

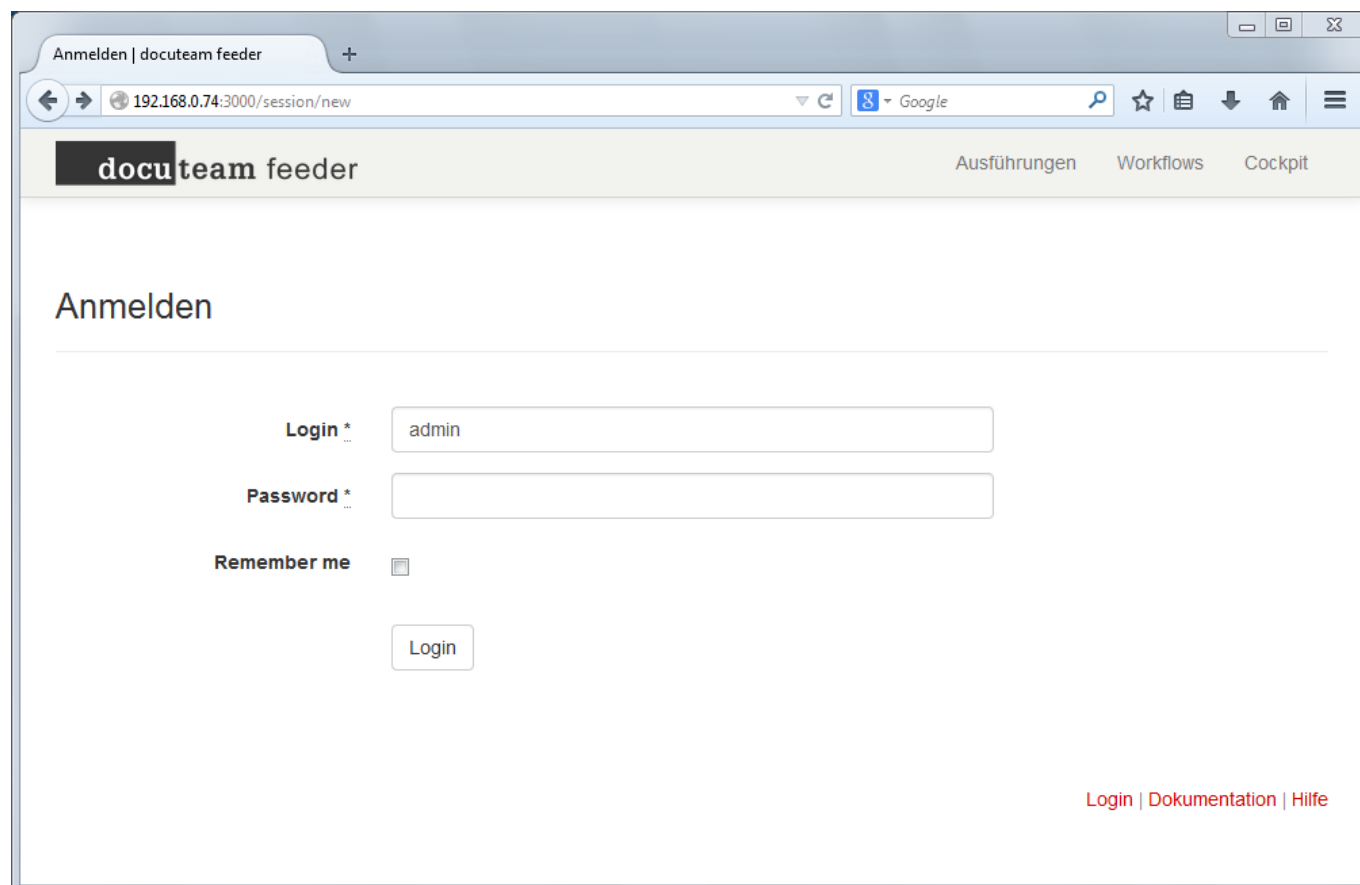
docuteam feeder 3.0.0

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 address bar displaying "192.168.0.74:3000/session/new". The page title is "Anmelden | docuteam feeder". The main heading is "Anmelden". Below the heading, there are three input fields: "Login *" with the value "admin", "Password *" which is empty, and a "Remember me" checkbox which is unchecked. A "Login" button is positioned below the password field. In the bottom right corner, there are links for "Login", "Dokumentation", and "Hilfe". The browser's address bar also shows "Google" as the search engine.

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

ID	Workflow	SIP	Status	Startzeit	Dauer
462	Quality Assurance	böütl	●●●●	Mittwoch, 01. April 2015, 12:30 Uhr	00:00:15 anzeigen
461	Quality Assurance	böütl	●●●●	Mittwoch, 01. April 2015, 12:29 Uhr	00:00:18 anzeigen

Nach erfolgreicher 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 durchgeführten, 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 Menüpunkt 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.

Ausführungen | docuteam feeder

192.168.0.74:3000/workflow_executions

docuteam feeder

Ausführungen Workflows Schritte Benutzer Cockpit

Ausführungen

alle | wartend | laufend | fertig | fehlgeschlagen

ID	Workflow	SIP	Status	Startzeit	Dauer	
462	Quality Assurance	böütli	● ● ● ● ●	Mittwoch, 01. April 2015, 12:30 Uhr	00:00:15	anzeigen
461	Quality Assurance	böütli	● ● ● ● ●	Mittwoch, 01. April 2015, 12:29 Uhr	00:00:18	anzeigen

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: failed with Code 235

Command:

```
cd %FEEDER_JAVA% && java -Djava.library.path=C:\docuteam\apps\feeder_java\dlls\32bit ch.docuteam.feeder.ingest.SIPFileMigrator "182_RTR_docx-xlsx_20141001_nu" false
```

Response:

```
INFO 2015-03-19T13:24:12.520 (SIPFileMigrator) Initializing SIP File Migration...
INFO 2015-03-19T13:24:15.681 (Document) Opening copy of SIP in mode: 'ReadWrite': '\\192.168.0.100\scratch\ingest\099-test\2_work\182_RTR_docx-xlsx_20141001_nu'
INFO 2015-03-19T13:24:18.494 (SIPFileMigrator) Migrating files in SIP: C:\docuteam\apps\feeder_java\temp\METSDocument_read_20150319132412582\182_RTR_docx-xlsx_20141001_nu
INFO 2015-03-19T13:24:18.509 (FileConverter) Converting file: C:\docuteam\apps\feeder_java\temp\METSDocument_read_20150319132412582\182_RTR_docx-xlsx_20141001_nu\2014-08-19 Seduta da direzziun -09\2014-08-19 Aggiunta 02.1.3_Preschientscha DIR occurrenzas RTR.xlsx
INFO 2015-03-19T13:24:18.509 (FileConverter) Initializing File Migration, from config/migration-config.xml...
INFO 2015-03-19T13:24:35.259 (FileConverter) Looking for PUID fmt/214...
INFO 2015-03-19T13:24:35.306 (FileConverter) Treating file as fmt/214...
INFO 2015-03-19T13:24:35.306 (FileConverter) Destination file: C:\docuteam\apps\feeder_java\temp\SIPFileMigrator\Convert\2014-08-19 Aggiunta 02.1.3_Preschientscha DIR occurrenzas RTR.pdf
INFO 2015-03-19T13:24:35.322 (FileConverter) Executing command line: [Class:ch.docuteam.converter.FileConverter$OOConverter, C:\docuteam\apps\feeder_java\temp\METSDocument_read_20150319132412582\182_RTR_docx-xlsx_20141001_nu\2014-08-19 Seduta da direzziun -09\2014-08-19 Aggiunta 02.1.3_Preschientscha DIR occurrenzas RTR.xlsx, C:\docuteam\apps\feeder_java\temp\SIPFileMigrator\Convert\2014-08-19 Aggiunta 02.1.3_Preschientscha DIR occurrenzas RTR.pdf]
INFO 2015-03-19T13:24:38.34 (FileConverter$OOConverter) Could not find the default OO application for converting files in folder: apps/OpenOffice.org 3
ERROR 2015-03-19T13:24:38.34 (SIPFileMigrator) Error converting file
The OOConverter shall not be used because the default installation is absent and no external OO application path was supplied
WARN 2015-03-19T13:24:38.34 (SIPFileMigrator) ...file migration returned error code: 235
WARN 2015-03-19T13:24:38.34 (SIPFileMigrator) Migration stopped due to errors.
```

Error:

```
ch.docuteam.converter.exceptions.OOConverterShallNotBeUsedException: The OOConverter shall not be used because the default installation is absent and no external OO application path was supplied
    at ch.docuteam.converter.FileConverter$OOConverter.convert2PDF(FileConverter.java:734)
    at ch.docuteam.converter.FileConverter$OOConverter.main(FileConverter.java:511)
    at sun.reflect.NativeMethodAccessorImpl.invoke0(Native Method)
    at sun.reflect.NativeMethodAccessorImpl.invoke(Unknown Source)
    at sun.reflect.DelegatingMethodAccessorImpl.invoke(Unknown Source)
    at java.lang.reflect.Method.invoke(Unknown Source)
    at ch.docuteam.converter.FileConverter.executeCommandLine(FileConverter.java:434)
    at ch.docuteam.converter.FileConverter.convertFile(FileConverter.java:245)
    at ch.docuteam.converter.FileConverter.convertFile(FileConverter.java:175)
    at ch.docuteam.feeder.ingest.SIPFileMigrator.migrateFile(SIPFileMigrator.java:229)
    at ch.docuteam.feeder.ingest.SIPFileMigrator.migrateAllFiles(SIPFileMigrator.java:183)
    at ch.docuteam.feeder.ingest.SIPFileMigrator.run(SIPFileMigrator.java:100)
    at ch.docuteam.feeder.ingest.SIPFileMigrator.main(SIPFileMigrator.java:85)
```

Duration: 00:00:28

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.

The screenshot shows a web browser window with the address bar displaying `192.168.0.74:3000/workflows`. The page title is "Workflows | docuteam feeder". The application header includes the "docuteam feeder" logo and navigation links: "Ausführungen", "Workflows", "Schritte", "Benutzer", and "Cockpit". The main content area is titled "Workflows" and contains a table with 16 workflow entries. Each entry has an ID, a Name, and four action buttons: "anzeigen", "starten", "bearbeiten", and "löschen".

ID	Name	anzeigen	starten	bearbeiten	löschen
1	Quality Assurance, Migration and Storage	anzeigen	starten	bearbeiten	löschen
2	Quality Assurance and Migration	anzeigen	starten	bearbeiten	löschen
3	Quality Assurance	anzeigen	starten	bearbeiten	löschen
4	Create DIP	anzeigen	starten	bearbeiten	löschen
5	Administration	anzeigen	starten	bearbeiten	löschen
6	Cleanup Working Copies	anzeigen	starten	bearbeiten	löschen
7	Extract, Migration	anzeigen	starten	bearbeiten	löschen
8	Execute	anzeigen	starten	bearbeiten	löschen
9	Create SIP from into hotbox	anzeigen	starten	bearbeiten	löschen
15	Convert ead to ead4atom	anzeigen	starten	bearbeiten	löschen
16	Validate BARSIP	anzeigen	starten	bearbeiten	löschen

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

The screenshot shows a web browser window with the address bar displaying `192.168.0.74:3000/workflows/1`. The page title is "Workflow: Quality Assurance, Migr...". The main header of the application is "docuteam feeder" with navigation links: "Ausführungen", "Workflows", "Schritte", "Benutzer", and "Cockpit".

Workflow: Quality Assurance, Migration and Storage

Schritt-Präfix `cd %FEEDER_JAVA%`

1. `java ch.docuteam.feeder.qualityassurance.SIPExtractor "${SIP}"`
2. `java ch.docuteam.feeder.qualityassurance.SIPFixityCheck "${SIP}"`
3. `java ch.docuteam.feeder.qualityassurance.SIPSubmissionAgreementCheck "${SIP}" false`
4. `java ch.docuteam.feeder.qualityassurance.SIPVirusCheck "${SIP}" false`
5. `java ch.docuteam.feeder.qualityassurance.SIPConfirmation "${SIP}"`
6. `java -Djava.library.path=C:\docuteam\apps\feeder_java\dlls\32bit ch.docuteam.feeder.ingest.SIPFileMigrator "${SIP}" false`
7. `java ch.docuteam.feeder.storage.METSValidator "${SIP}" true`
8. `java ch.docuteam.feeder.storage.FOXMLCreator "${SIP}"`
9. `java ch.docuteam.feeder.storage.FOXMLIngester "${SIP}" false`
10. `java ch.docuteam.feeder.ingest.CreateEADFile "${SIP}"`
11. `java net.sf.saxon.Transform -s:"//skylia/scratch/ingest/099-test/3_done/${SIP}/EAD.xml" -xsl:C:/docuteam/apps/feeder_java/resources/xslt/ead2atom.xsl -o:"//skylia/scratch/ingest/099-test/3_done/${SIP}/ead4atom.xml"`
12. `java ch.docuteam.feeder.ingest.Cleanup "${SIP}"`
13. `java ch.docuteam.feeder.ingest.SIPRemoveFromDropbox "${SIP}" //192.168.0.100/scratch/ingest/099-test/4_output`

Erstellt durch admin um 01. November, 16:13 Uhr
Zuletzt bearbeitet von admin um 14. Mai, 10:33 Uhr

[bearbeiten](#) | [zurück](#)

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

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.

Workflow bearbeiten | docuteam f...

192.168.0.74:3000/workflows/3/edit

Google

docuteam feeder

Ausführungen Workflows Schritte Benutzer Cockpit

Workflow bearbeiten

Name *

Quality Assurance

Schritt-Präfix

cd %FEEDER_JAVA%

Quality Assu	Parameter	
Quality Assu	Parameter	
Quality Assu	Parameter	false
Quality Assu	Parameter	false
Ingest: clean	Parameter	
	Parameter	
	Parameter	
	Parameter	
	Parameter	
	Parameter	
	Parameter	
	Parameter	
	Parameter	
	Parameter	
	Parameter	
	Parameter	
	Parameter	

Workflow aktualisieren

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

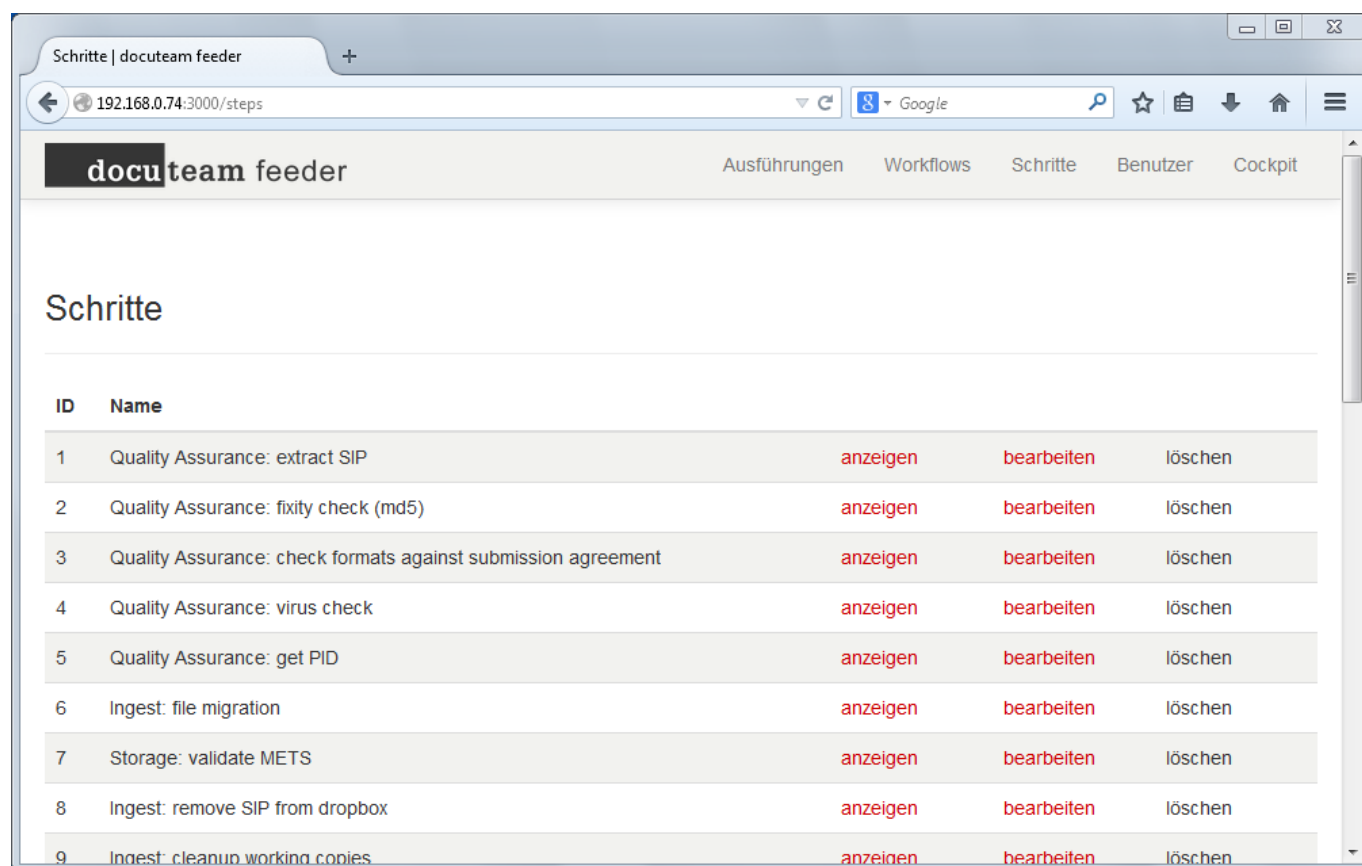
Für den Testbetrieb kann der Ingest einzelner SIPs auch manuell ausgelöst werden. Diesem Zweck dient der Befehl `start`. Im Formularfeld „SIP“, das nach der Aktivierung des Befehls erscheint,

wird der Name des SIP angegeben. **Wichtig:** der Name sollte normalerweise ohne die Endung „.zip“ angegeben werden.



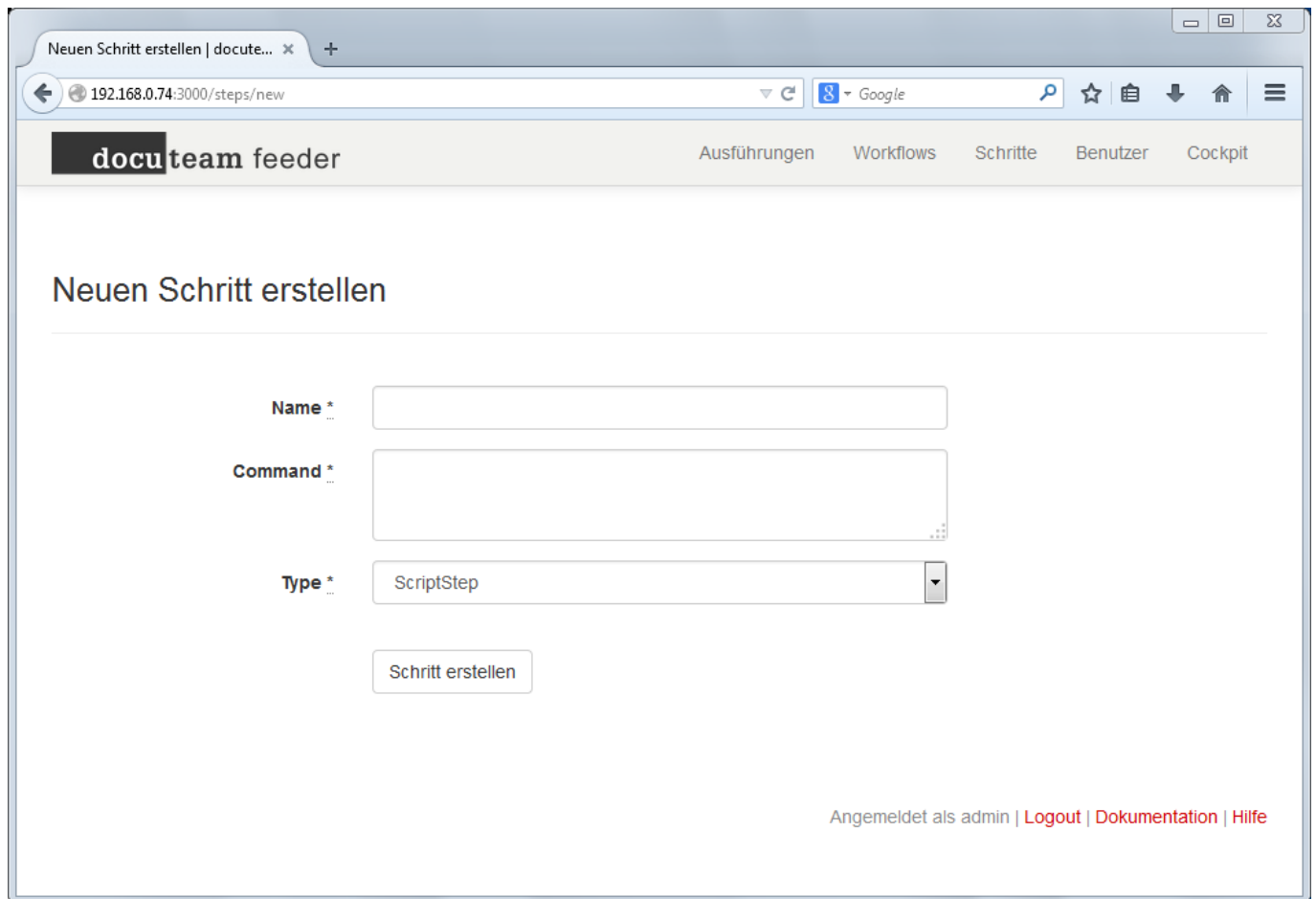
Menüpunkt "Schritte"

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



ID	Name			
1	Quality Assurance: extract SIP	anzeigen	bearbeiten	löschen
2	Quality Assurance: fixity check (md5)	anzeigen	bearbeiten	löschen
3	Quality Assurance: check formats against submission agreement	anzeigen	bearbeiten	löschen
4	Quality Assurance: virus check	anzeigen	bearbeiten	löschen
5	Quality Assurance: get PID	anzeigen	bearbeiten	löschen
6	Ingest: file migration	anzeigen	bearbeiten	löschen
7	Storage: validate METS	anzeigen	bearbeiten	löschen
8	Ingest: remove SIP from dropbox	anzeigen	bearbeiten	löschen
9	Ingest: cleanup working copies	anzeigen	bearbeiten	löschen

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



The screenshot shows a web browser window with the address bar displaying '192.168.0.74:3000/steps/new'. The page title is 'Neuen Schritt erstellen | docuteam feeder'. The navigation bar includes 'Ausführungen', 'Workflows', 'Schritte', 'Benutzer', and 'Cockpit'. The main content area has the heading 'Neuen Schritt erstellen' and a form with three fields: 'Name *', 'Command *', and 'Type *'. The 'Type *' field is a dropdown menu currently set to 'ScriptStep'. Below the form is a button labeled 'Schritt erstellen'. At the bottom right, there is a status bar indicating 'Angemeldet als admin' with links for 'Logout', 'Dokumentation', and 'Hilfe'.

Neuen Schritt erstellen

Name *

Command *

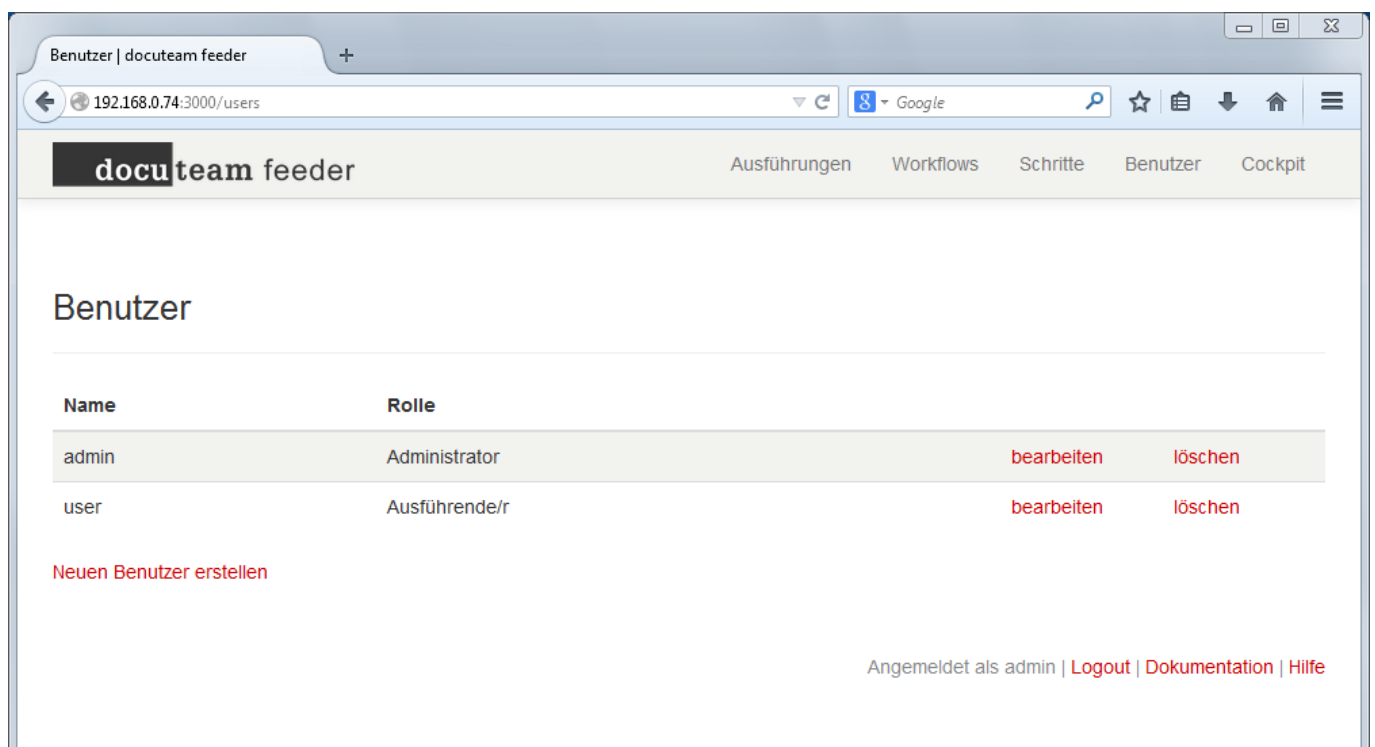
Type * ScriptStep

Schritt erstellen

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



Menüpunkt "Benutzer"



The screenshot shows a web browser window with the address bar displaying '192.168.0.74:3000/users'. The page title is 'Benutzer | docuteam feeder'. The navigation bar includes 'Ausführungen', 'Workflows', 'Schritte', 'Benutzer', and 'Cockpit'. The main content area has the heading 'Benutzer' and a table listing users. Below the table is a link 'Neuen Benutzer erstellen'. At the bottom right, there is a status bar indicating 'Angemeldet als admin' with links for 'Logout', 'Dokumentation', and 'Hilfe'.

Benutzer

Name	Rolle		
admin	Administrator	bearbeiten	löschen
user	Ausführende/r	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.

The screenshot shows a web browser window with the address bar displaying '192.168.0.74:3000/users/2/edit'. The page title is 'Benutzer bearbeiten | docuteam fe...'. The main header of the application is 'docuteam feeder' with navigation links: 'Ausführungen', 'Workflows', 'Schritte', 'Benutzer', and 'Cockpit'. The page content is titled 'Benutzer bearbeiten' and contains a form with the following fields:

- Name ***: Text input with 'user' entered.
- E-Mail ***: Text input with 'user@docuteam.ch' entered.
- Rolle**: Dropdown menu with 'Ausführende/r' selected.
- Passwort**: Empty text input.
- Bestätigung**: Empty text input.

Below the form is a button labeled 'Benutzer aktualisieren'. At the bottom right, there is a status bar showing 'Angemeldet als admin' and links for 'Logout', 'Dokumentation', and 'Hilfe'.



Menüpunkt "Cockpit"

Das Cockpit gibt Aufschluss über Speicherplatzverhältnisse, die Menge der unterschiedlichen Fileformate, die Anzahl der Ingests und der Objects.

The screenshot shows the DocuTeam Feeder web interface. The browser address bar displays '192.168.0.74:3000/repositories/test'. The page title is 'Test2015 Win2012 Fedora 3.8'. A green banner at the top indicates 'Erfolgreich angemeldet'. The main content area is divided into several sections:

- Speicherplatz**: A progress bar shows storage usage. Below it, a table lists storage details:

Speicherplatz	Größe
CH-333666-x	2,1 MB
test	540 MB
Total	540 MB
Frei	7,7 TB

- Letzte Integritätsprüfung**: A green banner indicates the last integrity check was performed on 'Freitag, 20. März 2015, 10:49 Uhr'.
- Basisdaten**: A table with repository metadata:

Basisdaten	
Name	Test2015 Win2012 Fedora 3.8
Basis-URL	http://192.168.0.74:8080/fedora
Version	3.8.0
PID Namespace	feeder
Support-Mail	support@docuteam.ch

- Häufigste Dateiformate**: A table showing the most common file formats and their counts/percentages:

Häufigste Dateiformate	Anzahl	Prozent
Acrobat PDF/A - Portable Document Format	406	52,52%
Tagged Image File Format	111	14,36%
Plain Text File	28	3,62%
JPEG File Interchange Format	26	3,36%
Portable Network Graphics	10	1,29%
Acrobat PDF 1.4 - Portable Document Format	6	0,78%
Acrobat PDF 1.3 - Portable Document Format	4	0,52%
Waveform Audio (PCM/WAVEFORMAT)	3	0,39%
Tagged Image File Format for Internet Fax (TIFF-FX)	3	0,39%
Acrobat PDF 1.5 - Portable Document Format	3	0,39%

- Ablieferungen**: A table showing the number of uploads for each repository:

Ablieferungen	Anzahl
CH-333666-x	3
test	83
Total	86

- Objekte**: A table showing the number of objects for each repository:

Objekte	Anzahl
CH-333666-x	9
test	764
Total	773

At the bottom right, the user is logged in as 'admin' with links for 'Logout', 'Dokumentation', and 'Hilfe'.

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

Test2015 Win2012 Fedora 3.8: Date...

192.168.0.74:3000/repositories/test/file_formats

Google

☆

📁

⬇

🏠

☰

docuteam feeder

AusführungenWorkflowsSchritteBenutzerCockpit

Test2015 Win2012 Fedora 3.8: Dateiformate

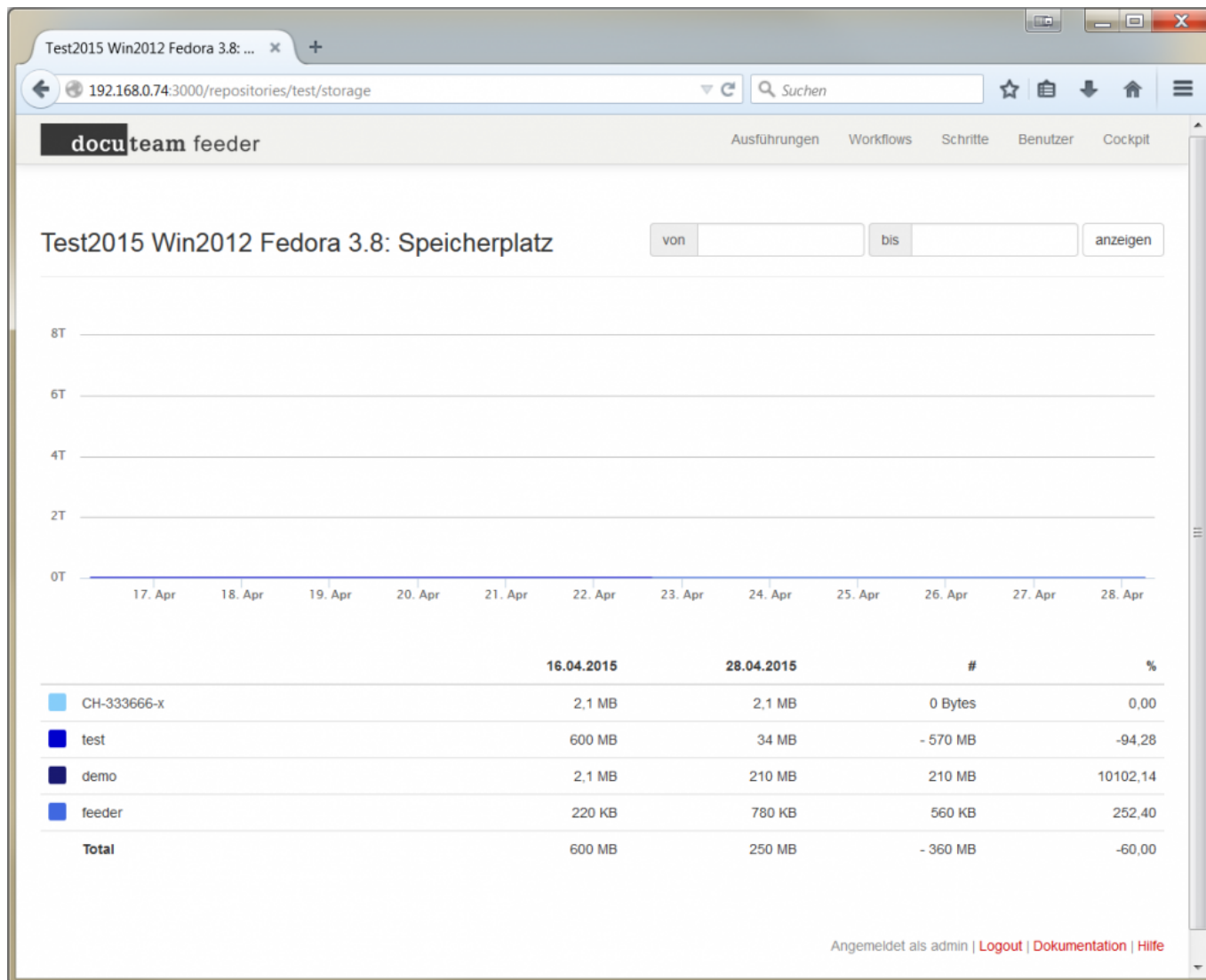
Total

CH-333666-x

test

Name	PUID	#	%
Acrobat PDF/A - Portable Document Format	fmt/95	366	47,35%
Tagged Image File Format	fmt/353	111	14,36%
Acrobat PDF/A - Portable Document Format	fmt/354	40	5,17%
Plain Text File	x-fmt/111	28	3,62%
JPEG File Interchange Format	fmt/43	26	3,36%
Portable Network Graphics	fmt/12	10	1,29%
Acrobat PDF 1.4 - Portable Document Format	fmt/18	6	0,78%
Acrobat PDF 1.3 - Portable Document Format	fmt/17	4	0,52%
Waveform Audio (PCMWAVEFORMAT)	fmt/141	3	0,39%
Tagged Image File Format for Internet Fax (TIFF-FX)	fmt/156	3	0,39%
Acrobat PDF 1.5 - Portable Document Format	fmt/19	3	0,39%
MPEG-4 Media File	fmt/199	3	0,39%
MPEG 1/2 Audio Layer 3	fmt/134	2	0,26%
MPEG-1 Video Format	x-fmt/385	2	0,26%
Acrobat PDF 1.6 - Portable Document Format	fmt/20	1	0,13%
Audio/Video Interleaved Format	fmt/5	1	0,13%
Tape Archive Format	x-fmt/265	1	0,13%
Quicktime	x-fmt/384	1	0,13%
Broadcast WAVE	fmt/1	1	0,13%

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



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

Permanent link:
https://wiki.docuteam.ch/doku.php?id=docuteam:feeder_300&rev=1430205628

Last update: **2019/01/07 11:16**

