

# docuteam feeder 3.8.0

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 bar 'Anmelden | docuteam feeder'. The address bar displays the URL '192.168.0.74:3000/session/new'. The main content area is titled 'Anmelden' (Login). It contains fields for 'Login \*' (with 'admin' entered) and 'Password \*'. There is also a 'Remember me' checkbox, which is checked, and a 'Login' button. At the bottom right of the form, there are links for 'Login | Dokumentation | Hilfe'.

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

ID	Workflow	SIP	Status	Startzeit	Dauer
462	Quality Assurance	böülli	•••• ●	Mittwoch, 01. April 2015, 12:30 Uhr	00:00:15 <a href="#">anzeigen</a>
461	Quality Assurance	böülli	•••• ●	Mittwoch, 01. April 2015, 12:29 Uhr	00:00:18 <a href="#">anzeigen</a>

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.
- **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
462	Quality Assurance	böütti	● ● ● ● ●	Mittwoch, 01. April 2015, 12:30 Uhr	00:00:15 <a href="#">anzeigen</a>
461	Quality Assurance	böütti	● ● ● ● ●	Mittwoch, 01. April 2015, 12:29 Uhr	00:00:18 <a href="#">anzeigen</a>

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_201401001_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_201401001_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_201401001_nu
INFO 2015-03-19T13:24:18.509 (FileConverter) Converting file: C:\\docuteam\\apps\\feeder_java\\temp\\METSDocument_read_20150319132412582\\182_RTR_docx-xlsx_201401001_nu\\2014-08-19 Seduta da direczin -09\\2014-08-19 Agiunta 02.1.3_Preschientscha DIR occurrentzas 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 Agiunta 02.1.3_Preschientscha DIR occurrentzas RTR.pdf
INFO 2015-03-19T13:24:35.322 (FileConverter) Executing command line: [Class:ch.docuteam.converter.FileConverter$OOConverter, C:\\docuteam\\apps\\feeder_java\\temp\\METSDocument_read_20150319132412582\\182_RTR_docx-xlsx_201401001_nu\\2014-08-19 Seduta da direczin -09\\2014-08-19 Agiunta 02.1.3_Preschientscha DIR occurrentzas RTR.xlsx, C:\\docuteam\\apps\\feeder_java\\.\\temp\\SIPFileMigrator\\Convert\\2014-08-19 Agiunta 02.1.3_Preschientscha DIR occurrentzas RTR.pdf]
INFO 2015-03-19T13:24:38.34 (FileConverter$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 titled "Workflows | docuteam feeder". The address bar displays the URL "192.168.0.74:3000/workflows". The main content area is titled "docuteam feeder" and contains a table titled "Workflows". The table has columns for "ID" and "Name". Each row in the table includes four red 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-based application for managing workflows. At the top, there's a header bar with the title "Workflow: Quality Assurance, Migr..." and a URL "192.168.0.74:3000/工作流/1". The main navigation menu includes "Ausführungen", "Workflows", "Schritte", "Benutzer", and "Cockpit". Below the header, the page title is "docuteam feeder". The main content area displays a workflow titled "Workflow: Quality Assurance, Migration and Storage". A section titled "Schritt-Präfix cd %FEEDER\_JAVA%" contains a numbered list of 13 Java commands. The footer provides creation and last modification details: "Erstellt durch admin um 01. November, 16:13 Uhr" and "Zuletzt bearbeitet von admin um 14. Mai, 10:33 Uhr". At the bottom right, there are links for "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.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.METSVvalidator "\${SIP}" true
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 net.sf.saxon.Transform -s:"//skylla/scratch/ingest/099-test/3\_done/\${SIP}/EAD.xml" -xsl:C:/docuteam/apps/feeder\_java/resources/xslt/ead2atom.xsl -o:"//skylla/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/工作flows/3/edit

Google

docuteam feeder

Ausführungen Workflows Schritte Benutzer Cockpit

## Workflow bearbeiten

Name \* Quality Assurance

Schritt-Prefix cd %FEEDER\_JAVA%

Quality Assu Parameter

Quality Assu Parameter

Quality Assu Parameter false

Quality Assu Parameter false

Ingest: clean 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 `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.



## Menüpunkt "Schritte"

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

ID	Name	anzeigen	bearbeiten	löschen
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.

Neuen Schritt erstellen | docute...

192.168.0.74:3000/steps/new

docuteam feeder

Ausführungen Workflows Schritte Benutzer Cockpit

### Neuen Schritt erstellen

Name \*

Command \*

Type \*

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



## Menüpunkt "Benutzer"

Benutzer | docuteam feeder

192.168.0.74:3000/users

docuteam feeder

Ausführungen Workflows Schritte Benutzer Cockpit

### Benutzer

Name	Rolle	bearbeiten	löschen
admin	Administrator	<a href="#">bearbeiten</a>	<a href="#">löschen</a>
user	Ausführende/r	<a href="#">bearbeiten</a>	<a href="#">löschen</a>

[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-based user management interface. At the top, there's a header bar with the title "Benutzer bearbeiten | docuteam fe...", a "+" button, and a search bar. Below the header, the URL "192.168.0.74:3000/users/2/edit" is visible. The main content area has a dark header "docuteam feeder" and a navigation menu with links like "Ausführungen", "Workflows", "Schritte", "Benutzer", and "Cockpit". The main form is titled "Benutzer bearbeiten" and contains the following fields:

- Name \*: Input field containing "user"
- E-Mail \*: Input field containing "user@docuteam.ch"
- Rolle: A dropdown menu set to "Ausführende/r"
- Passwort: An empty input field
- Bestätigung: An empty input field

A button labeled "Benutzer aktualisieren" is located at the bottom right of the form. At the very bottom of the page, there's a footer with the text "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.

Test2015 Win2012 Fedora 3.8 | doc...

192.168.0.74:3000/repositories/test Google

## docuteam feeder

Ausführungen Workflows Schritte Benutzer Cockpit

Erfolgreich angemeldet

### Test2015 Win2012 Fedora 3.8

Berichte ▾

Speicherplatz	
CH-333666-x	2,1 MB
test	540 MB
<b>Total</b>	<b>540 MB</b>
Frei	7,7 TB

Häufigste Dateiformate	
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 (PCMWAVEFORMAT)	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%

Letzte Integritätsprüfung	
Freitag, 20. März 2015, 10:49 Uhr	

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

Ablieferungen	
	CH-333666-x 3
	test 83
<b>Total</b>	<b>86</b>

Objekte	
	CH-333666-x 9
	test 764
<b>Total</b>	<b>773</b>

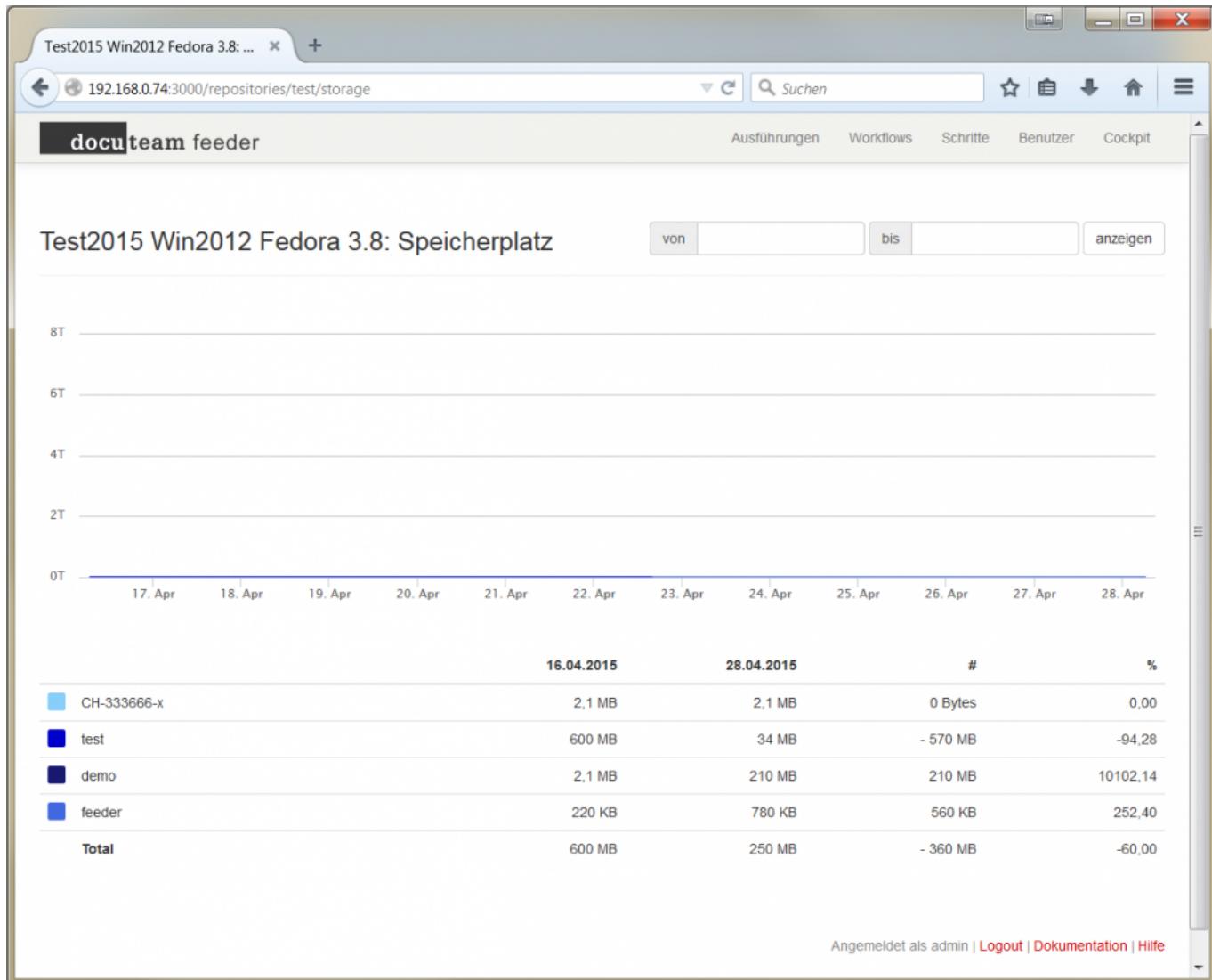
Angemeldet als admin | [Logout](#) | [Dokumentation](#) | [Hilfe](#)

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

Test2015 Win2012 Fedora 3.8: Dateiformate

Total	CH-333666-x	test	
<b>Name</b>		<b>PUID</b>	<b>#</b>
Acrobat PDF/A - Portable Document Format		fmt/95	366
Tagged Image File Format		fmt/353	111
Acrobat PDF/A - Portable Document Format		fmt/354	40
Plain Text File		x-fmt/111	28
JPEG File Interchange Format		fmt/43	26
Portable Network Graphics		fmt/12	10
Acrobat PDF 1.4 - Portable Document Format		fmt/18	6
Acrobat PDF 1.3 - Portable Document Format		fmt/17	4
Waveform Audio (PCMWAVEFORMAT)		fmt/141	3
Tagged Image File Format for Internet Fax (TIFF-FX)		fmt/156	3
Acrobat PDF 1.5 - Portable Document Format		fmt/19	3
MPEG-4 Media File		fmt/199	3
MPEG 1/2 Audio Layer 3		fmt/134	2
MPEG-1 Video Format		x-fmt/385	2
Acrobat PDF 1.6 - Portable Document Format		fmt/20	1
Audio/Video Interleaved Format		fmt/5	1
Tape Archive Format		x-fmt/265	1
Quicktime		x-fmt/384	1
Broadcast WAVE		fmt/1	1

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



From:

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

Permanent link:

[https://wiki.docuteam.ch/doku.php?id=docuteam:feeder\\_380&rev=1500037147](https://wiki.docuteam.ch/doku.php?id=docuteam:feeder_380&rev=1500037147)

Last update: **2019/01/07 11:12**

