KOST.07

Archivierung von Unterlagen aus Geschäftsverwaltungssystemen (AUGev)

Prozess- und Datenmodellierung

Erarbeitet von

Ursula Bausenhart, StABS

Georg Büchler, KOST

Olivier Debenath, StAZG

Martin Kaiser, KOST

Lambert Kansy, StABS

Martin Lüthi, StASG

Version	Datum	Änderungen
1.0	2007-09-10	

Inhalt

Zusammenfassung 4

1 Einleitung 5

1.1 Beschreibung des heutigen Archivierungsprozesses 5

1.2 Projektinhalt 6

1.2.1 Modellierung nach OAIS 6

1.2.2 Überblick über Software-Layer 8

1.2.3 AUGev-Modellierung 8

2 Prozessmodellierung 10

2.1 Use-Case-Modellierung 10

2.1.1 Lebenszyklus digitaler Unterlagen 10

2.1.2 Archiv-Prozesse 10

2.2 Unterlagen übernehmen 11

2.2.1 Anwendungsfallbeschreibung (Use Case Scenario) 11

2.2.2 Anwendungsfalldiagramm (Use Case Diagram) 12

2.2.3 Aktivitätsdiagramm (Activity Diagram) 14

2.3 Unterlagen anbieten 15

2.3.1 Anwendungsfallbeschreibung (Use Case Scenario) 15

2.3.2 Aktivitätsdiagramm (Activity Diagram) 16

2.4 Angebot prüfen 17

2.4.1 Anwendungsfallbeschreibung (Use Case Scenario) 17

2.4.2 Aktivitätsdiagramm (Activity Diagram) 18

2.5 Unterlagen abliefern 19

2.5.1 Anwendungsfallbeschreibung (Use Case Scenario) 19

2.5.2 Aktivitätsdiagramm (Activity Diagram) 19

2.6 Unterlagen entgegennehmen und prüfen 20

2.6.1 Anwendungsfallbeschreibung (Use Case Scenario) 20

2.6.2 Anwendungsfalldiagramm (Use Case Diagram) 21

2.6.3 Aktivitätsdiagramm (Activity Diagram) 22

2.7 Format erkennen und prüfen 23

2.7.1 Anwendungsfallbeschreibung (Use Case Scenario) 23

2.7.2 Anwendungsfalldiagramm (Use Case Diagram) 23

2.7.3 Aktivitätsdiagramm (Activity Diagram) 24

2.8 Formatkatalog bewirtschaften 25

2.8.1 Anwendungsfallbeschreibung (Use Case Scenario) 25

2.8.2 Anwendungsfalldiagramm (Use Case Diagram) 25

2.8.3 Aktivitätsdiagramm (Activity Diagram) 26

2.9 Formatkonvertierung bewirtschaften 27

2.9.1 Anwendungsfallbeschreibung (Use Case Scenario) 27

2.9.2 Aktivitätsdiagramm (Activity Diagram) 28

2.10 Unterlagen ordnen und verzeichnen 29

2.10.1 Anwendungsfallbeschreibung (Use Case Scenario) 29

2.10.2 Aktivitätsdiagramm (Activity Diagram) 30

2.10.3 Prototyp (Use Case Prototype) 31

2.11 Unterlagen magazinieren 31

2.11.1 Anwendungsfallbeschreibung (Use Case Scenario) 31

2.11.2 Aktivitätsdiagramm (Activity Diagram) 33

2.11.3 Anforderungen (Requirements) 34

2.12 Unterlagen erhalten 34

2.12.1 Anwendungsfallbeschreibung (Use Case Scenario) 34

2.12.2 Anwendungsfalldiagramm (Use Case Diagram) 35

2.12.3 Aktivitätsdiagramm (Activity Diagram) 36

2.12.4 Anforderungen (Requirements) 37

3 Datenmodellierung 39

3.1 Gegenstand der Archivierung 39

3.1.1 Ablieferung 39

3.1.2 Speicherung 39

3.1.3 Benutzung 39

3.2 Aktenplan - Archivtektonik - Signatur 39

3.3 Modellierung 41

3.4 Modellierung: SIP 41

3.4.1 Übernahme *41*
3.4.2 Mengenbeziehungen *42*
3.5 Modellierung: AIP 43
3.5.1 Verzeichniseinheit *43*
3.5.2 AIP (Archival Information Package) *43*
3.5.3 Objektrelationale Speicherung *44*
3.6 Archivjournal 44
3.6.1 Journal *44*
3.6.2 Akzessions-, Mutations- und Archivjournal *44*
3.7 Modellierung in XML 45
3.7.1 XML-Repräsentation des AIP *45*
3.7.2 XML-Repräsentation des SIP *46*
4 Digitales Magazin 48
4.1 Anforderungen 48
4.2 Speicherkonzepte 48
4.2.1 Speicherung in offener Struktur: *48*
4.2.2 Speicherung durch Migration von Blackbox zu Blackbox. *49*
4.3 Datensicherheit 49
4.4 Konsens im Projekt 49
5 Glossar 50
6 Anhang: Kurzdarstellung UML 55
6.1 Anwendungsfalldiagramm (Use Case Diagram) 55
6.2 Aktivitätsdiagramm (Activity Diagram) 55
6.3 Klassendiagramm 56
6.4 Objektdiagramm 57

Zusammenfassung

Im Rahmen des Arbeitsprogramms KOST.07 führen die Staatsarchive von Basel-Stadt, Zug und St. Gallen zusammen mit der KOST ein gemeinsames Modellprojekt zur Archivierung von Unterlagen aus Geschäftsverwaltungssystemen (AUGev) durch. Das Projekt folgt der klassischen Phasengliederung. Die Konzeptphase (Phase I) diente der gemeinsamen Erarbeitung von Prozessmodellierung und Anforderungsanalyse. In der Entwurfsphase werden Schnittstellen spezifiziert und funktionale Prototypen entwickelt. In der Konstruktions- und Übergabephase werden sich die Entwicklungslinien in den einzelnen Staatsarchiven zeitlich und nach Einbezug der jeweiligen Informatikdienstleister auch inhaltlich voneinander lösen.

Das vorliegende Papier ist der Schlussbericht der Phase I. Es umfasst die Modellierung der Archivierungsprozesse und Grundlagen der Datenmodellierung, ergänzt um Überlegungen zum digitalen Magazin, ein Glossar, eine Bibliografie und ein Anhang zur UML. Sein Inhalt entspricht dem Arbeitsstand des Projekts im August 2007.

Für die Modellierung dient das im Bereich der digitalen Archivierung weit verbreitete Referenzmodell für ein Open Archival Information System (OAIS) als Leitfaden. Das OAIS beschreibt einerseits die Prozesse, die ein digitales Archiv umfasst, anderseits die Schnittstellen zur Übergabe von Informationen zwischen Prozessen. Das Projekt AUGev konzentriert sich auf die vier Teilprozesse Ingest (Übernahme), Archival Storage (Speicherung), Data Management (Verwaltung) und Access (Zugang) und spezifiziert die Schnittstellen SIP (Schnittstelle zum GEVER-System), AIP (archivische Datenkapsel), DIP (Zugangsformat) und Descriptive Info (Schnittstelle zum Findmittel).

Zur Strukturierung und Darstellung wird die Modellierungssprache UML (Unified Modelling Language) verwendet. Diese erlaubt ein iteratives Vorgehen zur Prozessmodellierung: In einem ersten Durchgang wird die Interaktion eines Anwenders mit dem modellierten System mittels Anwendungsfällen beschrieben. Diese werden danach in weitere Anwendungsfälle unterteilt, in eine Abfolge von Aktivitäten aufgegliedert und detailliert beschrieben. So entsteht eine abstrakte Abbildung des gesamten Archivierungsprozesses auf einzelne Schritte und ihre Resultate. Ausgehend vom Lebenszyklus der Unterlagen identifiziert und untersucht das Projekt die Archivprozesse „Unterlagen übernehmen“, „Unterlagen magazinieren“ und „Unterlagen erhalten“.

Die Prozessmodellierung hält zudem fest, welcher Input und Output für jeden Schritt erforderlich ist. Diese Objekte werden im Rahmen der Datenmodellierung genauer spezifiziert, ebenfalls mit Hilfe der UML. Es handelt sich dabei einerseits um die Schnittstellenobjekte, die vom OAIS gefordert werden, anderseits um weitere Objekte, die aus der Prozessmodellierung hervorgehen. Im Verlauf der Arbeit hat sich gezeigt, dass die Spezifikation der Objekte sinnvollerweise direkt an Prototypen erprobt und verbessert wird. Dies ist ein Hauptbestandteil der Entwurfsphase (Phase II). Das vorliegende Dokument gibt deshalb nur die grundsätzlichen Überlegungen zur Datenmodellierung wieder.

1 Einleitung

Das Projekt AUGev befasst sich mit der Archivierung von Unterlagen aus Geschäftsverwaltungssystemen. Es umfasst die Modellierung der Prozesse und Objekte der Archivierung sowie Überlegungen zum digitalen Magazin. (Zu den verwendeten Begriffen siehe das **Glossar**.)

Eine kurze Beschreibung des heutigen, d.h. papierbasierten Archivierungsprozesses soll in die Thematik einführen. Danach wird der Projektinhalt abgegrenzt und auf das Referenzmodell für ein Open Archival Information System (**OAIS**) abgebildet.

Der besseren Lesbarkeit wegen wird nur die männliche Form verwendet. Selbstverständlich gilt das Gesagte jeweils auch für Archivarinnen, Archivinformatikerinnen etc.

1.1 Beschreibung des heutigen Archivierungsprozesses

Die vorarchivische Phase und die eigentliche Archivierung in papierbasierten Archivierungsprozessen lassen sich folgendermassen generalisieren. (Im Idealfall werden alle diese Schritte durchlaufen, in der Realität oft nicht. Die Reihenfolge ist nicht zwingend.)

- Angebot von Unterlagen an das Archiv durch ein Amt
- Bewertung der Unterlagen durch das Archiv
- Unterlagenaufbereitung im Amt
- Ablieferung der Unterlagen ins Archiv
- Erstellen eines Ablieferungsprotokolls (Journal im Archiv)
- Ablieferungsquittung an das Amt
- Eingangskontrolle im Archiv
- Umschachteln und konservatorische Bearbeitung
- Erschliessen (Ordnen & Verzeichnen)
- Nachbewertung im Archiv (Aussondern)
- Magazinierung
- Eintrag des Standortes im Findmittel
- Rückmeldung an das liefernde Amt (Kassation im Amt)
- Konservatorische Nachbearbeitung (Bestandeserhaltung)

Amt	**Archiv**
Angebot	->
<-	Bewertung
Unterlagenaufbereitung	
Ablieferung	->
<-	Ablieferungsprotokoll & Quittung
	Kontrolle
	Umschachteln
	Erschliessen
	Nachbewertung
	Magazinierung
	Bestandeserhaltung

Tabelle 1: Vorarchivische Phase und Archivierung

Die Phase der Benutzung gestaltet sich folgendermassen:

- Der Benutzer gelangt mit einem Informationsbedürfnis ans Archiv
- Der Benutzer sucht im Findmittel und findet die betreffenden Unterlagen

- Der Benutzer fordert Unterlagen an
- Bereitstellung der Unterlagen zur Einsicht
- Einsicht durch den Benutzer
- Rückschub der Unterlagen ins Magazin

Wir verweisen hier auf die auf die detaillierten Ergebnisse der KOSTPROBE-Arbeitsgruppe Benutzung: www.kost-ceco.ch projekte > kostprobe > ag benutzung

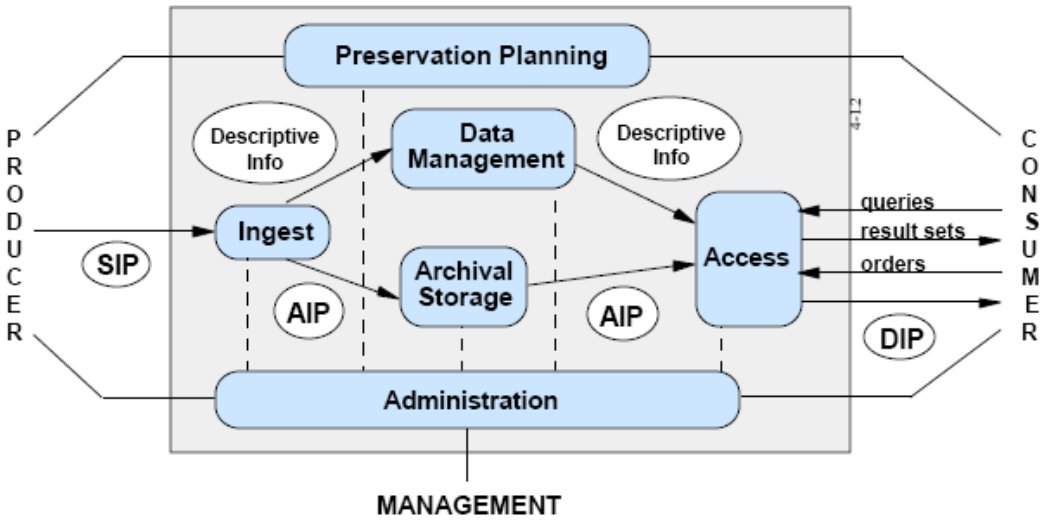
Benutzer	Archiv
Informationsbedürfnis	
Suchen	
Unterlagenanforderung	->
<-	Bereitstellen
Unterlageneinsicht	
->	Rückschub

Tabelle 2: Benutzung

1.2 Projektinhalt

1.2.1 Modellierung nach OAIS

Als Grundreferenz dient uns das Open Archival Information System OAIS. Das OAIS-Referenzmodell beschreibt ein digitales Archiv als Interaktion von Prozessen und Spezifikation von Schnittstellen, die dazu dienen, digitale Akten zu speichern und wieder verfügbar zu machen. Das OAIS-Modell gilt allgemein als anerkanntes Referenzmodell für eine digitale Archivierung und wurde 2003 zum internationalen Standard ISO 14721 erklärt.
Die Implementierung eines OAIS-konformen Archivs ist durch diese Empfehlungen jedoch nicht festgelegt.

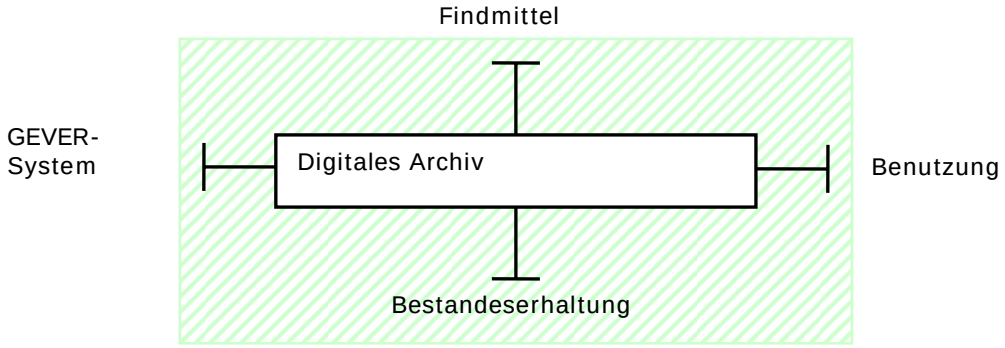


Grafik 1: Übersicht Referenz-Modell: Consultative Committee on Space Data Systems, Reference Model for an Open Archival Information System (OAIS), CCSDS 650.0-B-1, BLUE BOOK, January 2002. http://public.ccsds.org/publications/archive/650x0b1.pdf

SIP: Submission Information Package
AIP: Archival Information Package
DIP: Dissemination Information Package

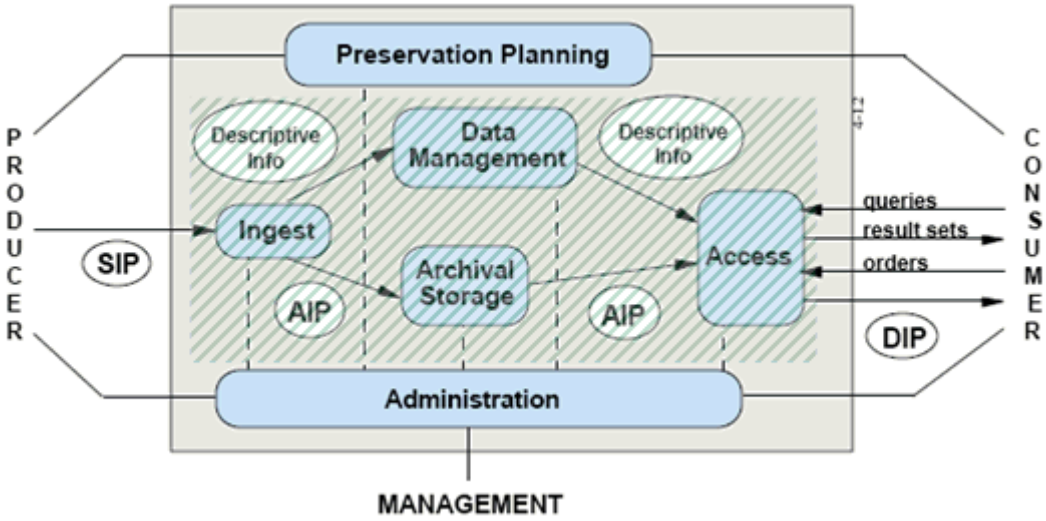
1.2.1.1 Abgrenzen der Aufgabenstellung

Im Projekt soll die Archivierung von Unterlagen aus Geschäftsverwaltungssystemen (GEVER-System) behandelt werden. Der Projekt-Fokus ist also eingeschränkt auf ein digitales Archiv inklusive Bestandeserhaltung mit Schnittstellen zu einem GEVER-System, zum archivinternen Findmittel und zur Benutzung. Daraus ergibt sich, dass der vorarchivische Bereich, die Ausprägung des Findmittels und die Benutzung nicht Teil des Projekts sind. Diese Bereiche müssen jedoch mitgedacht werden und fliessen in die Spezifikation der Schnittstellen ein. Im Folgenden eine Darstellung mit dem Digitalen Archiv im Zentrum und den Schnittstellen zu den angrenzenden Systemen, der Projektfokus ist im weitern jeweils grün schraffiert.



Grafik 2: Abgrenzung der Aufgabenstellung

Eine generische Lösung zur Langzeitarchivierung von Geschäftsdaten ist eher ein Konzept bzw. eine Menge von Spezifikationen als eine konkrete Softwarelösung. Dennoch sind Überlegungen zur Implementierung auch in der Modellierungsphase nicht unwesentlich, damit rechtzeitig technische Probleme erkannt und angegangen werden können. Zudem wird die Umsetzung der Modellierung in ein konkretes lauffähiges System in der zweiten Projektphase dadurch erleichtert.



Grafik 3: Kernprozesse der Archivierung: Ingest, Archival Storage, Data Management, Access

Auf das OAIS-Modells abgebildet bedeutet die Projektausrichtung eine Konzentration auf die vier Kernprozesse *Ingest* (Übernahme), *Archival Storage* (Speicherung), *Data Management* (Verwaltung), *Access* (Zugang). Daneben müssen die vier Schnittstellen *SIP* (Schnittstelle zum GEVER-SYSTEM), *AIP* (Datenkapsel), *DIP* (Zugangsformat) und *Descriptive Info* (Schnittstelle zum Findmittel) spezifiziert werden.

1.2.2 Überblick über Software-Layer

Ein erster Schritt der Konkretisierung des OAIS besteht in der Zuordnung der Funktionen zu den betroffenen Software-Layern. Die folgende Tabelle gibt eine Übersicht.

Software-Layer	Funktion und Beschreibung
Archival Management Software	Findmittel bzw. Katalog erlaubt den (strukturierten) Zugriff auf die archivierten Daten
Digital Archiving Software	Entgegennehmen, Speichern, Verwalten, und Ausliefern von digitalen Objekten
Storage Management Software	Verwalten des Speicherplatzes auf dem Speichersystem
Storage Hardware	Physisches Speichermedium

Tabelle 3: Die wichtigsten Software-Layer

Das Projekt AUGev konzentriert sich auf den Layer der "Digital Archiving Software". Die "Archival Management Software" ist, wie erwähnt, nur für die Spezifikation der entsprechenden Schnittstellen relevant. Die "Storage Management Software" sowie die darunter liegende Hardware entsprechen dem digitalen Magazin. Dieses liegt zwar innerhalb der Aufgabenstellung. Die genaue Spezifikation und Beschaffung eines digitalen Magazins wird im Projekt jedoch nur summarisch behandelt und wird Thema eines weiteren KOST-Projekts sein.

Die Teilprozesse nach OAIS lassen sich folgendermassen auf den Software-Layers verorten:

1.2.2.1 Preservation Planning und Administration

Die "Archival Management Software", auch "Archival Information System" genannt, kann die OAIS-Nebenprozesse *Preservation Planning* und *Administration* enthalten. Wir verzichten in diesem Projekt auf die Spezifizierung dieser Managementprozesse, da sie ausserhalb des Projektfokus liegen.

1.2.2.2 Ingest, Data Management, Archival Storage und Access

Dies sind die eigentlichen Prozesse des Software-Layers „Digital Archiving Software“. Hier wird ein digitales Objekt (Datei) entgegengenommen, gespeichert und wieder ausgeliefert. Üblicherweise sind dazu in Lösungen im kommerziellen Umfeld eine Datenbank, eine Archivsoftware und ein Speichersystem notwendig.
In der Datenbank speichert der Prozess *Data Management* Metadaten, Indexdaten und Verweise auf die archivierten Unterlagen. Dieser Aspekt wird im Modellprojekt teilweise bereits durch die existierende „Archival Management Software“ wahrgenommen. Falls wir die Verwaltung deskriptiver Informationen vollständig an diese Softwareschicht delegieren können (was in den Pilotimplementationen zu verifizieren ist), übernimmt der Prozess *Data Management* allein die Steuerung des Archivierungsprozesses und die Verwaltung der technischen Metadaten.
Für das Entgegennehmen und Ausliefern von Objekten muss je ein Schnittstellenprotokoll implementiert werden, das erlaubt, digitale Objekte an das Archiv zu schicken und von dort zu empfangen, etwa SOAP, XML-RPC (beides Protokolle zum Austausch von Daten zwischen Computersystemen). Das Schnittstellenprotokoll wird sinnvollerweise auch zur Kommunikation zwischen den einzelnen Prozessen verwendet. Die Schnittstellen selber, SIP, AIP, DIP, können sowohl als *Interface* (im Sinne der Objektmodellierung) wie auch als *Data Definition* spezifiziert werden.
Der Aspekt der Datenintegrität kann von der Archivsoftware oder vom darunter liegenden Speichersystem wahrgenommen werden.

1.2.3 AUGev-Modellierung

Die Spezifikation der Archivierung von Unterlagen aus Geschäftsverwaltungssystemen gliedert sich in drei Abschnitte:

Prozessmodellierung: Ein erster Abschnitt modelliert den Archivierungsprozess, indem er von einer globalen Übersicht zur Detailmodellierung von Teilprozessen fortschreitet. Dazu werden Use-Case- und Aktivitätsdiagramme der UML (Unified Modeling Language) benutzt: Die Use-Case-Diagramme erlauben den Überblick über die Prozesse und Akteure, die

Aktivitätsdiagramme detaillieren die einzelnen Schritte in den Prozessen und die dazu benötigten Objekte. (Zur UML siehe die kurze Einführung im Anhang.)

Datenmodellierung: Im zweiten Abschnitt richtet sich der Fokus auf die in den Prozessen verarbeiteten Daten und ihre Beziehungen. Dazu werden Klassen- und Objektdiagramme der UML benutzt.

Digitales Magazin: Ein kürzerer, dritter Teil enthält grundlegende Überlegungen zur eigentlichen Speicherung, hier „Digitales Magazin“ genannt.

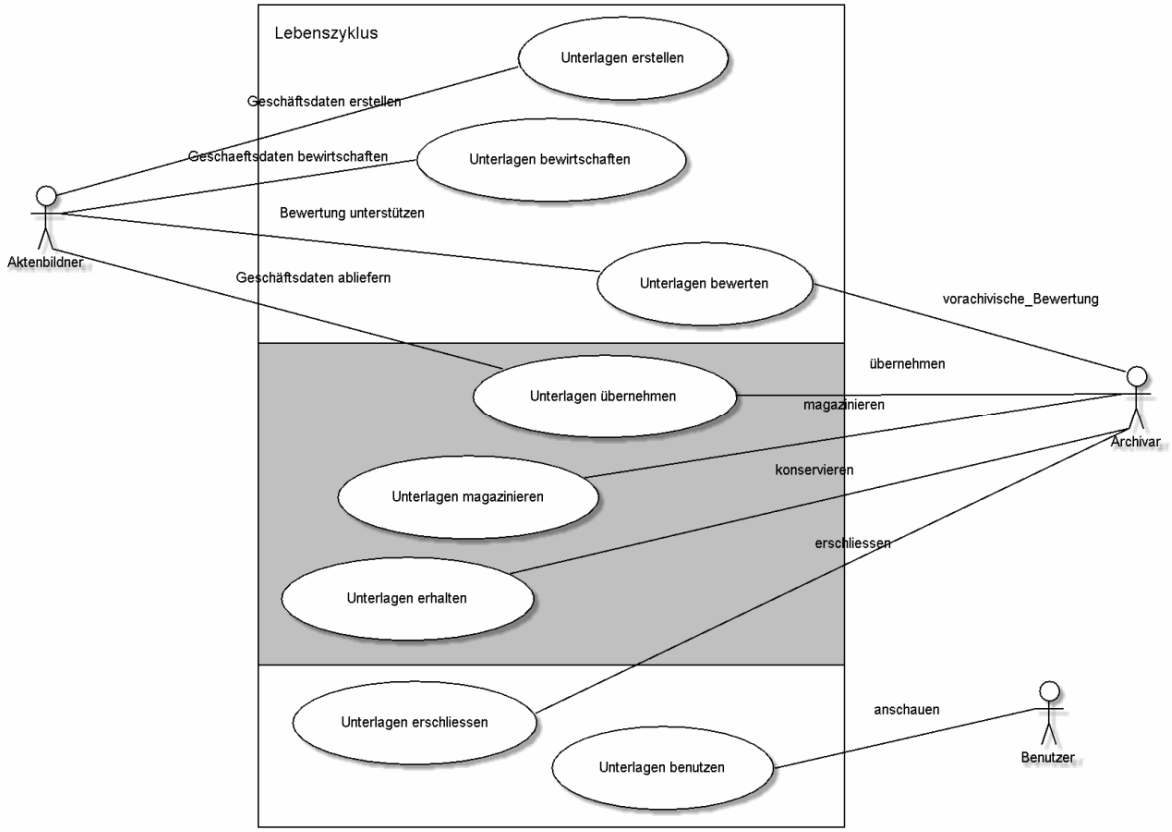
2 Prozessmodellierung

2.1 Use-Case-Modellierung

Die *Anwendungsfallmodellierung* konkretisiert das OAIS-Modell. Wir bestimmen hier die Begrifflichkeit für die weitere Arbeit. Die einzelnen Anwendungsfälle werden sukzessive detailliert und durch *Aktivitätsdiagramme, Anforderungskataloge* und *Anwendungsfall-Prototypen* genauer spezifiziert.

2.1.1 Lebenszyklus digitaler Unterlagen

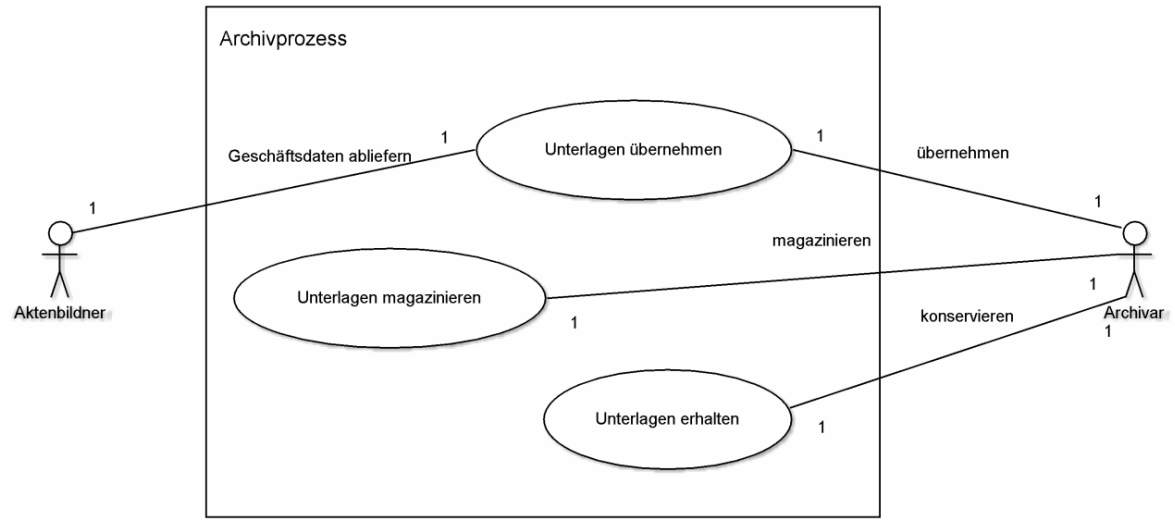
Das Use-Case-Diagramm *Lebenszyklus* bildet den *Lifecycle* digitaler Unterlagen aus Sicht des Archivs ab. Die Definition des Systems ist hier etwas künstlich, besteht der Lebenszyklus doch aus den Provenienz-, Archivierungs- und Benutzungsprozessen und ist weder eine organisatorische Einheit noch ein IT-System.



Grafik 4: Use Case Lebenszyklus mit Abgrenzung der Provenienz- und Benutzungs-Prozesse

2.1.2 Archiv-Prozesse

Das Use Case Diagramm *Archiv-Prozesse* schränkt auf den eigentlichen Projektfokus ein und umfasst nur noch die Anwendungsfälle der Archivierung im engeren Sinne: *Unterlagen übernehmen*, *Unterlagen magazinieren* und *Unterlagen erhalten*.



Grafik 5: Use Case Archivprozess

Die ebenfalls dem Akteur Archiv zugeordneten Use Cases *Unterlagen bewerten* und *Unterlagen erschliessen* sehen wir ausserhalb des eigentlichen Projektfokus:

Unterlagen bewerten beinhaltet Bewertung und Beratung in der vorarchivischen Phase, wenn sich die Unterlagen noch beim Amt befinden.

Bei *Unterlagen erschliessen* handelt es sich um archivische Nacherschliessung, insbesondere mit Angaben auf höheren Erschliessungsniveaus (Archiv, Bestände, Ablieferungen). Dieser Use Case unterscheidet sich nicht von der entsprechenden Arbeit für analoge Unterlagen. Im Gegenzug dazu umfasst *Unterlagen ordnen und verzeichnen* im Use Case *Unterlagen übernehmen* die automatisierte Einordnung in die Archivtektonik und die Übernahme der Metadaten.

Der Begriff „Unterlagen“ umfasst sowohl die Primärdaten wie auch die zugehörigen Metadaten. Ein Bestand ist eine Aggregation von Unterlagen, die erst bei der Erschliessung stattfindet (sog. Bestandbildung).

Im Glossar AUGev, Kapitel 5 werden diese und weitere Begriffe erläutert.

2.2 Unterlagen übernehmen

2.2.1 Anwendungsfallbeschreibung (Use Case Scenario)

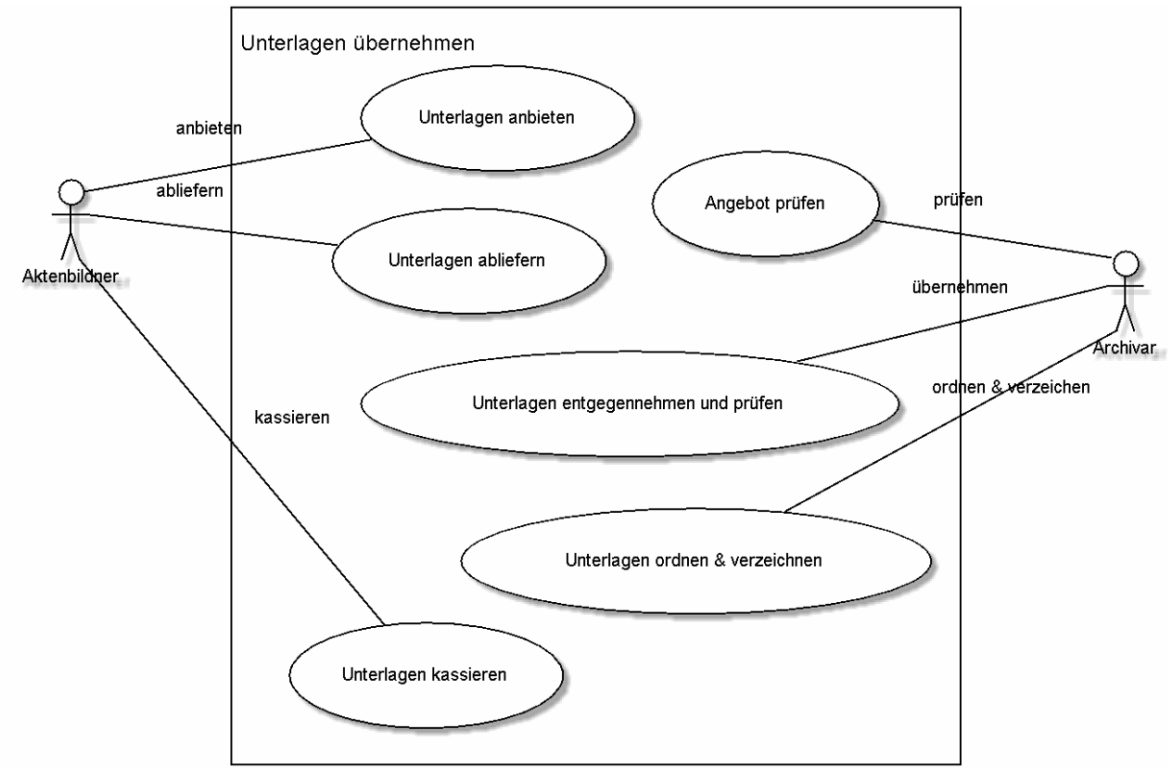
Name	Unterlagen übernehmen
Kurzbeschreibung	Export von Unterlagen eines Aktenbildners aus einer bestimmten Periode. Exportiert werden GEVER-Metadaten und Primärdaten, Anschliessend erfolgt die Übergabe der Daten an das Archiv, so wie es in der Schnittstelle SIP *(Submission Information Package)* definiert ist.
Akteure	Der Aktenbildner als Lieferant Der Archivar als Empfänger
Auslöser (äusseres Ereignis, das den Anwendungsfall auslöst)	Vereinbarung über Aktenlieferung ausserhalb des Systems
Vorbedingungen (vorausgesetzter innerer Zustand des Systems)	Die technischen Voraussetzungen für eine Ablieferung sind gegeben (Übereinkunft). Das Liefersystem kann die Schnittstellenanforderungen erfüllen.
Eingehende Daten (Benutzereingaben im Rahmen des Anwendungsfalls)	Eine Menge Primär- und Metadaten gegliedert nach Ablage, Ordnungssystem, Dossier und Dokument
Ergebnis(se) (was einem oder mehreren	Die übernommenen Primär- und Metadaten befinden sich wohlgeformt in einer Eingangszone (Landing Zone) des digitalen Archivs.

Akteuren geliefert worden ist)	
Nachbedingungen (erreichter innerer Zustand des Systems)	Die übernommenen Daten stehen für die Langzeitarchivierung bereit Der Lieferant hat keinen Zugriff auf die abgelieferten Primär- und Metadaten mehr
Essenzielle Schritte (Ablauf im Detail)	Abzuliefernde digitale Unterlagen werden dem Archiv übermittelt Formale Prüfung der gelieferten Daten Inhaltliche Prüfung durch Klassenbildung nach Anzahl Dokument pro Rubrik und Dossier Teilkassierung von überflüssigen Teilen Übernahmequittung an Anlieferer ausstellen

Tabelle 4: Anwendungsfallbeschreibung Unterlagen übernehmen

2.2.2 Anwendungsfalldiagramm (Use Case Diagram)

Der Anwendungsfall *Unterlagen übernehmen* lässt sich weiter detaillieren:



Grafik 6: Use Case Diagramm Unterlagen übernehmen detailliert

2.2.2.1 Unterlagen anbieten

Der Aktenbildner bietet dem Archiv Unterlagen an, die bleibend aufbewahrt werden sollen. Grundlage ist eine Ablieferungsvereinbarung, welche im Zuge der vorarchivischen Bewertung bereits festlegt, welche Akten als archivwürdig klassifiziert sind und in welchen Mengen und Zeiträumen abgeliefert wird.
Das Angebot kann auf verschiedenem Weg erfolgen (z.B. via Mail, SOAP, etc.).

2.2.2.2 Angebot prüfen

Das Archiv prüft die Menge und in Stichproben die Qualität der angebotenen Unterlagen und vergleicht das Angebot mit der vorarchivischen Ablieferungsvereinbarung. Besteht keine Ablieferungsvereinbarung mit dem Aktenbildner, muss eine solche manuell eingerichtet werden. Möglicherweise lehnt das Archiv aus formalen oder inhaltlichen Gründen das Angebot ab und

erarbeitet eine neue Ablieferungsvereinbarung mit dem Aktenbildner. Sonst bewertet das Archiv das Angebot und erstellt einen Übernahmeentscheid für dieses Angebot.

2.2.2.3 Unterlagen abliefern

Der Aktenbildner stellt die im Übernahmeentscheid aufgeführten Unterlagen zusammen mit ihren Metadaten in geeigneter Form zur Ablieferung bereit (SIP). Grundlage für die Ablieferung ist der Übernahmeentscheid (eigentliche eine Liste der zu liefernden Unterlagen). Die Ablieferung selbst kann auf unterschiedlichem Wege geschehen (z.B. CD, FTP, etc.).

2.2.2.4 Unterlagen entgegennehmen und prüfen

Das Archiv prüft, ob alle im Übernahmeentscheid referenzierten Unterlagen in der Ablieferung enthalten sind. Eine unvollständige oder zu grosse Ablieferung wird zurückgewiesen und der Prozess abgebrochen. Bei einer korrekten Ablieferung erfolgt eine mehrstufige formale Prüfung der gelieferten Primär- und Metadaten:

- Syntaktische Korrektheit und Vollständigkeit der gesamten Ablieferung prüfen.
- Prüfung der Metadaten auf syntaktische Korrektheit (XML-Syntaxprüfung, Schema-Validierung).
- Vollständigkeit im Verhältnis von Metadaten zu Primärdaten prüfen.
- Format-Validierung der angelieferten Primärdaten, den eigentlichen Dokumentdateien.
- Eine korrekte Ablieferung wird zur Weiterverarbeitung (digitale Langzeitarchivierung) übernommen.

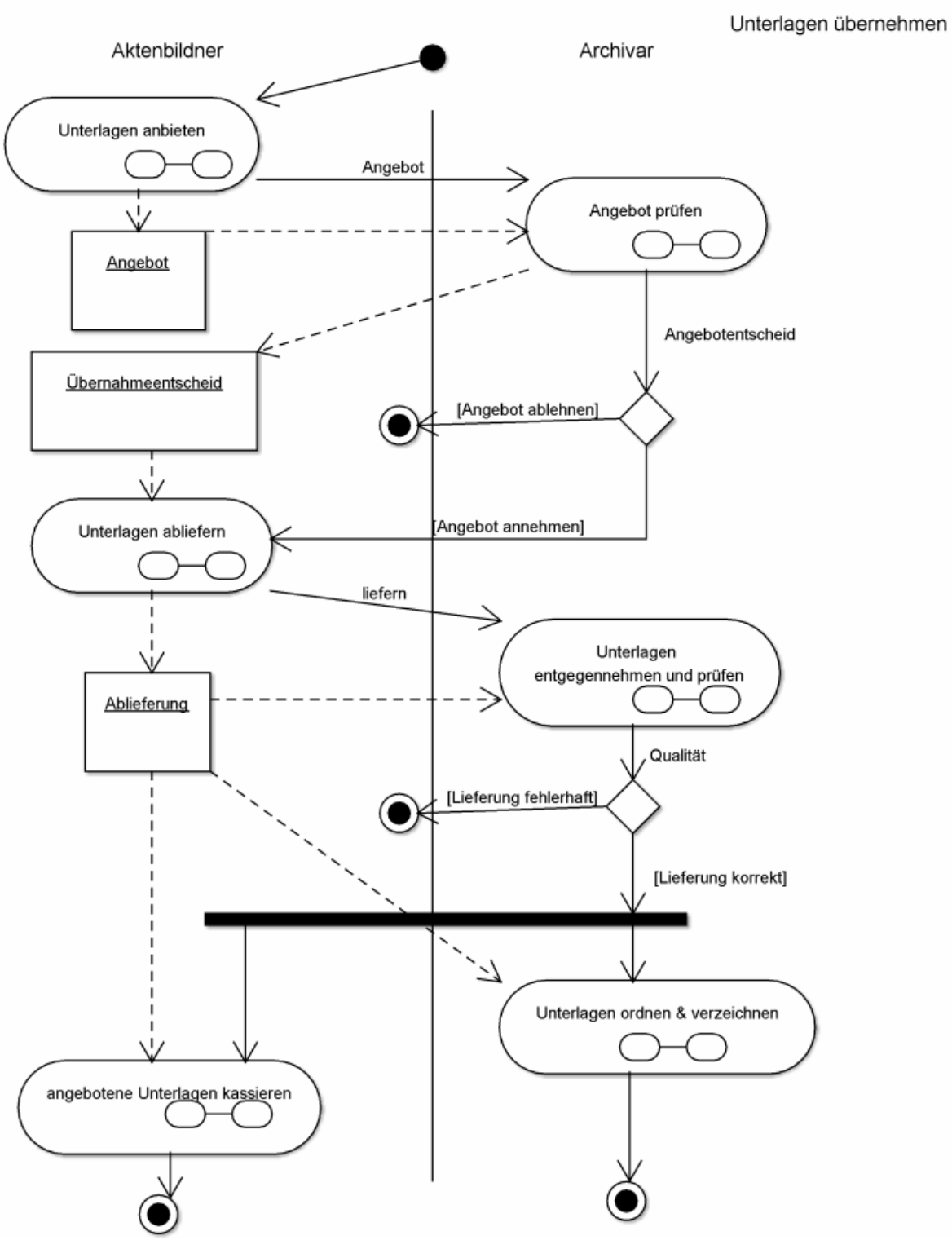
2.2.2.5 Unterlagen Ordnen und Verzeichnen

Aus den Unterlagen einer Ablieferung wird ein Bestand gebildet. Die Bestandsbildung, eine logische Gruppierung der Ablieferung nach Amt und Ablage, ist Grundlage beim Eintragen der Ablieferung ins Findmittel und steuert die Übernahme der GEVER-Metadaten ins Findmittel. Ordnen und Verzeichnen ist eigentlich Klassenbildung gemäss den Kriterien des Aktenplans und Abbilden des Aktenplanes auf den Archivplan. In diesem Prozess kann das Findmittel mit Metadaten beliefert und eine Referenz (Signatur) auf die archivierten Unterlagen erstellt werden. Die Signatur ist eigentlich eine durch den Verzeichnisort gebildete GUID (Globally Unique Identifier), durch die die archivierten Unterlagen zitierbar werden.

2.2.2.6 Unterlagen kassieren

Alle einmal angebotenen Unterlagen sollen bei der aktenbildenden Stelle zum Kassieren markiert werden. Die abliefernde Stelle erhält dazu eine Bestätigung für sämtliche angebotenen Unterlagen in einer definierten und damit maschinell zu verarbeitenden Form. Darin ist aufgeführt, welche Unterlagen archiviert wurden und welche nicht.

2.2.3 Aktivitätsdiagramm (Activity Diagram)



Grafik 7: Aktivitätsdiagramm Unterlagen übernehmen

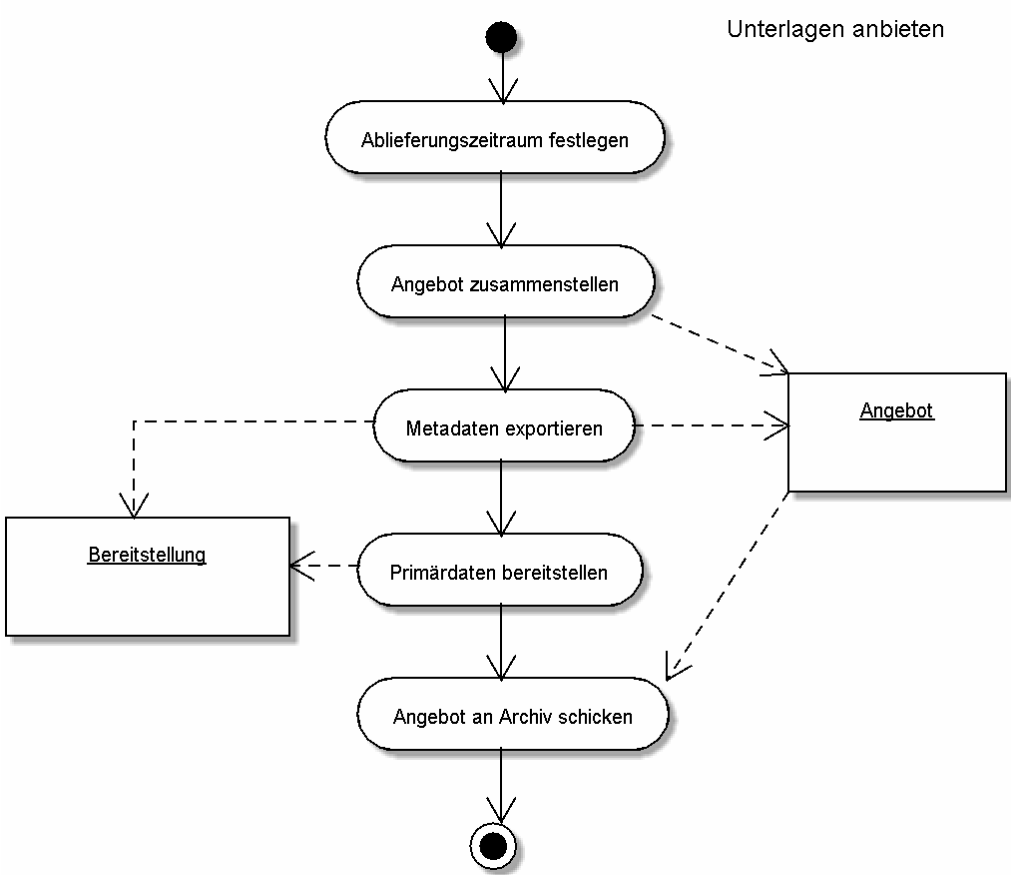
2.3 Unterlagen anbieten

2.3.1 Anwendungsfallbeschreibung (Use Case Scenario)

Name	Unterlagen anbieten
Kurzbeschreibung	Der Aktenbildner bietet dem Archiv Unterlagen an, die bleibend aufbewahrt werden sollen. Grundlage ist eine Ablieferungsvereinbarung, welche im Zuge der vorarchivischen Bewertung bereits festlegt, welche Akten als archivwürdig klassifiziert sind und in welchen Mengen und Zeiträumen abgeliefert wird.
Akteure	Der Prozess ist vollumfänglich unter Kontrolle des Aktenbildners.
Auslöser (äusseres Ereignis, das den Anwendungsfall auslöst)	Der Aktenbildner löst den Prozess der Ablieferung aus.
Vorbedingungen (vorausgesetzter innerer Zustand des Systems)	Die angesammelte Aktenmenge oder die verflossene Zeit rechtfertigen eine Ablieferung. Die abzuliefernden Unterlagen sind im GEVER-System als "zu archivieren" markiert.
Eingehende Daten (Benutzereingaben im Rahmen des Anwendungsfalls)	Aktenbildner legt den Zeitraum für das Angebot fest.
Ergebnis(se) (was einem oder mehreren Akteuren geliefert worden ist)	Ein Angebot, eine Liste aller abzuliefernden Unterlagen zusammen mit den zugehörigen Ordnungsinformationen und Metadaten, wird zusammengestellt und ans Archiv geschickt. Die Primärdaten werden für den Zugriff von aussen oder den Versand bereitgestellt.
Nachbedingungen (erreichter innerer Zustand des Systems)	Die im Angebot enthaltenen Unterlagen sind im GEVER-System als angeboten markiert.
Essenzielle Schritte (Ablauf im Detail)	Ablieferungszeitraum festgelegt oder vom GEVER-System ermittelt Angebot zusammenstellen (Liste von Unterlagen) Meta- und Primärdaten zum Angebot bereitstellen/erzeugen Angebot an Archiv schicken

Tabelle 5: Anwendungsfallbeschreibung Unterlagen anbieten

2.3.2 Aktivitätsdiagramm (Activity Diagram)



Grafik 8: Aktivitätsdiagramm Unterlagen anbieten

2.3.2.1 Ablieferungszeitraum festlegen

In der Regel beginnt der Ablieferungszeitraum beim Ende des Zeitraums der vorhergehenden Ablieferung und reicht bis maximal zum Termin, der um die Dauer der Aufbewahrungsfrist vor dem aktuellen Datum liegt.

2.3.2.2 Angebot zusammenstellen

Die Dossiers, die im Ablieferungszeitraum abgeschlossen wurden, bilden den Kern des Angebots. Manuelle Korrektur ist möglich.

2.3.2.3 Metadaten exportieren

Zu den anzubietenden Dossiers werden die GEVER-Metadaten exportiert. Die Metadaten bilden das Angebot, das einem definierten Format entsprechen muss.

2.3.2.4 Unterlagen bereitstellen

Die anzubietenden Unterlagen werden im System entsprechend markiert. Damit ist auch der Zugriff des Archivs möglich. Falls das System dies nicht zulässt, muss eine Kopie der anzubietenden Unterlagen an einem speziellen Ort bereitgestellt werden.

2.3.2.5 Angebot an Archiv schicken

Das Angebot wird auf eine vereinbarte Art (z.B. FTP, Mail, etc.) an das Archiv geschickt.

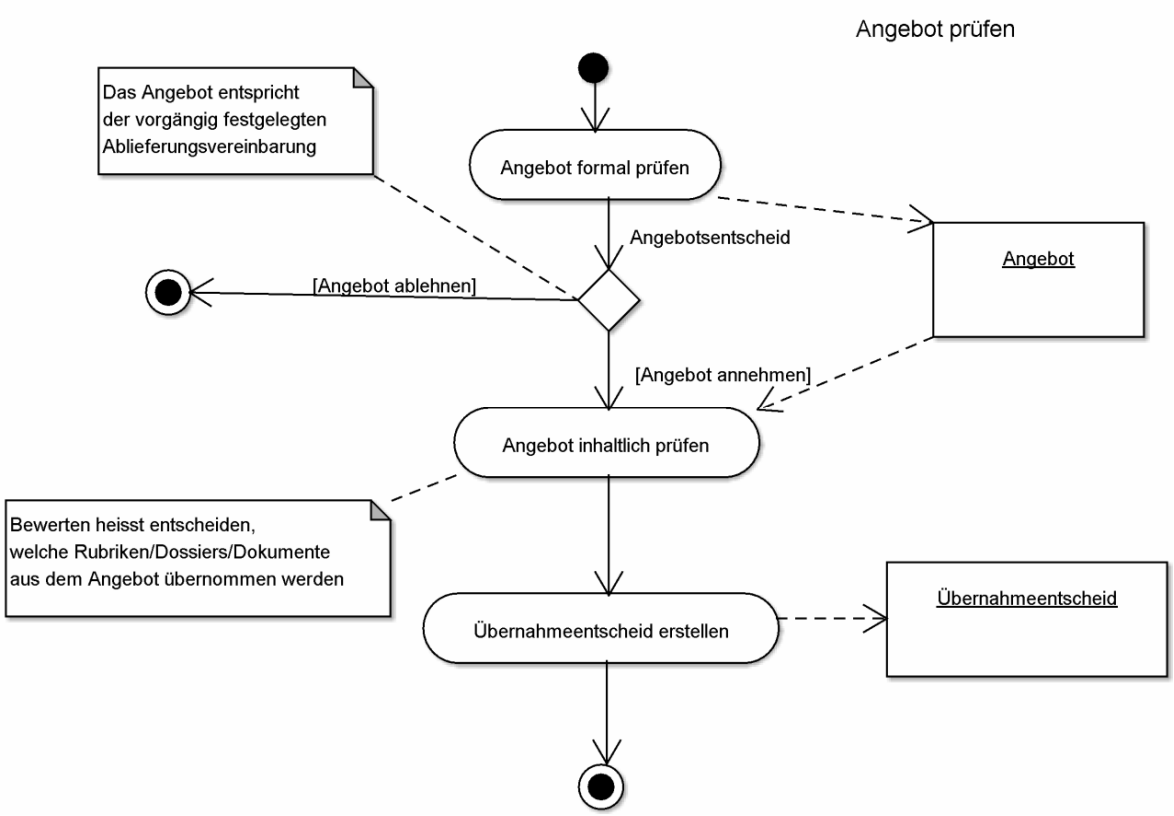
2.4 Angebot prüfen

2.4.1 Anwendungsfallbeschreibung (Use Case Scenario)

Name	Angebot prüfen
Kurzbeschreibung	Das Archiv prüft die Menge und in Stichproben die Qualität der angebotenen Unterlagen und vergleicht das Angebot mit der vorarchivischen Ablieferungsvereinbarung. Möglicherweise lehnt das Archiv aus formalen oder inhaltlichen Gründen das Angebot ab und erarbeitet eine neue Ablieferungsvereinbarung mit dem Aktenbildner. Sonst bewertet das Archiv das Angebot und erstellt einen Übernahmeentscheid für dieses Angebot.
Akteure	Der Archivar
Auslöser (äusseres Ereignis, das den Anwendungsfall auslöst)	Ein Angebot ist eingetroffen.
Vorbedingungen (vorausgesetzter innerer Zustand des Systems)	Das Angebot entspricht den formalen Vorgaben (XML-Schema o.ä.) und kann damit verarbeitet werden.
Eingehende Daten (Benutzereingaben im Rahmen des Anwendungsfalls)	Ein Angebot: eine Liste aller abzuliefernden Unterlagen zusammen mit den zugehörigen Ordnungsinformationen und Metadaten
Ergebnis(se) (was einem oder mehreren Akteuren geliefert worden ist)	Ein Übernahmeentscheid: eine Liste der abzuliefernden Unterlagen, d.h. eine Teilmenge der Angebotsliste
Nachbedingungen (erreichter innerer Zustand des Systems)	Die Bewertung für dieses Angebot wird vom Archiv festgehalten.
Essenzielle Schritte (Ablauf im Detail)	Formale Prüfung des Angebots Bewertung des Angebots (klassische Archivtätigkeit) Übernahmeentscheid erstellen

Tabelle 6: Anwendungsfallbeschreibung Angebot prüfen

2.4.2 Aktivitätsdiagramm (Activity Diagram)



Grafik 9: Aktivitätsdiagramm Angebot prüfen

2.4.2.1 Angebot formal prüfen

Dieser Schritt beinhaltet eine automatisierte Prüfung des Angebots an Hand der Vorgaben, im Speziellen der Ablieferungsvereinbarung.

2.4.2.2 Angebot inhaltlich prüfen

Das Archiv prüft das Angebot qualitativ und quantitativ. D.h., das Archiv ermittelt mit Stichproben, ob die angebotenen Unterlagen der Vereinbarung und den archivischen Belangen entsprechen. Bei seriellen Akten wird auch die Grösse und statistische Relevanz der Stichprobenmenge untersucht. Aufgrund dieser Bewertung wird ein Übernahmeentscheid erstellt, das ist eine Liste der zu liefernden Unterlagen. Diese kann das ganze Angebot, aber auch nur eine Teilmenge daraus umfassen.

Im Projekt wird davon ausgegangen, dass eine prospektive Bewertung bereits stattgefunden hat. Beim Schritt "Angebot inhaltlich prüfen" handelt es sich also nicht um eine Bewertung von Grund auf, sondern um eine Kontrolle der Anwendung von Bewertungsentscheiden, eine allfällige Korrektur dieser Entscheide und die konkrete Durchführung eines allfälligen Samplings.

2.4.2.3 Übernahmeentscheid erstellen

Der Übernahmeentscheid, eigentlich eine Liste der Dossiers und Dokumente, die aus Sicht des Archivs aus dem vorhergegangenen Angebot übernommen werden sollen, wird an den Aktenbildner zurückgeschickt. Der Übernahmeentscheid ist eine Teilmenge des Angebots.

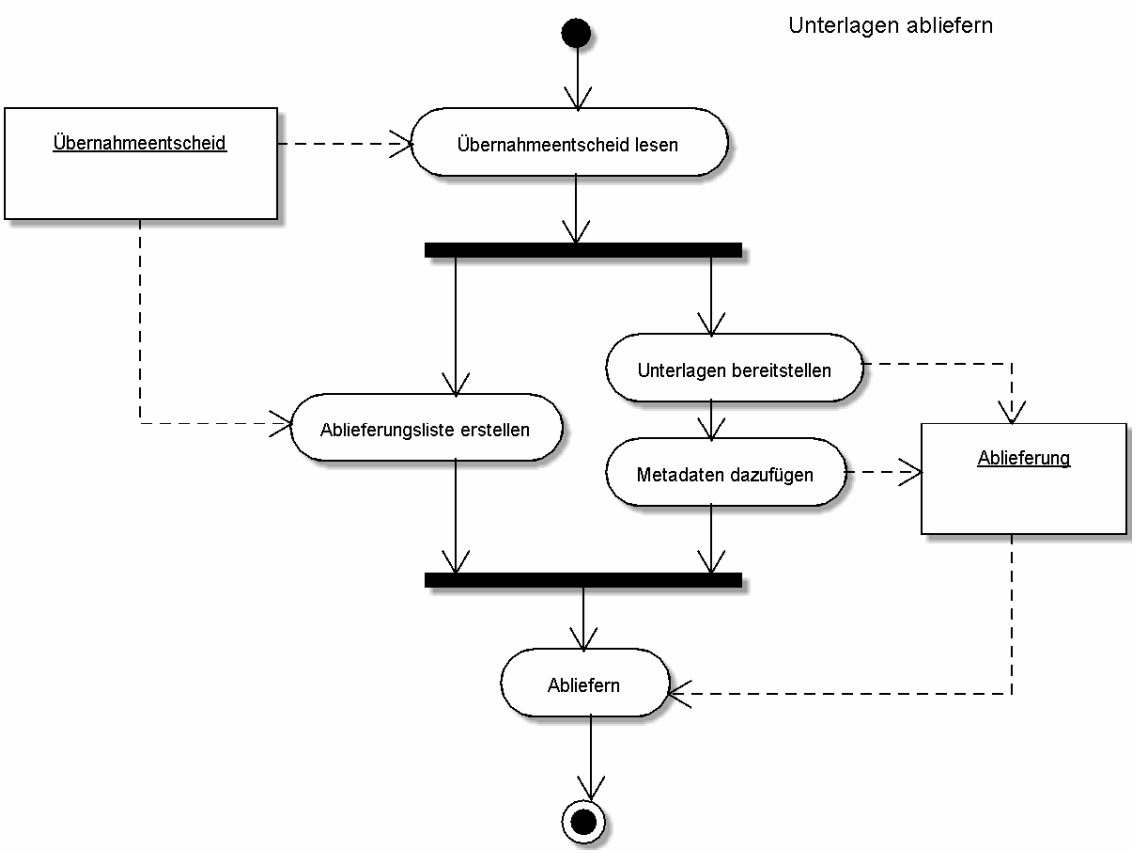
2.5 Unterlagen abliefern

2.5.1 Anwendungsfallbeschreibung (Use Case Scenario)

Name	Unterlagen abliefern
Kurzbeschreibung	Der Aktenbildner stellt die im Übernahmeentscheid aufgeführten Unterlagen (Primär- und Metadaten) in geeigneter Form zur Ablieferung bereit.
Akteure	Der Aktenbildner
Auslöser (äusseres Ereignis, das den Anwendungsfall auslöst)	Ein Übernahmeentscheid ist als Antwort auf ein Angebot eingetroffen.
Vorbedingungen (vorausgesetzter innerer Zustand des Systems)	Der Übernahmeentscheid kann einem eigenen Angebot zugeordnet werden. Alle abzuliefernden Unterlagen sind entsprechend im GEVER-System markiert (archivieren, zur Ablieferung bereit o.ä.).
Eingehende Daten (Benutzereingaben im Rahmen des Anwendungsfalls)	Ein Übernahmeentscheid: eine Liste der abzuliefernden Unterlagen
Ergebnis(se) (was einem oder mehreren Akteuren geliefert worden ist)	Die verlangten Unterlagen stehen in geeigneter Form zur Ablieferung bereit. Sie werden entweder vom Archiv abgeholt (pull: FTP Space) oder ans Archiv geschickt (push: Filetransfer o.ä.).
Nachbedingungen (erreichter innerer Zustand des Systems)	Im GEVER-System sind die abgelieferten Unterlagen als abgeliefert markiert.
Essenzielle Schritte (Ablauf im Detail)	Formale Prüfung des Übernahmeentscheids Zuordnung zu einem Angebot Primär- und Metadaten bereitstellen

Tabelle 7: Anwendungsfallbeschreibung Unterlagen abliefern

2.5.2 Aktivitätsdiagramm (Activity Diagram)



Grafik 10: Aktivitätsdiagramm Unterlagen abliefern

2.5.2.1 Übernahmeentscheid lesen

Der Aktenbildner prüft den Übernahmeentscheid formal und inhaltlich und ordnet ihn einem offenen Angebot zu.

2.5.2.2 Ablieferungsliste erstellen

Dieser Schritt umfasst die Aktualisierung der Informationen auf Seiten des Aktenbildners mit Hilfe des Übernahmeentscheids. Dabei muss der Status der abgelieferten Unterlagen im GEVER-System nachgetragen werden.

2.5.2.3 Unterlagen bereitstellen

Die Unterlagen werden aus dem eigenen GEVER-System im Ablieferungsformat exportiert.

2.5.2.4 Metadaten dazufügen

Vorgangs- und technische Metadaten werden zur Ablieferung hinzugefügt.

2.5.2.5 Ablieferung

Die Ablieferung wird an das Archiv übermittelt.

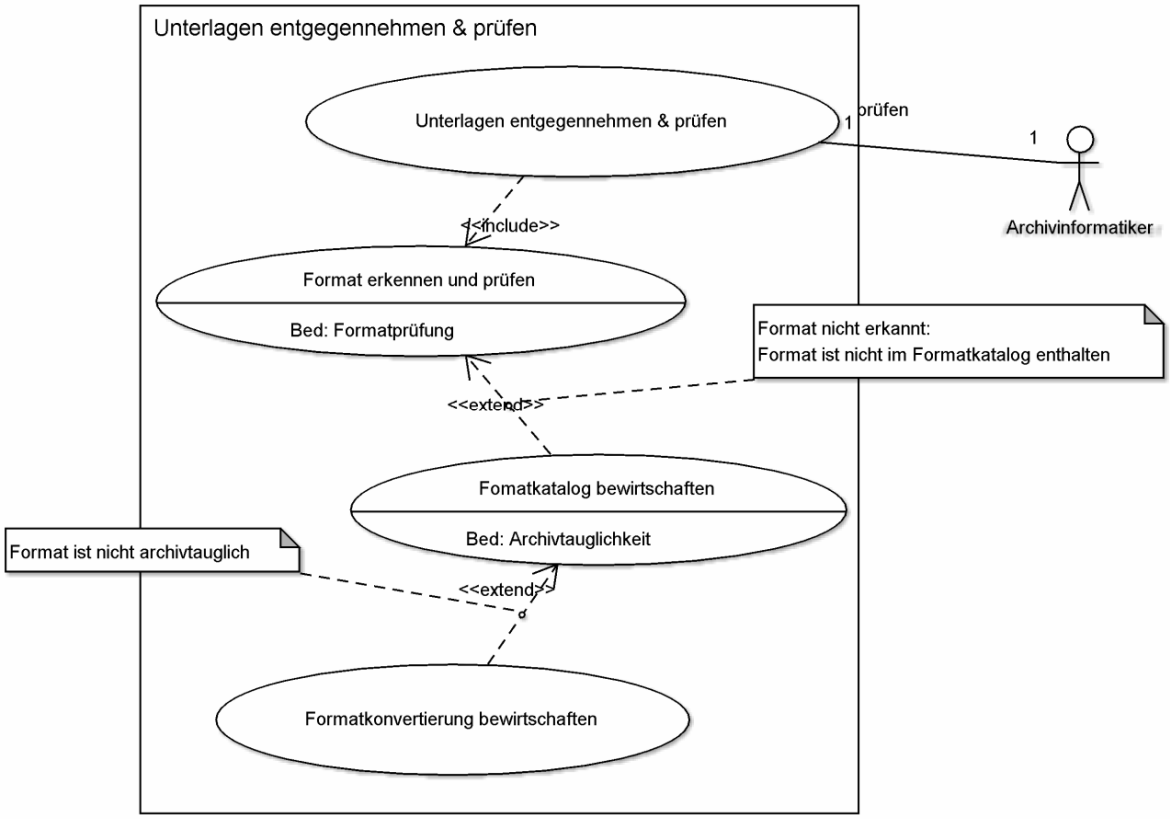
2.6 Unterlagen entgegennehmen und prüfen

2.6.1 Anwendungsfallbeschreibung (Use Case Scenario)

Name	Unterlagen entgegennehmen und prüfen
Kurzbeschreibung	Das Archiv prüft, ob alle im Übernahmeentscheid referenzierten Unterlagen in der Ablieferung enthalten sind. Eine unvollständige oder zu grosse Ablieferung wird zurückgewiesen. Anschliessend erfolgt eine formale Prüfung der gelieferten Primär- und Metadaten. Eine korrekte Ablieferung wird zur Weiterverarbeitung übernommen.
Akteure	Das Archiv
Auslöser (äusseres Ereignis, das den Anwendungsfall auslöst)	Eine Ablieferung ist im Archiv eingetroffen.
Vorbedingungen (vorausgesetzter innerer Zustand des Systems)	Die Ablieferung kann einem aktuellen Übernahmeentscheid zugeordnet werden.
Eingehende Daten (Benutzereingaben im Rahmen des Anwendungsfalls)	Die im Übernahmeentscheid verlangten Unterlagen stehen in geeigneter Form zur Ablieferung bereit.
Ergebnis(se) (was einem oder mehreren Akteuren geliefert worden ist)	Die verlangten Unterlagen sind entweder vom Archiv abgeholt (pull: FTP Space) oder ans Archiv geschickt worden (push: Filetransfer o.ä.) und stehen jetzt zur Weiterverarbeitung bereit.
Nachbedingungen (erreichter innerer Zustand des Systems)	Alle notwendigen Daten einer Ablieferung sind beim Archiv angelangt.
Essenzielle Schritte (Ablauf im Detail)	Vollständigkeit der Ablieferung prüfen Primär- und Metadaten formal prüfen

Tabelle 8: Anwendungsfallbeschreibung Unterlagen entgegennehmen und prüfen

2.6.2 Anwendungsfalldiagramm (Use Case Diagram)



Grafik 11: Use Case Diagramm Unterlagen entgegennehmen und prüfen

2.6.2.1 Unterlagen entgegennehmen und prüfen

Unterlagen entgegennehmen und prüfen ist der eigentliche, hier erläuterte Use Case. Er ruft den Use Case *Format erkennen und prüfen* auf.

2.6.2.2 Format erkennen und prüfen

Dieser Use Case wiederum wird durch *Formatkatalog bewirtschaften* erweitert, falls das zu prüfende Format unbekannt ist.

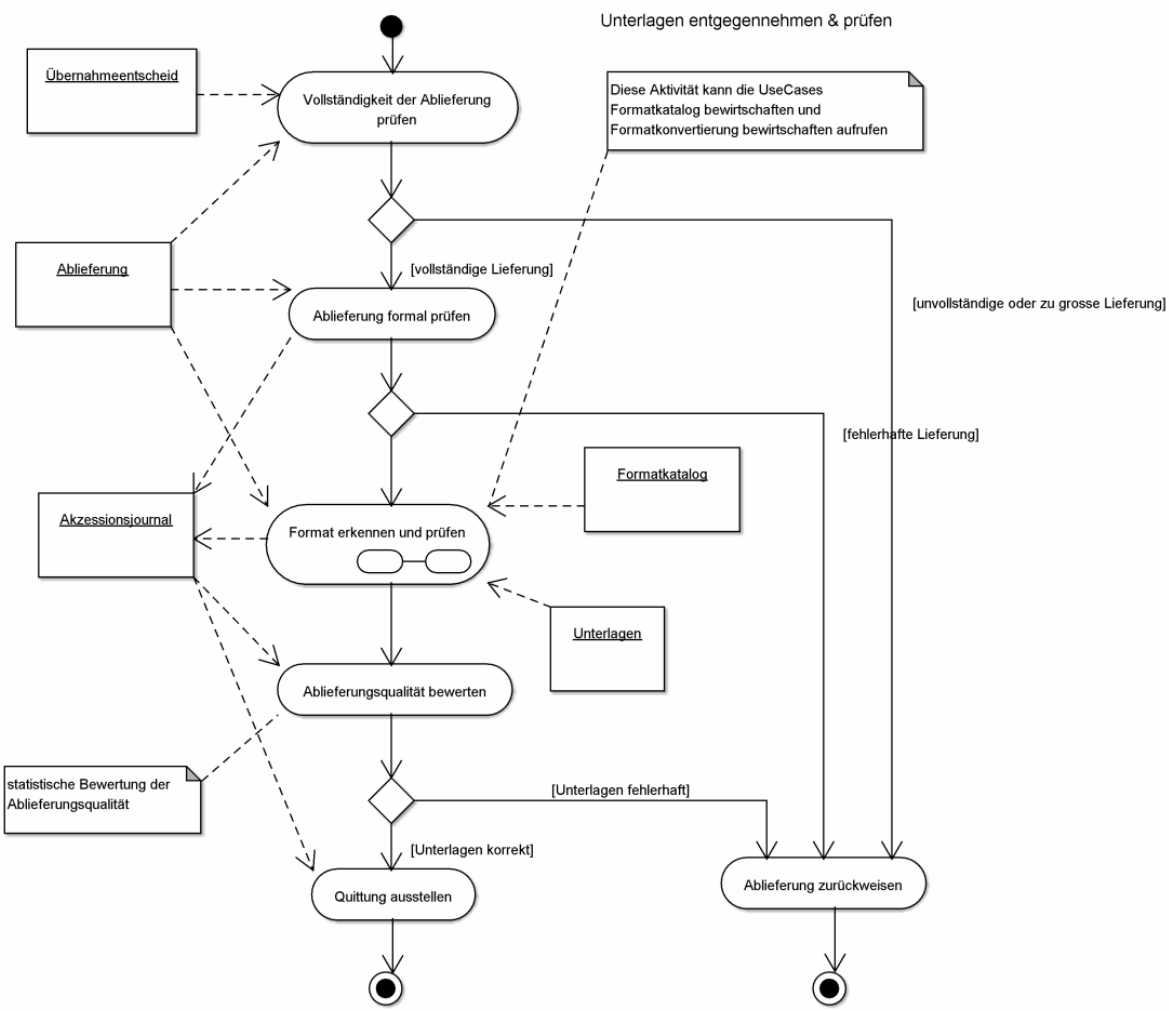
2.6.2.3 Formatkatalog bewirtschaften

Formatkatalog bewirtschaften wird durch *Formatkonvertierung bewirtschaften* erweitert, für den Fall, dass das neu aufgenommene Format nicht archivtauglich ist und ein Konvertierungspfad festgelegt werden muss.

2.6.2.4 Formatkonvertierung bewirtschaften

Hier wird für nicht archivtaugliche Formate ein Konvertierungspfad zu einem archivtauglichen Format festgelegt.

2.6.3 Aktivitätsdiagramm (Activity Diagram)



Grafik 12: Aktivitätsdiagramm Unterlagen entgegennehmen und prüfen

2.6.3.1 Vollständigkeit der Ablieferung prüfen

Die Ablieferung muss dem Übernahmeentscheid genau entsprechen.

2.6.3.2 Ablieferung formal prüfen

Die Ablieferung muss formal korrekt sein.

2.6.3.3 Format erkennen und prüfen

Für alle Primärdaten der Ablieferung wird das Format erkannt und validiert. Die nötigen Informationen dazu (Erkennungssoftware, Validierungstool, Archivtauglichkeit) finden sich im Formatkatalog. Die Resultate werden im Akzessionsjournal verzeichnet.
(Es wird darauf hingewiesen, dass eine Formatbewirtschaftung idealerweise bereits im GEVER-System erfolgen muss. Dies liegt jedoch ausserhalb des Projektfokus.)

2.6.3.4 Ablieferungsqualität bewerten

Die Informationen über die Formate der in der Ablieferung enthaltenen Dokumente erlauben eine statistische Aussage über die Qualität der Ablieferung. Es macht keinen Sinn und ist auch nicht durchführbar, eine Ablieferung zurückzuweisen, nur weil einige wenige Dokumente kein valides Format haben. Der Entscheid über die Annahme oder Rückweisung der Ablieferung muss

vielmehr auf Grund der Qualität der gesamten Ablieferung, d.h. vom Anteil der fehlerhaften Unterlagen, getroffen werden.

2.6.3.5 Quittung ausstellen

Die Quittung teilt dem Aktenbildner für sämtliche angebotenen Unterlagen mit, welche davon archiviert und welche vom Archiv nicht übernommen wurden. Er muss nun diese Unterlagen im eigenen System kassieren, d.h. physisch löschen.

2.6.3.6 Ablieferung zurückweisen

Eine unvollständige oder formal fehlerhafte Ablieferung führt auf jeden Fall zur Zurückweisung der Ablieferung. Mangelhafte Qualität der Ablieferung, beschädigte Dateien, Dateien in nicht erkennbaren oder konvertierbaren Dateiformaten in grosser Zahl, können ebenfalls zur Zurückweisung führen.

2.7 Format erkennen und prüfen

2.7.1 Anwendungsfallbeschreibung (Use Case Scenario)

Name	Format erkennen und prüfen
Kurzbeschreibung	Das Format einer Datei wird erkannt und die Datei gegen die Spezifikation des Formats validiert. Anschliessend wird die Archivtauglichkeit des Formats festgehalten. Die Resultate der Prüfungen werden im Akzessionsjournal festgehalten.
Akteure	Das Archiv
Auslöser (äusseres Ereignis, das den Anwendungsfall auslöst)	Eine Ablieferung ist formal geprüft worden.
Vorbedingungen (vorausgesetzter innerer Zustand des Systems)	Formatkatalog und Akzessionsjournal sind vorhanden.
Eingehende Daten (Benutzereingaben im Rahmen des Anwendungsfalls)	Die formal geprüfte Ablieferung gemäss Übernahmeentscheid ist im Ingest-System des Archivs vorhanden.
Ergebnis(se) (was einem oder mehreren Akteuren geliefert worden ist)	Angaben über das Format und die Validität der Datei sind ins Akzessionsjournal geschrieben worden.
Nachbedingungen (erreichter innerer Zustand des Systems)	
Essenzielle Schritte (Ablauf im Detail)	Mit Hilfe der Informationen im Formatkatalog wird das Format der Datei erkannt Mit einem geeigneten Validierungstool wird die Datei gegenüber der Formatspezifikation validiert; falls die Validierung scheitert, ist eine manuelle Abklärung nötig. Informationen über die Archivtauglichkeit des Formats werden aus dem Formatkatalog ausgelesen.

Tabelle 9: Anwendungsfallbeschreibung Format erkennen und prüfen

2.7.2 Anwendungsfalldiagramm (Use Case Diagram)

2.7.2.1 Format erkennen und prüfen

Format erkennen und prüfen ist der hier beschriebene Use Case. Dieser Use Case wird durch *Formatkatalog bewirtschaften* erweitert, falls das zu prüfende Format unbekannt ist.

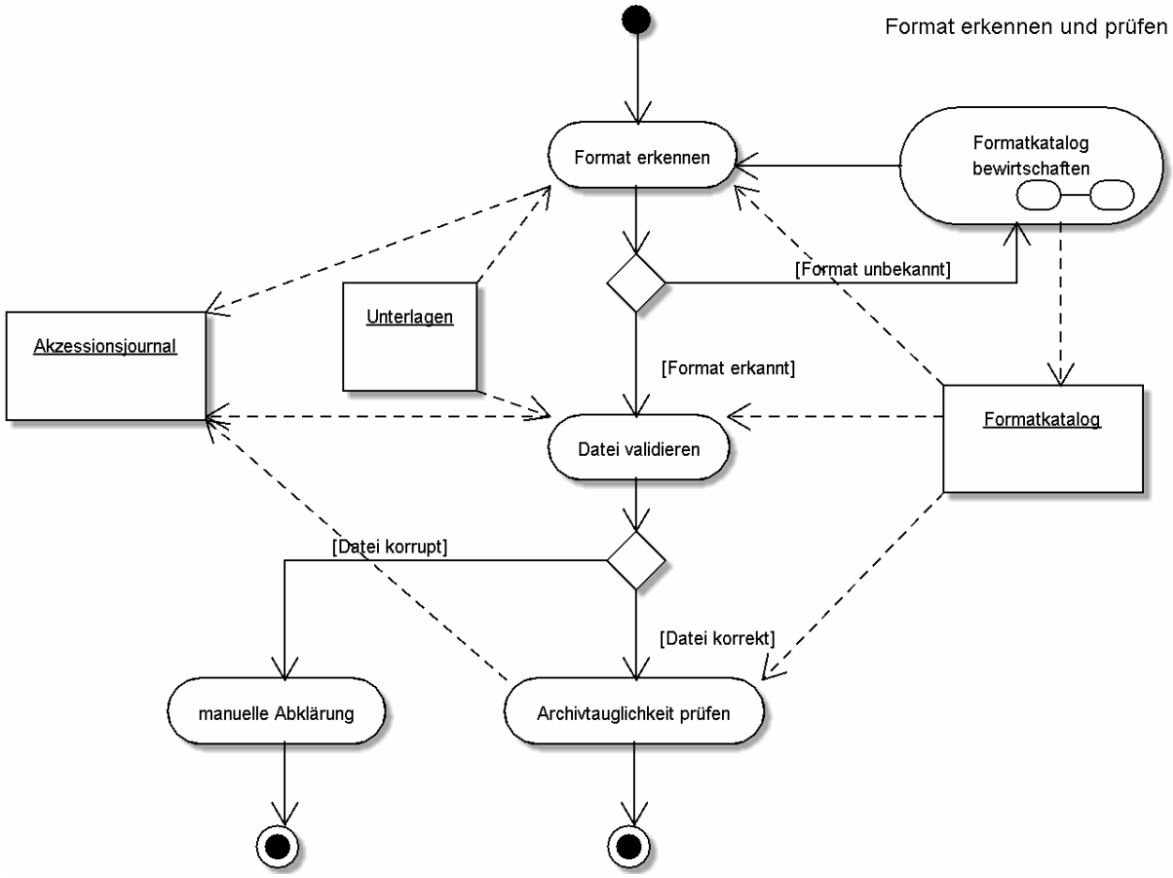
2.7.2.2 Formatkatalog bewirtschaften

Formatkatalog bewirtschaften wird durch *Formatkonvertierung bewirtschaften* erweitert, falls das neu aufgenommene Format nicht archivtauglich ist und ein Konvertierungspfad festgelegt werden muss.

2.7.2.3 Formatkonvertierung bewirtschaften

Hier wird für nicht archivtaugliche Formate ein Konvertierungspfad zu einem archivtauglichen Format festgelegt.

2.7.3 Aktivitätsdiagramm (Activity Diagram)



Grafik 13: Aktivitätsdiagramm Format erkennen und prüfen

2.7.3.1 Format erkennen

Im Formatkatalog sind Informationen verzeichnet, an welchen ein Format erkannt werden kann. Das Format wird ins Akzessionsjournal geschrieben. (Für die Formaterkennung existieren bereits verschiedene frei zugängliche Tools: FRED: A format registry demonstration, http://tom.library.upenn.edu/fred/; JHOVE - JSTOR/Harvard Object Validation Environment, http://hul.harvard.edu/jhove/; PRONOM, http://www.nationalarchives.gov.uk/pronom/, u.a.)

2.7.3.2 Datei validieren

Ebenfalls im Formatkatalog sind zu jedem Format ein oder mehrere Validierungstools angegeben. Diese lesen die ganze Datei und stellen fest, ob sie der Spezifikation des erkannten Formats entspricht. Das Resultat der Validierung wird ins Akzessionsjournal geschrieben. Falls die Datei der Formatspezifikation nicht entspricht, ist eine manuelle Abklärung erforderlich.

2.7.3.3 Archivtauglichkeit prüfen

Auf Grund der entsprechenden Angaben im Formatkatalog wird im Akzessionsjournal festgehalten, ob die Datei in einem archivtauglichen Format vorliegt.

2.8 Formatkatalog bewirtschaften

2.8.1 Anwendungsfallbeschreibung (Use Case Scenario)

Name	Formatkatalog bewirtschaften
Kurzbeschreibung	Für ein neues Format werden Name, Version, Erkennungsmethode, Validierungsprogramm, Viewer und Archivtauglichkeit ermittelt und im Formatkatalog eingetragen. Für ein bestehendes archivtaugliches Format wird der Viewer bewertet und allenfalls aufdatiert.
Akteure	Das Archiv
Auslöser (äusseres Ereignis, das den Anwendungsfall auslöst)	Bei der Formaterkennung wurde ein Datei in unbekanntem Format identifiziert.
Vorbedingungen (vorausgesetzter innerer Zustand des Systems)	
Eingehende Daten (Benutzereingaben im Rahmen des Anwendungsfalls)	Eine Datei in einem nicht im Formatkatalog verzeichneten Format liegt vor.
Ergebnis(se) (was einem oder mehreren Akteuren geliefert worden ist)	Die wichtigen Angaben zum neuen Format sind im Formatkatalog eingetragen.
Nachbedingungen (erreichter innerer Zustand des Systems)	Für alle Dateien im digitalen Magazin ist das Format vollständig beschrieben.
Essenzielle Schritte (Ablauf im Detail)	*(für ein unbekanntes Format)* Das Format wird manuell ermittelt; es wird ein Eintrag von Format und Version im Formatkatalog angelegt; dem Eintrag werden eine Formaterkennungsmethode, ein Validierungsprogramm und ein Viewer hinzugefügt; die Archivtauglichkeit des Formats wird bewertet. *(für ein bekanntes archivtaugliches Format)* Der Viewer wird bewertet und bei Bedarf ersetzt; die Archivtauglichkeit wird neu bewertet.

Tabelle 10: Anwendungsfallbeschreibung Formatkatalog bewirtschaften

2.8.2 Anwendungsfalldiagramm (Use Case Diagram)

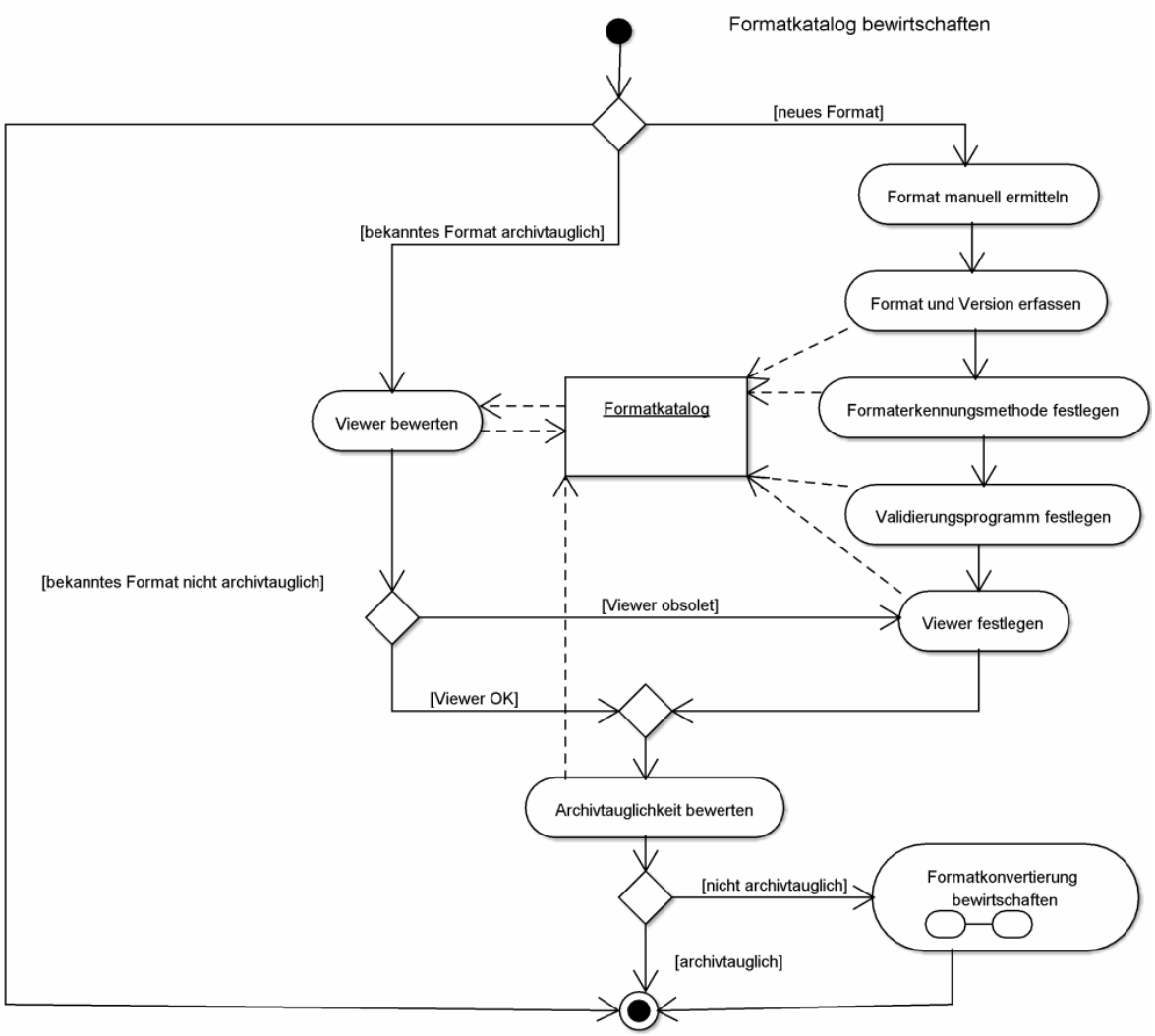
2.8.2.1 Formatkatalog bewirtschaften

Formatkatalog bewirtschaften ist der hier beschriebene Use Case. *Formatkatalog bewirtschaften* wird durch *Formatkonvertierung bewirtschaften* erweitert, für den Fall, dass das neu aufgenommene Format nicht archivtauglich ist und ein Konvertierungspfad festgelegt werden muss.

2.8.2.2 Formatkonvertierung bewirtschaften

Hier wird für nicht archivtaugliche Formate ein Konvertierungspfad zu einem archivtauglichen Format festgelegt.

2.8.3 Aktivitätsdiagramm (Activity Diagram)



Grafik 14: Aktivitätsdiagramm Formatkatalog bewirtschaften

2.8.3.1 Format manuell ermitteln

Wenn eine Datei in einem noch nicht registrierten Format abgeliefert wird, muss dieses im Formatkatalog verzeichnet werden. Dazu ist zuerst das Format zu ermitteln, im Normalfall auf der Basis bestehender Informationen aus der abliefernden Stelle.

2.8.3.2 Format und Version erfassen

Format und Version werden im Formatkatalog eingetragen.

2.8.3.3 Formaterkennungsmethode festlegen

Im Formatkatalog wird festgehalten, wie (d.h. an Hand welcher Merkmale) das Format erkannt werden kann.

2.8.3.4 Validierungsprogramm festlegen

Es wird ein geeignetes Programm identifiziert, welches eine Datei gegen die Formatspezifikation validieren kann.

2.8.3.5 Viewer festlegen / Viewer bewerten

Im Formatkatalog wird festgehalten, welches Programm zur Wiedergabe des Formates (in der entsprechenden Version) verwendet werden soll.

2.8.3.6 Archivtauglichkeit bewerten

Auf der Basis von Informationen wie dem *Katalog archivischer Dateiformate KaD* wird bewertet, ob das Format archivtauglich ist. Auch dieses Merkmal wird im Formatkatalog festgehalten. Falls das Format nicht archivtauglich ist, wird der Use Case *Formatkonvertierung bewirtschaften* angestossen.

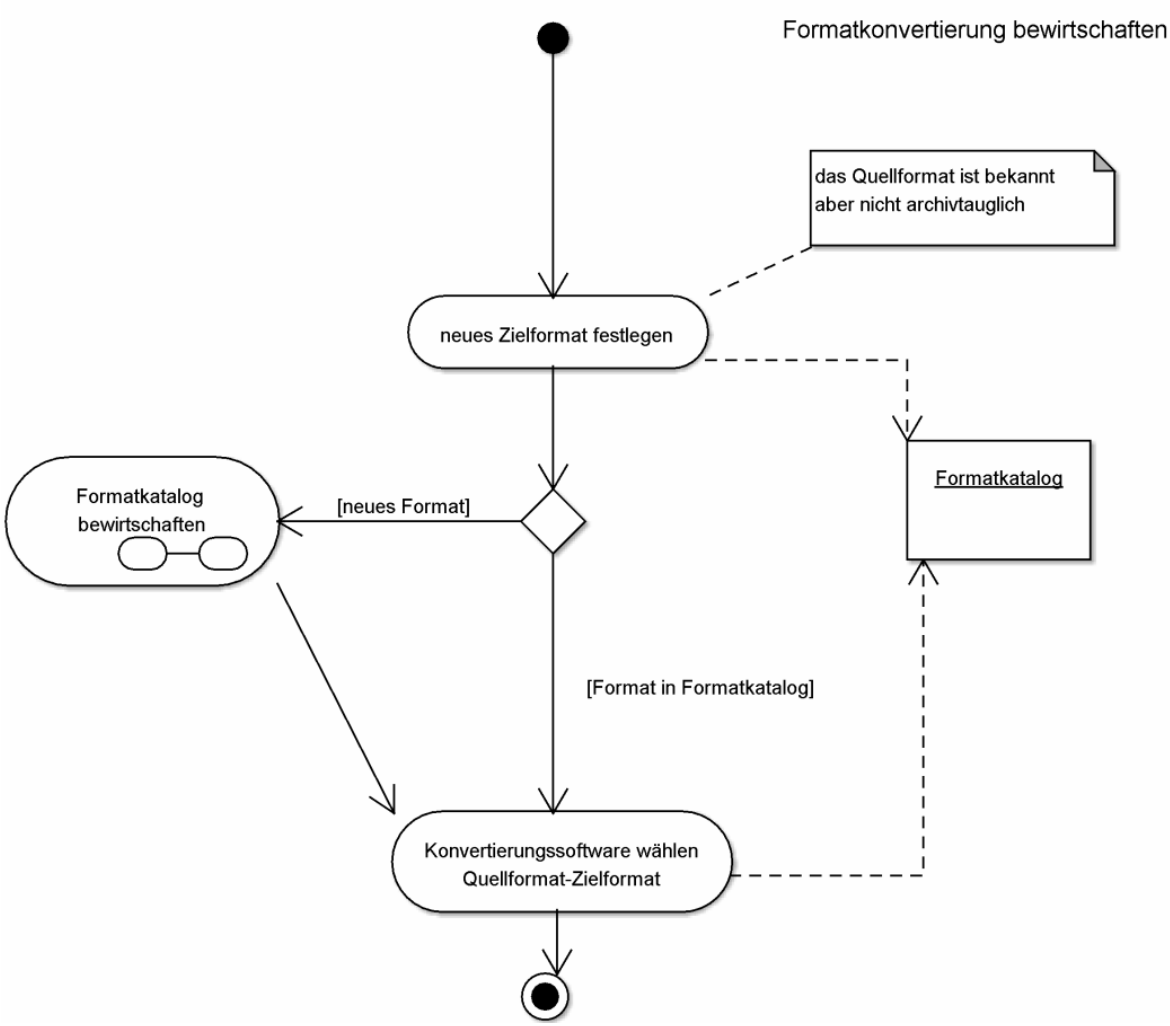
2.9 Formatkonvertierung bewirtschaften

2.9.1 Anwendungsfallbeschreibung (Use Case Scenario)

Name	Formatkonvertierung bewirtschaften
Kurzbeschreibung	Für ein nicht archivtaugliches Format wird ein Konvertierungspfad festgelegt.
Akteure	Das Archiv
Auslöser (äusseres Ereignis, das den Anwendungsfall auslöst)	Beim Bewirtschaften des Formatkatalogs wird ein neues oder bekanntes Format als nicht archivtauglich identifiziert.
Vorbedingungen (vorausgesetzter innerer Zustand des Systems)	
Eingehende Daten (Benutzereingaben im Rahmen des Anwendungsfalls)	Ein als nicht archivtauglich bewertetes Format
Ergebnis(se) (was einem oder mehreren Akteuren geliefert worden ist)	Ein neuer Konvertierungspfad wurde im Formatkatalog eingetragen.
Nachbedingungen (erreichter innerer Zustand des Systems)	Für alle im Formatkatalog verzeichneten nicht archivtauglichen Formate existiert ein Konvertierungspfad.
Essenzielle Schritte (Ablauf im Detail)	Für das nicht mehr archivtaugliche Quellformat wird ein Zielformat festgelegt. Falls das Zielformat noch nicht im Formatkatalog eingetragen ist, wird dies nachgeholt. Eine Konvertierungssoftware wird ausgewählt, die das Quellformat in das Zielformat überführt.

Tabelle 11: Anwendungsfallbeschreibung Formatkonvertierung bewirtschaften

2.9.2 Aktivitätsdiagramm (Activity Diagram)



Grafik 15: Aktivitätsdiagramm Formatkonvertierung bewirtschaften

2.9.2.1 Neues Zielformat festlegen

Es wird ein archivtaugliches Format festgelegt, in welches das nicht (mehr) archivtaugliche Quellformat konvertiert werden kann. (Anmerkung: Es ist auch denkbar, dass kein archivtaugliches Zielformat existiert, in welches das Quellformat direkt konvertiert werden kann. In diesem Fall muss der Umweg über ein Zwischenformat gewählt werden.)
Das Quell-/Zielformat-Paar wird im Formatkatalog eingetragen.
Falls das Zielformat noch nicht im Formatkatalog beschrieben ist, muss dies durch den Prozess *Formatkatalog bewirtschaften* durchgeführt werden.

2.9.2.2 Konvertierungssoftware wählen

Eine geeignete Software, die das Quellformat ins Zielformat konvertieren kann, wird ausgewählt und im Formatkatalog festgehalten. (Ein Beispiel für eine Formatkonvertierungssoftware ist "Conversion and Recommendation of Digital Object Formats (CRiB)" der Universität Minho, http://crib.dsi.uminho.pt/.)

2.10 Unterlagen ordnen und verzeichnen

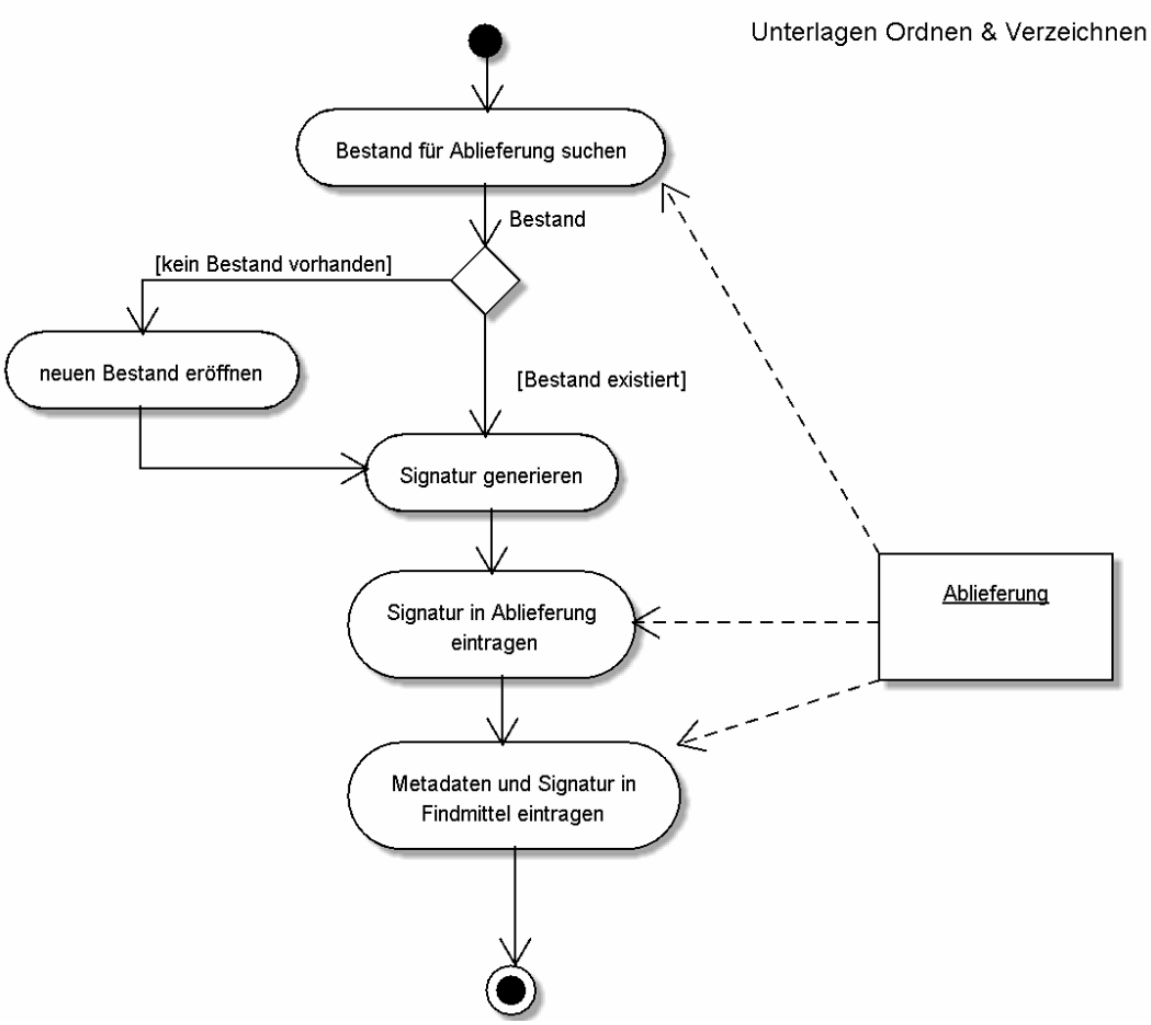
2.10.1 Anwendungsfallbeschreibung (Use Case Scenario)

Name	Unterlagen ordnen und verzeichnen
Kurzbeschreibung	Aus den Unterlagen einer *Ablieferung* wird ein *Bestand* gebildet. Die Bestandbildung, eine logische Gruppierung der Ablieferung nach *Amt* und *Ablage*, ist Grundlage beim Eintragen der Ablieferung ins Findmittel und steuert die Übernahme der GEVER-Metadaten ins Findmittel.
Akteure	Der Archivar bedient das System im Einzelschritt- oder Batch-Modus.
Auslöser (äusseres Ereignis, das den Anwendungsfall auslöst)	Eintreffen einer Ablieferung
Vorbedingungen (vorausgesetzter innerer Zustand des Systems)	Eine Prüfung auf formale Korrektheit der Daten war erfolgreich.
Eingehende Daten (Benutzereingaben im Rahmen des Anwendungsfalls)	Eine Menge von Primär- und zugehörigen Metadaten bilden zusammen eine Klasse, das ist ein *Bestand*.
Ergebnis(se) (was einem oder mehreren Akteuren geliefert worden ist)	Dossiersignatur und Dokumentsignatur, GEVER-Metadaten und technische Metadaten sind im Findmittel eingetragen.
Nachbedingungen (erreichter innerer Zustand des Systems)	
Essenzielle Schritte (Ablauf im Detail)	Die Menge der Ablieferung ist in einen hierarchischen Baum von Verzeichnungseinheiten gegliedert. Ein für den Eintrag der Ablieferung notwendiger Bestand ist im Findmittel bekannt. Zu jeder Verzeichungseinheit ist eine Signatur generiert und im Findmitteleingetragen. Metadaten sind im Findmittel eingetragen.

Tabelle 12: Anwendungsfallbeschreibung Unterlagen ordnen und verzeichnen

Bestand ist hier die Menge der *Ablieferungen* aus einem Amt/Departement o.ä, siehe Glossar AUGev.

2.10.2 Aktivitätsdiagramm (Activity Diagram)



Grafik 16: Aktivitätsdiagramm Unterlagen ordnen und verzeichnen

2.10.2.1 Bestand für Ablieferung suchen

Für die Ablieferung wird im Findmittel (automatisch) ein entsprechender Bestand gesucht. Für das Einbinden der Ablieferung in den Bestand gibt es in den Archiven zwei Methoden, die beide unterstützt werden sollen.
Im einen Fall wird die Ablieferung unverändert mit der darunter hängenden Struktur der Ablage im Amt angehängt. Ein Bestand besteht dann aus einer Menge von in sich strukturierten Ablieferungen (z.B. Schweizerisches Bundesarchiv).
Im zweiten Fall werden mehrere Ablieferungen in einem Bestand aggregiert, das heisst das Ordnungssystem der Ablieferung wird auf die Archivtektonik des Archivs gemappt, solange eine hinreichende Übereinstimmung besteht, und die Dossiers der neuen Ablieferung am entsprechenden Ort versorgt. Ändert sich das Ordnungssystem im Amt fundamental, wird im Archiv ein neuer Bestand eröffnet (z.B. Staatsarchiv Basel-Stadt).
Siehe zu diesem Mapping weiter unten Grafik 17: Prototyp Ordnen und Verzeichnen.

2.10.2.2 Neuen Bestand eröffnen

Existiert in der Archivtektonik noch kein entsprechender Eintrag, muss im Findmittel in der Archivtektonik ein neuer Bestand eröffnet werden.

2.10.2.3 Signatur generieren

Die Information Bestand, Ablieferung, Serie, Dossier mit fortlaufender Nummer und Dokument mit fortlaufender Nummer generiert die Signatur für jedes gelieferte Dossier bzw. jedes Dokument:

Archiv:[Abteilung]:[Bestand]:[Ablieferung]:[Serie]:Dossier-Nr.:Dokument-Nr.

Siehe dazu im Detail Kapitel 3.2, Aktenplan-Archivtektonik-Signatur.

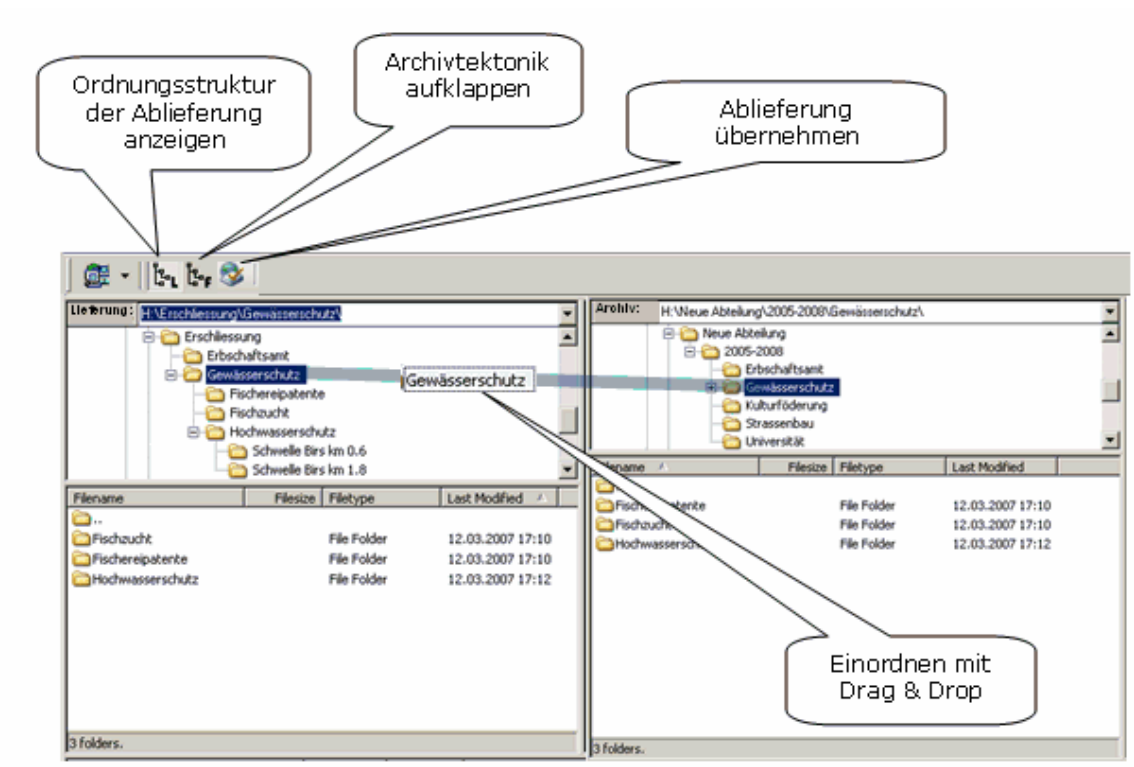
2.10.2.4 Signatur in Ablieferung eintragen

Die Signatur wird in die Ablieferung eingetragen. Die Signatur (die GUID von Dossier und Dokument) erlaubt den Zugriff im Benutzungsfall.

2.10.2.5 Metadaten und Signatur in Findmittel eintragen

Sämtliche Metadaten, inklusive die eben eingefügte Signatur, können nun in einem Vorgang ins Findmittel eingetragen werden. Dies geschieht automatisiert mit Hilfe des entsprechenden Importassistenten.

2.10.3 Prototyp (Use Case Prototype)



Grafik 17: Prototyp Ordnen und Verzeichnen

2.11 Unterlagen magazinieren

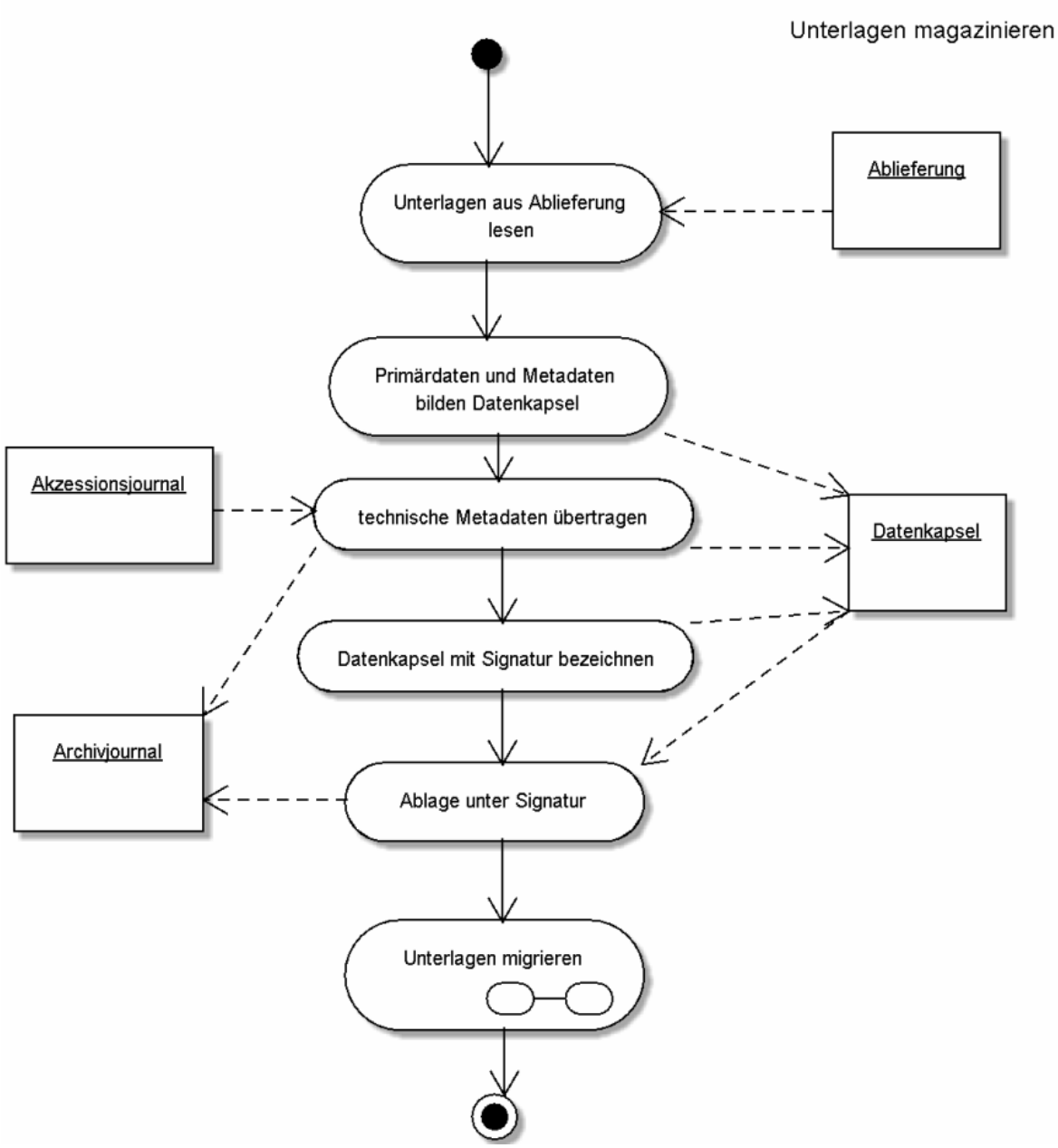
2.11.1 Anwendungsfallbeschreibung (Use Case Scenario)

Name	Unterlagen magazinieren
Kurzbeschreibung	Unterlagen magazinieren bedeutet das Ablegen der entgegengenommenen digitalen Primär- und Metadaten in ein Repositorium für digitale Daten. Das Repositorium, im folgenden digitales Magazin genannt, muss gewisse Anforderungen (functional & non-functional requirements) erfüllen.

Akteure	Der Archivar bedient das System im Einzelschritt- oder Batch-Modus.
Auslöser (äusseres Ereignis, das den Anwendungsfall auslöst)	Eintreffen von digitalen Primär- und Metadaten
Vorbedingungen (vorausgesetzter innerer Zustand des Systems)	Ordnen und Verzeichnen hat stattgefunden Signatur für das Objekt ist definiert Eine Prüfung auf formale Korrektheit der Daten war erfolgreich
Eingehende Daten (Benutzereingaben im Rahmen des Anwendungsfalls)	Primär- und zugehörige Metadaten bilden zusammen ein Datenobjekt.
Ergebnis(se) (was einem oder mehreren Akteuren geliefert worden ist)	Das Datenobjekt ist persistent auf einem Speichermedium abgelegt.
Nachbedingungen (erreichter innerer Zustand des Systems)	Die ursprünglichen Primär- und Metadaten können nicht mehr verändert werden.
Essenzielle Schritte (Ablauf im Detail)	Aus Primär- und Metadaten wird eine Datenkapsel gebildet. Die Datenkapsel wird mit der Signatur bezeichnet. Die Datenkapsel wird mit der Signatur als GUID im digitalen Magazin abgelegt.

Tabelle 13: Anwendungsfallbeschreibung Unterlagen magazinieren

2.11.2 Aktivitätsdiagramm (Activity Diagram)



Grafik 18: Aktivitätsdiagramm Unterlagen magazinieren

2.11.2.1 Unterlagen aus Ablieferung lesen

Der Gegenstand der Archivierung wird aus der gesamten Ablieferung extrahiert.

2.11.2.2 Primär- und Metadaten bilden Datenkapsel

Aus Primär- und Metadaten wird eine Datenkapsel gebildet. Metadaten sind dabei:

- Die gelieferten GEVER-Metadaten
- Generierte technische Metadaten aus der formalen Prüfung im Use Case *Unterlagen übernehmen*
- Die Migrationsgeschichte der Datenkapsel
- Die Packliste, die uns den Inhalt der Datenkapsel erschliesst

- Die Signatur

2.11.2.3 Technische Metadaten übertragen

Technische Metadaten zu den einzelnen Dokumenten/Dateien (Dateiformat etc.) können vom Use Case *Format erkennen und prüfen* übernommen werden. *Format erkennen und prüfen* schreibt diese Informationen für jede Datei ins Akzessionsjournal. Jetzt werden die Informationen aus dem Akzessionsjournal in das Archivjournal und die Datenkapsel übertragen.

2.11.2.4 Datenkapsel mit Signatur bezeichnen

Die Datenkapsel wird mit der im Use Case *Unterlagen ordnen & verzeichnen* generierten Signatur gekennzeichnet.

2.11.2.5 Ablage unter Signatur

Die Datenkapsel wird mit der Signatur als GUID im digitalen Magazin abgelegt:

Archiv:[Abteilung]:[Bestand]:[Ablieferung]:[Serie]:Dossier-Nr.:Dokument-Nr.

2.11.2.6 Unterlagen migrieren

Möglicherweise befinden sich in der magazinierten Datenkapsel Dateien in Dateiformaten, die nicht als archivwürdig erachtet werden. Deshalb wird hier zum Abschluss der Use Case Unterlagen migrieren aufgerufen.

2.11.3 Anforderungen (Requirements)

2.11.3.1 Funktionale Anforderungen

- Das Magazin muss Dateien speichern und unverändert wieder zurückliefern.
- Die Dateien müssen eindeutige und unveränderbare Schlüssel zugewiesen erhalten (GUID).
- Alle Zugriffe müssen protokolliert werden.
- Veränderung an Struktur und Daten müssen festgehalten werden.

2.11.3.2 Nicht funktionale Anforderungen

- Das System muss skalierbar (veränderbar) sein.
- Speicherplatz soll günstig verfügbar sein.
- Die Datenstruktur muss transparent und vom Produkt unabhängig sein, d.h. es muss möglich sein, Hardware/Software zu wechseln, ohne die logische Speicherstruktur zu ändern.
- Inhalt und Speicherungslogik müssen getrennt sein; d.h. die Speicherungslogik darf keine impliziten Inhaltsinformationen beinhalten.
- Die Spezifikationen müssen exakt erfüllt, aber nicht übererfüllt sein.

2.12 Unterlagen erhalten

2.12.1 Anwendungsfallbeschreibung (Use Case Scenario)

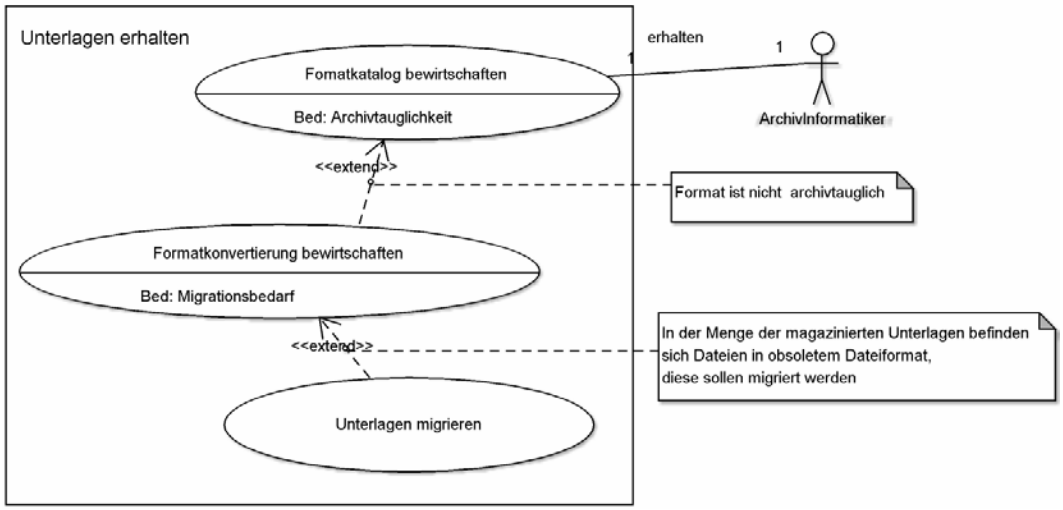
Name	Unterlagen erhalten
Kurzbeschreibung	Ist es aus technischen Gründen in absehbarer Zeit nicht mehr möglich, die inhaltliche Struktur der Daten zu lesen und/oder deren ursprüngliche Erscheinungsform wieder herzustellen, müssen die Daten in ein neues Dateiformat migriert werden.
Akteure	Archivinformatiker
Auslöser (äusseres Ereignis, das den Anwendungsfall auslöst)	Externe Erkenntnis: Ein Dateiformat wird in absehbarer Zeit nicht mehr lesbar sein.
Vorbedingungen (vorausgesetzter innerer	Die datenspeichernde Infrastruktur ist jederzeit funktionsfähig, das heisst, der Bitstrom ist unverändert lesbar.

Zustand des Systems)	Ein Migrationspfad für obsolete Dateiformate ist bekannt.
Eingehende Daten (Benutzereingaben im Rahmen des Anwendungsfalls)	Auslösen der Migration Allenfalls Simulation einer Migration ohne Schreiben der migrierten Dateien
Ergebnis(se) (was einem oder mehreren Akteuren geliefert worden ist)	Bestätigung der Formatmigration für eine Liste von Dateien (Anzahl, Stichproben)
Nachbedingungen (erreichter innerer Zustand des Systems)	Alle Dateien sind in einem archivisch akzeptablen Format, wenn auch allenfalls nicht in der aktuellsten Version.
Essenzielle Schritte (Ablauf im Detail)	Migrationsbedarf für eine bestimmte Primärdatei festlegen Datei nach Vorgabe Migrationspfad in das neue Format konvertieren Datenkapsel mit Originaldatei und migrierter Datei erneut archivieren

Tabelle 14: Anwendungsfallbeschreibung Unterlagen erhalten

2.12.2 Anwendungsfalldiagramm (Use Case Diagram)

Der Anwendungsfall *Unterlagen erhalten* lässt sich weiter aufgliedern in den eigentlichen Anwendungsfall *Unterlagen migrieren*, der durch Formatmigration die Verfügbarkeit der Daten in einem archivtauglichen Dateiformat gewährleistet und die beiden Bewirtschaftungsfälle *Formatkatalog bewirtschaften* und *Formatkonvertierung bewirtschaften*. Im Use Case *Formatkatalog bewirtschaften* wird die Notwendigkeit zur Migration erkannt und festgehalten. Anschliessend, bei Bedarf einer Migration, wird in *Formatkonvertierung bewirtschaften* ein Migrationspfad für das obsolete Format festgelegt. Auf dieser Basis kann die eigentliche Migration der Unterlagen erfolgen.



Grafik 19: Use Case Diagramm Unterlagen erhalten

2.12.2.1 Formatkatalog bewirtschaften

Formatkatalog bewirtschaften wird durch *Formatkonvertierung bewirtschaften* erweitert, für den Fall, dass ein Format als nicht mehr archivtauglich erkannt worden ist (externer Input).

2.12.2.2 Formatkonvertierung bewirtschaften

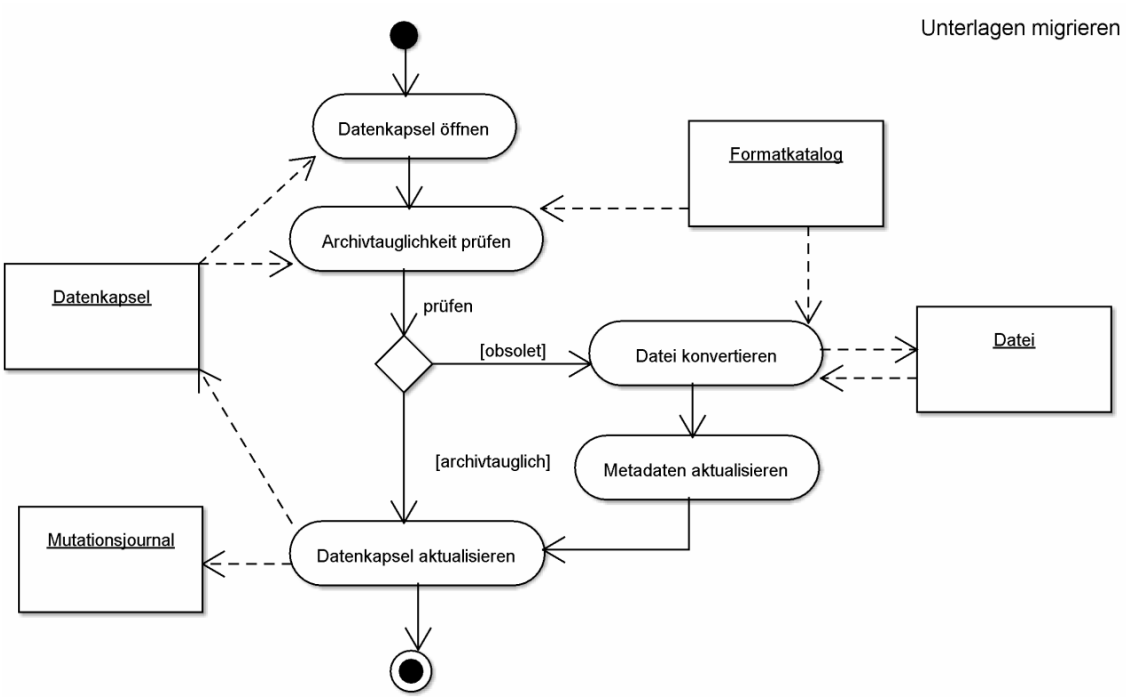
Hier wird für nicht archivtaugliche Formate ein Konvertierungspfad zu einem archivtauglichen Format festgelegt.

2.12.2.3 Unterlagen migrieren

Unterlagen migrieren ist der hier beschriebene Use Case. Dieser Use Case erweitert den Use Case *Formatkonvertierung bewirtschaften*. Für jeden neu definierten Konvertierungspfad müssen alle betroffenen Dateien im digitalen Magazin migriert werden.

2.12.3 Aktivitätsdiagramm (Activity Diagram)

Vorbedingung: Die datenspeichernde Infrastruktur ist jederzeit funktionsfähig, das heisst, der Bitstrom ist unverändert lesbar. Ziel ist es, dass die inhaltliche Struktur der Daten lesbar bleibt und deren ursprüngliche Erscheinungsform wieder herstellbar ist.



Grafik 20: Aktivitätsdiagramm Unterlagen migrieren

2.12.3.1 Unterlagen migrieren

Die Aktivität "Unterlagen migrieren" setzt eine Infrastruktur voraus, die durch den Use Case "Formaterkennung bewirtschaften" bereitgestellt wird. Diese Formaterkennungsstruktur, eine Tabelle, ist eigentlich archivunabhängig und kann von einer zentralen Stelle zur Verfügung gestellt werden. ***Archivtauglich* bedeutet, dass das Format der Dateiformatpolicy des jeweiligen Archivs oder einer gesamtschweizerischen Archivpolicy entspricht.** Ein archivtaugliches Format wird dann als *obsolet* bezeichnet, wenn eine Konvertierung in dieses Format nicht mehr empfehlenswert ist. Bereits vorhandene Dateien in diesem Format werden aber nicht konvertiert, solange das Format noch als archivtauglich klassifiziert ist (siehe oben).

Format	Version	archivtauglich	obsolet	Viewer
MSWord	2003	nein		Word-Viewer-2003
	97	nein		Word-Viewer-97/2000
	95	nein		Word-Viewer-97/2000
	6.0	nein		Word-Viewer-97/2000
PDF	1.6	ja	nein	PDF Reader 8.0
PDF	1.5	ja	ja	PDF Reader 6.0
PDF	1.4	ja	ja	PDF Reader 6.0
GIF	89a	nein		FireFox 1.5
...				

Tabelle 15: Formaterkennungsstruktur

Die Menge der zu migrierenden Dateien ist die Menge der Dateien im digitalen Magazin, welche in nicht mehr archivtauglichem Format sind. Die betroffenen Datenkapseln können aus dem Archivjournal, wo u.a. das Dateiformat für jede Datei im Magazin festgehalten ist, ermittelt werden.

2.12.3.2 Datenkapsel öffnen

Eine Datenkapsel wird aus dem digitalen Magazin gelesen.

2.12.3.3 Archivtauglichkeit prüfen

Die Archivtauglichkeit der in der Datenkapsel vorhandenen Dateien wird geprüft.

2.12.3.4 Datei konvertieren

Die eigentliche Formatkonvertierung benötigt Quell- und Zielformat und ein Konvertierungswerkzeug. Der Use Case *Formatkonvertierung bewirtschaften* stellt das Notwendige zur Verfügung.

Ursprungsformat	Archivformat	Konvertierungswerkzeug
MSWord 97	PDF/A	Acrobat PDF Maker
MSWord 95	PDF/A	Adobe PDF Printer
GIF	PNG	gif2png
...		

Tabelle 16: Konvertierungspfade

2.12.3.5 Metadaten aktualisieren

Die Migrationsgeschichte wird in den technischen Metadaten festgehalten.

2.12.3.6 Datenkapsel aktualisieren

Die aktualisierte Datenkapsel wird ins digitale Magazin zurückgeschrieben. Die Aktion mit der Datenkapsel wird im Archivjournal festgehalten.

2.12.4 Anforderungen (Requirements)

Langzeitarchivierung digitaler Daten beinhaltet nicht nur die stabile Aufbewahrung einmal archivierter Bitströme, sondern ebenso wichtig ist, dass die zum Verständnis dieser Bitströme notwendigen syntaktischen und semantischen Informationen mit erhalten werden. Ist es in absehbarer Zeit nicht mehr möglich, die inhaltliche Struktur der Daten zu lesen und deren ursprüngliche Erscheinungsform wieder herzustellen, müssen die Daten in ein neues Dateiformat migriert werden. Erhalten bedeutet also auch und vor allem die Migration obsoleter Formate. Zum Thema Formatmigration erarbeiten wir folgenden Policy-Vorschlag:

- Die eingelieferte Originaldatei wird in jedem Fall aufbewahrt.
- Ist die Originaldatei nicht in einem archivisch akzeptablen Format, wird sie in ein adäquates archivisches Format in der aktuellsten Version konvertiert.
- Lieferungen in archivisch akzeptablem Format in älteren Versionen werden nicht konvertiert.
- Archivische Formate werden erst migriert, wenn in absehbarer Zeit kein Viewer mehr für das Format zur Verfügung stehen wird.

Die Formatmigration beinhaltet immer ein gewisses Risiko von Informationsverlust. Das an und für sich richtige Ziel, die Anzahl der im Archiv vertretenen Formate zu reduzieren, muss also im vorarchivischen Bereich durch Absprache mit den aktenführenden Stellen angestrebt werden. Formatmigrationen in Archiv, nur um die Anzahl archivtauglicher Formate klein zu halten oder um alle Dateien vom gleichen Typ (z.B. PDF) in der gleichen Version zu halten, sind nicht zu befürworten.

Formatmigration bzw. Bestandeserhaltung bezieht sich hier in unseren Betrachtungen auf Primärdaten. Migration der XML-Metadaten würde eine Migration des digitalen Magazins als gesamtes bedeuten, denn davon wäre nicht nur der XML-Viewer betroffen, sondern es müssten alle Prozesse, welche im digitalen Magazin Metadaten lesen und schreiben, angepasst werden.

3 Datenmodellierung

3.1 Gegenstand der Archivierung

In den einzelnen Phasen des Archivierungsprozesses tritt jeweils eine andere Sicht auf die Objekte der Archivierung in den Vordergrund. Wir unterteilen für diese Analyse grob in drei Phasen: Ablieferung, Speicherung und Benutzung.

3.1.1 Ablieferung

Im Falle der **Ablieferung** bzw. **Übernahme** fokussieren wir auf das Ablieferungsobjekt. Das Ablieferungsobjekt ist die Menge der übernommenen Unterlagen zusammen mit dem Ordnungssystem (OS) der Provenienz im Zeitpunkt der Ablieferung. Eine Ablieferung können wir in diesem Falle als SIP bezeichnen. Damit ist also die gesamte archivwürdige Ablage des Aktenbildners gemeint. Das Archiv selber wird damit zu einer Sammlung von Ablieferungen.

3.1.2 Speicherung

Bei der **Speicherung** zerlegen wir das SIP. Aus dem SIP entsteht durch umformen (denormalisieren) auf die Ebene Verzeichnungseinheit das AIP. Das AIP umfasst folglich die Primär- und Metadaten einer Verzeichnungseinheit, die zugehörigen Ablageinformationen und die technischen Metadaten. Es stellt sich die Frage, ob die beim Ingest-Prozess gewonnenen technischen Metadaten ebenfalls im AIP gespeichert werden sollen. Die technischen Metadaten sind gewissermassen dem AIP inhärent, das heisst, sie können jederzeit wieder generiert werden. Wir müssen sie aber aus Effizienzgründen dennoch dort speichern.

3.1.3 Benutzung

Wir können vereinfachend zwei Arten der **Benutzung** unterscheiden: einerseits die Suche nach einzelnen Informationen (Primärdaten), und anderseits die Recherche, die eher auf der Ebene und im Zusammenhang der Dossiers angesiedelt ist. Heute erfolgt die Erschliessung je nach Dossiertyp in unterschiedliche Tiefe. Ausschlaggebend für die Erschliessungstiefe ist aber weniger Provenienz- oder Dossiertyp als vielmehr die Ressourcen, die dem Archiv zum Ordnen und Verzeichnen zur Verfügung stehen. Das Ordnen und Verzeichnen bei der Übernahme digitaler Archivalien sollte grundsätzlich automatisiert stattfinden, so dass ein Verzeichnen bis auf Dokumentebene nicht mit einem Mehraufwand verbunden ist. Zusammenfassend ergibt sich folgendes Bild:

Ablieferung -> **SIP** (die Ablieferung ist die Menge der übernommenen Unterlagen zusammen mit dem Ordnungssystem der Provenienz)

Speicherung -> **AIP** (das Ablieferungsobjekt heruntergebrochen auf Ebene Verzeichungseinheit)

Benutzung -> **DIP** (das einzelne Dokument mit Kontextinformationen aus dem Dossier)

3.2 Aktenplan - Archivtektonik - Signatur

Damit Unterlagen aus Geschäftsverwaltungssystemen geordnet übernommen werden können, müssen gewisse Anforderungen erfüllt sein. Wir stellen dazu einen kurzen Anforderungskatalog zusammen:

Die aktenbildende Stelle muss für ihre Aktenführung einen *Aktenplan* oder ein *Ordnungssystem* festlegen, d.h. die Unterlagen müssen strukturiert nach Ablage und Rubriken in Dossiers abgelegt sein.

Aufbewahrungsfrist muss hinterlegt sein. Die Aufbewahrungsfrist kann im Aktenplan, pro Dossier oder pro Dokumenttyp hinterlegt sein. Ebenfalls muss zu gegebener Zeit das Dossier abgeschlossen werden.

Klassifizierung muss erfolgt sein (archivwürdig oder kassieren nach Ablauf der Aufbewahrungsfrist im Amt). Die Klassifizierung kann ebenfalls auf Ebene Aktenplan, Dossier oder Dokumenttyp erfolgen.

Siehe dazu z. B.: Weisungen über die Aktenführung in der Bundesverwaltung vom 13. Juli 1999

Wenn dieser Anforderungskatalog von der Provenienz erfüllt wird, ist die Ablieferung eine strukturierte und klassifizierte Menge von Unterlagen. Die Struktur, das Ordnungssystem (OS), ist eine Baumstruktur mit der Wurzel *Ablage*, mehreren Zweigen vom Typ *Ordnungssystemposition* oder *Aufgabe*, welche in den Blättern *Dossiers* enden. Die Dossiers enthalten die Dokumente.

Eine solche strukturierte Menge von Unterlagen bildet nach dem Ablauf der Aufbewahrungsfrist und bei entsprechender Klassifizierung eine *Ablieferung* ans Archiv. Die Ablieferung wird mit dieser Struktur in die Archivtektonik, in den dieser Provenienz zugewiesenen Bestand, übernommen.

Ein Objekt im Archiv ist durch seine *Signatur* identifiziert und zitierbar. Bei den meisten Archiven widerspiegelt die Signatur in sprechender Syntax die Ablagestruktur im Archiv, mithin die Archivtektonik. Wir können bei der Archivierung aus Geschäftsverwaltungssystemen folgendermassen für die Signatur verallgemeinern:

Ebene	Sicht Amt	Sicht Archiv
Archivtektonik	Provenienz	Archiv
		[Bestand]
		[Ablieferung]
Aktenplan Geschäftsverwaltung	[OS-Position]	[Serie]
	Dossier-Nr.	Dossier-Nr.
	Dokument-Nr.	Dokument-Nr.

[…] sind fakultative Verzeichnungsebenen bei der Signaturbildung
Provenienz/Bestand und Ordnungssystemposition/Serie können rekursiv definiert werden

Grafik 21: Signaturstruktur

Eine minimale Signatur besteht also aus einem *Archivkennzeichen*, der *Dossiernummer* und der *Dokumentnummer*. Das Archivobjekt ist damit hinreichend qualifiziert. Ausführlichere Signatursysteme bilden eine eigentlich semantische GUID nach Ablageort oder Aufgabe der Provenienz. Die konkrete Signaturbildung im Archiv orientiert sich an jeweils andern Gepflogenheiten.

3.3 Modellierung

Die Objektmodellierung kann nun einerseits in Form von UML-Objektdiagrammen bzw. UML-Klassendiagrammen oder als XML-Datenmodell erfolgen. Wir verfolgen beide Wege:

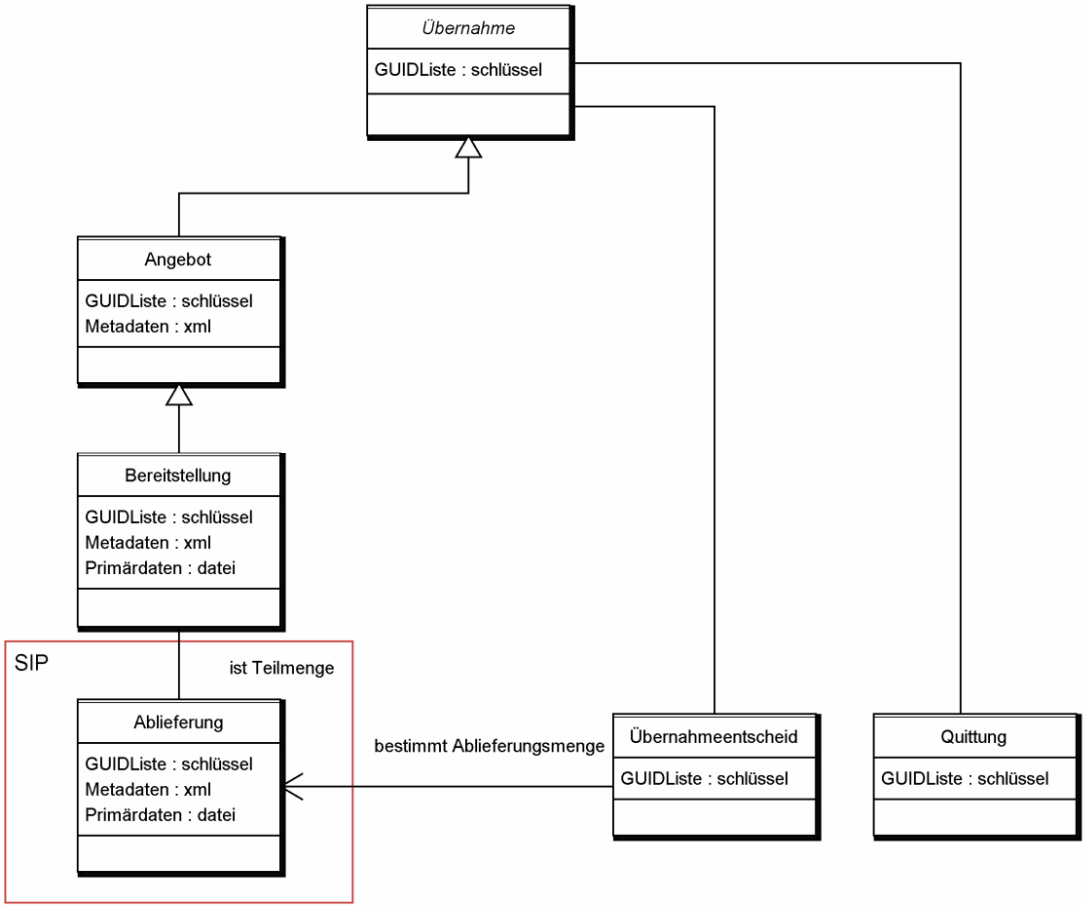
Das **UML-Klassendiagramm** zeigt die Beziehungen zwischen den Objekten und ihre Topographie in den folgenden Abschnitten: Modellierung SIP, Modellierung AIP und Archivjournal.

Wir ergänzen die Klassendiagramme durch **UML-Objektdiagramme**: Objektdiagramme dienen zur Anzeige einer Klasse zu einem best. Zeitpunkt; so lassen sich reale Beispiele und ihre Beziehungen darstellen.

Das **XML Schema** eignet sich hingegen besonders dazu, die Baumstruktur abzubilden: Datenmodellierung in XML

3.4 Modellierung: SIP

3.4.1 Übernahme



Grafik 22: Klassendiagramm SIP

Die Klasse **Übernahme** ist die abstrakte Menge aller Referenzen (GUID) auf die bei einer Ablieferung angebotenen digitalen Unterlagen. Es ist eine Liste von *GUID*.

Übernahmeentscheid und **Quittung** sind die konkreten Klassen zur *Übernahme*.

Übernahmeentscheid ist eine Teilmenge von *Angebot*, nämlich nur die vom Archiv wirklich aus dem *Angebot* übernommenen Unterlagen.

Quittung entspricht dem Umfang nach dem *Angebot*, da das Archiv der abliefernden Stelle eine Liste sämtlicher angebotenen Unterlagen zurückliefert, auf der vermerkt ist, ob diese Unterlagen archiviert wurden oder nicht.

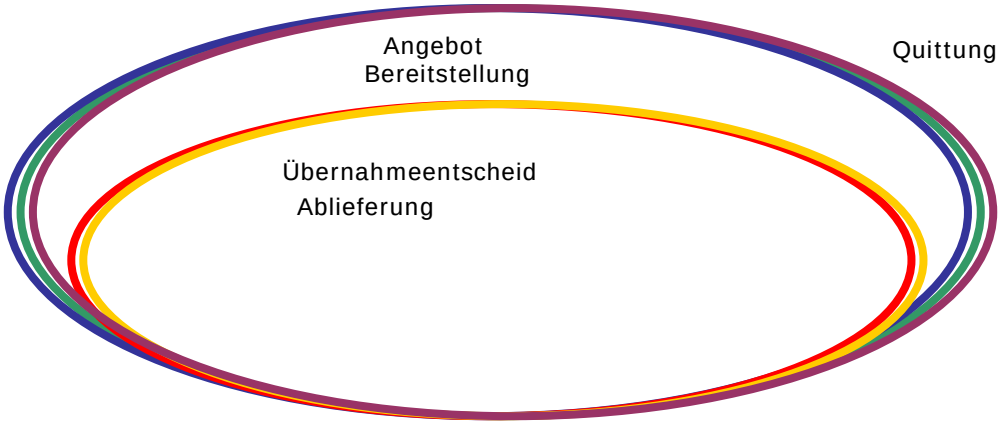
Angebot ist eine Spezialisierung von Übernahme insofern, als dass nicht nur eine Liste von GEVER-GUID, sondern in Form von Metadaten auch Ablagestruktur und Metainformationen zu so referenzierten Unterlagen enthalten sind.

Bereitstellung und **Ablieferung** stellen eine weitere Spezialisierung von *Angebot* dar. Hier sind die Unterlagen, das heisst Primär- und Metadaten mitenthalten.

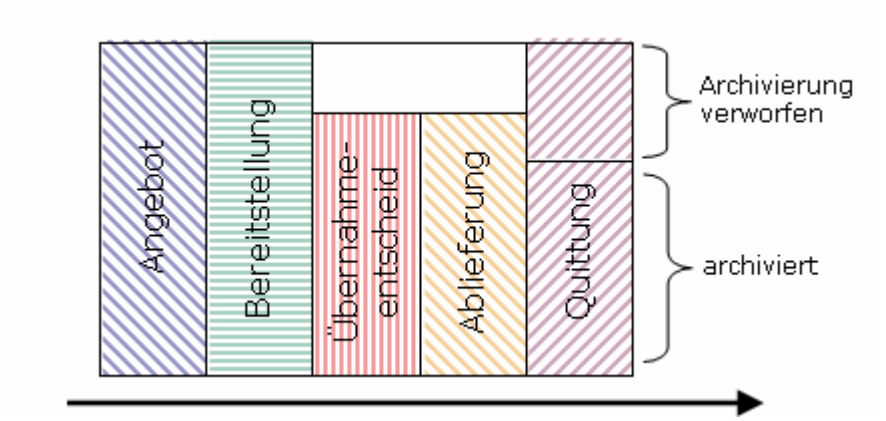
Bereitstellung ist die Menge aller angebotenen Unterlagen.

Ablieferung ist die Menge der dem Archiv nach einer inhaltlichen Bewertung wirklich gelieferten Unterlagen und entspricht der Menge in *Übernahmeentscheid*.

3.4.2 Mengenbeziehungen



Grafik 23: Mengenbeziehungen

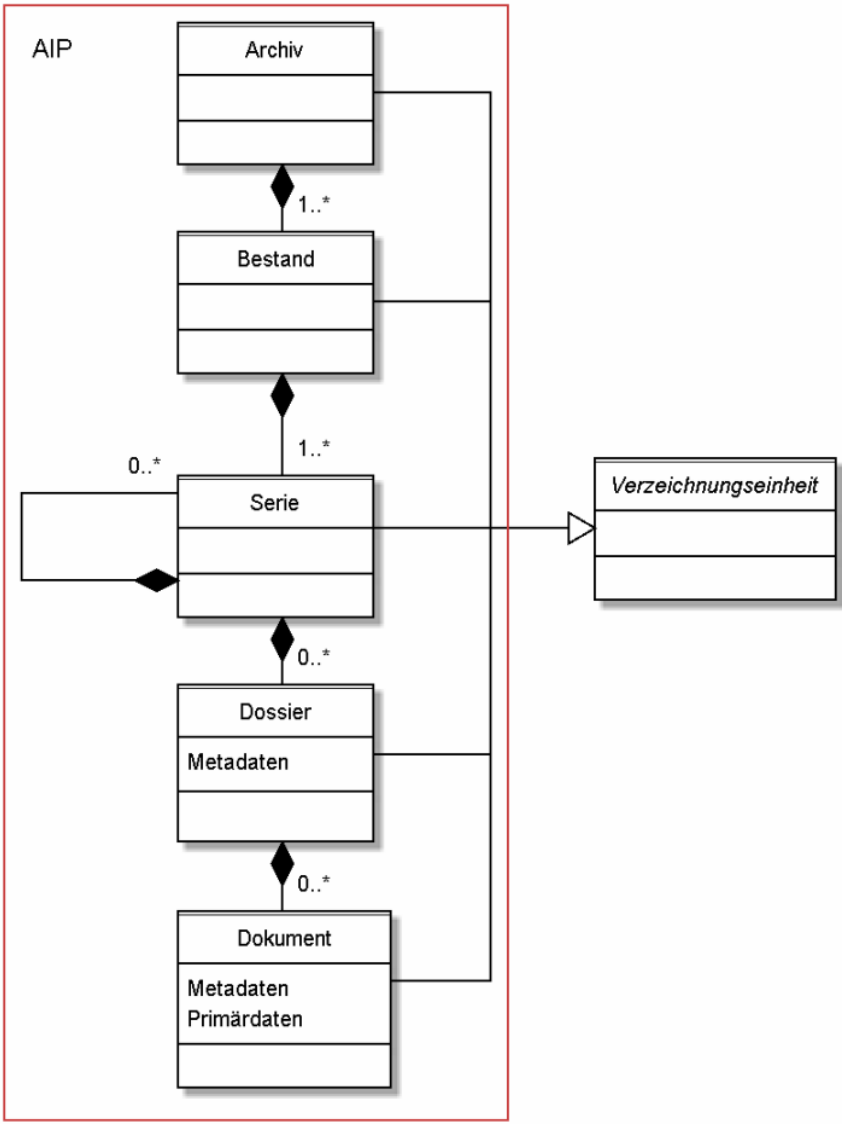


Grafik 24: Mengenbeziehungen im Prozess der Archivierung

Übernahme, Angebot, Bereitstellung und *Quittung* umfassen alle Unterlagen, *Übernahmeentscheid* und *Ablieferung* umfassen das nach der Bewertung durch das Archiv um die ausgeschiedenen Unterlagen reduzierte Angebot.

3.5 Modellierung: AIP

Wir bilden das *AIP (Archival Information Package)* als Komposition von *Verzeichniseinheiten* (Komposition: Ganze-Teile Hierarchie, bei der die Teile vom Ganzen existenzabhängig sind). Die Menge aller AIPs in einem Archiv bildet so die Baumsstruktur der Archivtektonik ab.



Grafik 25: Klassendiagramm AIP

3.5.1 Verzeichniseinheit

Eine **Verzeichniseinheit** ist ein Knoten oder Blatt in der Archivtektonik. Jede Verzeichniseinheit enthält die auf der betreffenden Ebene vorhandenen Meta- und/oder Primärdaten.

Die Gliederung aus Archivsicht in **Archiv**, **Bestand**, **Serie**, **Dossier** und **Dokument** folgt etwas vereinfacht der Ordnungsstruktur I017 - GEVER Metadaten. (siehe ISB Standards)

3.5.2 AIP (Archival Information Package)

Das **AIP** als Komposition von Verzeichniseinheiten besteht also aus der Archivtektonik oder Ordnungsstruktur die das entsprechende Archivobjekt beschreibt und der Menge der auf jeder Ebene dieses Baumes verfügbaren Meta- und Primärdaten.

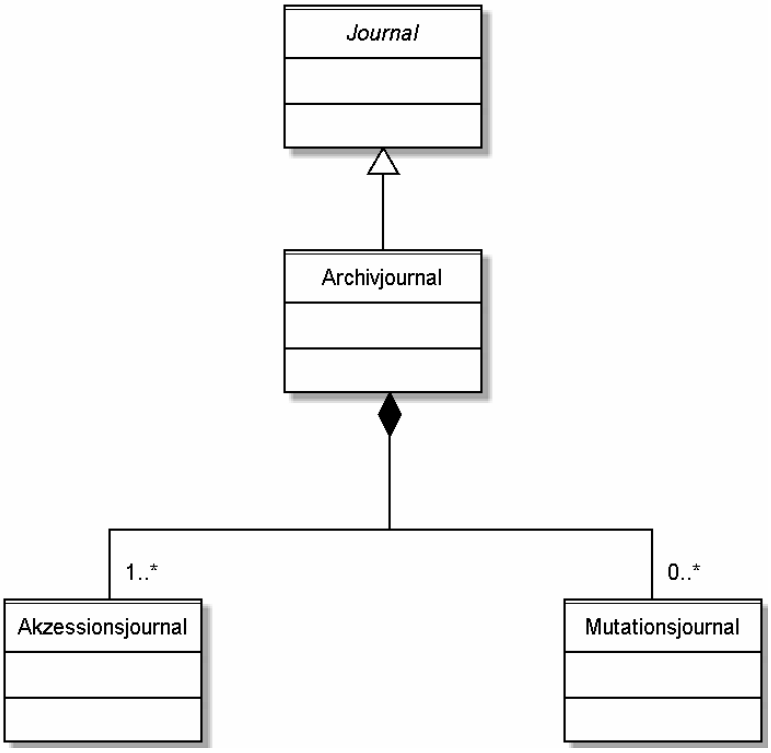
3.5.3 Objektrelationale Speicherung

Wenn wir das AIP als physische Datenkapsel betrachten, z.B. in Form einer XML-Datei, entsteht bei der Speicherung oder Übermittlung vieler AIPs eine enorme Datenredundanz. Das heisst, für die Speicherung ist eine Normalisierung im Sinne des relationalen Datenmodells bzw. die Abbildung in einem *Objektrelationalen Datenmodell* denkbar, welche allerdings gegen den Autonomiegrad (strukturell und strategisch) der Datenkapsel abgewogen werden muss.

3.6 Archivjournal

3.6.1 Journal

Die abstrakte Klasse *Journal* dient als Vorlage für Akzessions-, Mutations- und Archivjournal. Das Archivjournal repräsentiert in technischem Sinne eine Protokolldatei, das heisst das Journal besteht aus *Journaleinträgen* (Log-Zeilen) die jede Veränderung im digitalen Archiv festhalten. Durch geeignete Auswertung des Journals müsste sich also theoretisch ein korruptes digitales Archiv wieder herstellen lassen (vgl. Transaktionslog bei Datenbanken).



Grafik 26: Klassendiagramm Archivjournal

3.6.2 Akzessions-, Mutations- und Archivjournal

Das Akzessions- und Mutationsjournal zusammen bilden das **Archivjournal**. Das **Akzessionsjournal** enthält alle Zugänge ins digitale Magazin. Das Akzessionsjournal dient auch der Zwischenspeicherung von technischen Metadaten im Übernahmeprozess. Die technischen Metadaten sind zwar dem Archivobjekt inhärent, das Zwischenspeichern spart aber das jeweils erneute Generieren dieser Daten.

Das **Mutationsjournal** entsteht beim Prozess *Unterlagen erhalten*, wenn Dateien von einem Format ins andere migriert werden. Ob eine Medienmigration, z.B. von Band auf Disk, hier mitgeschrieben werden soll, ist noch offen.

Ein **Zugriffsjournal** könnte die Zugriffe auf das digitale Archiv protokollieren. Hier steht aber nicht die Wiederherstellung des digitalen Archivs im Zentrum, sondern die Dokumentation der Benutzungsgeschichte.

3.7 Modellierung in XML

Mit XML und Schema-Syntax können wir das AIP bzw. das SIP ebenso gut modellieren.

3.7.1 XML-Repräsentation des AIP

Die Archivtektonik ist im **AIP** als Schachtelung von Verzeichnungseinheiten modelliert. Die Verzeichnungseinheiten sind in hierarchischer Folge identifiziert (nummeriert) und benannt. Das Beispiel beschränkt sich auf das Wesentliche. Metadaten (z.B. in Form von Key-Value Properties) und Primärdateien (base64-codiert eingebettet oder referenziert) sind bewusst weggelassen.

Im digitalen Magazin des StAZG liegt in einer ersten Ablieferung des Amts für Raumplanung aus der Rubrik Bewilligungen das Dossier Kehrrichtverbrennung mit einer Datei Bewilligung.doc

```
  <?xml version="1.0" encoding="UTF-8" ?>
- <aip xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
    xsi:noNamespaceSchemaLocation="verzeichnungseinheit.xsd">
  - <verzeichnungseinheit id="1" name="Archiv">
      StAZG
    - <verzeichnungseinheit id="2" name="Provenienz">
        Amt für Raumlanung (ARP)
      - <verzeichnungseinheit id="3" name="Ablieferung">
          Ablieferung #1
        - <verzeichnungseinheit id="4" name="Serie">
            Bewilligungen
          - <verzeichnungseinheit id="5" name="Dossier">
              Kehrichtverbrennung
              <verzeichnungseinheit id="6" name="Dokument">Bewilligung.doc</verzeichnungseinheit>
            </verzeichnungseinheit>
          </verzeichnungseinheit>
        </verzeichnungseinheit>
      </verzeichnungseinheit>
    </verzeichnungseinheit>
  </aip>
```

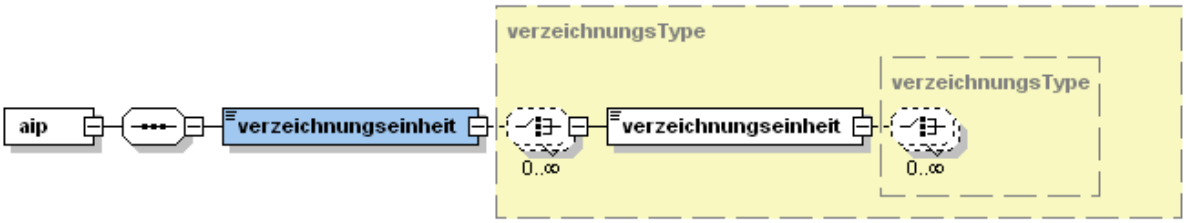
Grafik 27: verzeichnungseinheit.xml

Das zugehörige Schema ist rekursiv aufgebaut: Es ist ein Element „verzeichnungseinheit“ vom Typ „verzeichnungsType“ definiert, das wiederum zwischen null und beliebig vielen „verzeichnungseinheiten“ enthalten kann und zwei Attribute hat: ID und Name.

```
  <?xml version="1.0" encoding="UTF-8" ?>
- <xs:schema elementFormDefault="qualified"
    xmlns:xs="http://www.w3.org/2001/XMLSchema">
  - <xs:element name="aip">
    - <xs:complexType>
      - <xs:sequence>
          <xs:element name="verzeichnungseinheit" type="verzeichnungsType" />
        </xs:sequence>
      </xs:complexType>
    </xs:element>
  - <xs:complexType mixed="true" name="verzeichnungsType">
    - <xs:choice maxOccurs="unbounded" minOccurs="0">
        <xs:element name="verzeichnungseinheit" type="verzeichnungsType" />
      </xs:choice>
      <xs:attribute name="id" use="required" type="xs:integer" />
      <xs:attribute name="name" use="required" type="xs:string" />
    </xs:complexType>
  </xs:schema>
```

Grafik 28: verzeichnungseinheit.xsd

In der XML-Spy-Notation ist die Schachtelung besonders gut sichtbar:



Grafik 29: verzeichnungseinheit.xsd XML-Spy

3.7.2 XML-Repräsentation des SIP

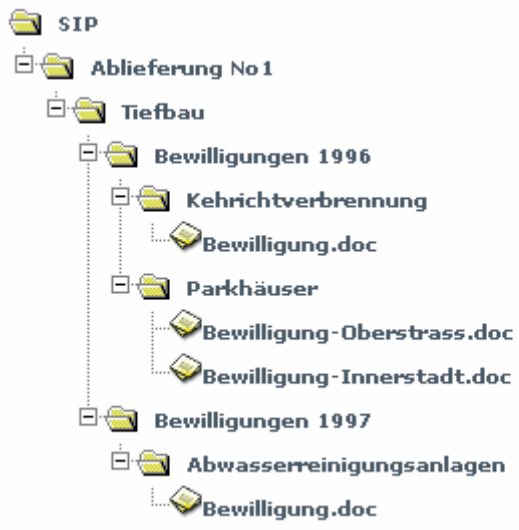
Mit demselben Schema kann eine **Ablieferung** dargestellt werden. Da die Ablieferung noch nicht Teil der Archivtektonik ist, entfallen die entsprechenden Ebenen Archiv und Provenienz.

Die Ablieferung No 1 aus dem Baudepartement Abteilung Tiefbauamt umfasst die Zeitspanne 1996 und 1997. Darin sind zwei Dossiers aus der Rubrik Bewilligungen 1969, Kerichtverbrennung und Parkhäuser. Das Dossier Parkhäuser umfasst zwei Dokumente. Die Rubrik Bewilligungen 1997 mit dem Dossier ARA enthält wieder ein Dokument Bewilligungen.doc.

Die Provenienz "Baudepartement" erscheint nicht in der Ablieferung, denn die Ablieferung stammt vom Baudepartement. Durch die Schachtelung gibt es keine Verwechslung mit dem Dokument Bewilligungen.doc im Dossier Kehrichtverbrennung.

```xml
<?xml version="1.0" encoding="UTF-8" ?>
- <aip xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:noNamespaceSchemaLocation="verzeichnungseinheit.xsd">
  - <verzeichnungseinheit id="1" name="Ablieferung">
      Ablieferung #1
    - <verzeichnungseinheit id="2" name="Rubrik">
        Tiefbau
      - <verzeichnungseinheit id="3" name="Rubrik">
          Bewilligungen 1996
        - <verzeichnungseinheit id="4" name="Dossier">
            Kehrichtverbrennung
            <verzeichnungseinheit id="5" name="Dokument">Bewilligung.doc</verzeichnungseinheit>
          </verzeichnungseinheit>
        - <verzeichnungseinheit id="4" name="Dossier">
            Parkhäuser
            <verzeichnungseinheit id="5" name="Dokument">Bewilligung-Oberstrass.doc</verzeichnungseinheit>
            <verzeichnungseinheit id="5" name="Dokument">Bewilligung-Innerstadt.doc</verzeichnungseinheit>
          </verzeichnungseinheit>
        </verzeichnungseinheit>
      - <verzeichnungseinheit id="3" name="Rubrik">
          Bewilligungen 1997
        - <verzeichnungseinheit id="4" name="Dossier">
            Abwasserreinigungsanlagen
            <verzeichnungseinheit id="5" name="Dokument">Bewilligung.doc</verzeichnungseinheit>
          </verzeichnungseinheit>
        </verzeichnungseinheit>
      </verzeichnungseinheit>
    </verzeichnungseinheit>
  </aip>
```

Grafik 30: ablieferung.xml



Grafik 31: ablieferung.html

4 Digitales Magazin

Wir verstehen unter digitalem Magazin das Zusammenspiel von *Storage Hardware* und *Storage Management Software* (siehe Tabelle 3: Die wichtigsten Software-Layer). Diese beiden Schichten sind für die Persistenz des gespeicherten Bitstromes zuständig. Eine Speicherinfrastruktur für digitale Daten muss unter Berücksichtigung der untenstehenden archivischer Anforderungen bereitgestellt werden.

4.1 Anforderungen

Die Speicherung digitaler Objekte muss folgende Anforderungen erfüllen:

- Dauernde Funktionsfähigkeit der datenspeichernden Infrastruktur, das heisst, der einmal gespeicherte Bitstrom ist jederzeit unverändert lesbar
- Speichern eines Objektes mit Angabe seiner Referenz
- Zugriff auf ein Objekt via Referenz
- Gewährleistung der Integrität des Objektes
- Wiederherstellungsfunktionalität (Recovery)
- Protokollierung von allen Veränderungen an Strukturen und Informationsobjekten
- Gewährleistung des Zugriffsschutzes.

4.2 Speicherkonzepte

Zwei Grundkonzepte beherrschen die Diskussion um digitale Langzeitarchivierung auf der Ebene *Storage Hardware* und *Storage Management Software*:

- Speicherung in offener Struktur. Die Integrität der Objekte wird durch die Tradierung der Information über die Speicherung gewährleistet. D.h. die technische Information über die Speicherung des Bitstroms und die logische Struktur der Daten im Bitstrom sind bekannt.
- Speicherung durch Migration digitaler Objekte von Blackbox zu Blackbox. Die Beherrschung des Migrationsprozesses garantiert die jederzeitige Verfügbarkeit und Integrität der Objekte.

Für beide Konzepte gibt es eine Reihe von Realisierungsmöglichkeiten:

4.2.1 Speicherung in offener Struktur:

- Speicherung des Bitstromes:
    - Bitstrom wird auf ein Magnetband mit bekanntem Format gespeichert (DAT, DLT etc.)
    - Bitstrom wird auf eine Optische Platte mit bekanntem Format gespeichert (CD, DVD etc.)
    - Bitstrom wird mit einem bekannten Decoder als 'bits on film' auf Mikrofilm geschrieben.
- Logische Struktur der Daten:
    - Speicherung der Objekte und Metadaten als Dateien in einem hierarchischen Dateisystem, der Dateipfad widerspiegelt dabei den Zugriffspfad, bzw. die GUID des Objektes.
    - Objekte und Metadaten sind willkürlich, zugriffsoptimiert organisiert. Software mit Hilfe einer Datenbank übernimmt die Verwaltung und den Zugriff auf die Archivobjekte.
    - Speichern der Objekte und Metadaten in grossen Dateien. Software mit Hilfe einer Datenbank übernimmt die Verwaltung dieser Dateien und unterhält die Zugriffspfade auf die Objekte.

- Speicherung aller Metadaten und Objekte in einer Datenbank. Die Datenbank übernimmt die gesamte Datenverwaltungsfunktion.
- Das digitale Archiv besteht aus einer einzigen XML-Datei. Die Datei widerspiegelt die Struktur des Archivs. Die ursprünglichen Dokumente sind in der XML-Datei eingebettet. Diese Lösung ist sehr akademisch und kaum praktikabel, würde jedoch eine einfache Überführung der SIP in ein AIP ermöglichen.

4.2.2 Speicherung durch Migration von Blackbox zu Blackbox.

- Speicherung der Objekte und Metadaten als digitale Objekte (sequenzieller Bytestrom) in einer transparenten "Blackbox". Die Box erfüllt die Anforderungen an die digitale Archivierung. "End of life" dieser proprietären Speicherbox wird ihr Inhalt in die nächste "Blackbox" kopiert.
- Verteilte Speicherung nach dem Modell von Distarnet (http://www.distarnet.ch).
- Speicherung der verschlüsselten digitalen Objekte bei mehreren kommerziellen Anbietern von Speicherplatz.

4.3 Datensicherheit

Gegen welche Ereignisse, die zu Datenverlust führen können, müssen und können wir uns schützen?

- Hardware-Schäden (z.B. durch Überspannung, Wackelkontakte oder Alterung/Verschleiss)
- Diebstahl
- Zerstörung des Speichermediums, z.B. durch Feuer, Fluten oder andere Naturgewalten
- versehentliches oder absichtliches Löschen
- Überschreiben
- logische Fehler innerhalb der Daten

Die Datensicherheit kann einerseits durch Kontrollmechanismen, welche die Integrität des Bitstromes gewährleisten, und andererseits durch mehrfache unabhängige Speicherung und/oder durch Anlegen von externen Kopien erreicht werden.

4.4 Konsens im Projekt

Im Projekt können wir einen vorläufigen Konsens folgendermassen umschreiben:

Mehrfachspeicherung zusammen mit der Auslagerung einer Sicherungskopie erfüllt die archivischen Anforderungen hinreichend. Für die Sicherungskopie wäre Mikrofilm grundsätzlich geeignet. Da sich aber noch kein Standard für digitalen Output auf Mikrofilm etabliert hat, scheint im Moment Magnetband noch das geeignetere Medium zu sein.

Wir möchten hier aber auf die weiterführende Veranstaltungsreihe der KOST zum Thema "archivtaugliche Speicherinfrastruktur" verweisen (Herbst 2007). In dieser Veranstaltungsreihe soll dieses Thema detailliert angegangen werden. Ziel ist es, am Schluss eine Entscheidungsmatrix erstellen zu können, die es den beteiligten Archiven erlaubt, die für ihre Bedürfnisse und äusseren Gegebenheiten richtige Speicherinfrastruktur bereitzustellen.

5 Glossar

zitierte Quellen

BAR 2006: BAR: Glossar http://virtor.bar.admin.ch/De/glo.aspx, Stand vom 14.Jan.2007

DAT III: A. Menne-Haritz, Schlüsselbegriffe der Archivterminologie, Marburg 1999
http://www.staff.uni-marburg.de/~mennehar/datiii/germanterms.htm

ISAD(G): Internationale Grundsätze für die archivische Verzeichnung, Zweite, überarbeitete Ausgabe, Marburg, 2002 http://www.icacds.org.uk/eng/ISAD(G)de.pdf

LandesarchivNRW: Landesarchiv Nordrhein-Westfalen, Glossar
http://www.archive.nrw.de/LandesarchivNRW/Glossar

OAIS: Consultative Committee on Space Data Systems, Reference Model for an Open Archival Information System (OAIS), CCSDS 650.0-B-1, BLUE BOOK, January 2002
http://public.ccsds.org/publications/archive/650x0b1.pdf

Wikipedia, die freie Enzyklopädie
http://de.wikipedia.org

Ablage

Ablage, früher Registratur: Aufbewahrung und Verwaltung von Schriftgut nach festgelegten Ordnungsmerkmalen.

Ablieferung

An das Archiv abgegebene Unterlagen. (DAT III)

Die Tätigkeit der Übergabe incl. Vorbereitung, Zusammenstellung und Transport. (DAT III)

Ablieferungsvereinbarung

Übernommen und dauernd aufbewahrt werden nur archivwürdige Unterlagen. Damit sind jene Unterlagen gemeint, die rechtlich, administrativ, politisch, wirtschaftlich, historisch, sozial oder kulturell wertvoll sind und/oder der Nachvollziehbarkeit staatlichen Handelns dienen. Die Archivwürdigkeit der Unterlagen wird im Einvernehmen mit der abliefernden Stelle beurteilt und in einer gegenseitig unterzeichneten *Ablieferungsvereinbarung* festgehalten. Das Archiv berät die Organe, wie die Unterlagen für die Ablieferung aufzubereiten sind.
Die Ablieferungsvereinbarung bezieht sich häufig nicht nur auf ein einzelnes Aktenangebot, sondern auch auf zukünftige Angebote ("prospektive Vereinbarung").
(Merkblatt zum Archivgesetz, StAZG, http://www.zug.ch/staatsarchiv/data/mb_archivgesetz.pdf
Merkblatt Überlieferungsbildung und Bewertung, StAZH 2003,
http://www.staatsarchiv.zh.ch/download/Merkblatt_Bewertung.pdf)

AIP (Archival Information Package)

Informationen können nicht als beliebig lange Datenströme übernommen und nutzbar gemacht werden. Sie müssen in handhabbare Pakete aufteilt werden. Ein solches Paket enthält zum einen den Informationsinhalt (*Content Information*) und zum anderen Informationen, die für die Erhaltung des Informationsinhalts notwendig sind (*Preservation Description Information*) zusammengefasst durch die *Packaging Information*. (Übersetzung OAIS)
Im Falle der *Archivierung von Unterlagen aus Geschäftsverwaltungssystemen* soll das AIP folgendes umfassen (AUGev):

- Primär- und Metadaten eines Dokumentes
- Ablageinformationen (Archivtektonik und Aktenplan)
- technische Metadaten

Angebot

Alle dem Archiv nach Massgabe der Ablieferungsvereinbarung anzubietenden Unterlagen, zusammengestellt in einem Ablieferungsverzeichnis. Siehe auch Anbietepflicht nach jeweiligem Archivgesetz (Archivierungsgesetz, BGA, SR 152.1)

Angebotsvereinbarung

Synonym für Ablieferungsvereinbarung

Batch-Modus

Im Gegensatz zum Einzelschritt-Modus werden im Batch-Modus Listen von Aufgaben vom Computer selbständig, ohne Interaktion mit einem Anwender, abgearbeitet. Man spricht auch von Stapelverarbeitung. (Wikipedia)

Bestand

Bestand (fonds) – Alle Unterlagen, unabhängig von Form und Trägermaterial, die auf organische Weise bei einer Person, Familie oder Körperschaft im Rahmen ihrer Tätigkeit und Funktion erwachsen und / oder von ihr zusammengestellt bzw. genutzt worden sind. (ISAD(G))

digitales Magazin

Digitale Speicherinfrastruktur, also die *Storage Hardware*, das eigentliche Speichersystem und die *Storage Management Software* zur Verwaltung des Speicherplatzes auf dem Speichersystem. Ein verwandter Begriff ist *digitales Repositorium*. Das digitale Magazin ist klar abgegrenzt von der *Digital Archiving Software* und stellt bildlich gesprochen nur das Magazin mit den Regalen dar, nicht aber das Archiv in seiner Gesamtheit. (AUGev)

DIP (Dissemination Information Package)

Im OAIS Referenzmodell werden die archivierten Daten (AIP) über die sogenannten *Dissemination Information Packages (DIP)* zugänglich gemacht. Diese können für bestimmte Nutzergruppen je nach Vorliegen bestimmter Bedürfnisse und rechtlicher Bedingungen generiert und zielgruppenorientiert zur Verfügung gestellt werden. (Übersetzung OAIS)

Dokument

Ein Dokument ist ein mehr oder weniger umfangreiches Schriftstück zusammen mit den zum Verständnis notwendigen Zusatzinformationen, i.d.R. eine Primärdatei und Metadaten (ein einzelnes Objekt vom Typ "Unterlagen"). Ein Dokument kann auch aus mehreren Primärdateien bestehen, (entstanden aus Versionierung im Entstehungsprozess) oder Migration bei der Erhaltung. (AUGev)

Dossier

Die Unterlagen zu einem Verwaltungsvorgang (Geschäft) werden in Dossiers zusammengefasst und abgelegt. Das Dossier stellt also eine inhaltliche und eine physische Einheit dar. Ein Dossier ist die Menge der Dokumente, Vorgänge und Beteiligten zu einem Geschäft, wobei Vorgänge und Beteiligte die Dossierprozessdaten bilden, also Metadaten zum Dossier. (AUGev)

Elektronische Signatur

Unter einer elektronischen Signatur versteht man Daten, mit denen man den Unterzeichner bzw. Signaturersteller identifizieren kann und sich die Integrität der signierten, elektronischen Daten prüfen lässt. Die elektronische Signatur erfüllt somit technisch gesehen den gleichen Zweck wie eine eigenhändige Unterschrift auf Papierdokumenten. (Wikipedia)

Findmittel

Ein Findmittel ist eine Suchhilfe für die Archivrecherche. Wir unterscheiden zwischen digitalen Findmitteln, die in elektronischer Form zur Verfügung stehen, und konventionellen Findmitteln auf Papier. (BAR 2006)
Oberbegriff für Erschließungsinstrumente für Archivbestände, wie z.B. Ablieferungslisten, Findkarteien, Findbücher, Inventare, Datenbanken und Beständeübersichten. (LandesarchivNRW)

FTP

Das File Transfer Protocol (engl. für "Dateiübertragungsverfahren", kurz FTP), ist ein 1985 spezifiziertes Netzwerkprotokoll zur Dateiübertragung über Netzwerke. (Wikipedia)

GUID

Globally Unique Identifier. Ein weltweit eindeutiger Identifikationsstring für eine Ressource. Häufig basieren GUID auf dem Prinzip des *namespace*, d.h. einem bestimmten Bereich (Institution, Land, etc.) entspricht ein erster Teil des GUID, ein zweiter Teil wird dann bereichsintern vergeben. Es sind auch mehrere solcher hierarchischer Schachtelungen möglich. (AUGev)

Kassieren

Unter **Kassation** versteht man im Archivwesen die datenschutzgerechte Vernichtung nicht archivwürdig bewerteter Unterlagen. Behörden haben aufgrund gesetzlicher Vorschriften die Auflage, alle Unterlagen (Akten, aber auch Daten), die sie nicht mehr benötigen, dem zuständigen Archiv anzubieten. (Wikipedia)

Magazin

Gebäude, Gebäudeteil oder Raum, in dem Archivgut unter besonderen klimatischen Bedingungen aufbewahrt wird. (LandesarchivNRW)
Siehe digitales Magazin

Metadaten

Metadaten sind diejenigen Informationen, die vom Geschäftsverwaltungssystemen zu einem Dokument oder Dossier zusätzlich verwaltet werden. Sie werden in der Regel in einer Datenbank gehalten und erst für den Datenexport in eine XML-Datei umgewandelt. (AUGev)
Siehe Dokument und Primärdaten.

Primärdaten

Unter Primärdaten verstehen wir die Datei (oder Dateien), welche beim Anlegen eines Dokuments mit den eingesetzten Office-Werkzeugen (Textverwaltung, Tabellenkalkulation, etc.) angelegt werden. In der Regel sind die Primärdaten die Grundlage für das ausgedruckte Dokument. (AUGev)
Siehe Dokument und Metadaten.

Schnittstelle

Eine Schnittstelle ist ein Teil eines Systems und dient der Kommunikation. Eine Schnittstelle wird durch eine Menge von Regeln beschrieben, der Schnittstellenbeschreibung. Softwareschnittstellen oder softwareseitige Datenschnittstellen sind logische Berührungspunkte in einem Softwaresystem: sie definieren, wie Kommandos und Daten zwischen verschiedenen Prozessen und Komponenten ausgetauscht werden. (Wikipedia)

Signatur

Kennzeichen zur Identifikation einzelner Einheiten von Archivgut und zu ihrer Verknüpfung mit den Findmitteln. Archivische Signaturen geben die Reihenfolge der Lagerung im Magazin an. Sie sind aus früheren Lokaturen oder Aktenkennzeichen entstanden, die zusätzlich auch eine inhaltliche Bedeutung (ähnlich einer Notation bei einer Klassifikation) hatten. Sie finden sich deshalb in vielfältigen, z.T. komplizierten Varianten bei alten Beständen. (DAT III)

Die elektronische Signatur ist ein Ersatz für Unterschriften. Sie soll Fälschungssicherheit und Authentizität von elektronischen Aufzeichnungen begründen. (DAT III)

SIP (Submission Information Packages)

Das SIP ist das Einlieferungspaket für die zu archivierende Daten im OAIS Referenzmodell. Die Standardisierung des SIP ermöglicht, die einkommenden Pakete mit einer begrenzten Menge unterschiedlicher Methoden zu entpacken und zu überprüfen. Anschliessend wird aus dem SIP ein Archiv-Informationspaket (*Archival Information Package*, AIP) erzeugt, das den archivinternen Regeln (z.B. für Datenformat, Datendokumentation) entspricht. (Übersetzung OAIS / AUGev)

SOAP

SOAP (ursprünglich für Simple Object Access Protocol) ist ein Netzwerkprotokoll, mit dessen Hilfe Daten zwischen Systemen ausgetauscht werden können. SOAP stützt sich auf die Dienste anderer Standards: XML zur Repräsentation der Daten und Internet-Protokolle der Transport- und Anwendungsschicht zur Übertragung der Nachrichten. Die gängigste Kombination ist SOAP über HTTP und TCP. (Wikipedia)

Software-Layer

Software-Layer, zu deutsch Softwareschicht, ist ein Begriff aus der Systementwicklung und meint die Kapselung der Komplexität einer Hardware oder Software durch das vollständige Umschliessen durch eine andere Softwareschicht. Jede Schicht löst ein spezifisches Problem mit der jeweils notwendigen Komplexität. Grundsätzlich lässt sich sagen: Je weiter weg die Software-Schichten von der Hardware sind, desto besser sind die Anwendungsvielfalt und Schnittstelle zum Menschen hin optimiert. (AUGev)

Übernahmeentscheid

Nicht zu verwechseln mit Übernahmevereinbarung bzw. Ablieferungsvereinbarung. Die vom Archiv getätigte Auswahl, welche Dossiers und Dokumente bei einer Ablieferung (Akzession) übernommen werden.
(AUGev)

UML

UML ist die Abkürzung für *Unified Modeling Language*. Die UML ist eine von der Object Management Group (OMG) standardisierte Notation und Semantik zur Visualisierung, Konstruktion und Dokumentation von Modellen für die objektorientierte Softwareentwicklung. Siehe www.uml.org

Unterlagen

In den Archivgesetzen eingeführter Begriff für Schriftgut aus Verwaltungsbehörden unabhängig von seinem Trägermaterial, das archivreif ist und zur Aussonderung und Anbietung an das zuständige Archiv ansteht. (DAT III)
Mit Unterlagen sind alle aufgezeichneten Informationen, unabhängig vom Informationsträger, gemeint, die auf Grund einer Geschäftstätigkeit entstanden sind. Zu den Unterlagen gehören auch die Hilfsmittel und die ergänzenden Informationen, die für das Verständnis der Unterlagen und deren Nutzung notwendig sind. (Zit. sinngemäss BAR 2006)
siehe auch hier: Dokument

Verzeichniseinheit

Eine Verzeichniseinheit ist ein Knoten oder Blatt in der Archivtektonik. Jede Verzeichniseinheit enthält die auf der betreffenden Ebene vorhandenen Meta- und/oder Primärdaten. (AUGev)

Viewer

Ein Viewer (Dateibetrachter) ist ein Computerprogramm, mit dem sich Computerdateien anschauen, aber nicht verändern lassen. (AUGev)

XML

Die Extensible Markup Language (engl. für "erweiterbare Auszeichnungssprache"), abgekürzt XML, ist eine Auszeichnungssprache zur Darstellung hierarchisch strukturierter Daten in Form von Textdateien. XML wird bevorzugt für den Austausch von Daten zwischen unterschiedlichen IT-Systemen eingesetzt. Die vom World Wide Web Consortium (W3C) herausgegebene XML-Spezifikation definiert eine Metasprache, auf deren Basis durch strukturelle und inhaltliche Einschränkungen anwendungsspezifische Sprachen definiert werden. www.w3.org/XML

XML-RPC

XML-RPC (Extensible Markup Language Remote Procedure Call) ist eine Definition zum Methodenaufruf in verteilten, durch ein Netzwerk verbundenen Systemen. Für den Transport der Daten wurde auf das Hypertext Transfer Protocol (HTTP) zurückgegriffen, während die Darstellung der übertragenen Daten in der Extensible Markup Language (XML) geschieht. Historisch gesehen stellt XML-RPC den Vorgänger zu SOAP dar, ist im Gegensatz zu diesem jedoch wesentlich schlanker und schneller zu verstehen. (Wikipedia)

6 Anhang: Kurzdarstellung UML

"Die Unified Modeling Language (UML) ist eine von der Object Management Group (OMG) entwickelte und standardisierte Sprache für die Modellierung von Software und anderen Systemen. ... Die UML ist heute eine der dominierenden Sprachen für die Modellierung von betrieblichen Anwendungssystemen (Softwaresystemen)." Wikipedia 2007.

Wir verwenden für unsere Modellierung vier der 16 UML 2.0 Diagrammtypen: Use Case Diagramm, Aktivitätsdiagramm, Klassendiagramm und Objektdiagramm. In dieser Kurzdarstellung werden nur die explizit in der Modellierung verwendeten syntaktischen Elemente erläutert.

Eine knappe und gut verständliche Einführung von Boris Schäling findet sich unter: http://www.highscore.de/uml/

Etwas ausführlicher in Buchform: M. Jeckle, C. Rupp, J. Hahn, B. Zengler, S. Queins: UML 2 Glasklar, Hanser, 2003, ISBN: 3-446-22575-7.

Die UML 2 Spezifikation der Object Management Group (OMG) findet sich unter UML 2.0 OCL.

6.1 Anwendungsfalldiagramm (Use Case Diagram)

Das Anwendungsfalldiagramm beschreibt die Interaktion zwischen einem Anwender und einem (geschäftlichen/Software) System. Ein Akteur initiiert einen Anwendungsfall (grafisch durch eine Ellipse dargestellt). Der Anwendungsfall führt zu einem wahrnehmbaren Ergebnis.

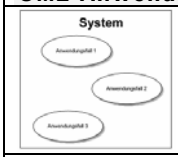
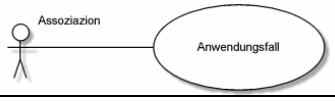
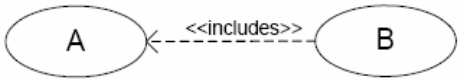
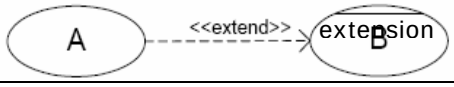
UML-Anwendungsfalldiagramm	
	*System* Die Einheit, die das Verhalten realisiert, das durch den Anwendungsfall beschrieben wird
	Akteur Der Akteur repräsentiert die Rolle, die von außen mit dem System interagiert.
	Anwendungsfall (Use Case) Ein Anwendungsfall beschreibt in einem zusammenhängenden Arbeitsablauf die Interaktion mit einem System.
	*Assoziation* Beschreibt die Beziehung zwischen einem Akteur und einem Anwendungsfall
	*<<include>> Beziehung* Mehrere Anwendungsfälle (B) importieren das Verhalten eines andern Anwendungsfalls (A).
	*<<extend>> Beziehung* Bei gewissen Bedingungen (extension points) erweitert Anwendungsfall (A) den Anwendungsfall (B).

Tabelle 17: UML-Anwendungsfalldiagramm

6.2 Aktivitätsdiagramm (Activity Diagram)

Das Aktivitätsdiagramm beschreibt einen Ablauf, definiert durch Knoten, die durch Objekt- und Kontrollflüsse miteinander verbunden sind.

UML Aktivitätsdiagramm	
	Startknoten bzw. Endknoten Start und Ende des Ablaufs

	Verzweigung bzw. Zusammenführung Auf Grund von Bedingungen wird entschieden welcher von mehreren Kontrollflüssen fortgesetzt wird.
	Teilung (Splitting) bzw. Synchronisation Der Kontrollfluss wird in mehrere nebenläufige Kontrollflüsse aufgeteilt, respektive mehrere nebenläufige Kontrollflüsse werden zu einem Kontrollfluss zusammengeführt.
Aktion	*Aktion* Ein Einzelschritt (Knoten) innerhalb der Verarbeitung
Aktivität	*verschachtelte Aktion* Eine verschachtelte Aktion enthält eine komplette Aktivität.
Objekt	*Objekt* repräsentiert und transportiert Daten und Werte in der Aktivität
	Übergang (Kontrollfluss) Kanten sind gerichtete Übergänge zwischen den Knoten
Swimlane	*Verantwortlichkeitsbereiche (Partitionen)* Die Aktivitäten können nach Verantwortlichkeitsbereich der Akteure vertikal gegliedert werden.

Tabelle 18: UML-Aktivitätsdiagramm

6.3 Klassendiagramm

Das Klassendiagramm ist das wichtigste UML Diagramm der objektorientierten Modellierung. Es beschreibt die statische Struktur des abzubildenden Systems. Eine Klasse ist eine Menge von Objekten im Sinne einer Abstraktion. In der Klasse werden bestimmte Eigenschaften (Attribute), Operationen und die Semantik der Objekte definiert.

UML Klassendiagramm	
Klassenname -attribut +operation()	*Klasse* Eine Menge von Objekten die dieselben Eigenschaften, Einschränkungen und Semantik haben *Attribut* Eigenschaften einer Klasse *Operation* Verhalten von Objekten
Klasse A 1 0..* Klasse B	Assoziation Die Assoziation dient zur Darstellung von fachlichen Beziehungen zwischen Klassen (Bedeutung: hat, gehört zu). Die Multiplizität gibt an, mit wie vielen Objekten der gegenüberliegenden Klasse ein Objekt assoziert sein kann
Klasse A Klasse B	Aggregation Die Aggregation ist ein spezieller Fall einer Assoziation mit der Bedeutung "besteht aus" (A besteht aus B)
Klasse A 1 * Klasse B	Komposition Eine strenge Form der Aggregation, fasst eine Menge von Einzelkomponenten zu einer zusammengehörigen übergeordneten Komponente zusammen
Klasse A Klasse B1 Klasse B2	Generalisierung Die Generalisierung extrahiert aus den beiden Unterklassen B1 und B2 die gemeinsamen Charakteristika der Oberklasse A. Die Generalisierung ist die Grundlage der Vererbung in der Objektorientierten Modellierung.

Tabelle 19: UML-Klassendiagramm

6.4 Objektdiagramm

Objektdiagramme entsprechen syntaktisch den Klassendiagrammen. Sie zeigen aber die zu einem bestimmten Zeitpunkt aus den Klassen instanzierten (erzeugten) Objekten. So lassen sich reale Beispiele und ihre Beziehungen darstellen.

Grafik 32: Beispiel Wikipedia 2007 "Objektdiagramm"