

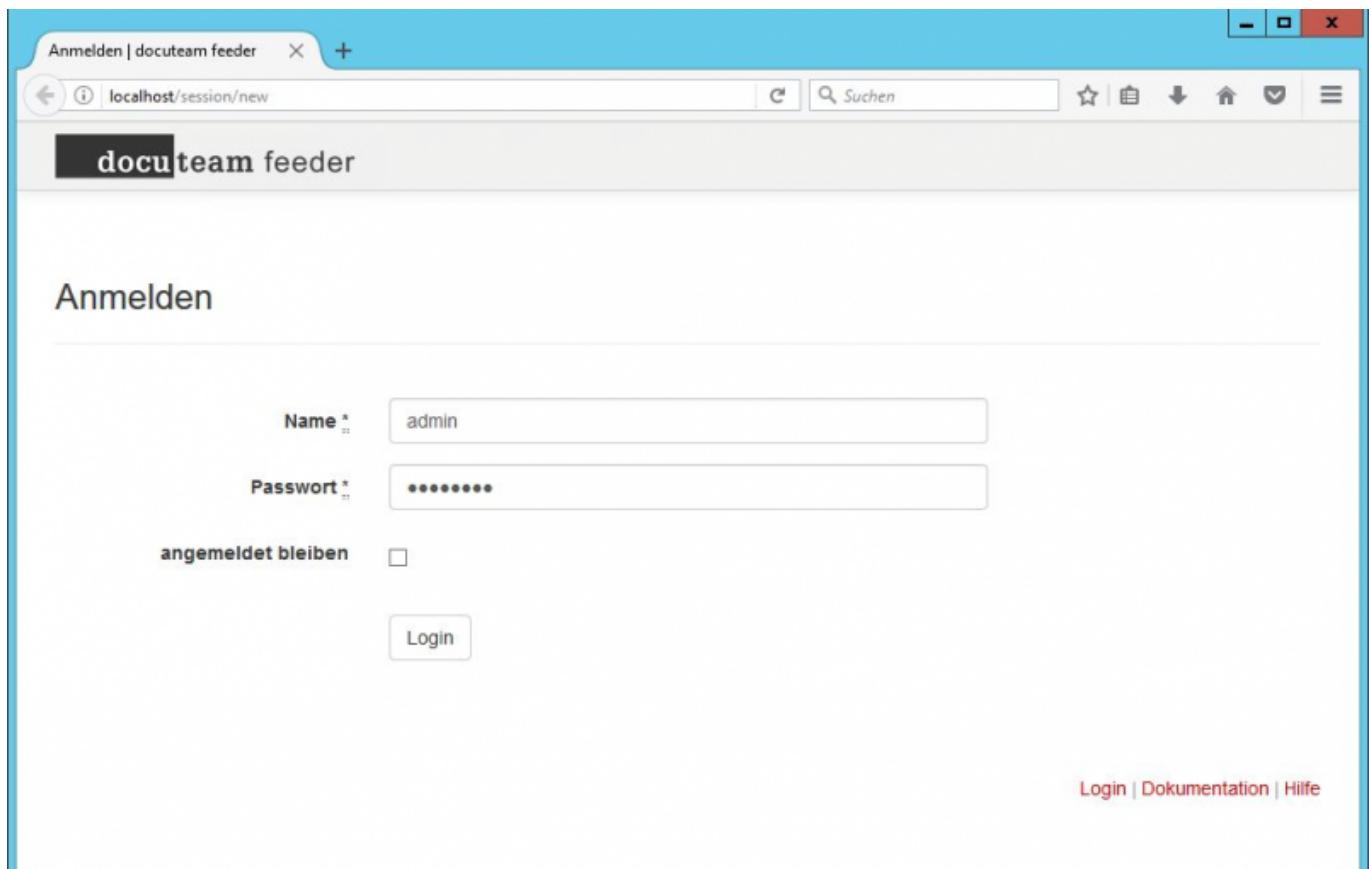
docuteam feeder 3.4.0

Im Normalbetrieb werden in einem Archiv die Ingest-Vorgänge für digitale Ablieferungen nicht manuell durchgeführt, sondern automatisch. Sobald ein neues SIP in einem dafür vorgesehenen Ordner eingetroffen ist, wird ein entsprechender Ingest-Workflow gestartet. Dieser Workflow verarbeitet das SIP in mehreren Schritten. Die Verarbeitung wird protokolliert und kann online überwacht werden.

Die Interaktion mit docuteam feeder erfolgt über ein Web-Frontend.

Anmeldung ans System

Beim Aufruf von docuteam feeder im Browser wird ein Login-Fenster angezeigt.



The screenshot shows a web browser window with the title 'Anmelden | docuteam feeder'. The address bar displays 'localhost/session/new'. The main content area is titled 'Anmelden' (Login). It contains two input fields: 'Name*' with the value 'admin' and 'Passwort*' with the value '*****'. Below these fields is a checkbox labeled 'angemeldet bleiben' (Remember me) which is unchecked. At the bottom of the form is a 'Login' button. In the bottom right corner of the page, there are links for 'Login | Dokumentation | Hilfe'.

Bei aktivierter Checkbox `Remember me` werden bei späteren Anmeldungen die eingegebenen Daten vorgeschlagen.

ID	Workflow	SIP	Status	Startzeit	Dauer
80	Quality Assurance, Migration and Storage	Test_Wi_02_zip	•••• •••• ••••	Freitag, 07. Juli 2017, 10:56 Uhr	00:00:32 anzeigen löschern
79	Quality Assurance	Test_Wi_02_zip	•••• •	Freitag, 07. Juli 2017, 10:55 Uhr	00:00:06 anzeigen löschern

Nach erfolgter Anmeldung befinden Sie sich im Bereich Ausführungen. Wenn Sie sich als Administrator angemeldet haben, erscheint rechts oben das Hauptmenu mit fünf Bereichen: Ausführungen, Workflows, Schritte, Benutzer und Cockpit:

Sind Sie als Benutzer angemeldet, dann stehen Ihnen drei Bereiche zur Verfügung:

- **Ausführungen:** Dieser Bereich zeigt die Zusammenfassung aller bislang durchgeföhrten, laufenden und bevorstehenden Ausführungen. Als Ausführung wird die konkrete Verarbeitung eines SIP anhand eines spezifischen Workflows bezeichnet.
- **Workflows:** Als Workflow wird eine geordnete Abfolge von definierten Verarbeitungsschritten bezeichnet, die während einer Ausführung auf ein SIP angewendet wird. In diesem Bereich werden die Ingest-Workflows definiert, verwaltet, gestartet und ggf. gelöscht.
- **Schritte:** Ein Schritt ist die Einbindung und Konfiguration von Softwarekomponenten und Diensten, die gekapselt als einzelne Verarbeitung in einem Ingest-Workflow verwendet werden können. Dieser Menüpunkt ist nur für Administratoren sichtbar.
- **Benutzer:** Benutzerverwaltung. Auch dieser Menupunkt ist nur für Administratoren sichtbar.
- **Cockpit:** Liefert Informationen über Speicherplatz, verwendete File-Formate usw.



Menüpunkt "Ausführungen"

Der Menüpunkt „Ausführungen“ gibt eine Übersicht über alle bisherigen (laufenden, abgeschlossenen und auch fehlgeschlagenen) Workflow-Abläufe.

ID	Workflow	SIP	Status	Startzeit	Dauer	
80	Quality Assurance, Migration and Storage	Test_Wi_02_zip	● ● ● ●	Freitag, 07. Juli 2017, 10:56 Uhr	00:00:32	zeigen löschchen
79	Quality Assurance	Test_Wi_02_zip	● ● ● ●	Freitag, 07. Juli 2017, 10:55 Uhr	00:00:06	zeigen löschchen

In diesem Bereich können die Ausführungen überwacht werden. Der Bereich enthält eine Liste aller SIPs,

- die gerade verarbeitet werden
- deren Verarbeitung abgeschlossen ist
- deren Verarbeitung aufgrund eines Fehlers abgebrochen wurde.

Die Liste besteht aus mehreren Spalten, die Informationen zur Verarbeitung enthalten:

- **ID** zeigt die eindeutige Verarbeitungsnummer.
- **Workflow** zeigt den Namen des Workflows an, der das SIP verarbeitet. Wenn Sie weitere Informationen zu diesem Workflow benötigen, wechseln Sie in den Bereich **Workflows** und klicken in der Zeile des entsprechenden Workflows auf **zeigen**.
- **SIP** zeigt das SIP, das in Verarbeitung ist.
- Die Spalte **Status** zeigt für jeden Verarbeitungsschritt des betreffenden Workflows einen Punkt. Aus der Farbe der Punkte können Sie auf den Status des entsprechenden Schritts schliessen:
 - **braun**: Dieser Schritt steht noch bevor.
 - **gelb**: Dieser Schritt wird gerade durchgeführt.
 - **grün**: Dieser Schritt wurde erfolgreich abgeschlossen.
 - **rot**: Dieser Schritt ist fehlgeschlagen.
 - **schwarz**: Dieser Schritt wurde übersprungen.
- **Startzeit** zeigt den Zeitpunkt, an dem der Verarbeitungsprozess gestartet wurde.
- **Dauer** zeigt die seit dem Start vergangene Zeit.

Bitte beachten Sie: Die Informationen beziehen sich auf den Zeitpunkt, an dem die Seite aufgerufen worden ist. Wenn Sie den aktuellen Status abrufen wollen, dann muss die Seite aktualisiert werden.

Um die Details zu einer Verarbeitung aufzurufen, kann der Link **zeigen** auf der Zeile, deren Detailinformationen abgerufen werden sollen, angeklickt werden.

Je nach eingestelltem Log-Level werden mehr oder weniger Informationen zu den einzelnen Workflow-Schritten angezeigt. Wenn ein Schritt erfolgreich beendet ist, dann wird ein grünes „beendet mit Code 0“ angezeigt und der nächste Schritt startet. Wenn ein Schritt fehlschlägt, dann wird dies in roter Schrift mit einem entsprechenden Fehlercode angezeigt. Im untenstehenden konnte eine File-

Migration nicht korrekt durchgeführt werden.

Ingest: file migration

Status: fehlgeschlagen mit Code 1

Befehl:

```
cd %FEEDER_JAVA% && java -Djava.library.path=C:\docuteam\apps\feeder_java\dlls\64bit ch.docuteam.feeder.ingest.SIPFileMigrator "400_Testdaten1" false
```

Antwort:

Antwort:

```
ERROR 2016-07-13T16:16:58.124 (FileConverter$PDFToolsConverter) 37817352
ERROR 2016-07-13T16:16:58.124 (FileConverter$PDFToolsConverter) Some parts of the document could not be processed
ERROR 2016-07-13T16:17:47.982 (FileConverter$PDFToolsConverter) 37817352
ERROR 2016-07-13T16:17:47.982 (FileConverter$PDFToolsConverter) Some parts of the document could not be processed
ERROR 2016-07-13T16:18:01.852 (FileConverter$PDFToolsConverter) -2143223551
ERROR 2016-07-13T16:18:01.852 (FileConverter$PDFToolsConverter) The file couldn't be opened.
ERROR 2016-07-13T16:18:01.867 (FileConverter$PDFToolsConverter) -2143222271
ERROR 2016-07-13T16:18:01.867 (FileConverter$PDFToolsConverter) The analysis has been stopped.
ERROR 2016-07-13T16:18:33.07 (FileConverter$PDFToolsConverter) 37817352
ERROR 2016-07-13T16:18:33.07 (FileConverter$PDFToolsConverter) Some parts of the document could not be processed
```

Dauer: 00:04:11

Auch hier gilt: Um den aktuellen Status abzurufen, muss die Seite aktualisiert werden.

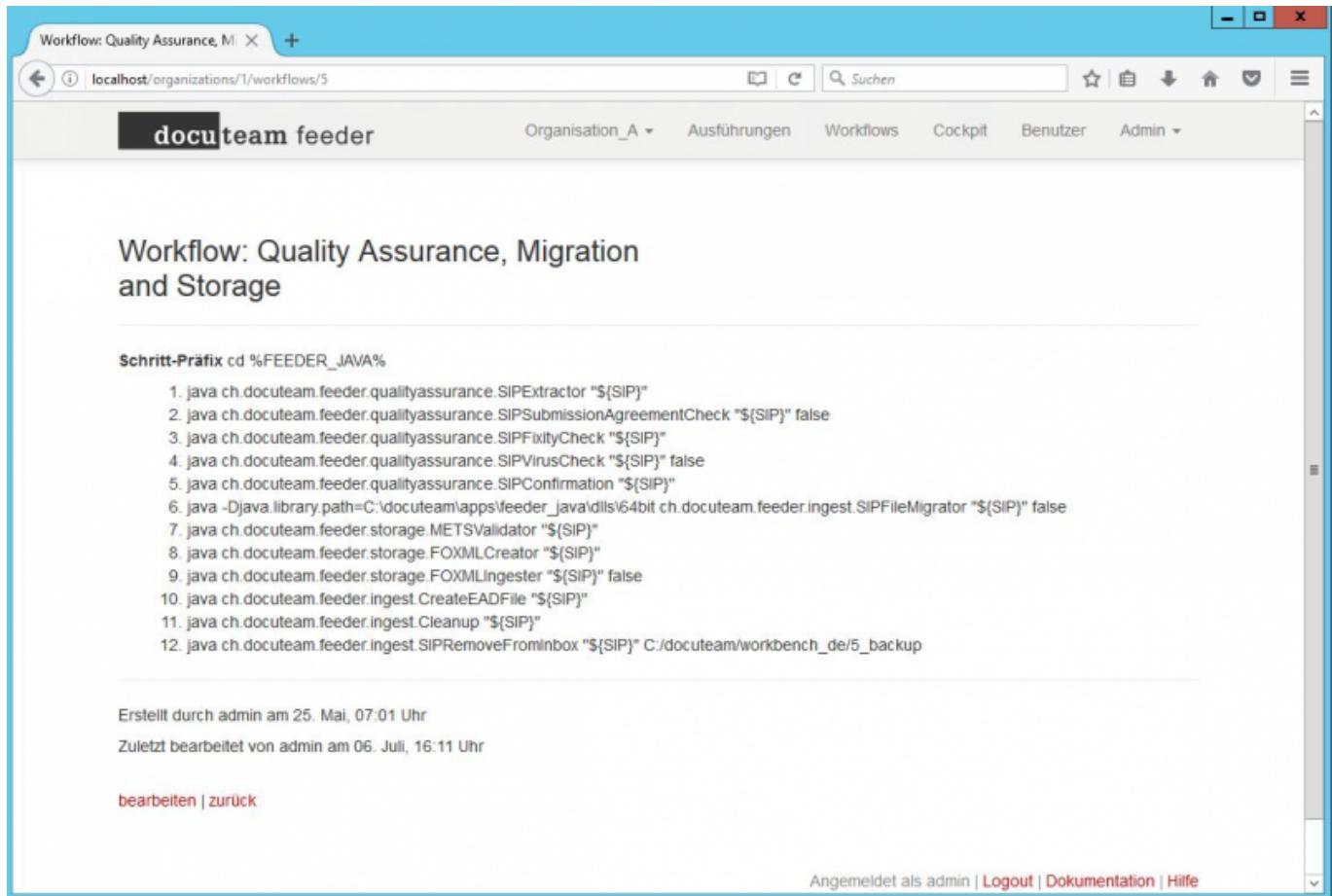


Menüpunkt "Workflows"

Der Menüpunkt „Workflows“ zeigt alle definierten Ingest-Workflows an, die auf ein konkretes SIP angewendet werden können.

ID	Name	starten	verwalten
1	Administration		
2	Run		
3	Quality Assurance		
4	Check SA and Migration into 3_storage		
5	Quality Assurance, Migration and Storage		
6	Cleanup		
7	Check Checksums		
8	Preservation: Deliver DIP by PUID		
9	Preservation: Update Fedora Objects		

Mit „Zeigen“ werden die einzelnen Schritte des jeweiligen Workflows angezeigt.



The screenshot shows a web browser window with the title "Workflow: Quality Assurance, M...". The address bar shows "localhost/organizations/1/workflows/5". The page header includes "docuteam feeder", "Organisation_A", "Ausführungen", "Workflows", "Cockpit", "Benutzer", and "Admin". The main content area displays the title "Workflow: Quality Assurance, Migration and Storage". Below the title is a code block labeled "Schritt-Präfix cd %FEEDER_JAVA%". The code lists 12 numbered steps in Java, which are part of a SIP processing pipeline. At the bottom of the code block, it says "Erstellt durch admin am 25. Mai, 07:01 Uhr" and "Zuletzt bearbeitet von admin am 06. Juli, 16:11 Uhr". A red link "bearbeiten | zurück" is visible. The bottom right corner shows a login status: "Angemeldet als admin | Logout | Dokumentation | Hilfe".

```
Schritt-Präfix cd %FEEDER_JAVA%
1. java ch.docuteam.feeder.qualityassurance.SIPExtractor "${SIP}"
2. java ch.docuteam.feeder.qualityassurance.SIPSubmissionAgreementCheck "${SIP}" false
3. java ch.docuteam.feeder.qualityassurance.SIPFixityCheck "${SIP}"
4. java ch.docuteam.feeder.qualityassurance.SIPVirusCheck "${SIP}" false
5. java ch.docuteam.feeder.qualityassurance.SIPConfirmation "${SIP}"
6. java -Djava.library.path=C:\docuteam\apps\feeder_java\libs\64bit ch.docuteam.feeder.ingest.SIPFileMigrator "${SIP}" false
7. java ch.docuteam.feeder.storage.METSValidator "${SIP}"
8. java ch.docuteam.feeder.storage.FOXMLCreator "${SIP}"
9. java ch.docuteam.feeder.storage.FOXMLIngestor "${SIP}" false
10. java ch.docuteam.feeder.ingest.CreateEADFile "${SIP}"
11. java ch.docuteam.feeder.ingest.Cleanup "${SIP}"
12. java ch.docuteam.feeder.ingest.SIPRemoveFromInbox "${SIP}" C:/docuteam/workbench_de/5_backup
```

Unter „Bearbeiten“ können bestehende Workflows angepasst werden. Ein Workflow ist immer ein linearer Ablauf: Das SIP wird gewissermassen in eine Pipeline hineingeschickt, in welcher ein Schritt nach dem anderen abgearbeitet wird. Im Falle eines Fehlers stoppt die Ausführung und lässt die allenfalls noch folgenden Schritte aus.

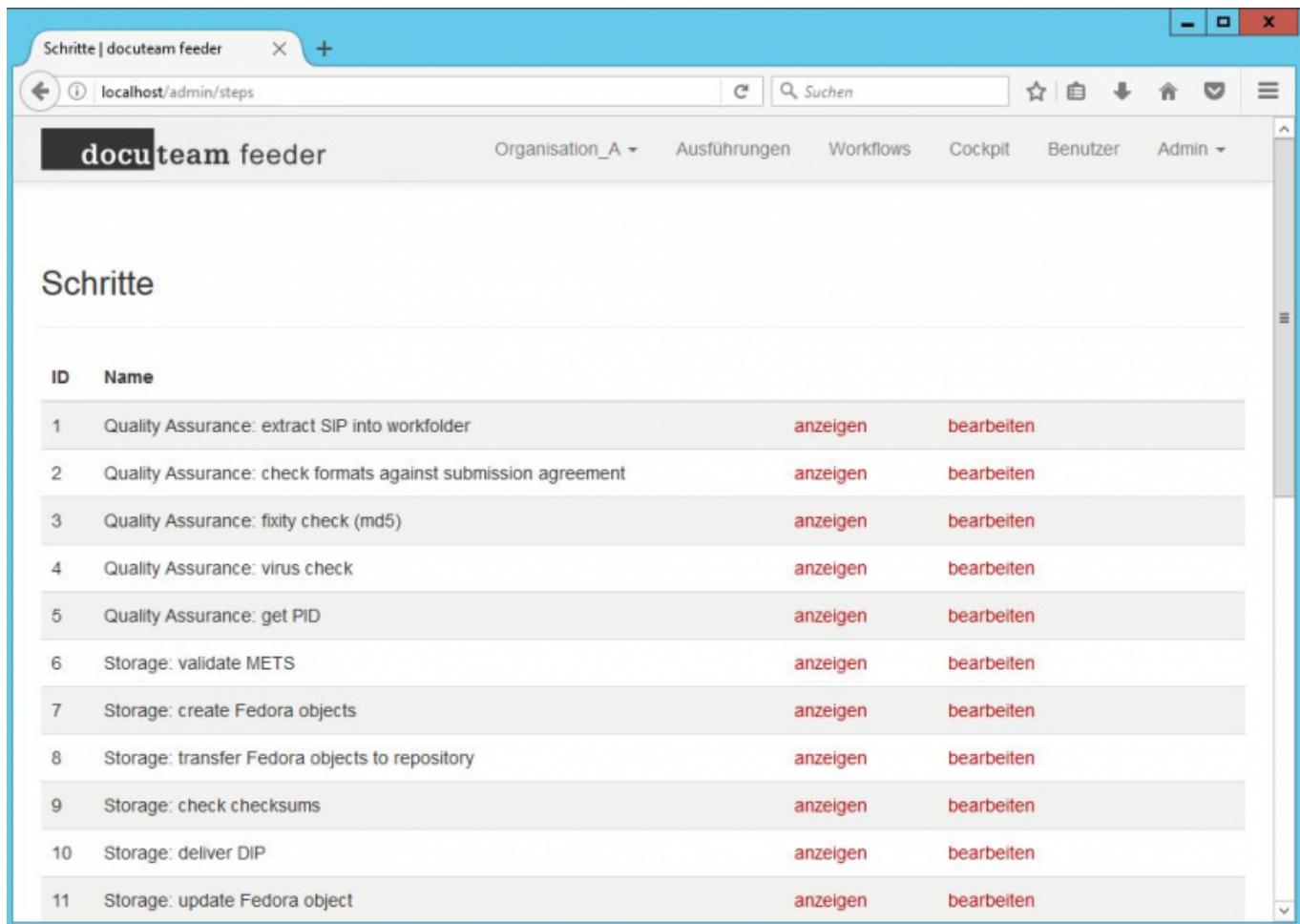
Für den Testbetrieb kann der Ingest einzelner SIPs auch manuell ausgelöst werden. Diesem Zweck dient der Befehl `starten`. Im Formularfeld „SIP“, das nach der Aktivierung des Befehls erscheint, wird der Name des SIP angegeben. **Wichtig:** der Name sollte normalerweise ohne die Endung „.zip“ angegeben werden.

The screenshot shows a web browser window with the following details:

- Title Bar:** Neue Ausführung für Quality As... X
- Address Bar:** localhost/organizations/1/workflow_executions/new?workflow_id=5
- Header:** docuteam feeder, Organisation_A, Ausführungen, Workflows, Cockpit, Benutzer, Admin
- Section:** Neue Ausführung für Quality Assurance, Migration and Storage
- Form Fields:**
 - Objekt-ID:
 - SIP:
 - oder
 - Manuelle Eingabe:
 -
- Footer:** Angemeldet als admin | [Logout](#) | [Dokumentation](#) | [Hilfe](#)

Menüpunkt "Schritte"

Der Menupunkt „Schritte“ zeigt sämtliche verfügbaren Workflowschritte an, die in docuteam feeder hinterlegt sind und aus denen Ingest-Workflows zusammengestellt werden können.



The screenshot shows a software application window titled "docuteam feeder". The window has a blue header bar with the title and a navigation bar with links like "Organisation_A", "Ausführungen", "Workflows", "Cockpit", "Benutzer", and "Admin". The main content area is titled "Schritte" (Steps) and contains a table with 11 rows of data. The table has columns for "ID", "Name", "anzeigen" (display), and "bearbeiten" (edit). The steps listed are:

ID	Name	anzeigen	bearbeiten
1	Quality Assurance: extract SIP into workfolder	anzeigen	bearbeiten
2	Quality Assurance: check formats against submission agreement	anzeigen	bearbeiten
3	Quality Assurance: fixity check (md5)	anzeigen	bearbeiten
4	Quality Assurance: virus check	anzeigen	bearbeiten
5	Quality Assurance: get PID	anzeigen	bearbeiten
6	Storage: validate METS	anzeigen	bearbeiten
7	Storage: create Fedora objects	anzeigen	bearbeiten
8	Storage: transfer Fedora objects to repository	anzeigen	bearbeiten
9	Storage: check checksums	anzeigen	bearbeiten
10	Storage: deliver DIP	anzeigen	bearbeiten
11	Storage: update Fedora object	anzeigen	bearbeiten

Es können auch neue Schritte hinzugefügt werden, etwa um weitere Migrationswerkzeuge einzubinden oder den Ingestprozess um zusätzliche Funktionalitäten zu erweitern.

Neuen Schritt erstellen | docuteam feeder X +

localhost/admin/steps/new

Suchen

docuteam feeder Organisation_A Ausführungen Workflows Cockpit Benutzer Admin

Neuen Schritt erstellen

Name *

Command *

Type *

Angemeldet als admin | [Logout](#) | [Dokumentation](#) | [Hilfe](#)



Menüpunkt "Benutzer"

Benutzer | docuteam feeder X +

localhost/organizations/1/users

Suchen

docuteam feeder Organisation_A Ausführungen Workflows Cockpit Benutzer Admin

Benutzer

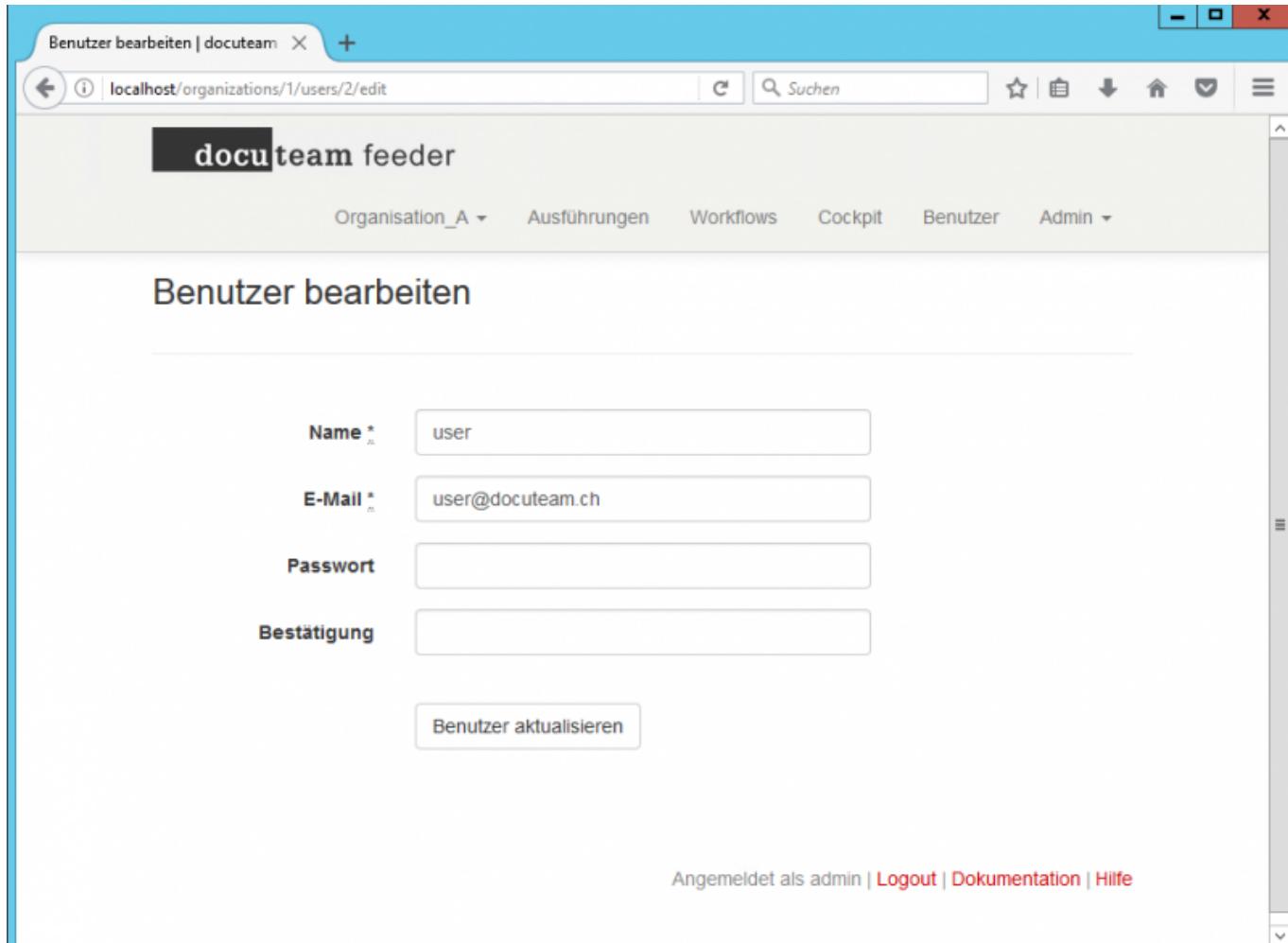
Name			
admin	bearbeiten	löschen	
peter	bearbeiten	löschen	

[Neuen Benutzer erstellen](#)

Angemeldet als admin | [Logout](#) | [Dokumentation](#) | [Hilfe](#)

Für einen neuen Benutzer muss ein Benutzername, eine E-Mail-Adresse und ein Passwort vergeben werden. Zudem muss die Rolle definiert werden, die der Benutzer übernehmen soll:

- **Ausführende/r** kann Ingest-Workflows lediglich durchführen und verfolgen, aber nicht neu erstellen oder verändern.
- **Administrator** kann Workflows ausführen wie auch neu erstellen und ändern. Zudem kann ein Administrator die Benutzer verwalten.



Benutzer bearbeiten | docuteam

localhost/organizations/1/users/2/edit

Suchen

docuteam feeder

Organisation_A Ausführungen Workflows Cockpit Benutzer Admin

Benutzer bearbeiten

Name * user

E-Mail * user@docuteam.ch

Passwort

Bestätigung

Benutzer aktualisieren

Angemeldet als admin | Logout | Dokumentation | Hilfe



Menüpunkt "Cockpit"

Das Cockpit gibt Aufschluss über Speicherplatzverhältnisse, die Menge der unterschiedlichen Fileformate, die Anzahl der Ingests und der Objects. Das Cockpit zieht Informationen aus dem Fedora Commons-Repository zusammen (mehrheitlich über die Schnittstellen von [docuteamOAIS](#)).

Demo | docuteam feeder X

localhost/organizations/1/repositories/demo C Suchen

Organisation_A Ausführungen Workflows Cockpit Benutzer Admin

docuteam feeder

Demo

Berichte ▾

Speicherplatz

	de 58 MB
	test 950 MB
Total	1000 MB
Frei	59 GB

Letzte Integritätsprüfung - Erfolgreich

Dienstag, 11. Juli 2017, 14:29 Uhr

Basisdaten

Name	Demo
Basis-URL	http://localhost:8080/fedora
Version	3.8.1
PID Namespace	docuteam
Support-Mail	support@docuteam.ch

Häufigste Dateiformate

Microsoft Word Document	905	49,02 %
Acrobat PDF 1.4 - Portable Document Format	147	7,96 %
Acrobat PDF 1.3 - Portable Document Format	147	7,96 %
JPEG File Interchange Format	92	4,98 %
Graphics Interchange Format	89	4,82 %
Extensible Markup Language	78	4,23 %
Tagged Image File Format	66	3,58 %
Microsoft Excel 97 Workbook (xls)	61	3,30 %
Acrobat PDF 1.2 - Portable Document Format	51	2,76 %
Hypertext Markup Language	38	2,06 %

Ablieferungen

	de 2
	test 17
Total	19

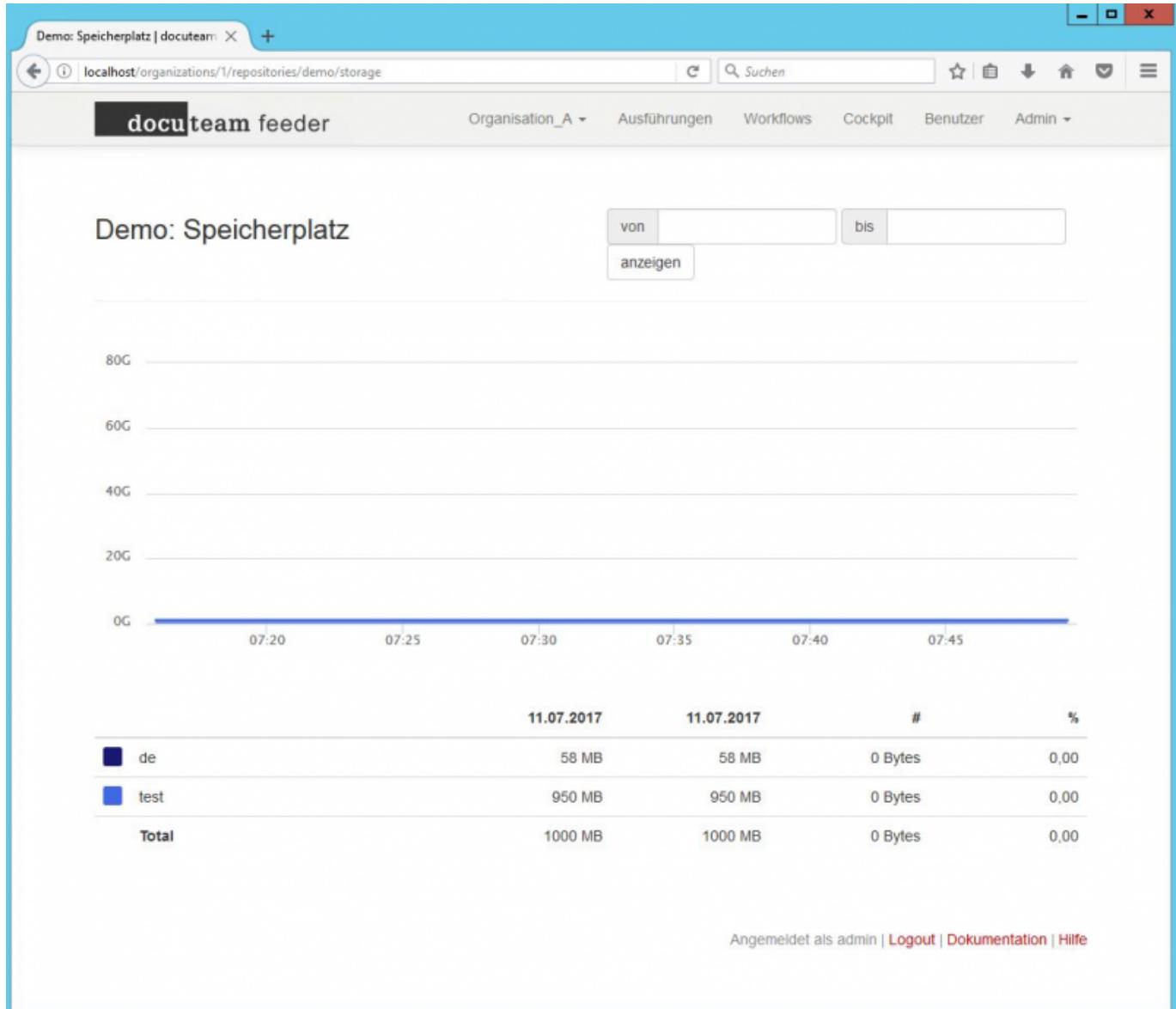
Objekte

	de 9
	test 2231
Total	2240

Angemeldet als admin | [Logout](#) | [Dokumentation](#) | [Hilfe](#)

Es können detailliertere Berichte für die Dateiformate und den Speicherplatz aufgerufen werden.

Demo: Dateiformate					
Total		de	test		
Name		PUID	#	%	
Microsoft Word Document		fmt/40	883	47,83 %	
Acrobat PDF 1.3 - Portable Document Format		fmt/17	147	7,96 %	
Acrobat PDF 1.4 - Portable Document Format		fmt/18	147	7,96 %	
Extensible Markup Language		fmt/101	78	4,23 %	
JPEG File Interchange Format		fmt/43	78	4,23 %	
Graphics Interchange Format		fmt/4	66	3,58 %	
Tagged Image File Format		fmt/353	66	3,58 %	
Microsoft Excel 97 Workbook (xls)		fmt/61	61	3,30 %	
Acrobat PDF 1.2 - Portable Document Format		fmt/16	51	2,76 %	
Rich Text Format		fmt/50	27	1,46 %	
Hypertext Markup Language		fmt/96	27	1,46 %	
Acrobat PDF 1.5 - Portable Document Format		fmt/19	25	1,35 %	
Log File		x-fmt/62	24	1,30 %	
Graphics Interchange Format		fmt/3	23	1,25 %	
Microsoft Word Document		fmt/39	22	1,19 %	
Plain Text File		x-fmt/111	11	0,60 %	
JPEG File Interchange Format		fmt/44	10	0,54 %	
Microsoft Excel 5.0/95 Workbook (xls)		fmt/59	10	0,54 %	
Thumbs DB file		fmt/682	9	0,49 %	
Hypertext Markup Language		fmt/100	9	0,49 %	
Windows Bitmap		fmt/116	8	0,43 %	



From:
<https://wiki.docuteam.ch/> - **docuteam wiki**



Permanent link:
https://wiki.docuteam.ch/doku.php?id=docuteam:feeder_330&rev=1499776741

Last update: **2019/01/07 11:17**