

Strategien und Vorgehen beim Aufbau eines digitalen Langzeitarchivs

Tobias Wildi, t.wildi@docuteam.ch

VSA AAS, Zyklus Archivpraxis Schweiz, 19.4.2012

Vorgehen

1. Ziele und Grenzen des OAIS formulieren
2. Grobkonzept, Verbündete suchen, Projektantrag
3. Genehmigungsprozess, Finanzierung sicherstellen
4. Projektgruppe zusammenstellen. Anbindung an RM, IT und Legal sicherstellen
5. Feinkonzept: Abläufe, Verantwortlichkeiten, Werkzeuge, Infrastruktur
6. Umsetzung, Testphase
7. Initial-Ingest
8. Übergang zum Normalbetrieb

1. Ist-Situation, Ziele, Grenzen

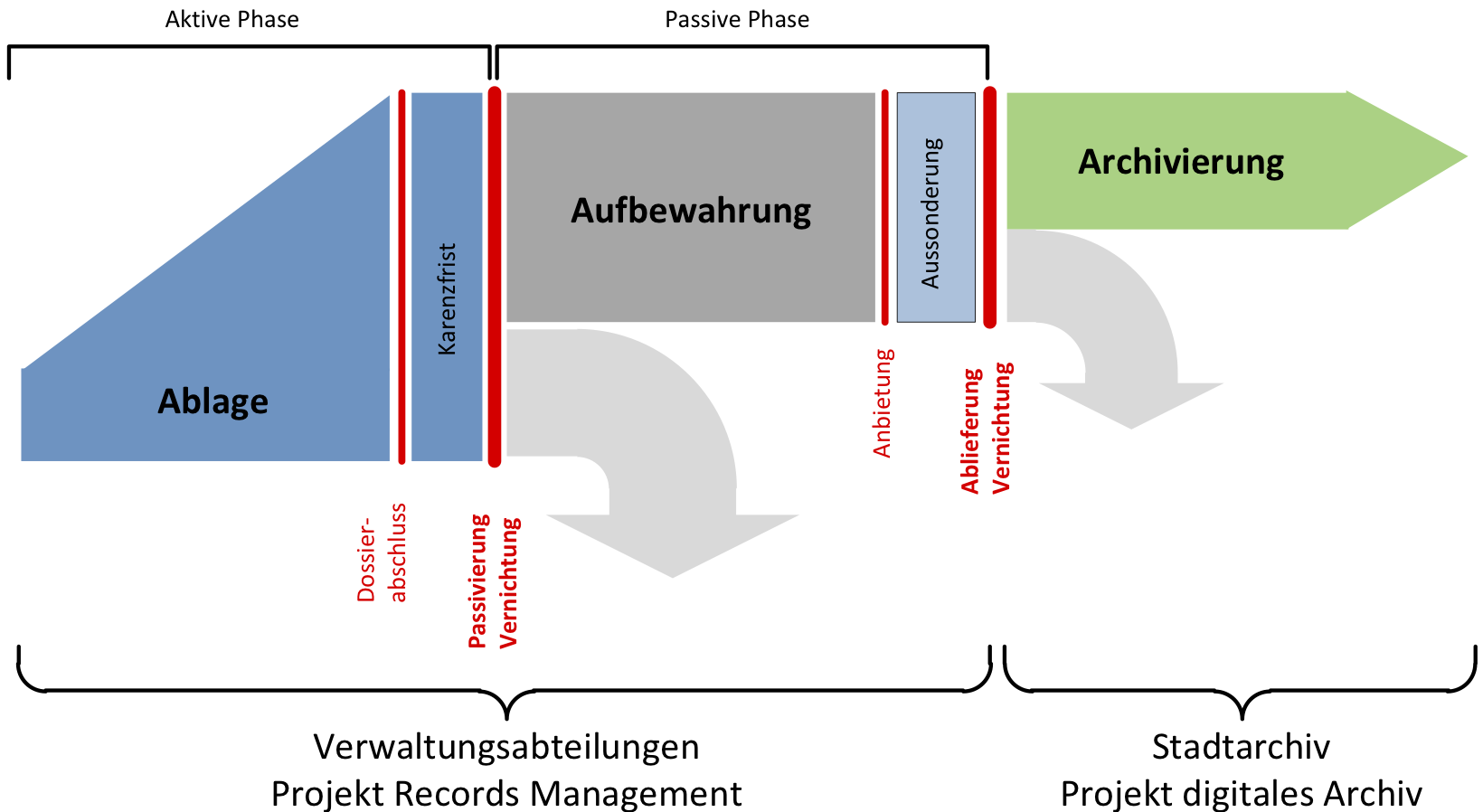
- Warum brauchen Sie ein digitales Langzeitarchiv? Welche Arten von Unterlagen können Sie nicht auf Papier archivieren?
- Handelt es sich um ein historisches Archiv oder bildet das Unternehmen zukünftig auch gesetzliche Aufbewahrungsfristen darin ab?
- Existiert ein ECM-Konzept oder RM-Policies? Wird dabei der Übergang zum digitalen Masterdossier angedacht?
- Welche führenden Systeme existieren zur Abbildung von Prozessen und über welche Schnittstellen verfügen sie zu Ihrem potentiellen LZA?
- Welche IT-Infrastruktur besteht bereits und kann genutzt werden?

2. Grobkonzept

- Begriffe festlegen, Dokumenten-Lifecycle definieren (Systeme, Zuständigkeiten, Systemgrenzen)
- Abbildung der Aufgaben eines OAIS:
 - **Sammeln:** Bewertung, Datenübernahme,
 - **Archivieren:** Datenaufbereitung, Ingest-Workflow, Erschliessung
 - **Speichern:** Repository, Bitsream-Preservation, Storage-Infrastruktur
 - **Zugang:** «digitaler Lesesaal», in welcher Form können die Datenlieferanten wieder auf ihre Daten zugreifen?
- Rollen, Aufgaben, Verantwortlichkeiten, Arbeitsteilung:
Technisch vs. inhaltlich / redaktionell arbeitende Archivare

Lifecycle definieren

In welchem grösseren Zusammenhang steht das Archiv?



Business Services

Erstellung und Ablage

Struktur u. Logik

Abgeschlossene Dokumente

Zwischenarchiv (ZA)

Struktur u. Logik

Dokumente für unbestimmte Zeit archivieren

Langzeitarchiv (LA)

Struktur u. Logik

Application Services

Organisation, Logik & Management

Ordnungssystem = OS
Abbildung von Retention Schedules, Zugriffsrechten, Vertraulichkeit/Datenschutz
Abbildung gesetzlicher Vorgaben (GeBü, MWST, ...)

Sharepoint Benutzt die Integration des Archivsystem, wird als Userinterface „Bearbeitung, Dynamik von Daten“ genutzt. Sharepoint benutzt den Connector, welcher vom Archivsystem zur Verfügung gestellt wird.

- Workflowengine
- Enterprise Portal
- Collaboration
- Userinterface
- Metatags Management
- Application Integration
- Documentmanagement
- Searchengine

Ordnungssystem

Speichern, Dynamik, Betrieb

Löschen von nicht dauerhaft archivwürdigen Unterlagen

Verschieben von dauerhaft archivwürdigen Unterlagen

Zugriff auf ZA

Duet

Fachanwendung

Fachanwendung

SAP Fachanwendung

Funktionen

Langzeitarchiv
Dauerhafte Archivierung von historisch wertvollen Unterlagen nach ISO 14721 (OAIS, Open Archival Information System).

Ingest-Werkzeuge (Fachanwendung Docuteam)

- Qualitätsprüfung
- Datenmigration
- Metadaten-Extraktion
- Datenorganisation

scopeArchiv Fachanwendung für die Verwaltung von beschreibenden Metadaten, Katalogsystem sowohl für das digitale wie auch für das konventionelle Papierarchiv.

Base Services

Filenet P8
Eine integrierte Plattform für Content, Process und Compliance

Ordnungssystem

Archivmanagement

Archivlink

Archivwürdigen Unterlagen +Unterlagen mit einer Archivierungsfrist

Ablieferung von dauerhaft zu archivierenden Daten + Metadaten nach Ablauf der gesetzlichen und administrativen Fristen.

Recordmanagement

Federation und Integration

Compliance Management

Web content management

Documentmanagement

Imagemanagement

Forms management

Report management

Data capture

Data classification

Applications Services

ECM Infrastruktur

Fedora Commons

Digitales LZ Repository

Fedora Commons: Verwalten von digitalen Objekte. Speicherung direkt auf der Storage-Infrastruktur (WORM).

- Verwaltung digitale Objekte
- Auslieferung via APIs
- Versionierung
- Indizierung
- Integritätssicherung

Platform Services

Storage Software

- Filesystem
- Archivfunktion
- Filelevelvirtualisation
- Scale out/ Gridsoftware
- Volumemanager
- Backupfunktion
- Blocklevelvirtualisation

Storage Hardware

- SAN
- NAS
- Highperformance Disk
- Low cost Disk
- Tape

Protected / WORM (EMC Centera)

2. Grobkonzept: Was kostet ein OAIS?

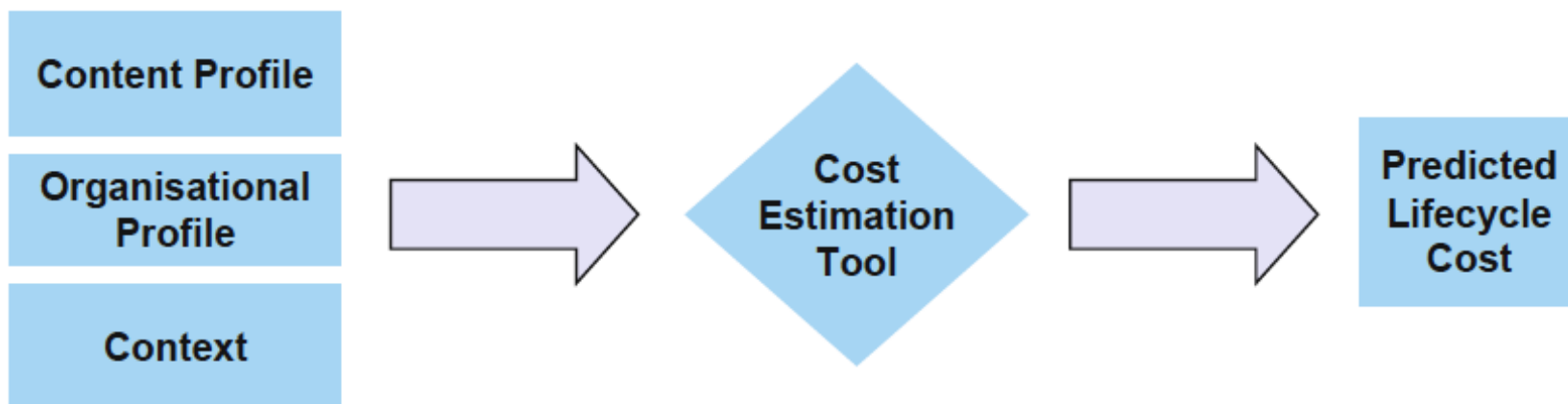
- Unterscheiden Sie zwischen
 - Investitionskosten
 - Betriebskosten
- LIFE-Modell für die Kostenkalkulation ermittelt die Kostentreiber, die bei der digitalen Archivierung entstehen
- Link für das Kostenberechnungs-Werkzeug in Excel (Case-Studies):
<http://www.life.ac.uk/2/documentation.shtml>



LIFE³: Estimating preservation costs

■ The LIFE³ Project:

- Aim: To develop the ability to estimate preservation costs across the digital lifecycle
- The Project is developing:
 - A series of costing models for each stage and element of the digital lifecycle
 - An easy to use costing tool
 - Support to enable easy input of data
 - Integration to facilitate use of the results



LIFE3 costing tool inputs

Content Profile

Organisational
Profile

Context

- **Content Profile**
 - Type of Content
 - File format
 - Complexity
 - Volume

- **Organisational Profile**
 - Existing infrastructure
 - Preservation policy
 - Legal constraints

- **Context**
 - Inflation
 - Hardware costs and trends
 - Staff costs

LIFE³ costing tool outputs – estimated costs

Lifecycle Stage	Creation or Purchase	Acquisition	Ingest	Bit-stream Preservation	Content Preservation	Access
	...	Selection	Quality Assurance	Repository Admin	Preservation Watch	Access Provision
Lifecycle Elements	...	Submission Agreement	Metadata	Storage Provision	Preservation Planning	Access Control
	...	IPR & Licensing	Deposit	Refreshment	Preservation Action	User Support
	...	Ordering & Invoicing	Holdings Update	Backup	Re-ingest	
	...	Obtaining	Reference Linking	Inspection	Disposal	
	...	Check-in				

Sammeln

Archivieren

Speichern

Zugang

3. Genehmigungsprozess, Finanzierung sicherstellen

- Mögliche Realisierungsstrategien:
 - Klein anfangen, OAIS nach und nach komplettieren und skalieren
 - Mit dem gesamten OAIS in den politischen Prozess gehen
- Bis zu welcher Grössenordnung können Sie Projekte selber vergeben, ab wann müssen Sie ein Ausschreibungsverfahren starten?
- Deklarieren Sie das digitale Archiv als eine neue Fachanwendung oder als neue Infrastruktur, resp. Querschnittsaufgabe?
- **Die Chance im Lift:** Erklären Sie in 30 Sekunden, warum es ein digitales Langzeitarchiv braucht.
- Verzweifeln Sie nicht über lange Genehmigungsprozesse. Nutzen Sie die Zeit, um konzeptuelle Arbeiten voranzutreiben.

4. Projektorganisation für die Umsetzung

- Legen Sie beim Projektstart Wert auf einheitliche Begriffsdefinitionen.
- Erarbeiten Sie eine Kurzschulung für Neumitglieder der Projektgruppe. Die Zusammensetzung solcher Gruppen ändert immer wieder.
- Legen Sie eine strukturierte und für alle Gruppenmitglieder zugängliche Projektablage an. Bauen Sie eine Dokumentation auf.
- Suchen Sie Verbündete bei...
 - IT-Infrastruktur und -Architekten
 - IT-Security, Business Continuity
 - Rechtsabteilung, Legal & Compliance
 - QM, PM
- Umschiffen Sie pessimistische Besserwisser

5. Feinkonzept, Konkretisierung

- Definieren Sie Schnittstellen für SIP, AIP, DIP und verändern Sie diese möglichst nicht mehr (oder nur abwärtskompatibel).
- Nutzen Sie für Metadaten weit verbreitete Standards und Formate:
 - ICA-Normen: ISAD(G), ISAAR(CPF), ISDF
 - EAD, EAC
 - METS (Metadata Encoding and Transmission Standard), XFDU
 - Premis für technische/administrative Metadaten
- **Die Nutzung weit verbreiteter Normen und Standards senkt die Kosten in der Umsetzung.** Denn es gibt dafür bereits zahlreiche Werkzeuge, die Sie nutzen können.
- Bauen sie auf bereits bestehender Infrastruktur auf (z. B. Speicherinfrastruktur, Rechenzentrum, schnelle Netzwerke) und nutzen Sie diese für Ihr OAIS.

METS / Premis / EAD in Docupack

<http://wiki.docuteam.ch>

The screenshot shows the Docupack application window with the following components:

- File Tree (Left):** A hierarchical view of folders and files. The 'Ablieferung' folder is selected, containing subfolders A, B, and C, each with further subfolders (A1, A2, B1, B2, C1, C2, C3, C4). Files are listed with their names and sizes.
- Info Panel (Right):** A metadata editor with tabs for 'Info' and 'Vorschau'. The 'Info' tab is active, showing a table of 'Eigenschaften:' (Properties) and 'Ereignisse:' (Events). The 'Stufe:' (Level) property is circled in red.

Eigenschaften:	
AkzessionsNr:	
Obj:	
Stufe:	Serie
Name:	Ablieferung
Pfad:	Ablieferung
Typ:	rootfolder
MIME-Typ:	
Format:	
Format-ID:	
Größe [kB]:	5568
Größe [%]:	100
Kinder:	3
Nachkommen:	14
Ereignisse:	2

Ereignisse:		
Zeitpunkt:	Typ:	Ergebnis:
2012-02-06T14:28:16	Creation	Success
2012-02-06T14:28:47	Renaming	Success

Ereignisdetails:	
Zeitpunkt:	2012-02-06T14:28:47
Typ:	Renaming
Detail:	Manually
Ergebnis:	Success
Ergebnisdetails:	

The screenshot shows the Docupack application window with the following components:

- File List (Left):** A table showing a list of files and folders. The file 'BBC_5276.jpg' is selected.
- Preview Panel (Right):** A large area showing a preview of the selected file, 'BBC_5276.jpg', which is a black and white photograph of a cityscape.

...	...	G...	...	Größe/%:
3	15	6'172	...	
2	5	5'568	90	
2	2	0	0	
4	5	604	9	
0	0	0	0	
0	0	0	0	
0	0	0	0	
1	1	604	9	
0	0	604	9	

EAD / EAC in ICA AtoM

<http://ica-atom.org>

English | français | español | Nederlands | português | polski | slovenščina home | about | help | my profile | log out

ICA AtoM

Advanced search

archival descriptions | authority records | archival institutions | functions | subjects | places | digital objects

Add | Manage | Import | Admin

View archival description

Fonds Add. MSS. 1399 - Townley, Matheson and Partners fonds (draft)

Identity area

Reference code	CA CVA Add. MSS. 1399
Title	Townley, Matheson and Partners fonds
Date(s)	1897-1990 (Creation)
Level of description	Fonds
Extent and medium	ca. 5000 architectural drawings 653 photographs 44 cm textual material

Context area

Name of creator	Townley, Matheson and Partners
-----------------	--------------------------------

Administrative history


Townley, Matheson and Partners was a major Vancouver architectural firm throughout most of the 20th century. It produced designs for over 1100 discrete projects including numerous landmark buildings in the Vancouver, most notably Vancouver City Hall. The partnership was formed in 1919 by Fred L. Townley (1887-1966) and Robert T. Matheson (1887-1935).

Townley was the son of former Vancouver Mayor T.O. Townley. Townley and Matheson were both raised in Vancouver, and received their architectural training at the University of Pennsylvania, graduating together in 1911.

Archival institution
City of Vancouver Archives

Creator(s)
Townley, Matheson and Partners

Digital object



Famous Cloak and Suit Com...

Fonds

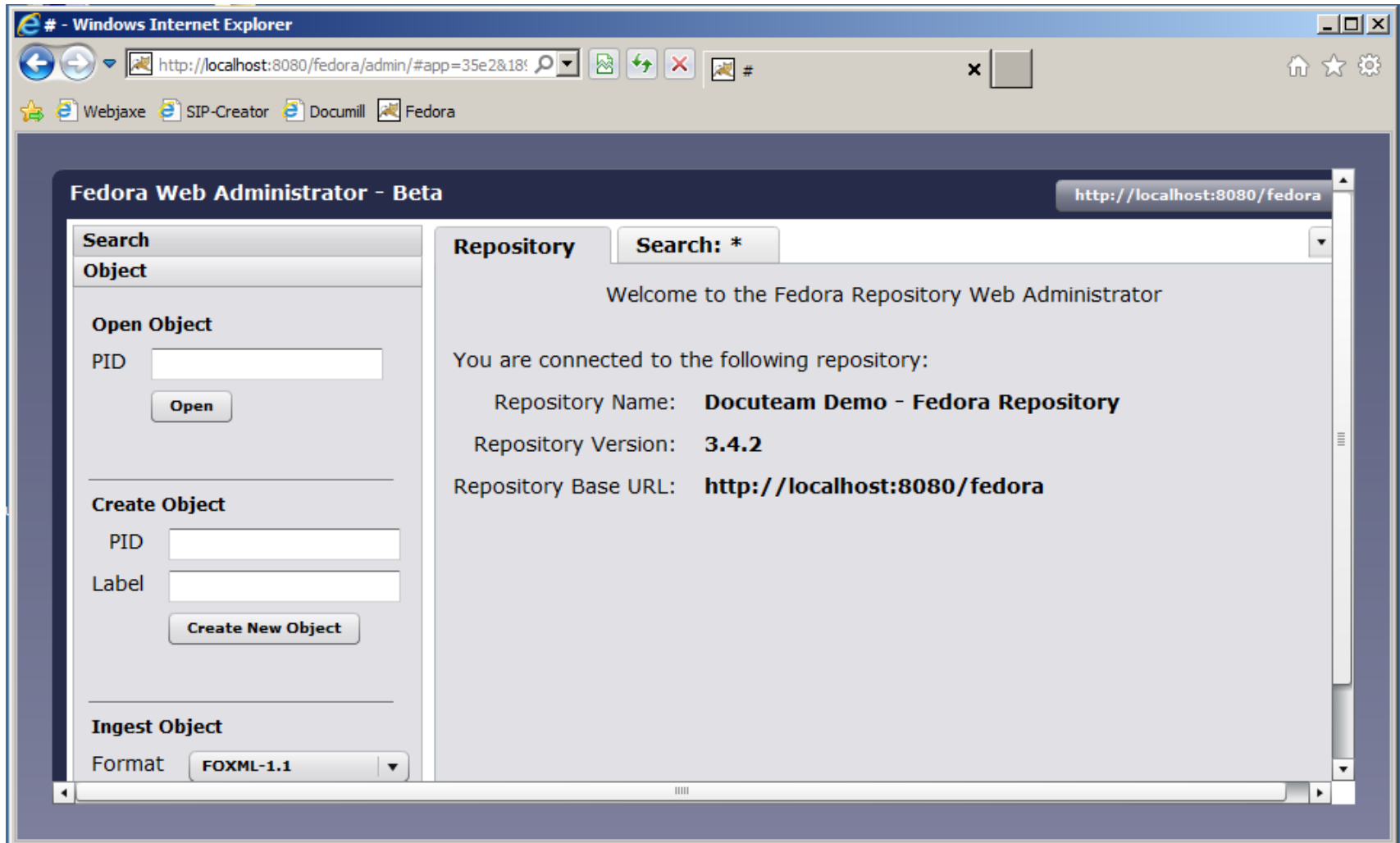
- Fonds Add. MSS. 1399 - Townley, Matheson and Pa...
- Series 12 - Architectural drawings
- Series 13 - Project photographs
- Series 14 - Miscellaneous records

Export

- Dublin Core 1.1 XML
- EAD 2002 XML

METS / DC in Fedora Commons

<http://www.fedora-commons.org>



The screenshot shows a Windows Internet Explorer browser window displaying the Fedora Web Administrator interface. The browser's address bar shows the URL `http://localhost:8080/fedora/admin/#app=35e2&18!`. The interface is titled "Fedora Web Administrator - Beta" and includes a search bar with the text "Search: *".

The main content area displays the following information:

- Welcome to the Fedora Repository Web Administrator
- You are connected to the following repository:
- Repository Name: **Docuteam Demo - Fedora Repository**
- Repository Version: **3.4.2**
- Repository Base URL: **http://localhost:8080/fedora**

The left sidebar contains three sections:

- Search Object**: Includes an "Open Object" section with a "PID" input field and an "Open" button.
- Create Object**: Includes a "PID" input field, a "Label" input field, and a "Create New Object" button.
- Ingest Object**: Includes a "Format" dropdown menu currently set to "FOXML-1.1".

6. Umsetzung

- Haupttätigkeit aus Sicht Archivar: Bewertung von Datenbeständen und der Aushandlung von Ablieferungsvereinbarungen
- Aufbau der Plattformen (Rechenleistung, Speicher, Netzwerkdienste)
- Einrichten, gegebenenfalls Programmierung von Ingest-Werkzeugen und Metadatenverwaltung
- Aufbau des digitalen Lesesaals. Kann zunächst sehr einfach gehalten werden und dann nach sich abzeichnenden Benutzerbedürfnissen erweitert werden.

6. Testphase

- Testen Sie möglichst unterschiedliche Fileformaten, Schnittstellen, Ursprungsapplikationen.
Es wird Überraschungen geben!
- Testen Sie Ihre Migrationswerkzeuge...
 - Gegen aktuelle Validatoren (Beispiel PDF/A)
 - Gegen die File-Registry Pronom mit Hilfe von Droid
- Testen Sie Ihre Repository-Umgebung...
 - Ingest von grossen Dateien (Videos)
 - Ingest von grossen Mengen von Dateien (Websites)
 - Hohe Zugriffsrate bei Webapplikationen

7. Betriebsstart, Initial-Ingest

- Überführung von bereits bestehenden digitalen Bestände in die neue Archiv-Infrastruktur.
- Konfiguration von Fachapplikationen, die ihren Output neu ebenfalls im OAIS ablegen sollen (z. B. Belegarchivierung aus SAP)
- Übernahme von Daten aus flüchtigen Medien (z. B. CD-Stapel aus Digitalisierungsprojekten)
- Stellen Sie sicher, dass bisher auf Papier geführte Serien im OAIS nahtlos weitergeführt werden.
- Der Initial-Ingest muss nicht sofort geschehen, sondern Altdaten können nach und nach übernommen werden.

8. Übergang zum Normalbetrieb

- Jährlicher Aufwand für Ingest/Erschliessung auf der Basis der Ablieferungsvereinbarungen und eines Ablieferungsfahrplans.
- Ausserordentliche Datenübernahmen aus Systemen, die «End of Life» sind.
- Erweiterung und Öffnung des digitalen Lesesaals gegen aussen
- Betrieb der Infrastruktur, insbesondere Überwachung der Speichermenge und gegebenenfalls Ausbau.

Aufgabe für die Gruppenarbeit

- **Ausgangslage:** Sie sind in einem Staatsarchiv für die Erarbeitung eines Grobkonzepts für ein digitales Archiv verantwortlich.
- Das Grobkonzept nimmt Stellung zu folgenden Punkten:
 - Lifecycle; Aufgabenteilung und Synergien zwischen Records Management und OAIS.
 - Grundsätze zur Archivstrategie und zur Bewertung
 - Archivische Standards und Formate, auf denen Sie aufbauen
- Konzipieren Sie die Hauptaufgaben des OAIS:
 - **Sammeln:** Bewertung, Datenübernahme,
 - **Archivieren:** Datenaufbereitung, Ingest-Workflow, Erschliessung
 - **Speichern:** Repository, Bitsream-Preservation, Storage-Infrastruktur
 - **Zugang:** Aufgabe am Nachmittag