

VdW-Arbeitskreis „Elektronische Archivierung“ (AKEA)

Anforderungen an ein elektronisches Archivsystem

1. Einführung

Ausgehend von der Beschäftigung mit dem OAIS-Konzept (OAIS steht für *Open Archival Information System*), dem Metadatenstandard Premis und dem Nestor-Kriterienkatalog „Vertrauenswürdige Archivierung“ sind im VdW-Arbeitskreis „Elektronische Archivierung“ (AKEA) Anforderungen an elektronische Archivsysteme formuliert worden. Damit möchten wir insbesondere denjenigen, die vor der Entscheidung stehen, Archivsoftware beschaffen zu müssen, Hinweise zur Entscheidungsfindung geben. Es muss allerdings betont werden, dass es DAS EINE System nicht gibt, mit dessen Anschaffung man die Probleme der elektronischen Archivierung rundum gelöst hätte. Vielmehr kann nach heutigem Stand elektronische Archivierung nur unter Einsatz einer modular aufgebauten Lösung erfolgen, wie sie z.B. das OAIS-Konzept vorsieht. Dabei müssen stets auch die Rahmenbedingungen jedes einzelnen Archivs berücksichtigt werden, die ihrerseits direkt und unmittelbar Einfluss auf die Entscheidung für oder gegen eine bestimmte Vorgehensweise bedingen.

Die Module ergeben sich aus dem OAIS-Konzept, zumeist setzen Archive bereits ein Erschließungsmodul zur Verzeichnung ihrer Unterlagen ein, das bislang vor allem analog überlieferte Dokumente erfasst, mit dem nun aber auch zunehmend elektronische Dokumente erfasst werden müssen.

Inwieweit das gewählte Erschließungssystem auch in einem elektronischen Archiv zum Einsatz kommen kann, ist jeweils im Einzelfall und unter Berücksichtigung der je besonderen Verhältnisse zu prüfen.

Die genannten Anforderungen stellen aus unserer Sicht Mindestanforderung dar und erheben nicht den Anspruch auf Vollständigkeit.

Zu jedem Modul werden zunächst einige Hinweise zu Anforderungen genannt. In der Spalte „Wertigkeit“ kann jedes Archiv individuell den Stellenwert der genannten Anforderung beurteilen. Die Tabellen zu den einzelnen Modulen sollen gegenüber Software-Anbietern, Kollegen aus Rechenzentren, aber auch aus den Fachabteilungen, in denen die Daten entstehen, die archivarischen Anforderungen verdeutlichen. Sie können dabei helfen, über einzelne Aspekte ins Gespräch zu kommen und gemeinsam nach einer guten Lösung für das jeweilige Archiv zu suchen.

Es sind zunächst Anforderungen formuliert worden für die Module

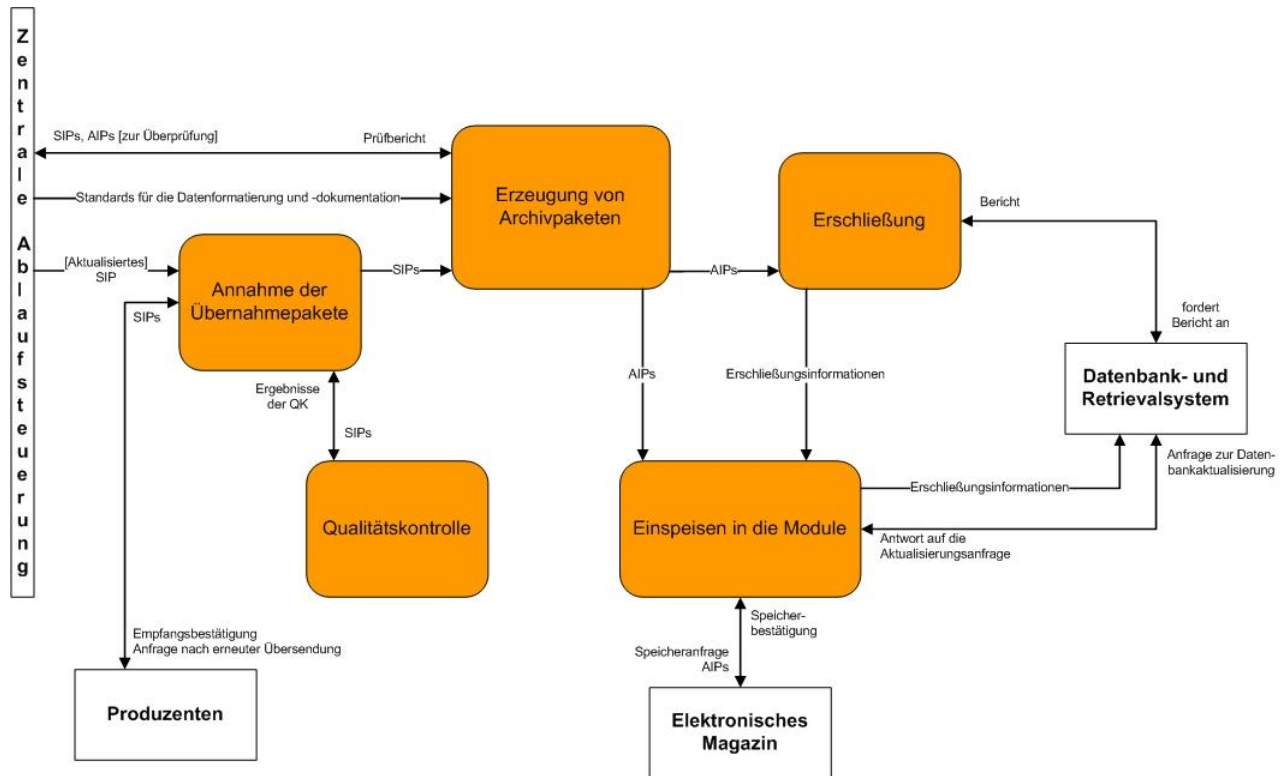
Übernahme
Elektronisches Magazin
Erschließung
Bestandserhaltung
Nutzung

Anforderungen an das Gesamtsystem wurden bislang nicht formuliert.

2. Anforderungen an die Module nach OAIS

2.1 Anforderungen an das Übernahmemodul

Abb.: Übernahmemodul nach OAIS



Das Übernahmemodul sorgt für die Annahme der Übergabepakete SIPs und ihre Umwandlung in AIPs. Mit Hilfe des Übernahmemoduls werden elektronische Daten und dazugehörige Metadaten von außen in das Digitale Archiv übergeben, Metadaten können im Zuge der Übernahme auch durch das Archiv selbst generiert werden. Das Übernahmemodul hat Schnittstellen zum Elektronischen Magazin und zum Erschließungsmodul, in welches Metadaten möglichst automatisiert hineinlaufen. Das Übernahmemodul überprüft die übergebenen Daten. Gegebenenfalls findet bei der Übernahme eine erste Umwandlung (Migration) statt, etwa die Umwandlung von Nutzungsformaten in archivgeeignete Formate.

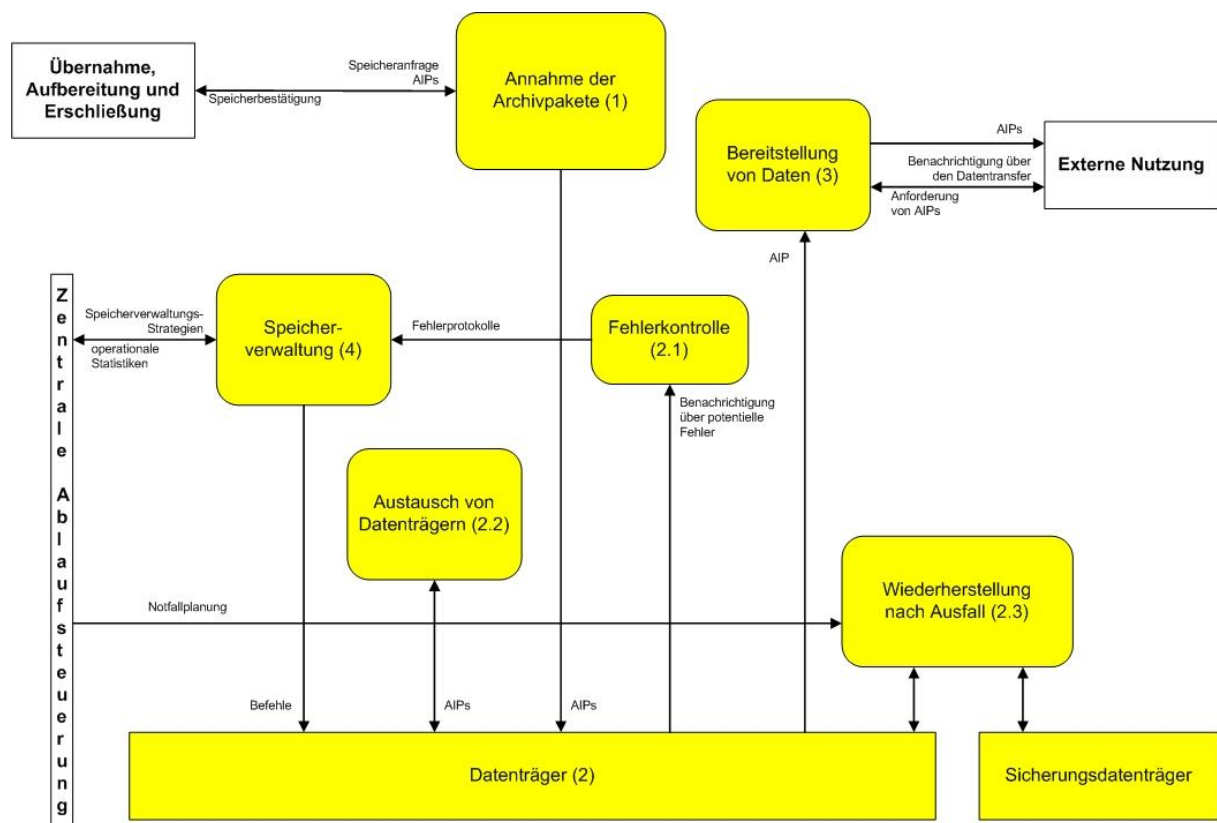
Nr.	Anforderung	Wertigkeit	Kommentar
1	Das Übernahmemodul verfügt über anpassbare Schnittstellen nach außen.		Bei immer wiederkehrenden Übernahmen können verschiedenen Schnittstellen standardisiert vorbereitet werden, die dann jeweils angesteuert werden können.
2	Das Übernahmemodul verfügt über eine standardisierte Schnittstelle zum Elektronischen Magazin.		
3	Das Übernahmemodul überprüft Formate und Qualität der angelieferten Daten.		

3.1	Das Übernahmemodul nimmt Migrationen in festgelegte Archivformate vor.		Für jeden Datentyp legt ein Archiv die zu verwendenden Archivformate fest.
3.2	Falls die Qualität der Daten nicht ausreicht, steuert das Übernahmemodul eine Migration.		
3.3	Alle im Übernahmemodul vorgenommenen Veränderungen an den Daten und Migrationen werden dokumentiert.		
4	Bereits im Vorfeld der Übergabe wurden vom Archiv zusammen mit der Fachabteilung die Metadaten festgelegt, die zusammen mit den Daten übernommen werden müssen. Dieser Metadatensatz umfasst Metadaten, die automatisch übernommen und solche, die jeweils individuell ergänzt werden müssen. Diese Metadaten werden in einem XML-Datensatz erfasst. Siehe 5.		
5	Das Übernahmemodul generiert ein SIP, das folgende Inhalte umfasst: Übernahmedaten (Primärdaten), migrierte oder umformatierte Daten, Informationen zu den im Moment der Übernahme vorgenommenen Veränderungen, Metadaten, soweit mit den Daten übergeben.		Den genauen Inhalt des SIP legt jedes Archiv selbst fest.
5.1	Das SIP hat eine eindeutige ID.		
5.2	Das Übernahmemodul generiert aus dem SIP und dem XML-Metadatensatz ein Archivpaket (AIP) zur Ablage im elektronischen Magazin.		
6	Das Übernahmemodul benachrichtigt die abgebende Abteilung und das Archiv nach Abschluss der Übernahme automatisch über deren Verlauf.		
7	Bei notwendig werdenden Veränderungen an den im elektronischen Magazin abgelegten AIPs, etwa aus technischen Gründen notwendigen Refreshings oder Migrationen, wird das neu erstellte migrierte Datenpaket als neue Übernahme betrachtet und als solche erneut ins Digitale Archiv übernommen. Der Zusammenhang zum Primärdokument bleibt dabei erhalten. Damit werden alle Veränderungen an dem Dokument in nachvollziehbarer Weise dokumentiert.		Man kann sich das wie einen „Boxenstopp“ der Daten vorstellen.
8	Alle an der Übernahme beteiligten Akteure und ihre Aktionen werden dokumentiert.		
9	Das Übernahmemodul liefert Statistiken zu durchgeführten Übernahmen zur Auswertung, z.B. durch das Archiv und seinen Träger.		
10	Die Akteure und ihre Handlungen werden protokolliert		

2.2 Anforderungen an das Elektronische Magazin (EMag)

Das Elektronische Magazin umfasst den digitalen Speicher sowie seine Organisation. Es dient der Ablage von Archivpaketen (AIPs). Archivpakete sind Container, die sich aus verschiedenen Einzelkomponenten zusammensetzen. Ihr Inhalt ist selbsterklärend und sowohl menschen- als auch maschinenlesbar. Die detaillierte Konzeption eines Archivpakets bestimmt jedes Digitale Archiv für sich selbst. Bei einem Totalverlust der technischen Infrastruktur kann ein Digitales Archiv mit Hilfe der gesicherten AIPs auf einer neuen technischen Plattform rekonstruiert werden.

Abb. : Aufgaben des Elektronischen Magazins



Das Elektronische Magazin stellt folgende Dienstleistungen und Funktionen bereit:

Nr.	Anforderung	Wertigkeit	Kommentar
1	Annahme der Archivpakete		
1.1	Das EMag verfügt über eine standardisierte Schnittstelle zum Import von Archivpaketen.		
1.2	Das EMag liefert eine automatische Speicherbestätigung.		
1.3	Das EMag generiert eine eindeutige		Die ID dient zur dauerhaften maschinellen Identifizierung

Nr.	Anforderung	Wertigkeit	Kommentar
	ID für jedes Archivpaket.		und zum Metadatenaustausch.
1.4	Das EMag kann von jeder Datei einen Hashwert ermitteln.		
1.5	Das EMag kann für jede Datei eine elektronische Signatur vergeben.		
2	Datenträger Das EMag bietet eine mehrfach redundante, bitgenaue Speicherung der Archivpakete.		Redundanz meint, das EMag (mehrfach) parallel anzulegen, damit beim Ausfall einer Komponente die anderen die Speicherung der Archivpakete weiterhin gewährleisten. ¹
2.1	Fehlerkontrolle Das EMag stellt einen Standardmechanismus bereit, der regelmäßig alle Archivpakete auf Vollständigkeit und Unversehrtheit überprüft und ggf. eine Fehlermeldung generiert.		
2.2	Austausch von Datenträgern Regelmäßige Überprüfung und ggf. Austausch der Datenträger.		
2.3	Wiederherstellung nach Ausfall Wiederherstellung der Daten nach Ausfall gemäß der Notfallplanung der Zentralen Ablaufsteuerung.		
3.	Bereitstellung von Daten		
3.1	Das EMag verfügt über eine standardisierte Schnittstelle zum Export von Archivpaketen (für Migrationen oder für die Externe Nutzung). Archivpakete können in einen Zwischenspeicher übertragen und/oder auf jeden beliebigen Datenträger kopiert werden (CD-ROM, USB-Stick).		
4.	Speicherverwaltung		
4.1	Das System steuernde Metadaten		Sollte das Archivsystem nicht

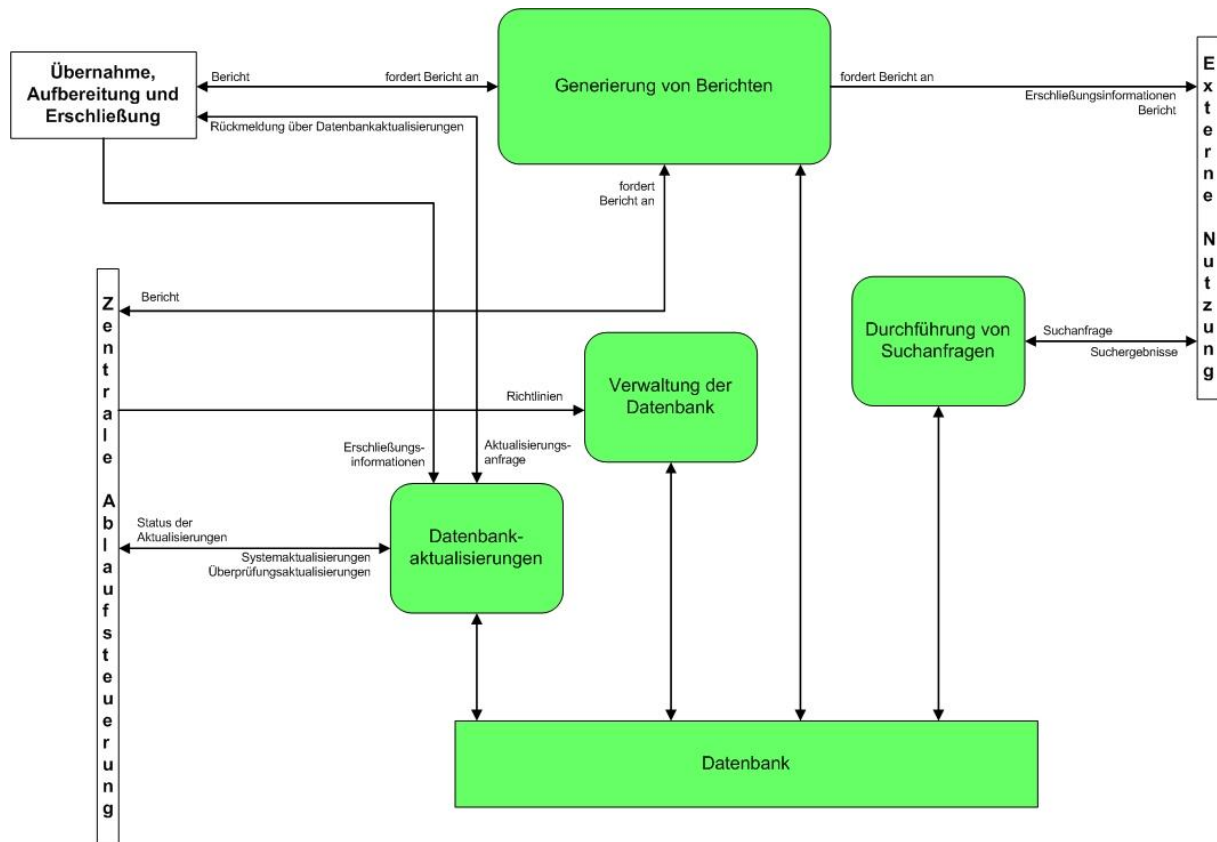
¹ Einen Überblick über verschiedene Datenträger und Speicherverfahren bieten Rolf Däßler: Datenträger und Speicherverfahren für die digitale Langzeitarchivierung http://nestor.sub.uni-goettingen.de/handbuch/artikel/nestor_handbuch_artikel_410.pdf; Hubert Salm.; nestor Informationsblatt 18 „Auswahl der Speichertechnologie“ http://files.d-nb.de/nestor/sheets/18_speichemedien.pdf.

Nr.	Anforderung	Wertigkeit	Kommentar
	können in einer Datenbank, die Archivpakete müssen in einem Dateisystem abgelegt werden. ²		mehr verwendbar sein, wären dennoch alle Archivpakete vorhanden und aufrufbar.
4.2	Das EMag ermöglicht die Vergabe und Verwaltung von Zugriffsrechten.		
4.3	Das EMag liefert Statistiken etwa über die Speicherkapazität und -auslastung oder die Bereitstellung von Daten.		

² S. Christian Keitel: Digitale Archivierung beim Landesarchiv Baden-Württemberg, in: Archivar 1/2010, S. 21.

2.3 Anforderungen an das Erschließungsmodul

Abb.: Erschließungsmodul nach OAIS



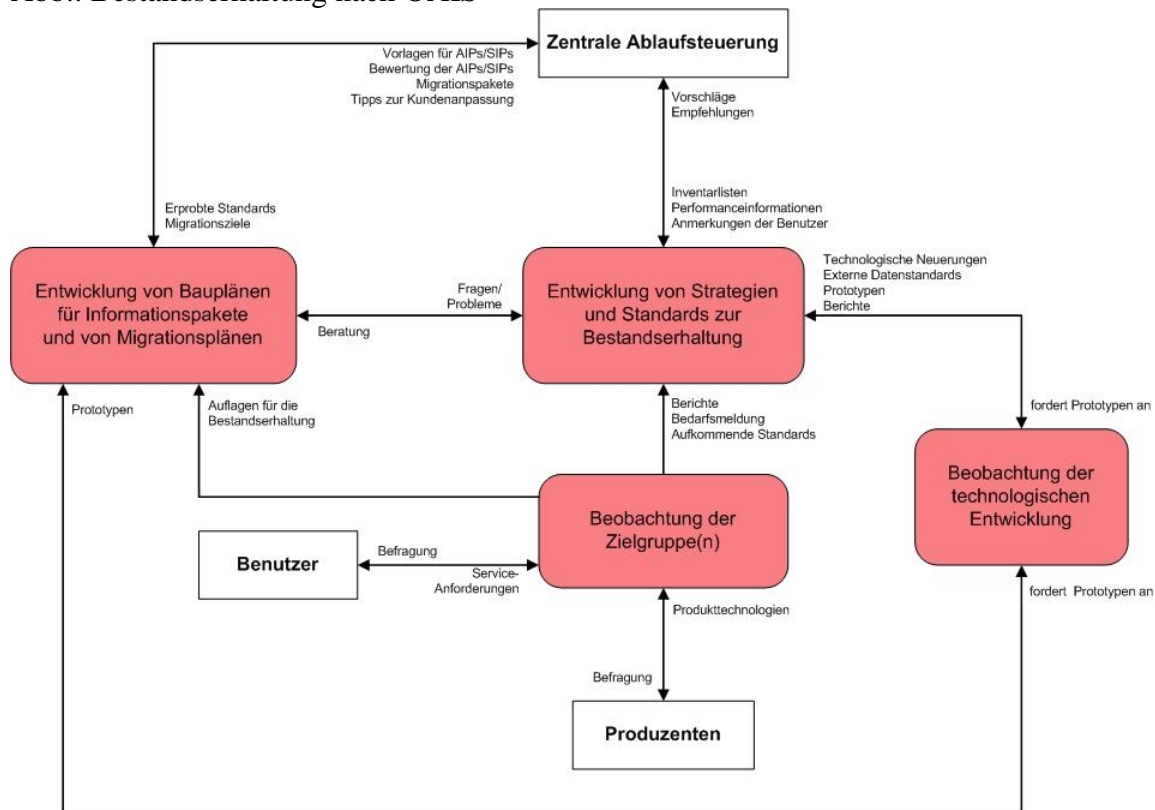
Mit Hilfe des Erschließungsmoduls werden die dem Archiv übergebenen Dokumente erfasst, und sind für eine spätere Nutzung recherchierbar.

Nr.	Anforderung	Wertigkeit	Kommentar
1	Das Erschließungsmodul stellt die Möglichkeit bereit, die spezifischen Metadaten (Erschließungsmetadaten) zu einzelnen Dokumenten unterschiedlicher Archivalien-gattungen zu erfassen.		Das Archiv stellt die Liste der gewünschten Metadaten individuell zusammen.
2	Die Erschließungsinformation ist im AIP mit den Primärdaten und anderen Metadaten verbunden.		
3	Das Erschließungsmodul ist gegen un-berechtigten Zugriff und unberechtigte Veränderungen geschützt.		
4	Autorisierte Personen authentifizieren sich über ein Berechtigungssystem.		Die Art und Weise der Authentifizierung und den Grad des Schutzes bestimmt das Archiv.

5	Veränderungen an den Einträgen werden protokolliert.		
6	Jeder Eintrag im Erschließungsmodul besitzt eine eigenen ID.		
7	Die dauerhafte Erhaltung der Erschließungsdaten und ihre Verknüpfung mit den jeweils zugehörigen Primärdaten ist gewährleistet.		
8	Alle von Menschen oder Maschinen vorgenommenen Veränderungen werden unter Nennung der Akteure und Ereignisse protokolliert.		

2.4 Anforderungen an die Bestandserhaltung

Abb.: Bestandserhaltung nach OAIS



Die Bestandserhaltung sorgt für die dauerhafte Verfügbarkeit der im elektronischen Magazin gelagerten Daten unter Wahrung ihrer Integrität (Unversehrtheit, Vollständigkeit) und Authentizität (Echtheit und Glaubwürdigkeit).

Planung und Wahl der zur Erreichung dieses Ziels geeigneten Maßnahmen erfolgen im Rahmen der Bestandserhaltung. Im Einzelnen müssen Strategien entwickelt werden für den Umgang mit den Informationspaketen (SIP, AIP, DIP), für Standards zur Bestandserhaltung sowie für die Notfallplanung des Digitalen Archivs.

Dafür muss das technische Umfeld, in dem das Digitale Archiv existiert, permanent und intensiv beobachtet werden, Fachabteilungen, also die Daten abgebenden Stellen, und ihr technisches Umfeld müssen ebenso in den Blick genommen werden wie Anforderungen der Zielgruppen (Nutzer) des Digitalen Archivs, um zukünftige Anforderungen frühzeitig zu erkennen und entsprechende Maßnahmen zu ihrer Erfüllung vorbereiten und umsetzen zu können.

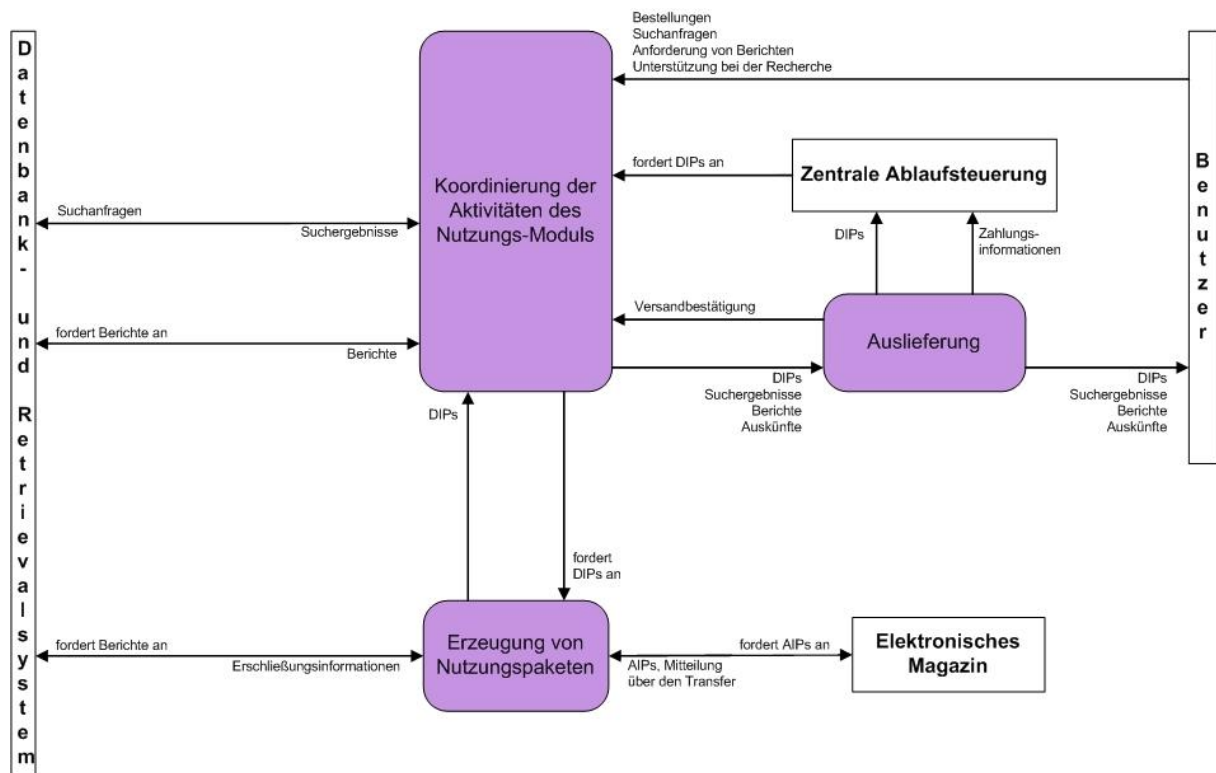
Nr.	Anforderung	Wertigkeit	Kommentar
1	Sofern ein Rechenzentrum (RZ) vorhanden ist, gibt es einen engen Austausch zwischen diesem und dem Archiv.		
2	Das Archiv verfügt über die notwendigen Mittel und das Archivpersonal über die notwendigen Kenntnisse zur Bestandserhaltung.		

3	Die Erstellung der Informationspakete erfolgt standardisiert.		
3.1	Das Archiv bestimmt die zur Archivierung geeigneten Formate für die AIPs (z.B. PDF/A für Textformate)		
3.2	Die Speicherung der Informationspakete, vor allem der AIPs erfolgt in geeigneten Formaten.		
3.3	Gegebenenfalls werden AIPs bei notwendigem Formatwechsel massenhaft in das neue geeignete Format migriert.		
4	Das Modul ist offen für neu entstehende Formate, für notwendige Software-Updates und Veränderungen an der Hardware.		
5	Die Speicherung der AIPs erfolgt auf geeigneten Trägern, z.B. RAID.		
5.1	Wenn möglich, erfolgt eine zweite, redundante Speicherung an einem anderen Ort, um bei einer eventuellen Havarie keinen Datenverlust zu erleiden.		
5.2	Wartungsroutinen und Backups werden regelmäßig durchgeführt.		
5.3	Die Ergebnisse werden protokolliert und dem Archiv mitgeteilt, insbesondere, wenn es ein Problem gab.		
5.4	Über Änderungen am Speichersystem wird das Archiv vorab informiert und greift beratend ein.		
5.5	Die Speicherung erfolgt Beweiswert erhaltend und Compliance-fähig mit Hilfe von fixity-Information und Hash-Werten.		Hier muss das Archiv in Abstimmung mit seinem Träger entsprechende Maßnahmen festlegen.
6	Die Daten produzierenden Stellen informieren das Archiv über neue operative Systeme.		
6.1	Das Archiv überprüft neue operative Systeme seiner Daten produzierenden Stellen auf die Möglichkeit, Daten daraus zur Archivierung übernehmen zu können.		Die Bedeutung der Vorfeldarbeit wird an diesem Punkt besonders deutlich. Möglichst frühzeitig ist sicher zu stellen, dass Daten überhaupt übernehmbar sind.
7	Trends zur künftigen Nutzung des Archivguts werden beobachtet. Dies umfasst Wünsche externer Nutzer ebenso wie Anforderungen, die sich durch den Archivträger oder das Archiv selbst ergeben.		

7.1	Maßnahmen aus neuen Nutzungsbedürfnissen werden eingeleitet.		
8	Über das Modul zur Bestandserhaltung erfolgen statistische Auswertungen zu Migrationen, Refreshings u.a. Maßnahmen.		
9	Akteure dieser Maßnahmen werden dokumentiert.		

2.5 Anforderungen an das Nutzungsmodul

Abb. Nutzungsmodul nach OAIS



Das Nutzungsmodul ermöglicht Recherchen durch externe und interne Nutzer des Digitalen Archivs. Es nimmt Anfragen und Bestellungen entgegen, leitet Anfragen an die Datenbank und das Erschließungsmodul weiter. Nach einer Überprüfung der Berechtigungen der anfragenden Person, wird das Suchergebnis generiert und die Bereitstellung der Informationspakete koordiniert.

Aus dem Suchergebnis werden im Nutzungsmodul ein oder mehrere Nutzungspakete (DIPs) generiert und ausgeliefert. Der Nutzer erhält dabei mit den Inhaltsdaten auch die zur Interpretation notwendigen Angaben, um die Informationen entsprechend auswerten zu können. Dies kann eine automatisierte Zusammenstellung von Daten zu einer bestimmten Anfrage sein, die unter Berücksichtigung der für den spezifischen Nutzer vorliegenden Berechtigungen vorgenommen wird.

Nr	Anforderung	Wertigkeit	Kommentar
1	Schnittstelle nach außen Das Nutzungsmodul verfügt über eine Schnittstellen nach außen zur Entgegennahme von Anfragen und Bestellungen.		
2	Schnittstelle nach innen Das Nutzungsmodul verfügt über eine Schnittstelle nach innen zum Erschließungsmodul zur Durchführung der Recherche		

3	Berechtigungen Das Nutzungsmodul überprüft bei jeder Anfrage, ob die anfragende Person die notwendigen Berechtigungen zur Nutzung des Systems besitzt.		
3.1	Das Nutzungsmodul überprüft, ob die anfragende Person berechtigt ist, das Rechercheergebnis zur Kenntnis zu erhalten		
3.2	Sperrfristen Das Nutzungsmodul prüft Sperrfristen		
4	Rechercheergebnis Das Nutzungsmodul generiert aus der Erschließungsdatenbank heraus das Rechercheergebnis: Liste der Archivsignaturen und zugehörige Metadaten		
5	Das Nutzungsmodul verfügt über eine Schnittstelle zum elektronischen Magazin		
5.1	Über die Schnittstelle zum elektronischen Magazin werden die Repräsentationen der gesuchten Archivalien bereitgestellt und dabei auf ihre Korrektheit (Fehlerfreiheit) überprüft.		
5.2	Die Bereitstellung erfolgt als Nutzungspaket (DIP), das die Repräsentation der gewünschten Archivdatei und Metadaten umfasst.		
5.3	Über die Ausgabeschnittstelle (s.o.) wird das DIP dem Nutzer, der sich entsprechend authentifiziert hat, zur Verfügung gestellt.		
5.4	Das DIP ist durch den Nutzer nicht veränderbar. Es kann einen Kopierschutz enthalten.		
6	Das Nutzungsmodul kann Statistiken zur Nutzung erzeugen: Liste der Nutzer, Liste der genutzten Archivalien, Liste der recherchierten Metadaten und Archivalien...		
7	Das Nutzungsmodul kann virtuelle Nutzerfächer für einzelne Nutzer anlegen, die zeitlich begrenzt für die Nutzung zur Verfügung stehen.		Das ist möglicherweise praktisch bei der Betreuung der Nutzer im Archiv, so muss nicht immer wieder eine gleich lautende Abfrage vorgenommen werden, sondern das Ergebnis bleibt zeitlich begrenzt im virtuellen Nutzungsfach gespeichert.
8	Die Akteure und ihre Handlungen werden protokolliert.		