeitarchivierung im Allgemeinen und im digitalen Videoarchiv im Speziellen meistern zu können? 2. Wie könnte ein Feedback-Mechanismus aus der Designated Community aussehen? **Antwort:** **1. Kooperationen** • National: VMS, VSA, KOST-CECO, Memoriav • International: ICM, NESTOR, IASA • …. **2. Feedback-Mechanismus** • E-Mail, Social Media, Umfragen, usw. • Feedback-Formulare (Website, online-Katalog)

2. Digital Object Management

2.1. Data Integrity and Authenticity

R7	Die Donatoren erklären sich im Vertrag damit einverstanden, dass das KM Daten verändern oder (bei fehlender Relevanz) vernichten kann. Mittels Sammlungsinformationen und Dokumentation können Donatoren und Interessierte summarisch die vorgenommenen Bearbeitungsschritte und Veränderungen nachvollziehen. Das KM schützt die ihm anvertrauten Daten vor fahrlässiger, absichtlicher und / oder böswilliger Manipulation sowie unerwünschten technisch bedingten Veränderungen, Defekten oder Datenverlust. Die Verantwortlichkeit beginnt ab dem Zeitpunkt des Transfers der Daten in die Systeme des KM. Das KM stellt die Integrität der Daten ab dem Zeitpunkt des Transfers, über den Pre-Ingest bis zum Ingest in die Cloud sicher. Nach dem Ingest in die Cloud werden die Daten von vom Cloud-Anbieter überwacht und verwaltet, das KM ist jedoch weiterhin für die Daten verantwortlich. Zur Gewährleistung der Integrität und Authentizität stellt das KM in Kooperation mit dem Cloud-Anbieter insbesondere Folgendes sicher: • **Dokumentation:** Die Daten werden beim Transfer dokumentiert. Zur Dokumentation gehören insbesondere Datenumfang, technische Prüfsummen, Formate und eingebettete Metadaten, die extrahiert werden. Im weiteren Verlauf der Bearbeitung überprüft das KM die Daten regelmässig auf unerwünschte Veränderungen und protokolliert allfällige bearbeitungsbedingte Veränderungen lückenlos oder anlässlich definierter Meilensteine. • **Schutz vor Veränderung:** Die Daten werden während des gesamten Arbeitsprozesses vor der Bearbeitung durch nicht autorisierte Personen und vor Veränderungen der Integrität, und Authentizität geschützt. Konkret stellt das KM mit Cloud-Anbieter sicher, dass

	nur geschulte KM-Mitarbeitende (oder vom KM autorisierte Personen) Daten bearbeiten können und dass die Hardware- und Softwaresysteme laufend auf dem neuesten Stand der Technik gehalten sowie optimal für die Herausforderungen der digitalen Langzeitarchivierung konfiguriert werden. • **Kontrolle der Integrität:** Hierzu werden die anlässlich des Transfers bzw. des Ingests erhobenen Mengengerüste und Checksummen vornehmlich mit technischen Mitteln laufend auf Veränderungen überprüft. • **Kontrolle der Authentizität:** Die als signifikant definierten Eigenschaften jeden Museumsobjekts bleiben erhalten. Hierzu wird insbesondere sichergestellt, dass Normalisierungs- und/oder Migrationsmassnahmen nur dann bzw. in einer Weise angewendet werden, dass die als signifikant definierten Eigenschaften erhalten bleiben. Die Überprüfung erfolgt durch repräsentative audiovisuelle Stichproben sowie durch Tools, welche die maschinell überprüfbaren signifikanten Eigenschaften vor und nach einer autorisierten Konvertierung / Veränderung vergleichen. • **Massnahmen bei Feststellung unerwünschter Veränderungen:** Bei allfälligen unerwünschten Veränderungen kann der Ursprungszustand aus einem Backup bzw. von einer Kopie wiederhergestellt werden. **Frage:** **1.** Was wurde bisher im Falle des Videoarchivs Yvan Kohler sichergestellt? 2. Wie kann das KM die Arbeitsvorgänge des Cloud-Anbieters überprüfen? **Antwort:** 1. Die bisher erhobenen Daten (Sichtungsbericht und Inventar) geben einigen Aufschluss über Entstehungszusammenhänge, über das Mengengerüst, die vorhandenen Formate, die Digitalisierung, deskriptive Metadaten, bisherige Erhaltungsmassnahmen, Prüfsummen, Dekodierbarkeit. 2. Stichproben, Begutachtung der Infrastruktur, enger Kontakt mit viel Rückfragen, Stärkung der eigenen Kompetenz. Audits durch externe Firmen

2.2. Appraisal

R8	**Vorarchivische Bewertung:** Vor jeder Datenübernahme prüft das KM, ob das angebotene Material seinem Sammlungsprofil und den Anforderungen für die (digitale) Langzeitarchivierung entspricht. Zudem dokumentiert es die angebotenen Daten hinsichtlich Umfang, Formate, technische Einschränkungen, Struktur und / oder bestehenden Metadaten. Je nach Ergebnis empfiehlt das KM vor dem Transfer zusätzlich eine Beratung, weitere (Vor)Ordnungsarbeiten oder das Beiziehen eines externen Dienstleisters. Das KM bespricht vor einem Transfer mit dem Donator fallbezogen die Löschung konkrete Aufbereitung der Daten. Das KM kann Daten, welche seine Kriterien nicht erfüllen zurückweisen oder ihre Übernahme mit weiteren Bedingungen verknüpfen. Auch Material, welches nicht den Beurteilungskriterien entspricht, kann aufgenommen werden, sofern es trotzdem bearbeitbar ist. **Fragen:** 1. Ergänzen Sie die folgenden Punkte mit spezifischem Blick auf das digitale Videoarchiv Yvan Kohler: Beurteilungskriterien für die Übernahme: Das KM übernimmt nur Daten a) die seinem Sammlungsprofil entsprechen (Kriterien: Geographie: ……; Zeitraum: ……. Themen: ……. Medientypen (z.B. Schriftgut?): ……. b) für die es über die Besitz- und Verwertungsrechte verfügt bzw. erhält. c) die folgenden Kriterien der Überlieferungswürdigkeit genügen: ……. d) die folgenden Kriterien der Überlieferbarkeit (Erhaltungszustand) genügen: digitale Daten in möglichst unkomprimierten Formaten und Codecs, für die Konvertierungstools in DLZA-Formate existieren. 2. Wie bewerten Sie die Überlieferbarkeit des digitalen Videoarchivs Yvan Kohler? **Antwort:** **1. Beurteilungskriterien für die Übernahme:** • die seinem Sammlungsprofil entsprechen (Kriterien: Geographie: Grossraum ; Zeitraum: keine Begrenzung; Themen: Stadtentwicklung, Alltag, Kultur, Kunst; Medientypen: audiovisuelle Quellen, Objekte, Schriftgut (inkl. Karten, Plakate) • für die es über die Besitz- und Verwertungsrechte verfügt bzw. erhält. • die folgenden Kriterien der Überlieferungswürdigkeit genügen: hoher Evidenzwert für die abliefernde Stelle, hoher Informationswert für die Nachwelt, Unikatscharakte, … • die folgenden Kriterien der Überlieferbarkeit (Erhaltungszustand) genügen: digitale Daten in möglichst unkomprimierten Formaten und Codecs, für die Konvertierungstools in DLZA-Formate existieren.

	2. Überlieferbarkeit Es muss überprüft werden, ob komprimierte Codecs in den MOV und DV-Dateien (Original Masters) verwendet wurde. Je nachdem ist über mehrere Migrationszyklen hinweg in den Preservation Masters und Access Masters mit Artefaktbildung zu rechnen.

2.3. Documented Storage Procedures

R9 **Sicherungsprozesse vor Ingest in der Cloud**

Bis zum Ingest in die Cloud lagern die Daten in den Systemen des KM auf einem eigens eingerichteten Akzessionsspeicher. Dieser dient als zentraler Speicher und verfügt über einen Backup-Akzessionsspeicher. Von digital generierten Daten wird frühestmöglich eine Kopie im Akzessionsspeicher erstellt. Die Daten werden mittels standardisiertem Vorgehen von Ihrem Träger getrennt, dabei wird auch sichergestellt, dass die Originalversion auf dem Träger erhalten bleibt.

Dokumentation und Management relevanter Prozesse

Die relevanten Prozesse und Zuständigkeiten sind in den unter R12 ausgelisteten Dokumenten festgelegt.

Sicherheitsstufen

In sämtlichen Prozessen gelten 3 Berechtigungsstufen. Es gibt derzeit keine Sicherheitsstufen im Sinn einer Differenzierung von Leserechten nach Bestand.

	Alle	Mitarbeiter	Power User
Create	-	X	X
Delete	-	-	X
Change	-	X	X
Read	X	X	X

Speicher- und Backupkonzept

Im Akzessionsspeicher werden Daten auf 2 örtlich getrennten Festplatten gespeichert. Nach erfolgreichem Ingest werden die Daten auf dem Akzessionsspeicher gelöscht. Im Cloudspeicher werden Daten standardmässig in mindestens 3 Kopien auf in unterschiedlichen Rechenzentren gelagert.

Risikomanagement

Zum Risikomanagement siehe Punkt R16.

Konsistenzprüfung

	Abgesehen von Prüfungen der md5-Checksummen sind bisher keine Konsistenzprüfungen vorgesehen. **Speichermedienprüfung** Das Speichersystem umfasst auch eine Management-Software, die periodisch sämtliche Festplatten auf ihre Funktionstüchtigkeit prüft. **Fragen:** 1. Sollte das KM zur Sicherheit die Daten auch noch selbst irgendwo speichern oder kann es die Speicherung vollkommen dem Cloud-Anbieter vertrauen? 2. Am 20.2.2020 entdeckt man, dass in der Cloud eine Datei versehentlich gelöscht wurde. Wie lange sollte eine Wiederherstellung durch den Cloud-Anbieter gewährleistet sein? **Antwort:** 1. Eine weitere Speicherung ist natürlich nicht vollkommen falsch. Allerdings löst sich damit das Sparpotenzial der Cloud auf und damit ein zentraler Mehrwert. Traut man einem Cloud-Anbieter die Datensicherheit nicht zu, sollte ohnehin der Anbieter oder sogar das Vorhaben überdacht werden. 2. Grundsätzlich sollte das Löschen und das Verändern von Ingests in der Archiv-Cloud nicht oder nur schwer möglich sein. Wie lange solche Löschungen/Veränderungen rückgängig gemacht werden können ist eine Kostenfrage und muss vom Archiv abgewogen werden.

2.4. Preservation Plan

R10	**Einverständnis der Donatoren** Die Donatoren erklären sich im Vertrag mit folgenden Punkten einverstanden: • Das überantwortete Museumsgut unterliegt allfälligen Bearbeitungen, Konversionen sowie Kassationen durch das KM und kann eingesehen und genutzt werden. • Ab Transfer der Daten liegt die Verantwortung für das digital(isiert)e Museumsgut beim KM. Der Transfer erfolgt über Festplatten oder ftp. **Preservation Level** Für alle in Cloud eingespiesenen Daten gilt eine unbegrenzte (ewige) Aufbewahrungsfrist.

Plan der Erhaltungsmassnahmen
Die Erhaltungsmassnahmen folgen den unter R0 beschriebenen Prämissen. Während die Übernahmeformate (Original Master) nur aufbewahrt werden, werden die Zielformate für (neue) Erhaltungsmaster (Preservation Master) und Nutzungskopien (Access Master) periodisch auf ihre Eignung hin überprüft und gegebenenfalls migriert. Ein entsprechender strategischer Erhaltungsplan (inkl. Community Watch und Technology Watch) ist institutionalisiert. Dadurch können Synergien zwischen den verschiedenen Fachspezialisten geschaffen werden. Strategische Entscheide werden dokumentiert.

Zuständigkeiten für Preservation
Erhaltungsmassnahmen werden durch den Cloud-Anbieter Anbieter in Rücksprache mit dem KM durchgeführt.

Fragen:
1. Spielen Sie den Preservation Workflow inkl. Zuständigkeiten für das digitale Videoarchiv für den Fall durch, dass das gegenwärtig favorisierte DLZA-Format für Videos (.mov) durch ein neues ersetzt werden muss.
2. Wie soll das KM die Datei-Formate Preservation festlegen? Gibt es sinnvolle universelle Vorlagen?
3. Ist es bei Videoarchiven absolut notwendig die Original-Dateien zu speichern

Antwort:
1. Technology Watch (PRONOM und andere)
2. Regelmässige Prüfung der Daten auf Dekodierbarkeit
3. Frühzeitige Signalisierung obsoleter Formate / Codecs
4. Planung entsprechender Erhaltungsmassnahmen
5. Export zu behandelnden Daten
6. Tests Erhaltungsmassnahmen (Authentizität, Lesbarkeit)
7. Entscheid Erhaltungsmassnahmen
8. Durchführung Migration
9. Prüfung Authentizität, Dekodierbarkeit
10. Technische Dokumentation der Erhaltungsmassnahmen
11. Re-Ingest in ARCHIV-CLOUD
12. Löschung (?) obsoleter Preservation oder Access Masters

Zuständigkeiten:

KM: 4, 6, 7, 9, 10, 12
Cloud-Anbieter: 1, 2, 3, 4, 5, 7, 8, 10, 11, 12

2.5. Data Quality

R11

Das KM verfügt für jedes Museumsobjekt über eine Dokumentation über den Zustand zum Zeitpunkt des Transfers sowie aller Bearbeitungsschritte oder Meilensteine. Für jedes Objekt wird die Entstehungs-und Überlieferungsgeschichte, die Relevanz der vorhandenen Daten sowie eine historische Kontextualisierung festgehalten.
Das KM versteht die Vergabe und Verifizierung von Metadaten als seine Kernaufgabe und betrachtet somit fehlende Metadaten anlässlich des Transfers nicht als problematisch. Es achtet jedoch darauf, dass nachvollziehbar bleibt, ob Metadaten Teil der Originalüberlieferung waren oder von ihm selbst zu einem späteren Zeitpunkt hinzugefügt wurden.
Der Hauptteil der beschreibenden Metadaten befindet sich im Museumskatalogsystem, wobei das KM vorsieht alle beschreibenden Metadaten auch in der Cloud einzuspeisen, wo die technischen Metadaten anlässlich des Ingest erhoben werden.

Frage:
Was bedeutet die oben geschilderte Anforderung hinsichtlich des Zusammenspiels zwischen Museumskatalogsystem und der Cloud? Sehen Sie mögliche Alternativen?

Antwort:
Sämtliche im Museumskatalogsystem erfassten Metadaten müssen anlässlich des Ingests (und danach) über eine Schnittstelle in ARCHIV-CLOUD übertragen und im Fall einer Änderung zwischen den beiden Systemen synchronisiert werden können, wobei das Museumskatalogsystem bei deskriptiven und administrativen Daten das Daten-Mastersystem ist.
Alternativ könnte man sich folgende Arbeitsteilung zwischen ARCHIV-CLOUD und Museumskatalogsystem überlegen:

Metadaten	Museumskatalog	DATAREP
Deskriptive	X	nur Titel
Administrative	X	X
Technische	Datenumfang, allgemeine Formatangaben	X

2.6. Workflows

R12

Die Regelung der Arbeitsprozesse zur Übernahme, Bewertung, Sicherung, Erhaltung, Zugänglichmachung und Nutzung digital(isiert)en Archivguts ist in folgenden Dokumenten festgehalten.

- Leitfaden / Checkliste zur Digitalisierung von KM-Museumsgut
- Leitfaden / Checkliste zur Sichtung von KM-Museumsgut
- Leitfaden / Checkliste Transfer digital(isiert)en Museumsguts
- Leitfaden / Checkliste Akzession digital(isiert)en Museumsguts

	• Leitfaden / Checkliste Analyse digital(isiert)en Museumsguts (DROID, Duplicate Cleaner, mediainfo, EXIF) • Leitfaden Katalogisierung digital(isiert)en Museumsguts • Leitfaden / Checkliste Ingest Cloud • Leitfaden / Checkliste Digital Preservation in der Cloud • Leitfaden Aufbereitung und Auslieferung von Museumsgut aus der Cloud Die Dokumentation der Arbeitsprozesse zur Übernahme, Bewertung, Sicherung, Erhaltung, Zugänglichmachung und Nutzung digital(isiert)en Archivguts erfolgt in folgenden Dokumenten. • Report zu Digitalisierungsprojekten • Sichtungsbericht und Sichtungsverzeichnis • Vertrag mit Donatoren • Übernahmebericht und Übernahmeverzeichnis • Akzessionseintrag und Analysebericht • Museumskatalogsystem (LIDO-Export) • Bericht zum Ingest in die Cloud • Bericht zu Erhaltungsmassnahmen in der Cloud • GEVER Nutzung Museumsgut **Frage:** Fallen Ihnen weitere Punkte ein, die noch nicht abgedeckt sind? **Antwort:** Nein

2.7. Data Discovery and Identification

R13	Obwohl sämtliche Daten aus dem Netz direkt angesteuert werden könn(t)en, sollen externe BenutzerInnen diese immer via Webclient des Museumskatalogsystems ansteuern. Von dort aus führen je ein Link zu den Access Masters (viewer, Stream etc.) resp. zu den Preservation bzw. Original Masters in der ARCHIV-CLOUD (Download-DIP). Jedes Museumsobjekt erhält im Katalogsystem und in der ARCHIV-CLOUD eine Referenz (Digital Object Identifier DOI vgl. www.data-cite.org) und eine persistente ID (PID) im Sinne einer zugewiesenen Lokatur im digitalen Magazin. Diese und der DOI dienen als Referenz-ID zwischen ARCHIV-CLOUD und Museumskatalogsystem. Der DOI-Name setzt sich aus der Institutions-Nummer sowie eindeutigen (aus der Katalogsignatur derivierten) technischen Signatur zusammen, während die PID in der ARCHIV-CLOUD als proprietäre Systemnummer vergeben wird. Die Identifizierung eines Museum-sobjekts kann in den involvierten Systemen somit über die (technische) Signatur, den DOI oder die PID erfolgen. **Frage:** Welche Zitiersignatur soll für ein digital(isiert)es KM-Museumsobjekt angegeben werden, wenn die Katalog-Signatur KM 4/2 lautet, die DOI 10.98765_KM-4-2 und die PID IE485692789? Begründen Sie Ihre Wahl. **Antwort zur Diskussion:** Die DOI ist in diesem Konzept ein Derivat aus der Museumskatalog-Signatur und kann als alternative Zitier-signatur verwendet werden. Vorteil: aus der DOI ist ersichtlich, dass ein digital(isiert)es Museumsobjekt zitiert wird. Die Persistenz der PID ist demgegenüber mit Vorsicht zu geniessen.

2.8. Data Reuse

R14	Das Museumskatalogsystem, die Workflows zum Pre-Ingest sowie der Ingest verarbeiten Metadaten im LIDO-Format. Nach dem Ingest in die Cloud werden die Metadaten in folgenden Datasets gespeichert. • LIDO: Deskriptive und administrative Metadaten aus der Katalogisierung • METS: Administrative und technische Metadaten, die, soweit nicht bereits vorhanden, beim Ingest gewonnen wurden. • Dublin Core: Metadaten des Systems In der Cloud können Informationsobjekte in verschiedenen Repräsentationen bereitgestellt werden. Ausgehend vom Originalformat werden zusätzliche Repräsentationen Preservation Master und Access Master angelegt. Die Access Master Formate werden regel-mässig auf ihre Tauglichkeit und Verbreitung in der Designated Community geprüft und gegebenenfalls in neue Access Master Formate migriert.

	Ein Auslieferungsinformationspaket (DIP) aus dem Langzeitmagazin umfasst eines oder mehrere Museumsobjekte, standardmässig mit den Metadaten im XML-Format (METS) und LIDO und den Primärdaten (Original Master, Preservation Master und Access Master). **Fragen:** 1. Eine Forscherin möchte nicht transkribierte Zeitzeugen-Videos im Rahmen eines digital humanities Projekts mit maschinellen Mitteln semiautomatisch transkribieren. Welches Format wird sie benötigen (Original, Preservation oder Access Master)? 2. Was fehlt in der Beschreibung des obengenannten Auslieferungsinformationspaket DIP? 3. Gäbe es noch weitere Möglichkeiten zum Datenbezug die ausser einem DIP-Download Forschenden angeboten werden sollten? **Antwort:** 1. [Keine Antwort] 2. Das DIP muss zusätzliche Infos enthalten (Zeitpunkt der Erstellung) und als solches erkennbar sein. 3. API, IIIF, Linked Open Data, usw.

3. Technology

3.1. Technical Infrastructure

R15	Das KM setzt zur Bearbeitung von digital generierten Daten auf eine modulare Infrastruktur. Als Basis nutzt es die von der Abteilung Informatik und Logistik bereitgestellten und / oder angebotenen Infrastrukturen und Dienste. Zusätzlich verwendet es Tools für die digitale Archivierung und hat einige Tools für besondere Aufgaben im Einsatz. Das KM entwickelt keine eigenen Systeme / Applikationen, adaptiert aber bestehende Tools für seine Zwecke. Zudem kooperiert es mit der Abteilung Informatik und Logistik und verschiedenen Dienstleistern bei der Entwicklung von Tools. Eine Dokumentation zur Infrastruktur, den Diensten und Tools ist vorhanden. **Frage:** Geben Sie ein paar Beispiele für technische Werkzeuge im Videobereich an, die also dokumentiert werden müssten. **Antwort:** Mediainfo MediaArea (Mediaconch) Adobe Premiere ffmpeg

3.2. Security

R16	Das KM und der Cloud-Anbieter sind sich folgender Risiken bewusst:

	IT-spezifisch • Bitsprünge, Schäden an Hardware (Speichersysteme, Träger) und Software • Cyber-Angriffe: Port-Scanning, Brute-Forcing, Spoofing, Denial-of-Service, Spamming, Phishing, Malware (Würmer, Trojanische Pferde), Spear Phishing, Drive By Attacken, Watering Hole Attacken, Advanced Persistent Threat (Bsp. ItaDuke) • … Allgemein • Faktor Mensch: Fehler in der Handhabung, «Ich wollte noch rasch» • Stromunterbruch, Wassereinbruch, Verschmutzung • Unwetter (Blitzeinschlag), Hochwasser, Erdbeben • Verkehrsunglück, Kernschmelze • politische Unruhen • EMP • … Das KM und der Cloud-Anbieter begegnen diesen Risiken mit folgenden präventiven Massnahmen: IT-spezifisch • Patching • Regelmässige Überwachung, Tests und zeitnahe Behebung technischer Schwachstellen (Hardware, Speichermedien, Betriebssysteme, Drittprodukte, Webapplikationen) • Schulung und Sensibilisierung der Mitarbeitenden • Firewall • Datensicherung und Backup • …. Allgemein • Schulung, Prüfung und Sensibilisierung von Personen mit Schreibrechten am Archivgut • Notfallplanung Dokumente: • Dokumentation der Backups und Recovery • Dokumentation der Security Aspekte betr. Schreiben und Lesen von Daten (v.a. intern) • Notfallplanung

	Frage: 1. Welche (an sich naheliegende) Präventionsmassnahme fehlt in der obengenannten Liste IT-spezifisch? 2. Welche Vor- und Nachteile bei der Sicherheit entstehen durch die Nutzung einer Cloud. **Antwort:** 1. Virenscan bei Akzessionen! 2. Der grosse Vorteil der Cloud liegt darin, dass davon auszugehen ist, dass der Cloud-Anbieter aufgrund seiner Grösse mehr Erfahrung und Ressourcen hat für die Sicherheit zu sorgen. Der Nachteil ist, dass man zu einem gewissen Grad Kontrolle abgibt und teilweise auf den Anbieter vertrauen muss.