

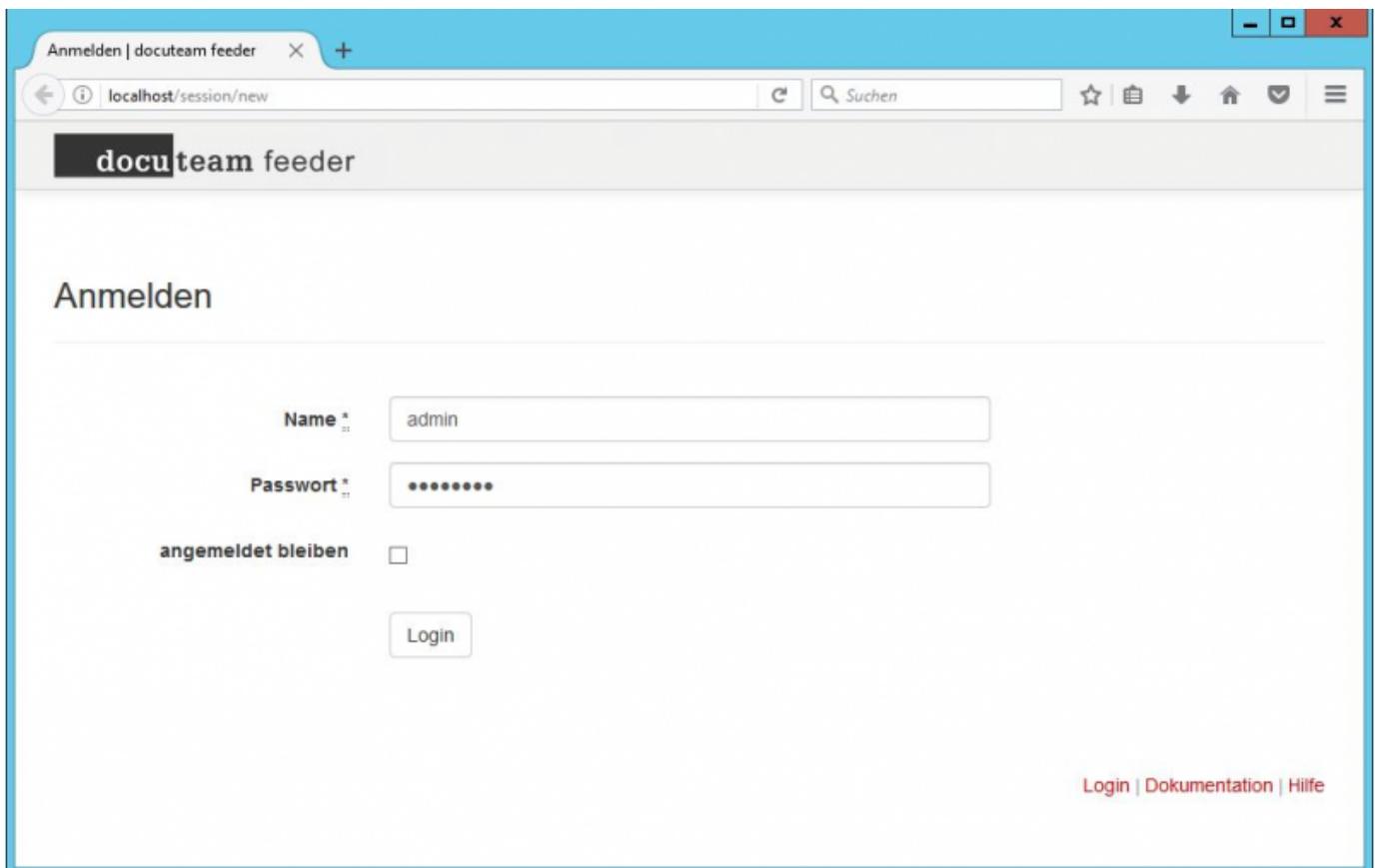
# docuteam feeder 3.4.0

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.



Anmelden | docuteam feeder

localhost/session/new

docuteam feeder

Anmelden

Name \* admin

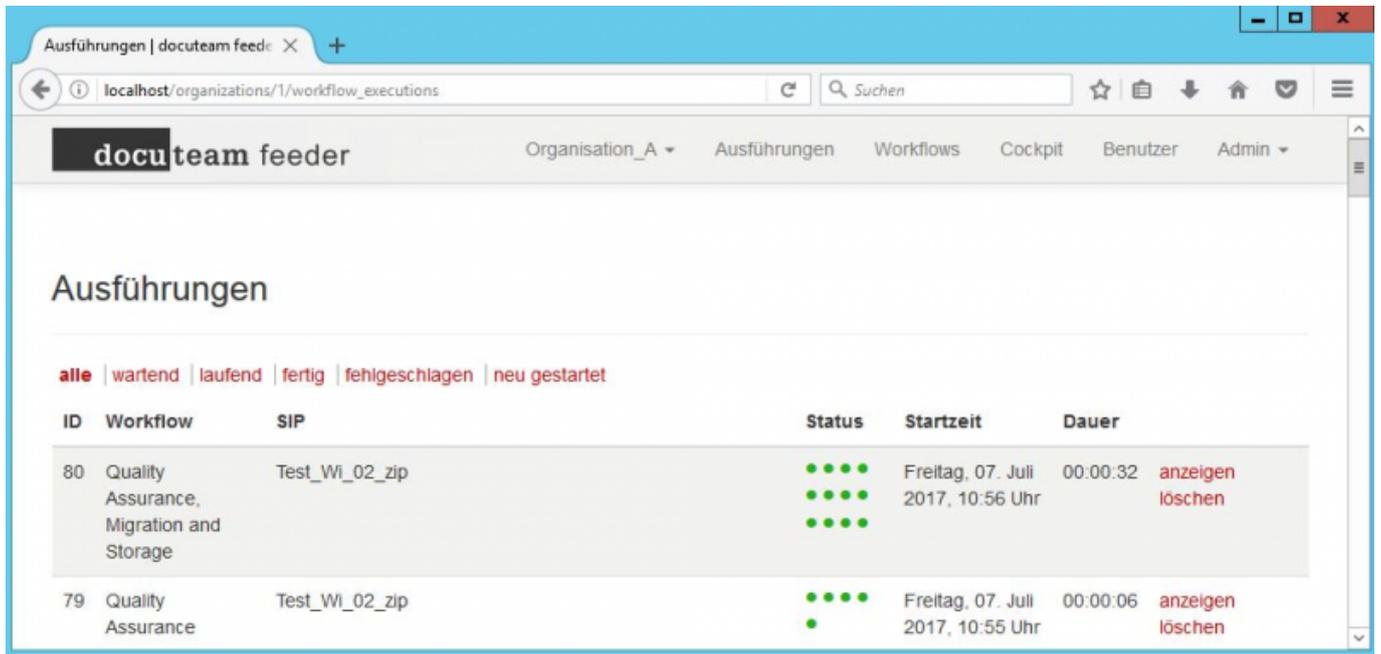
Passwort \* .....

angemeldet bleiben

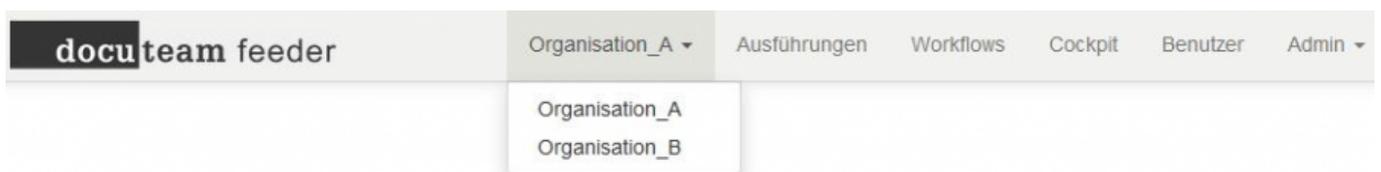
Login

Login | Dokumentation | Hilfe

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



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:

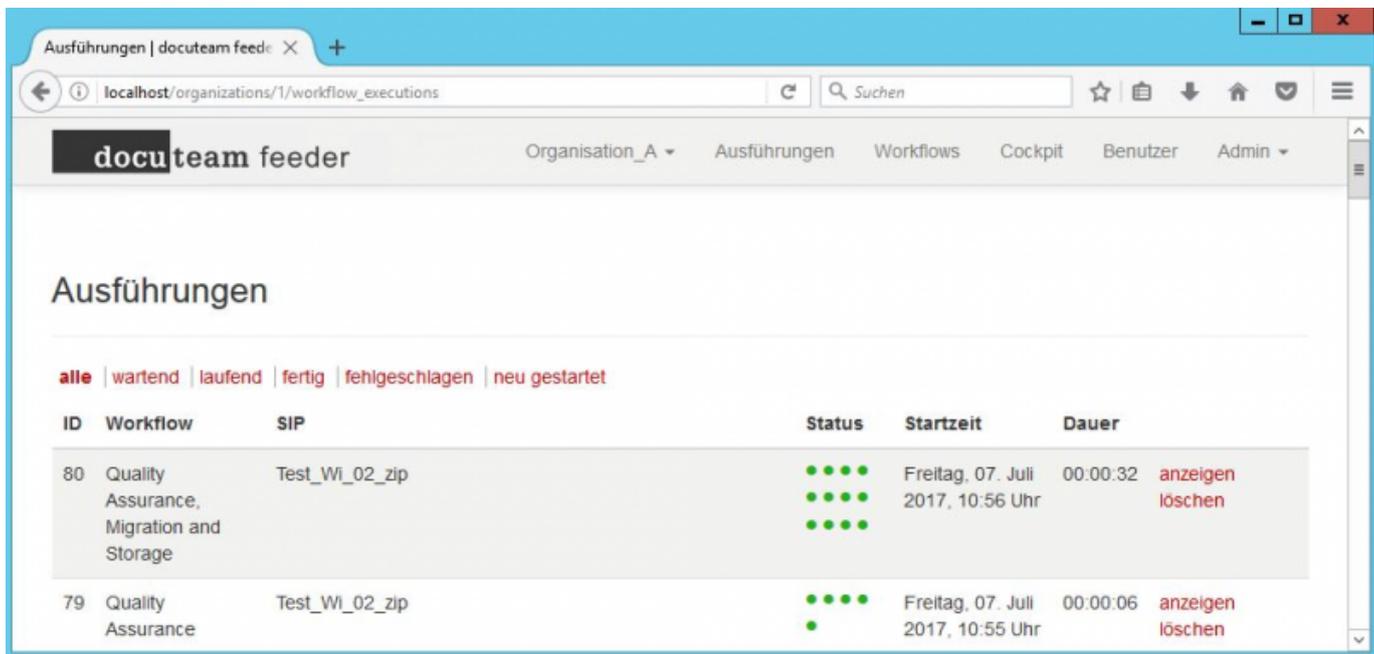


- **Ausführungen:** Dieser Bereich zeigt die Zusammenfassung aller bislang durchgeführten, laufenden und bevorstehenden Ausführungen. Als Ausführung wird die konkrete Verarbeitung eines SIP anhand eines spezifischen Workflows bezeichnet.
- **Workflows:** Als Workflow wird eine geordnete Abfolge von definierten 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.



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$PDFToolsConverter) -214322271
ERROR 2016-07-13T16:18:01.867 (FileConverter$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
```

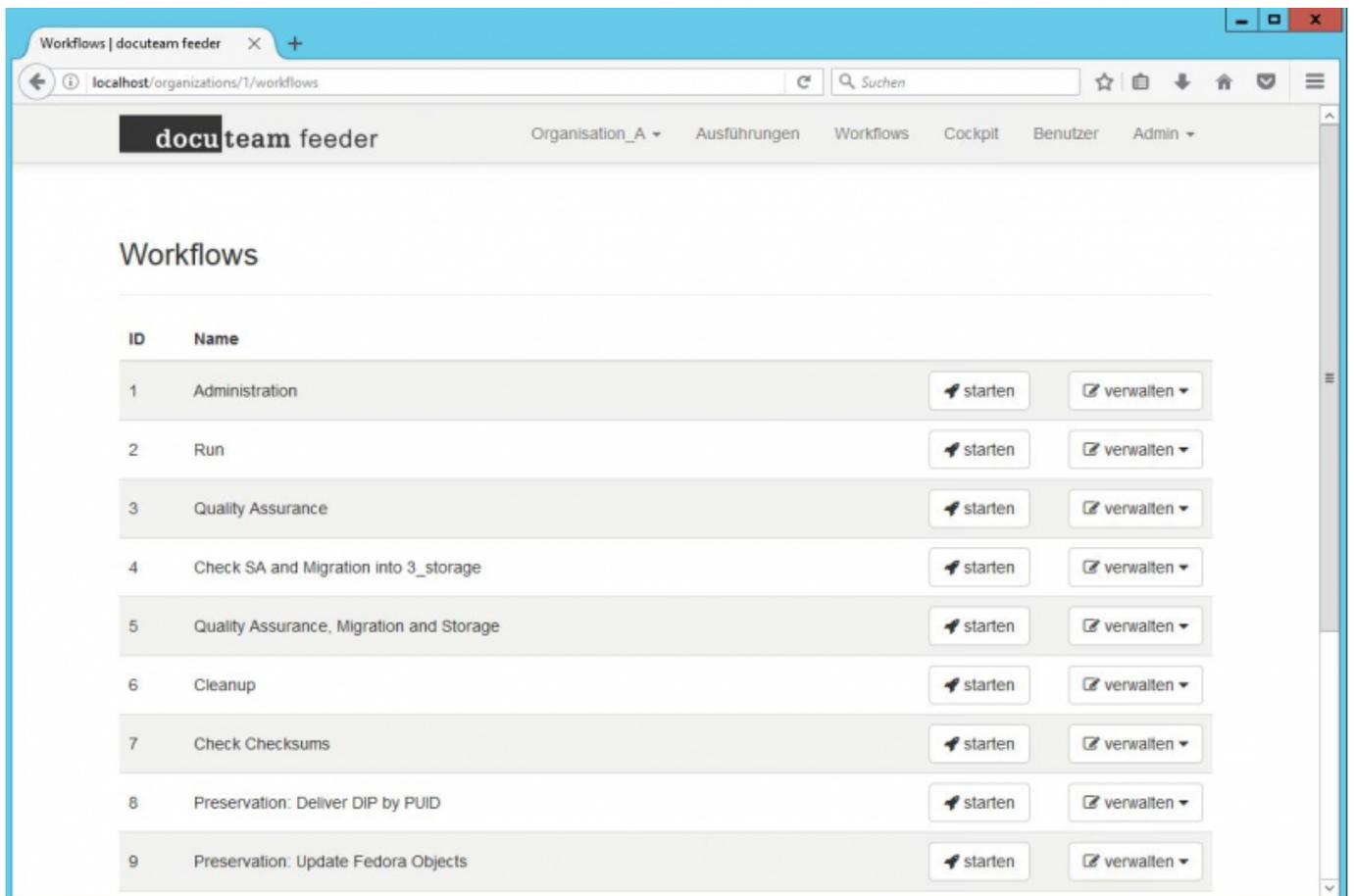
Dauer: 00:04:11

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

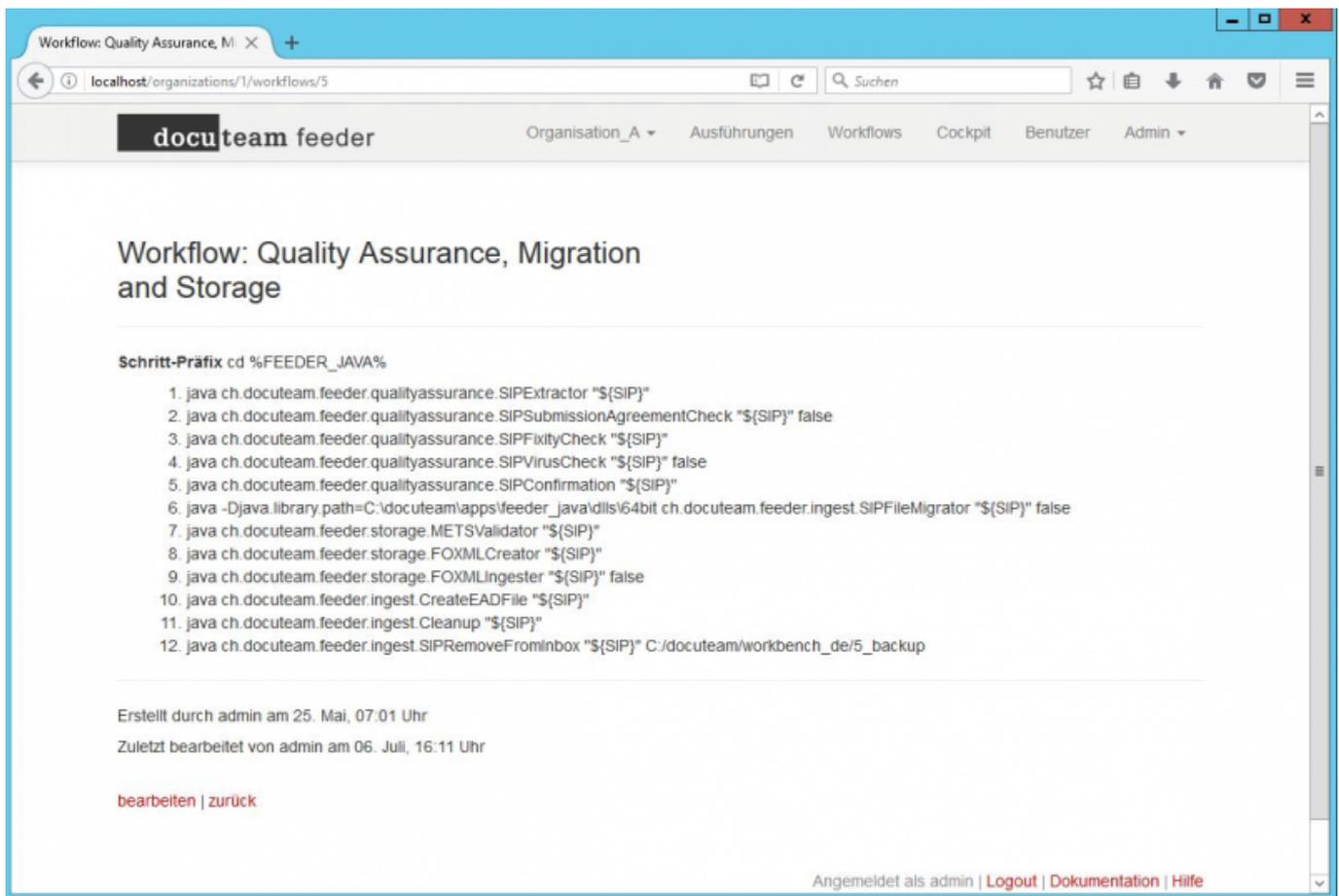


## Menüpunkt "Workflows"

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



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



The screenshot shows a web browser window with the URL `localhost/organizations/1/workflows/5`. The page title is "Workflow: Quality Assurance, Migration and Storage". The breadcrumb navigation shows "Organisation\_A > Ausführungen > Workflows > Cockpit > Benutzer > Admin". The main content area displays the workflow steps:

**Schritt-Präfix** `cd %FEEDER_JAVA%`

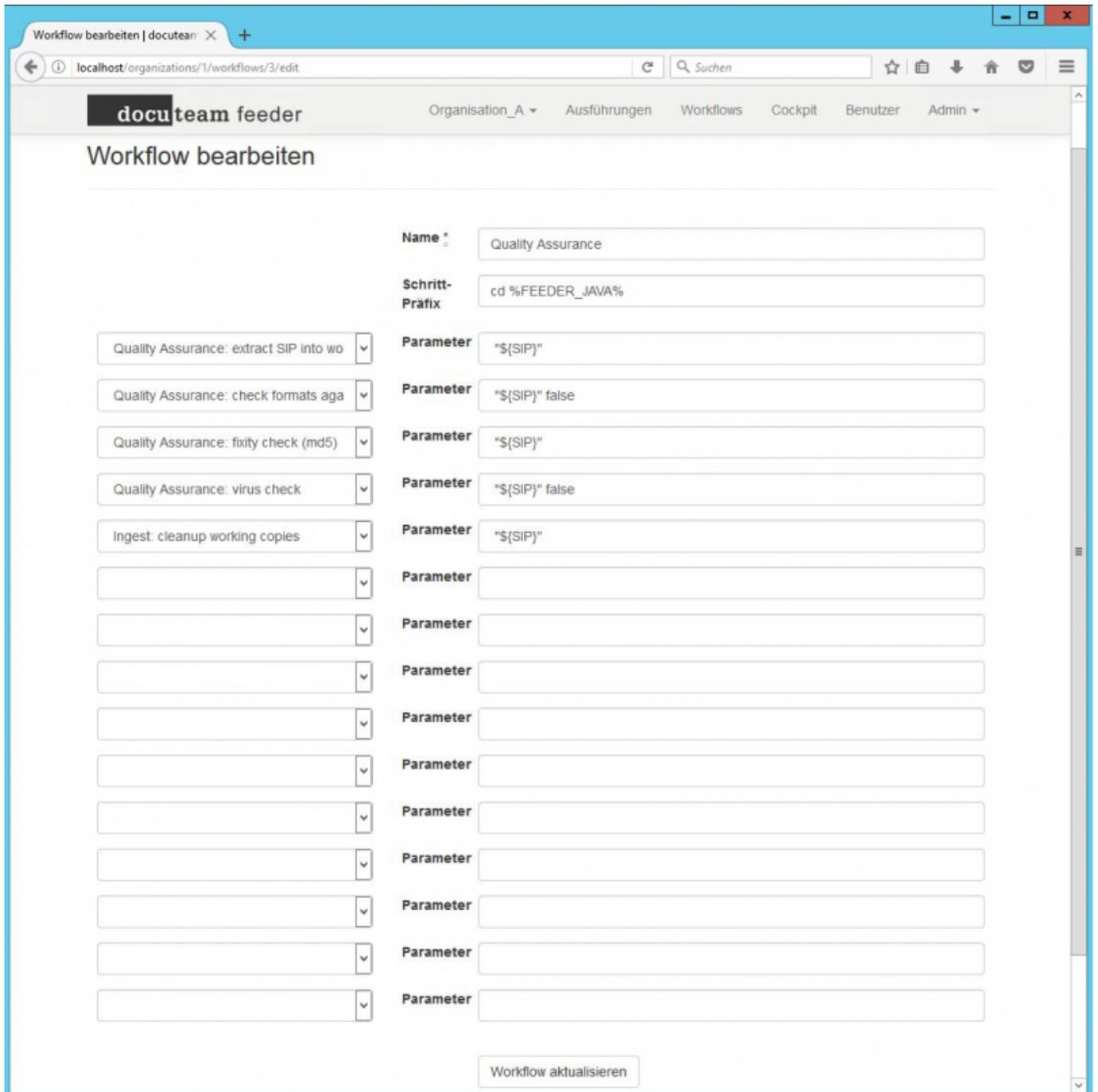
1. `java ch.docuteam.feeder.qualityassurance.SIPExtractor "${SIP}"`
2. `java ch.docuteam.feeder.qualityassurance.SIPSubmissionAgreementCheck "${SIP}" false`
3. `java ch.docuteam.feeder.qualityassurance.SIPFixityCheck "${SIP}"`
4. `java ch.docuteam.feeder.qualityassurance.SIPVirusCheck "${SIP}" false`
5. `java ch.docuteam.feeder.qualityassurance.SIPConfirmation "${SIP}"`
6. `java -Djava.library.path=C:\docuteam\apps\feeder_java\dlls\64bit ch.docuteam.feeder.ingest.SIPFileMigrator "${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`

Erstellt durch admin am 25. Mai, 07:01 Uhr  
Zuletzt bearbeitet von admin am 06. Juli, 16:11 Uhr

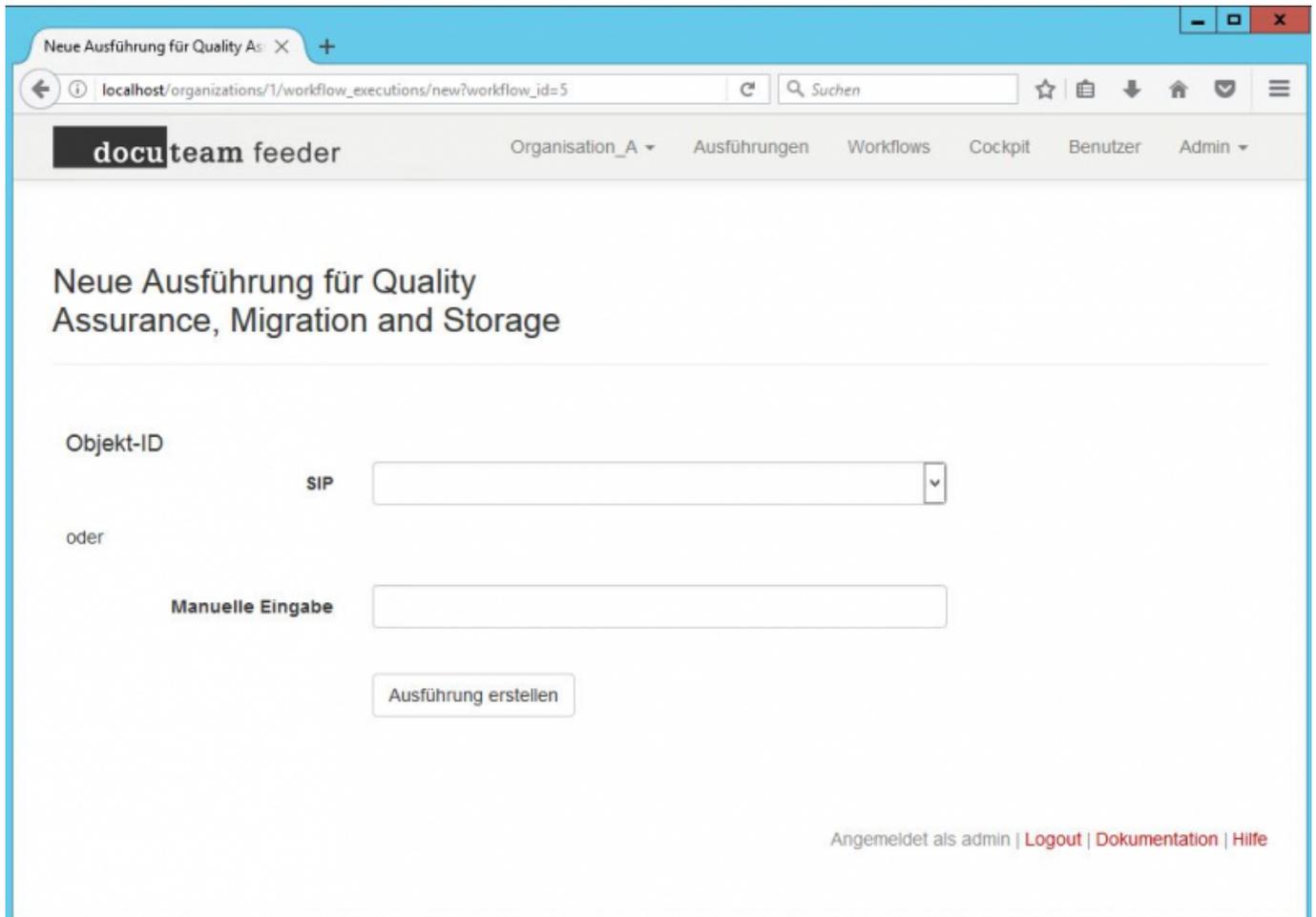
[bearbeiten](#) | [zurück](#)

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.

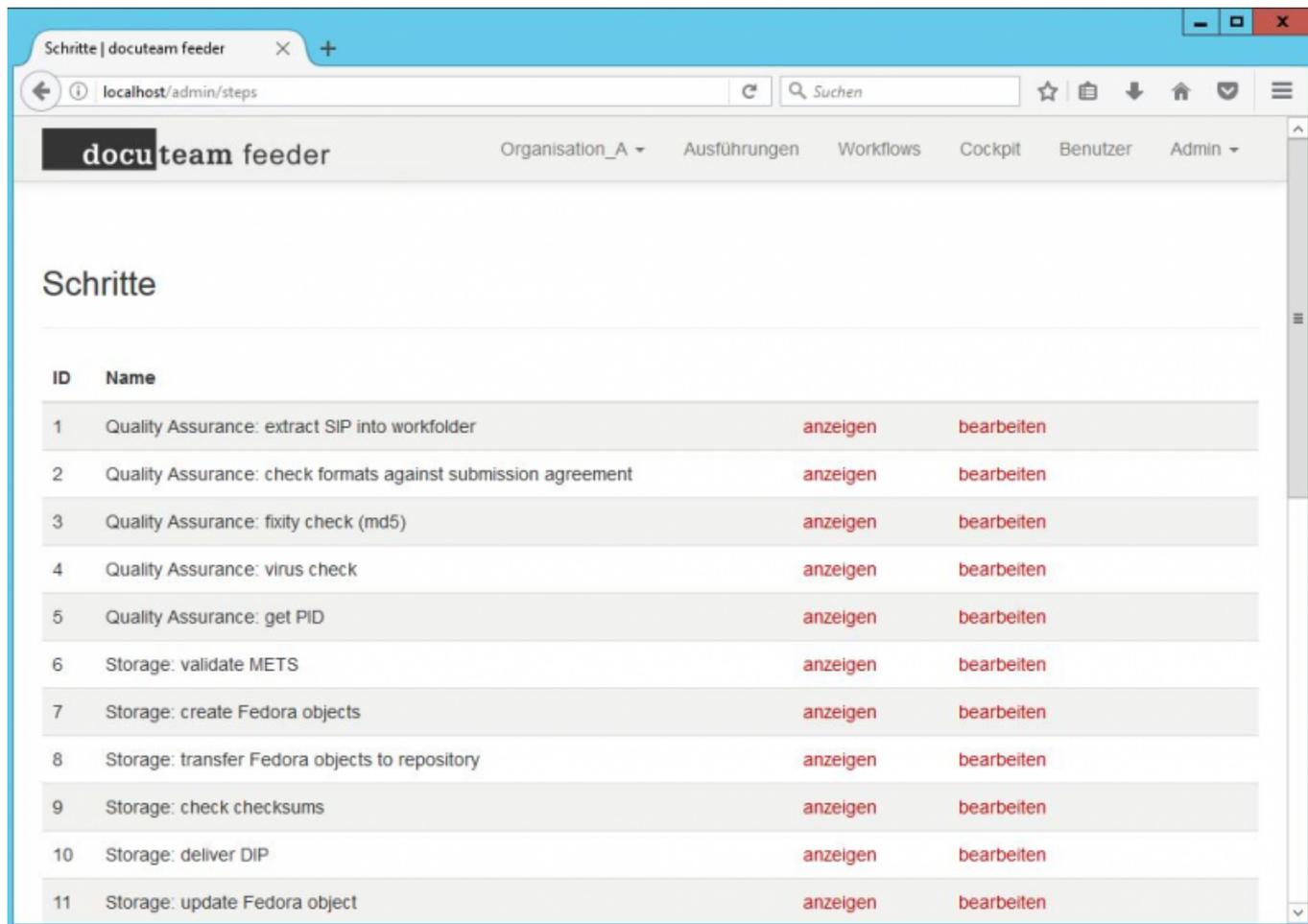


Für den Testbetrieb kann der Ingest einzelner SIPs auch manuell ausgelöst werden. Diesem Zweck dient der Befehl `starten`. Im Formularfeld „SIP“, das nach der Aktivierung des Befehls erscheint, wird der Name des SIP angegeben. **Wichtig:** der Name sollte normalerweise ohne die Endung „.zip“ angegeben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



## Menüpunkt "Schritte"

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



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 URL `localhost/admin/steps/new`. The page title is "Neuen Schritt erstellen | docuteam feeder". The navigation bar includes "Organisation\_A", "Ausführungen", "Workflows", "Cockpit", "Benutzer", and "Admin". The main content area has the heading "Neuen Schritt erstellen" and a form with three fields: "Name" (text input), "Command" (text area), and "Type" (dropdown menu set to "ScriptStep"). A "Schritt erstellen" button is located below the form. At the bottom right, it says "Angemeldet als admin | [Logout](#) | [Dokumentation](#) | [Hilfe](#)".



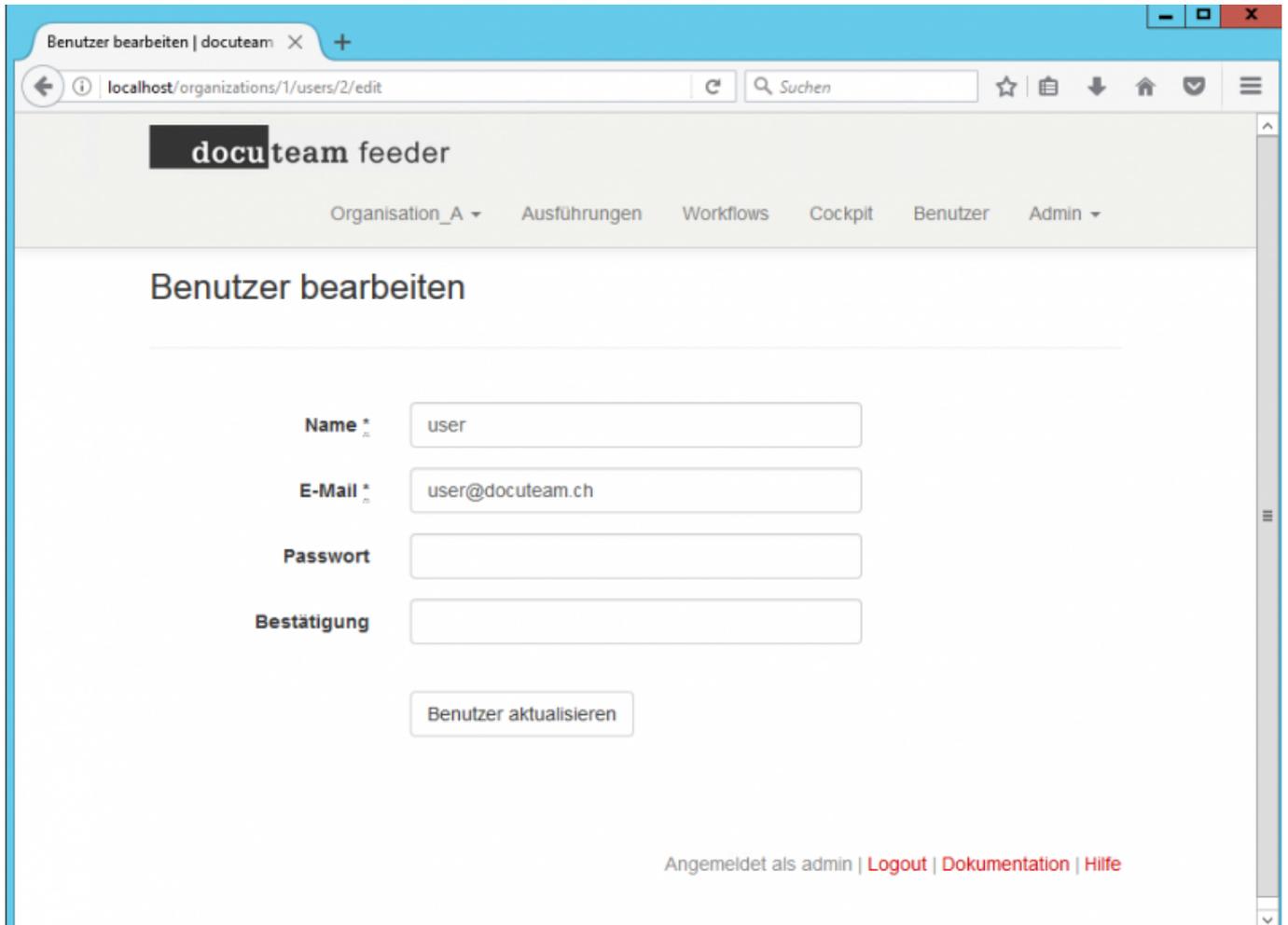
## Menüpunkt "Benutzer"

The screenshot shows a web browser window with the URL `localhost/organizations/1/users`. The page title is "Benutzer | docuteam feeder". The navigation bar includes "Organisation\_A", "Ausführungen", "Workflows", "Cockpit", "Benutzer", and "Admin". The main content area has the heading "Benutzer" and a table of users. Below the table is a link "Neuen Benutzer erstellen". At the bottom right, it says "Angemeldet als admin | [Logout](#) | [Dokumentation](#) | [Hilfe](#)".

Name		
admin	<a href="#">bearbeiten</a>	<a href="#">löschen</a>
peter	<a href="#">bearbeiten</a>	<a href="#">löschen</a>

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.



## 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](#)).

The screenshot shows the 'docuteam feeder' web interface. The browser address bar shows 'localhost/organizations/1/repositories/demo'. The page title is 'Demo'. A navigation bar includes 'Organisation\_A', 'Ausführungen', 'Workflows', 'Cockpit', 'Benutzer', and 'Admin'. A 'Berichte' dropdown menu is visible in the top right.

The main content area is divided into several sections:

- Speicherplatz:** A progress bar shows storage usage. A table lists usage for 'de' (58 MB) and 'test' (950 MB), with a total of 1000 MB and 60 GB free.
- Letzte Integritätsprüfung:** Shows a status of '-'.
- Basisdaten:** A table with fields: Name (Demo), Basis-URL (http://localhost:8080/fedora), Version (3.8.1), PID Namespace (docuteam), and Support-Mail (support@docuteam.ch).
- Häufigste Dateiformate:** A table listing file formats and their counts/percentages.
- Ablieferungen:** A table showing counts for 'de' (2) and 'test' (17), totaling 19.
- Objekte:** A table showing counts for 'de' (9) and 'test' (2231), totaling 2240.

At the bottom right, there is a footer: 'Angemeldet als admin | [Logout](#) | [Dokumentation](#) | [Hilfe](#)'.

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

The screenshot shows a web browser window with the URL `localhost/organizations/1/repositories/demo/file_formats`. The page title is "Demo: Dateiformate". Below the title, there are tabs for "Total", "de", and "test", with "Total" selected. The main content is a table listing file formats, their PUIDs, counts, and percentages.

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 %

Browser window: localhost/organizations/1/repositories/demo/storage

### docuteam feeder

Organisation\_A | Ausführungen | Workflows | Cockpit | Benutzer | Admin

## Demo: Speicherplatz

von  bis   
[anzeigen]

	11.07.2017	11.07.2017	#	%
de	58 MB	58 MB	0 Bytes	0,00
test	950 MB	950 MB	0 Bytes	0,00
<b>Total</b>	<b>1000 MB</b>	<b>1000 MB</b>	<b>0 Bytes</b>	<b>0,00</b>

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



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

Permanent link: [https://wiki.docuteam.ch/doku.php?id=docuteam:feeder\\_330&rev=1499773688](https://wiki.docuteam.ch/doku.php?id=docuteam:feeder_330&rev=1499773688)

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

