



Digitales Stadtarchiv Baden auf docuteam cosmos

Tobias Wildi, t.wildi@docuteam.ch

Memoriav Fachtagung + Workshops 19.3.2019 Digitale Archivspeicher/Archivage numérique



Docuteam GmbH, Baden Informationsmanagement und Archivdienstleistungen

- 2003 gegründet von den beiden Inhabern Tobias Wildi und Andreas Steigmeier
- 35 Mitarbeitende (26 Vollstellen)
- Standorte Baden und Yverdon
- Wir decken den ganzen Lebenszyklus von Unterlagen ab digital und analog.
- Seit 2016 «docuteam cosmos», Cloud-Lösung für digitale Archivierung

Informationsmanagement

 Analyse, Konzeption und Umsetzung von Information Lifecycle Management (Records Management, GEVER)

Beratung

Archivinformatik

- Konzeption von digitalen Archiven
- Entwicklung und Integration von Software für Langzeitarchive
- Betrieb Hostingplattform docuteam cosmos

Beratung/Informatik

Archivbewirtschaftung

- Bewertung und Erschliessung von Archivbeständen
- Bewirtschaftung von Archiven in Dauermandaten

Praktische Archivarbeit analog und digital

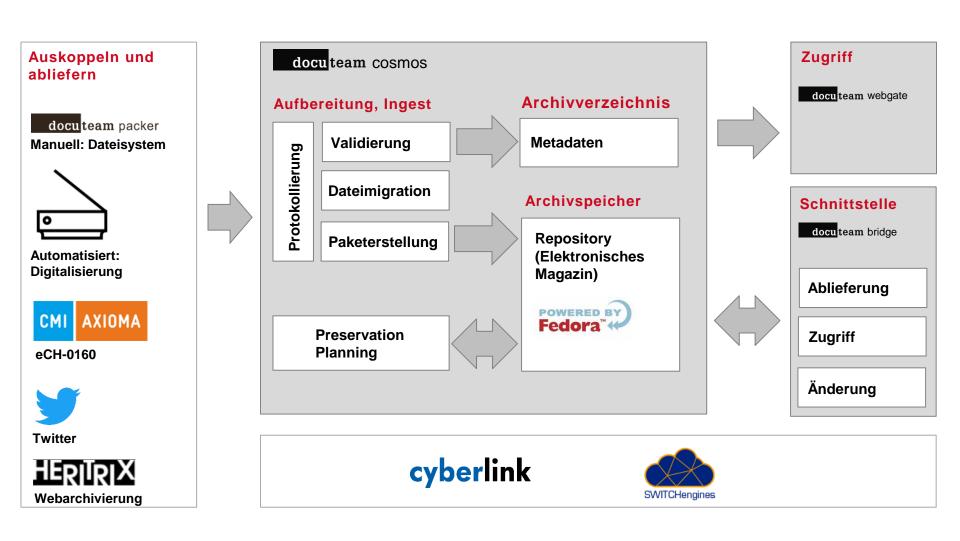


Digitales Stadtarchiv Baden

- Das Stadtarchiv Baden wird seit 2003 durch docuteam betreut.
- 2006: Vorstudien, Ideen zum Aufbau eines digitalen Archivs
- 2007/2008 Pilotprojekt, Machbarkeitsstudie
- 2011 Inbetriebnahme des produktiven Systems basierend auf:
 - docuteam curator (= internes AIS von docuteam)
 - docuteam packer f
 ür Vorbereitung der Ablieferungen
 - docuteam feeder f
 ür Ingest und Preservation Planning
 - Fedora Commons als Repository-System
- 2016 Migration der Infrastruktur auf eine Cloud-basierte Lösung (docuteam cosmos)



Digitales Stadtarchiv Baden auf docuteam cosmos

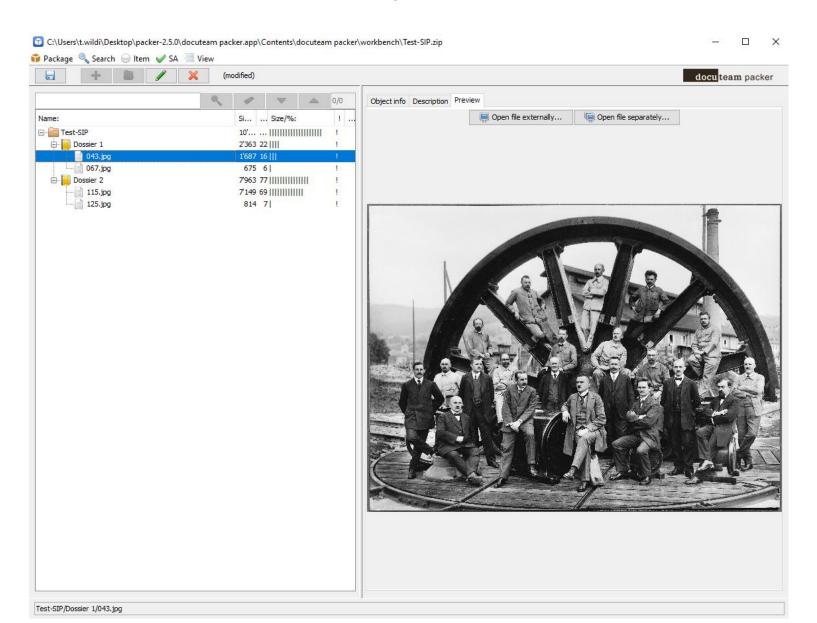




ABLIEFERN

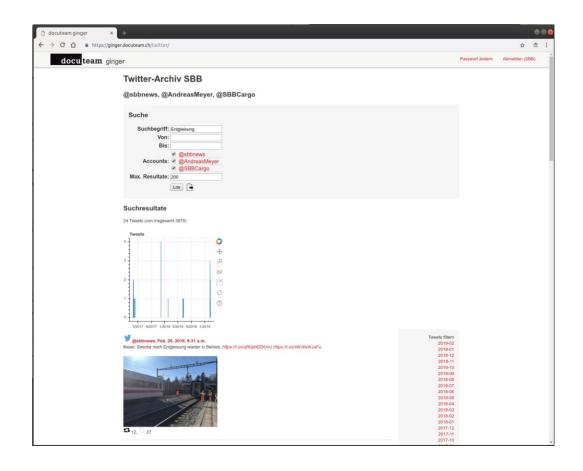


docuteam packer Abliefern aus dem Dateisystem

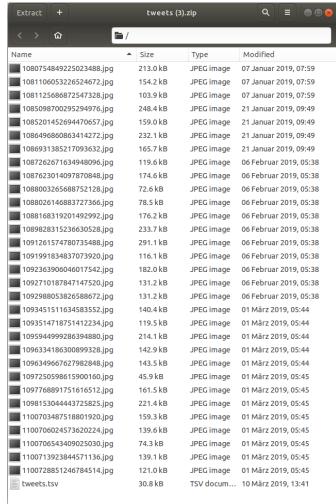


docuteam ginger: Twitter-Archivierung





- Post-Inhalt und Metadaten einer bestimmten Zeitperiode in einer Tabelle
- Bilder über relative Pfade verlinkt





docuteam feeder: Automatisierte Übernahme aus Digitalisierung

https://cosmos-ingest.docuteam.ch/organizations/6/workflow_executions

₩.

docu team feeder

Baden -

Ausführungen

Workflows

Cockpit

Benutzer

Admin ▼

Ausführungen

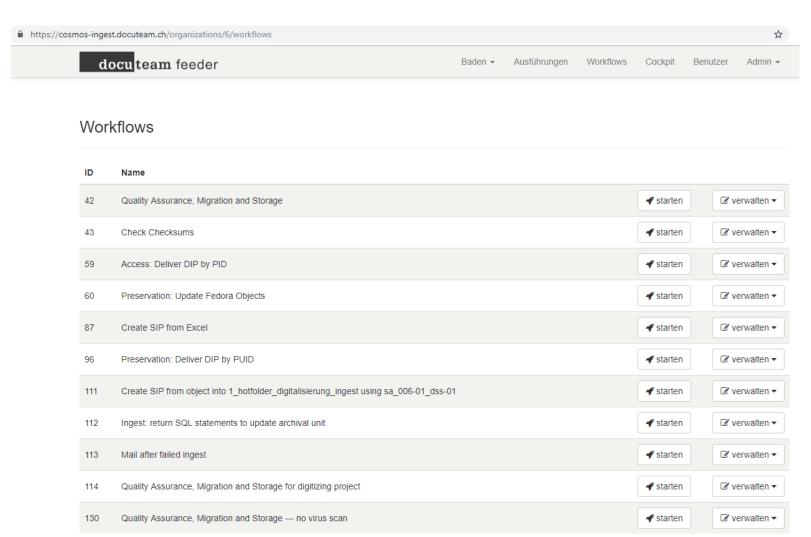
alle wartend laufend fertig fehlgeschlagen neu gestartet						
ID	Workflow	SIP	Status	Startzeit	Dauer	
79393	Quality Assurance, Migration and Storage for digitizing project	Q.01.18045	•••••	Mittwoch, 06. März 2019, 16:16 Uhr	00:01:06	anzeigen löschen
79392	Quality Assurance, Migration and Storage for digitizing project	Q.01.13871C	•••••	Mittwoch, 06. März 2019, 16:15 Uhr	00:01:02	anzeigen löschen
79391	Quality Assurance, Migration and Storage for digitizing project	Q.01.13871B	•••••	Mittwoch, 06. März 2019, 16:14 Uhr	00:01:01	anzeigen löschen
79390	Quality Assurance, Migration and Storage for digitizing project	Q.01.24102C	•••••	Mittwoch, 06. März 2019, 16:13 Uhr	00:00:58	anzeigen löschen
79389	Quality Assurance, Migration and Storage for digitizing project	Q.01.13871A	•••••	Mittwoch, 06. März 2019, 16:12 Uhr	00:00:59	anzeigen löschen
79388	Quality Assurance, Migration and Storage for digitizing project	Q.01.11139D	•••••	Mittwoch, 06. März 2019, 16:11 Uhr	00:01:00	anzeigen löschen
79387	Quality Assurance, Migration and Storage for digitizing project	Q.01.11139C	•••••	Mittwoch, 06. März 2019, 16:10 Uhr	00:00:59	anzeigen löschen
79386	Quality Assurance, Migration and Storage for digitizing project	Q.01.11139B	•••••	Mittwoch, 06. März 2019, 16:09 Uhr	00:00:58	anzeigen löschen
79385	Quality Assurance, Migration and Storage for digitizing project	Q.01.11139A	•••••	Mittwoch, 06. März 2019, 16:08 Uhr	00:00:58	anzeigen löschen



VALIDIEREN UND AUFBEREITEN



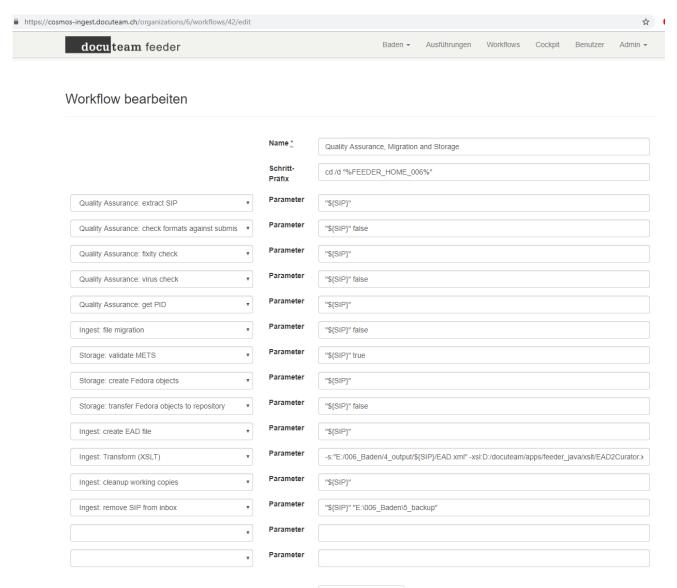
Ingest mit docuteam feeder: Validierung, Dateimigration, Speicherung



Neuen Workflow definieren



docuteam feeder: Bearbeitung einer Workflow-Pipeline



docu team

docuteam feeder: Überwachung der Ingest-Prozesse

https://cosmos-ingest.docuteam.ch/organizations/6/workflow_executions?page=225

docuteam feeder

Baden
Ausführungen

Workflows

Cockpit

Benutzer

Admin
A

Ausführungen

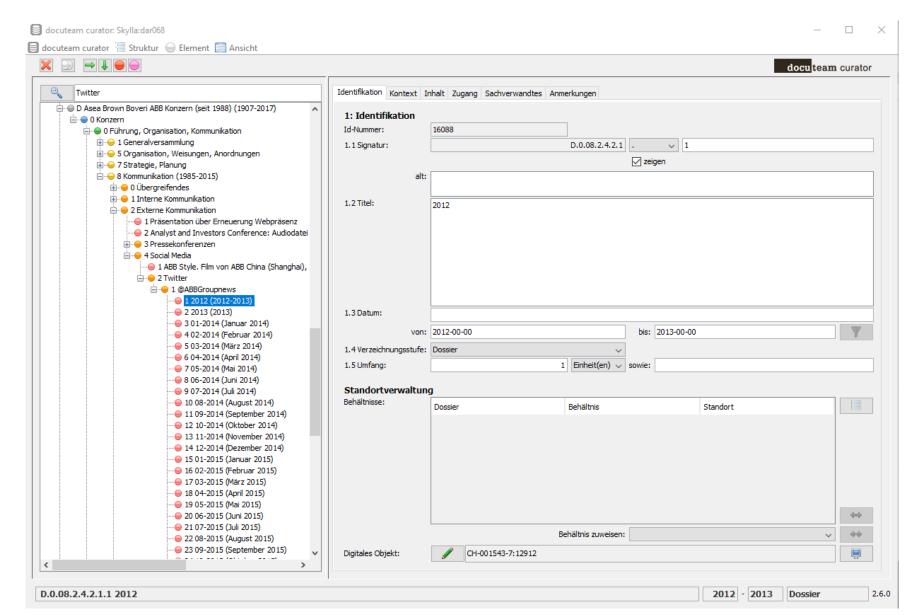
alle wartend laufend fertig fehlgeschlagen neu gestartet						
ID	Workflow	SIP	Status	Startzeit	Dauer	
59777	Create SIP from object into 1_hotfolder_digitalisierung_ingest using sa_006-01_dss-01	E:/006_Baden/1_inbox/Q.01.112.tif	•	Donnerstag, 06. Dezember 2018, 17:22 Uhr	00:00:13	anzeigen löschen
59742	Create SIP from object into 1_hotfolder_digitalisierung_ingest using sa_006-01_dss-01	E:/006_Baden/1_inbox/Q.01.163.tif	•	Dienstag, 04. Dezember 2018, 21:51 Uhr	00:00:05	anzeigen löschen
59596	Quality Assurance, Migration and Storage for digitizing project	Q.01.163	•••••	Dienstag, 04. Dezember 2018, 18:17 Uhr	00:00:19	anzeigen löschen
59595	Create SIP from object into 1_hotfolder_digitalisierung_ingest using sa_006-01_dss-01	E:/006_Baden/1_inbox/Q.01.163.tif	•	Dienstag, 04. Dezember 2018, 18:17 Uhr	00:00:04	anzeigen löschen
59566	Quality Assurance, Migration and Storage	006_Q-01-12399-A_20181204_we	•••••	Dienstag, 04. Dezember 2018, 11:21 Uhr	00:00:24	anzeigen löschen
59561	Quality Assurance, Migration and Storage	006_Q-01-11875_20181204_we	•••••	Dienstag, 04. Dezember 2018, 09:57 Uhr	00:00:24	anzeigen löschen
9088	Access: Deliver DIP by PID	CH-000249-X:681210	•	Dienstag, 13. November 2018, 15:25 Uhr	00:00:20	anzeigen löschen
1210	Preservation: Update Fedora Objects	CH-000249-X_666089	••••	Donnerstag, 08. November 2018, 12:11 Uhr	00:00:27	anzeigen löschen
1209	Access: Deliver DIP by PID	CH-000249-X:666089	•	Donnerstag, 08. November 2018, 12:03 Uhr	00:00:19	anzeigen löschen
1208	Access: Deliver DIP by PID	CH-000249-X:666089	•	Donnerstag, 08. November 2018, 08:21 Uhr	00:00:10	anzeigen löschen
937	Preservation: Deliver DIP by PUID	x-fmt/111	•	Montag, 15. Oktober 2018, 09:39 Uhr	00:00:12	anzeigen löschen
936	Preservation: Deliver DIP by PUID	x-fmt/111	•	Montag, 15. Oktober 2018, 05:55 Uhr	00:00:04	anzeigen löschen
935	Preservation: Deliver DIP by PUID	x-fmt/111	•	Montag, 15. Oktober 2018, 05:53 Uhr	00:00:07	anzeigen löschen
932	Preservation: Deliver DIP by PUID	x-fmt/111	•	Montag, 15. Oktober 2018, 05:37 Uhr	00:00:08	anzeigen löschen



ERSCHLIESSEN, METADATENIMPORT

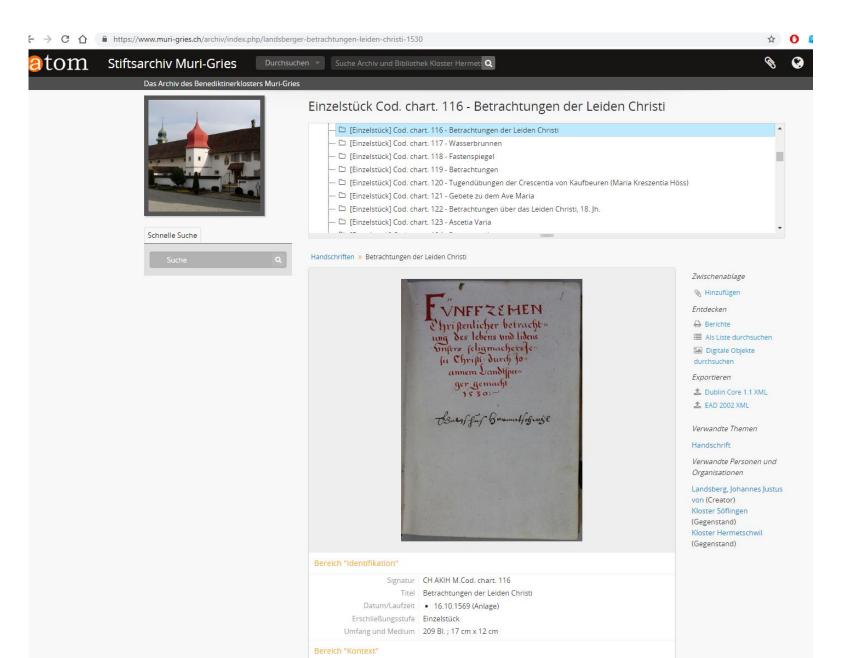


docuteam curator: Archivinformationssystem



Alternative: AtoM (Access to Memory, Artefactual)



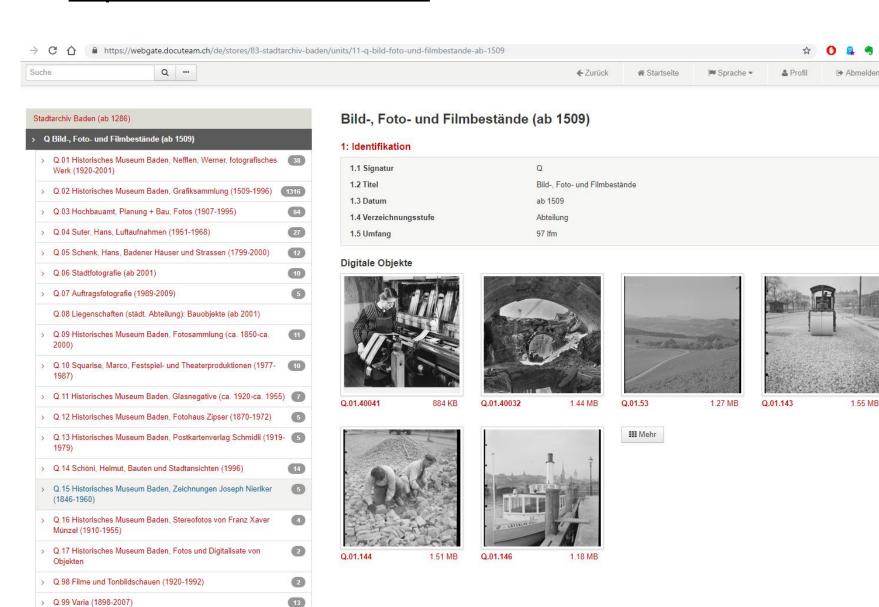


Benutzung, Recherche: http://baden.docuteam.ch



O Hilfe

1.55 MB



Zugriff, Benutzerrechte, Schutzfristen: Abgelieferte Zeugnisdossiers im digitalen Archiv

docu team

In diesem Beispiel: Öffentlich zugänglich ist die Information, dass Zeugnisdossiers sich im Stadtarchiv befinden. Geschützt ist die Dossierliste mit Personendaten. Geschützt sind auch die Nutzdaten (die Zeugnisdokumente selbst).





ÜBERWACHEN, PRESERVATION PLANNING

Baden ▼

docu team feeder

Stadtarchiv Baden

Berichte ▼

Speicherplatz			
test-006	170 Bytes		
CH-000249-X	2,4 TB		
Total	2,4 TB		
Frei	430 GB		

Letzte Integritätsprüfung		
	_	

Basisdaten	
Name	Stadtarchiv Baden
Basis-URL	https://cosmos-fedora006.docuteam.ch:8443/fedora
Version	3.8.1
PID Namespace	CH-000249-X
Support-Mail	support@docuteam.ch

Häufigste Dateiformate		
JP2 (JPEG 2000 part 1)	355554	57,83 %
Tagged Image File Format	141696	23,05 %
Extensible Markup Language	70339	11,44 %
Acrobat PDF/A - Portable Document Format	45302	7,37 %
Exchangeable Image File Format (Uncompressed)	1307	0,21 %
Plain Text File	98	0,02 %
Digital Video	61	0,01 %
Waveform Audio (WAVEFORMATEX)	59	0,01 %
Hypertext Markup Language	57	0,01 %
MPEG-4 Media File	52	0,01 %

Ablieferungen		
test-006	1	
CH-000249-X	7717	
Total	7718	

Objekte			
test-006	2		
CH-000249-X	623303		
Total	623305		



IST DAS SICHER?



Ist Datenhosting in der Cloud sicher?

- Für das Datenhosting nutzen wir SWITCHengines, die Cloudinfrastruktur von SWITCH (https://www.switch.ch/). Wir binden diesen stabilen und redundanten Speicher an unsere Repository-Systeme an.
- SWITCH wurde 1987 als privatrechtliche, nicht gewinnorientierte Stiftung vom Bund und den damals acht Hochschulkantonen gegründet.
- Die Infrastruktur von SWITCHengines (https://www.switch.ch/engines/) läuft in zwei SWITCH-eigenen Rechenzentren in Zürich und Lausanne.
- Der gesamte Technologie-Stack ist bekannt und dokumentiert, das schafft Vertrauen.
- SWITCH bietet Verträge mit langen Kündigungsfristen und Laufzeiten an. Kurze Kündigungsfristen stellen bei anderen Cloudanbietern oft ein Problem für Archivanwendungen dar.



Ist archivieren bei einem privaten Unternehmen sicher?

- Die in docuteam cosmos archivierten Daten und Metadaten gehören dem Archiv und nicht docuteam. Wir liefern Archivpakete aus dem Archivspeicher in standardisierter Form wieder aus.
- Wir bieten für Schnittstellen an, über welche Archive die Daten zusätzlich in Ihr eigenes Backupsystem integrieren können. So liegt eine zusätzliche Kopie der Daten und Metadaten lokal.
- Wir wollen keine Abhängigkeit von unserem System schaffen. Die exportierten Archivpakete basieren auf den internationalen Standards METS, EAD, PREMIS, so dass sie auch in andere Systeme eingelesen werden können.
- Die Software von docuteam steht unter einer **Open Source-Lizenz** und manche Institutionen entwickeln sie bereits heute eigenständig weiter.



Fazit

- Wir haben im Stadtarchiv Baden vergleichsweise früh mit einer ersten Umsetzung des digitalen Archivs begonnen und haben so bereits Erfahrungen mit mehreren Generationen von Software und Speicher sammeln können.
- Automatisieren, was immer möglich ist: damit haben wir eher zu lange zugewartet.
- Alles basiert auf internationalen Standards. Wir sind selber auch Mitglied verschiedener Standardisierungsgremien.
- Im Mittelpunkt stehen Daten- und Metadatenformate und Schnittstellendefinitionen.
 Werkzeuge kommen und gehen.
- Eine günstige Schweizer Cloud wie SWITCHengines bietet ein riesiges Synergiepotential.