

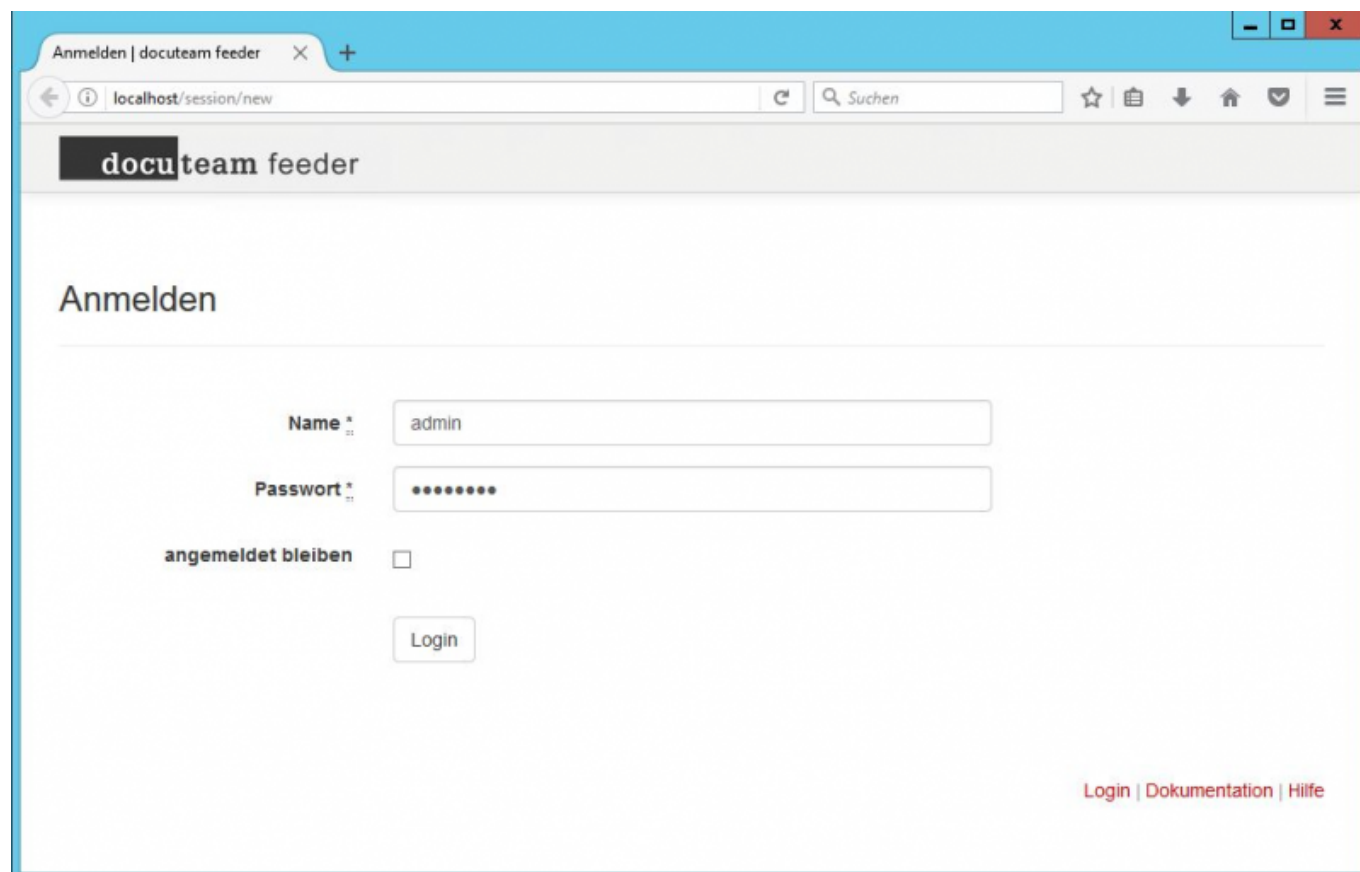
docuteam feeder 3.4.0

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 shows 'localhost/session/new'. The page has a header with the 'docuteam feeder' logo. Below the header, the title 'Anmelden' is displayed. The login form contains the following elements:

- A 'Name' field with a red asterisk and a red underline, containing the text 'admin'.
- A 'Passwort' field with a red asterisk and a red underline, containing masked characters '*****'.
- A checkbox labeled 'angemeldet bleiben'.
- A 'Login' button.

At the bottom right of the page, there are links: 'Login | Dokumentation | Hilfe'.

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

Ausführungen | docuteam feeder

Organisation_A | Ausführungen | Workflows | Cockpit | Benutzer | Admin

Ausführungen

alle | wartend | laufend | fertig | fehlgeschlagen | neu gestartet

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 löschen
79	Quality Assurance	Test_Wi_02_zip	● ● ● ●	Freitag, 07. Juli 2017, 10:55 Uhr	00:00:06	anzeigen löschen

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:

docuteam feeder

Organisation_A | Ausführungen | Workflows | Cockpit | Benutzer | Admin

Organisation_A
Organisation_B

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

docuteam feeder

Ausführungen | Workflows | Cockpit

- **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.

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
79	Quality Assurance	Test_Wi_02_zip	● ● ● ● ●	Freitag, 07. Juli 2017, 10:55 Uhr	00:00:06

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.

The screenshot shows the 'docuteam feeder' web application. The top navigation bar includes 'Organisation_A', 'Ausführungen', 'Workflows' (selected), 'Cockpit', 'Benutzer', and 'Admin'. The main content area is titled 'Workflows' and contains a table with the following data:

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 address bar displaying `localhost/organizations/1/workflows/5`. The page title is "Workflow: Quality Assurance, Migration and Storage". The breadcrumb navigation shows "Organisation_A" > "Ausführungen" > "Workflows". The main content area is titled "Workflow: Quality Assurance, Migration and Storage" and lists 12 steps under the prefix "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
```

Below the steps, it shows the creation and last modification dates:

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

At the bottom, there are links for "bearbeiten" (in red) and "zurück", and a footer with "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

localhost/organizations/1/workflows/3/edit
Suchen

docuteam feeder
Organisation_A ▾ Ausführungen Workflows Cockpit Benutzer Admin ▾

Workflow bearbeiten

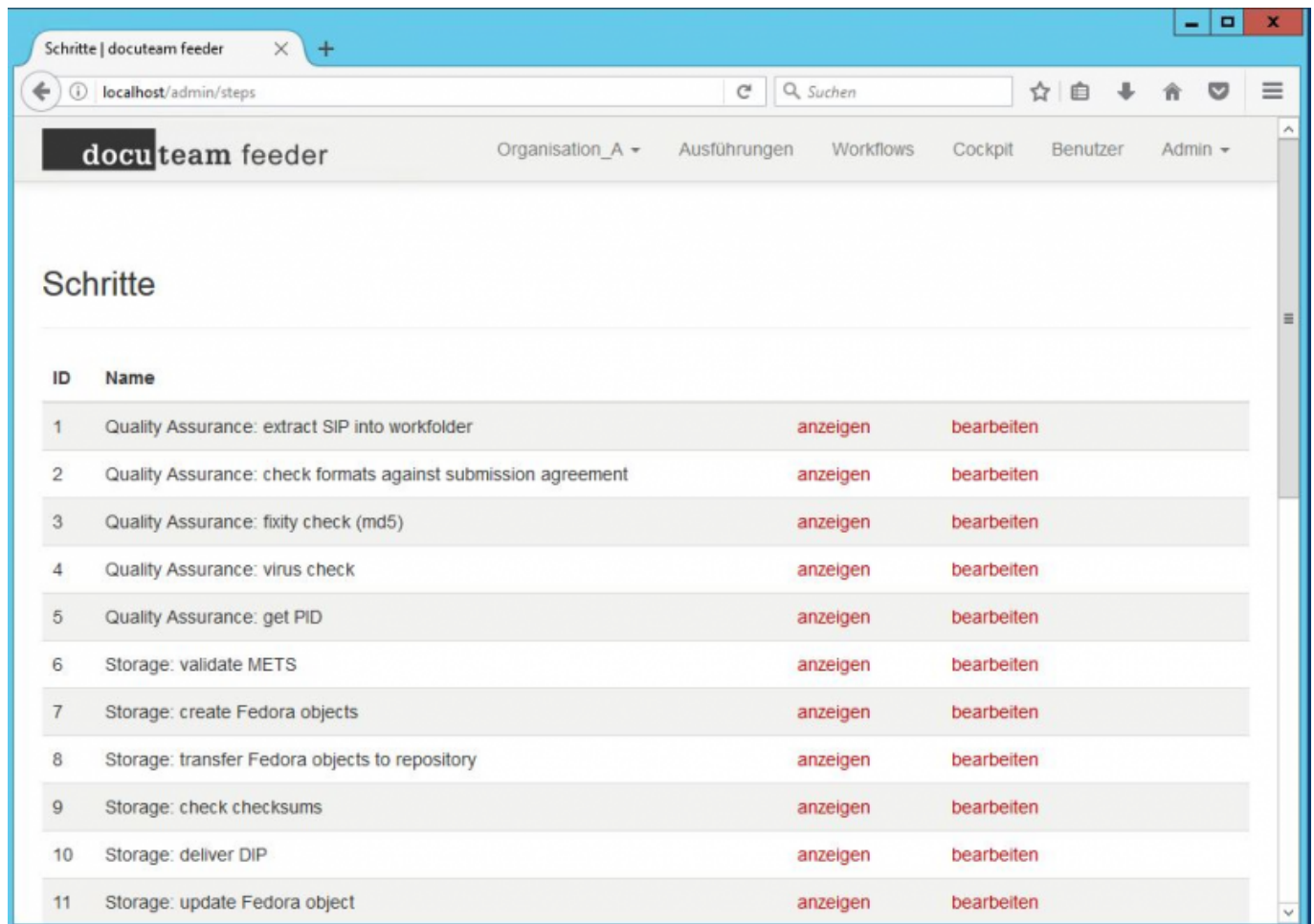
<div style="border: 1px solid #ccc; padding: 2px; margin-bottom: 5px;">Quality Assurance: extract SIP into wo ▾</div> <div style="border: 1px solid #ccc; padding: 2px; margin-bottom: 5px;">Quality Assurance: check formats aga ▾</div> <div style="border: 1px solid #ccc; padding: 2px; margin-bottom: 5px;">Quality Assurance: fixity check (md5) ▾</div> <div style="border: 1px solid #ccc; padding: 2px; margin-bottom: 5px;">Quality Assurance: virus check ▾</div> <div style="border: 1px solid #ccc; padding: 2px; margin-bottom: 5px;">Ingest: cleanup working copies ▾</div> <div style="border: 1px solid #ccc; padding: 2px; margin-bottom: 5px;">▾</div> <div style="border: 1px solid #ccc; padding: 2px; margin-bottom: 5px;">▾</div> <div style="border: 1px solid #ccc; padding: 2px; margin-bottom: 5px;">▾</div> <div style="border: 1px solid #ccc; padding: 2px; margin-bottom: 5px;">▾</div> <div style="border: 1px solid #ccc; padding: 2px; margin-bottom: 5px;">▾</div> <div style="border: 1px solid #ccc; padding: 2px; margin-bottom: 5px;">▾</div> <div style="border: 1px solid #ccc; padding: 2px; margin-bottom: 5px;">▾</div> <div style="border: 1px solid #ccc; padding: 2px; margin-bottom: 5px;">▾</div> <div style="border: 1px solid #ccc; padding: 2px; margin-bottom: 5px;">▾</div> <div style="border: 1px solid #ccc; padding: 2px; margin-bottom: 5px;">▾</div> <div style="border: 1px solid #ccc; padding: 2px; margin-bottom: 5px;">▾</div>	<p>Name * <input type="text" value="Quality Assurance"/></p> <p>Schritt-Präfix <input type="text" value="cd %FEEDER_JAVA%"/></p> <p>Parameter <input style="font-family: monospace; font-size: small; color: gray; background-color: #f0f0f0;" type="text" value="\${SIP}"/></p> <p>Parameter <input false="" style="font-family: monospace; font-size: small; color: gray; background-color: #f0f0f0;" type="text" value="\${SIP}"/></p> <p>Parameter <input style="font-family: monospace; font-size: small; color: gray; background-color: #f0f0f0;" type="text" value="\${SIP}"/></p> <p>Parameter <input false="" style="font-family: monospace; font-size: small; color: gray; background-color: #f0f0f0;" type="text" value="\${SIP}"/></p> <p>Parameter <input style="font-family: monospace; font-size: small; color: gray; background-color: #f0f0f0;" type="text" value="\${SIP}"/></p> <p>Parameter <input style="font-family: monospace; font-size: small; color: gray; background-color: #f0f0f0;" type="text" value=""/></p> <p>Parameter <input style="font-family: monospace; font-size: small; color: gray; background-color: #f0f0f0;" type="text" value=""/></p> <p>Parameter <input style="font-family: monospace; font-size: small; color: gray; background-color: #f0f0f0;" type="text" value=""/></p> <p>Parameter <input style="font-family: monospace; font-size: small; color: gray; background-color: #f0f0f0;" type="text" value=""/></p> <p>Parameter <input style="font-family: monospace; font-size: small; color: gray; background-color: #f0f0f0;" type="text" value=""/></p> <p>Parameter <input style="font-family: monospace; font-size: small; color: gray; background-color: #f0f0f0;" type="text" value=""/></p> <p>Parameter <input style="font-family: monospace; font-size: small; color: gray; background-color: #f0f0f0;" type="text" value=""/></p> <p>Parameter <input style="font-family: monospace; font-size: small; color: gray; background-color: #f0f0f0;" type="text" value=""/></p> <p>Parameter <input style="font-family: monospace; font-size: small; color: gray; background-color: #f0f0f0;" type="text" value=""/></p> <p>Parameter <input style="font-family: monospace; font-size: small; color: gray; background-color: #f0f0f0;" type="text" value=""/></p> <p>Parameter <input style="font-family: monospace; font-size: small; color: gray; background-color: #f0f0f0;" type="text" value=""/></p>
---	---

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 URL `localhost/organizations/1/workflow_executions/new?workflow_id=5`. The page title is "Neue Ausführung für Quality Assurance, Migration and Storage". The application header includes the "docuteam feeder" logo and a navigation menu with "Organisation_A", "Ausführungen", "Workflows", "Cockpit", "Benutzer", and "Admin". The main form contains two input fields: "Objekt-ID" with a "SIP" dropdown menu, and "Manuelle Eingabe". A "Ausführung erstellen" button is located below the "Manuelle Eingabe" field. The footer indicates the user is logged in as "admin" and provides links for "Logout", "Dokumentation", and "Hilfe".

Menüpunkt "Schritte"

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



The screenshot shows a web browser window with the address bar at `localhost/admin/steps`. The page title is "Schritte | docuteam feeder". The navigation bar includes "Organisation_A", "Ausführungen", "Workflows", "Cockpit", "Benutzer", and "Admin". The main heading is "Schritte". Below it is a table with 11 steps, each with an ID, a name, and two action buttons: "anzeigen" and "bearbeiten".

ID	Name		
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

localhost/admin/steps/new

Suchen

docuteam feeder

Organisation_A

Ausführungen

Workflows

Cockpit

Benutzer

Admin

Neuen Schritt erstellen

Name *

Command *

Type *

ScriptStep

Schritt erstellen

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



Menüpunkt "Benutzer"

Benutzer | docuteam feeder

localhost/organizations/1/users

Suchen

docuteam feeder

Organisation_A

Ausführungen

Workflows

Cockpit

Benutzer

Admin

Benutzer

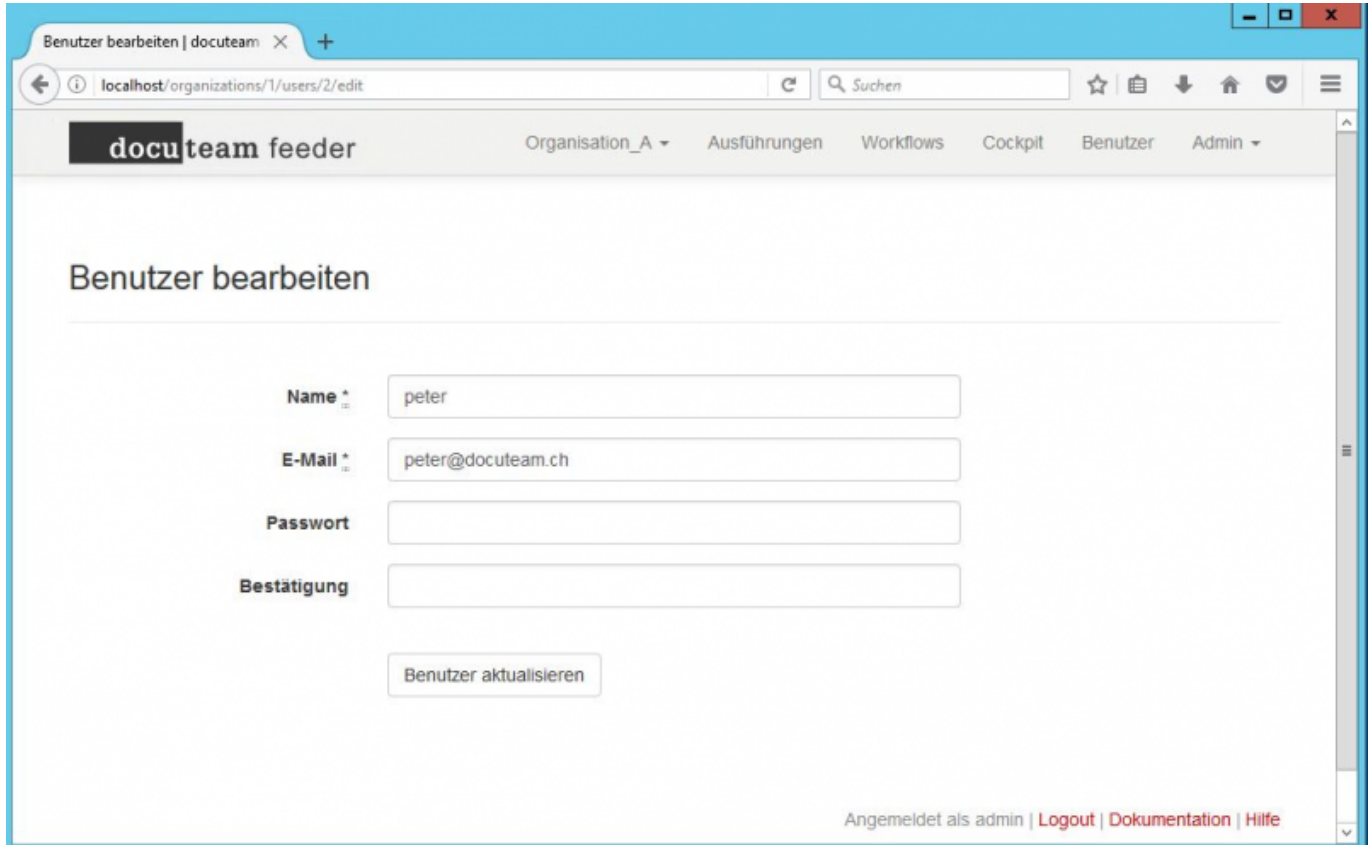
Name		
admin	bearbeiten	löschen
peter	bearbeiten	löschen

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.

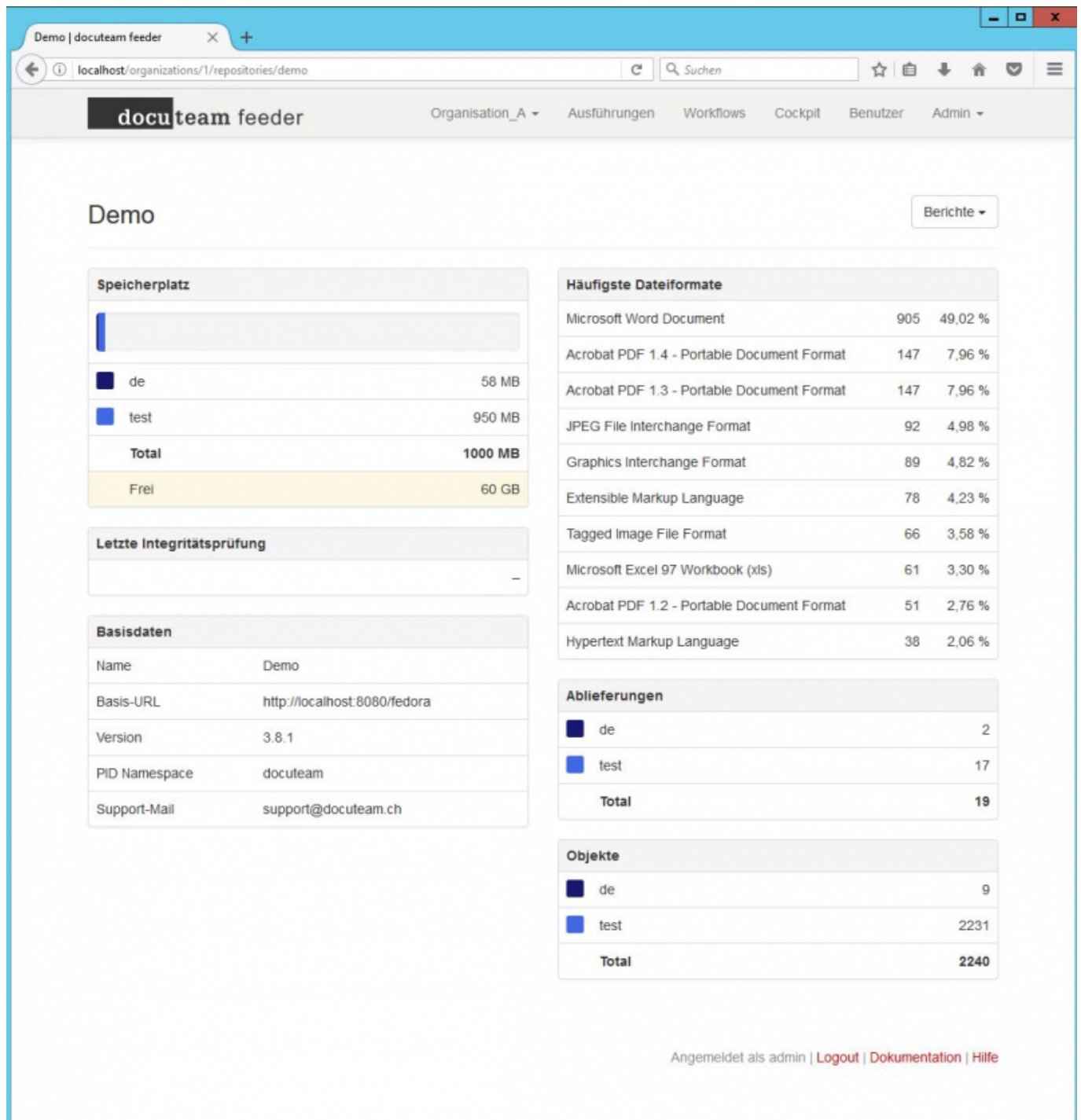


The screenshot shows a web browser window with the address bar displaying 'localhost/organizations/1/users/2/edit'. The page title is 'Benutzer bearbeiten | docuteam'. The navigation bar includes 'Organisation_A', 'Ausführungen', 'Workflows', 'Cockpit', 'Benutzer', and 'Admin'. The main heading is 'Benutzer bearbeiten'. Below it, there are four input fields: 'Name *' with the value 'peter', 'E-Mail *' with the value 'peter@docuteam.ch', 'Passwort', and 'Bestätigung'. A button labeled 'Benutzer aktualisieren' is at the bottom. The footer shows '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](#)).



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

Demo: Dateiformate | docuteam: X +

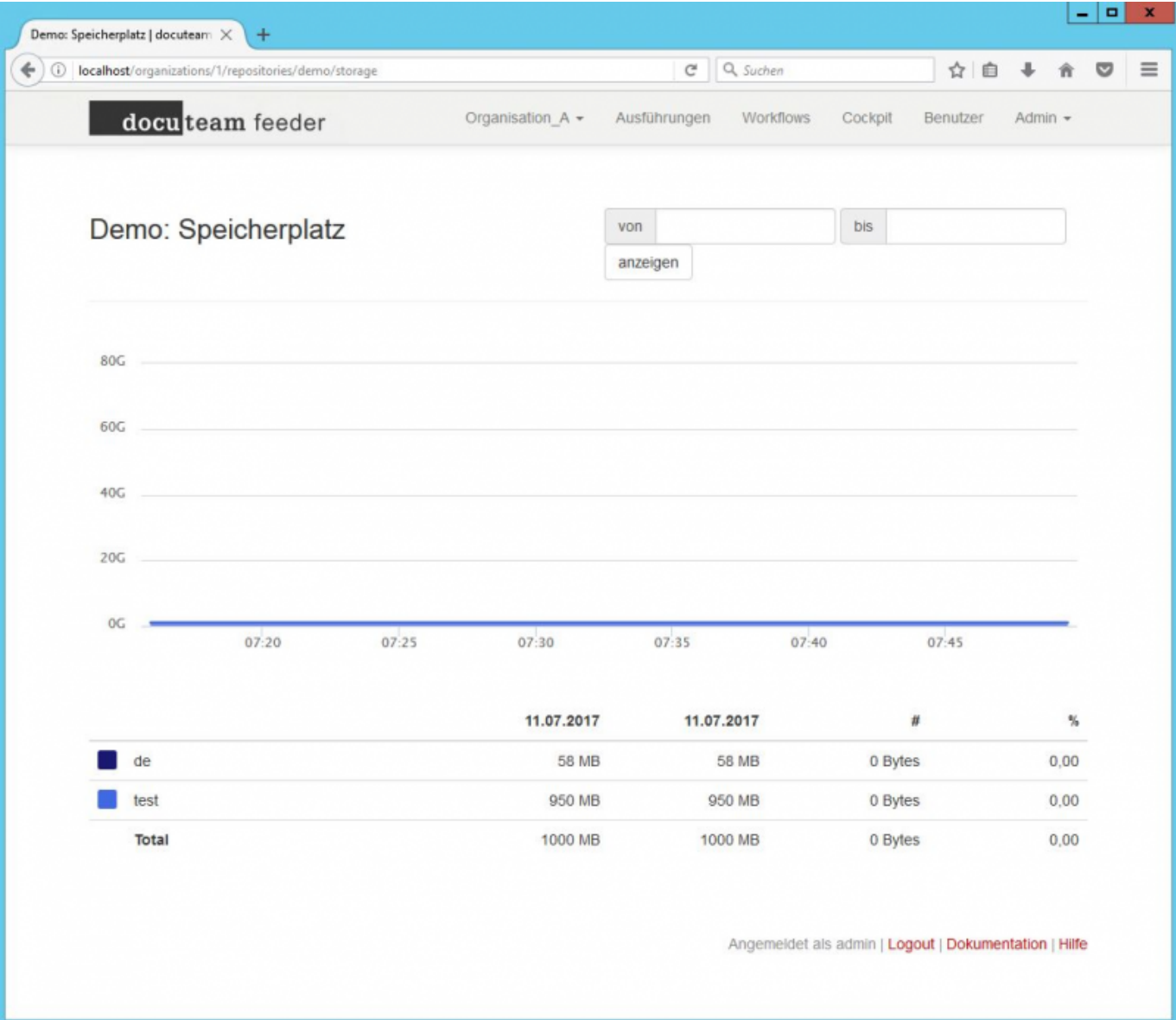
localhost/organizations/1/repositories/demo/file_formats Suchen

docuteam feeder Organisation_A Ausführungen Workflows Cockpit Benutzer Admin

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=1499773423

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

