

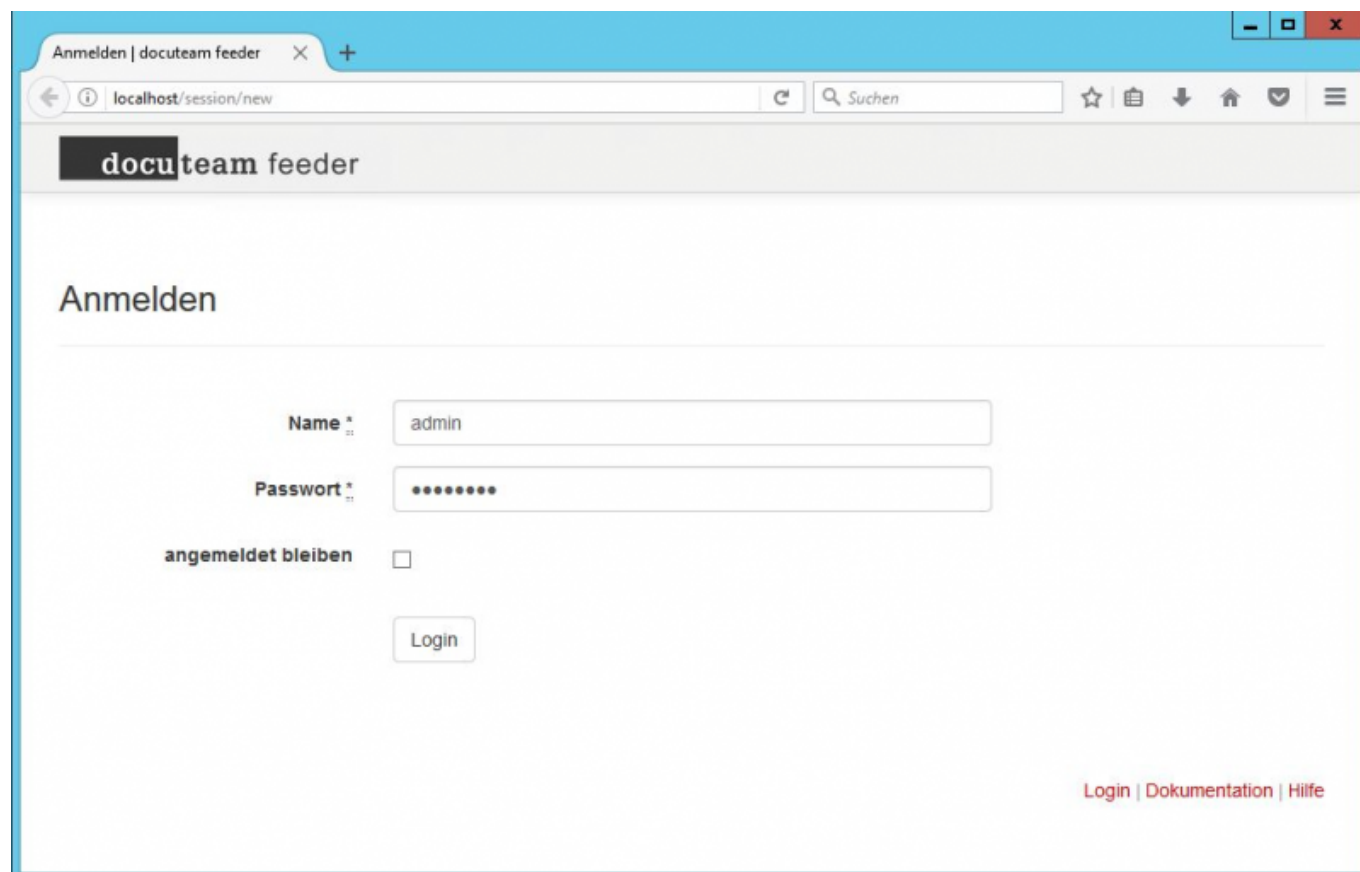
# docuteam feeder 3.4.0

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 search bar with the placeholder 'Suchen'. The main heading is 'docuteam feeder'. Below it, the section is titled 'Anmelden'. There are two input fields: 'Name \*' with the value 'admin' and 'Passwort \*' with masked characters '.....'. Below these is a checkbox labeled 'angemeldet bleiben' which is unchecked. A 'Login' button is positioned below the checkbox. In the bottom right corner, 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<br>löschen |
| 79 | Quality Assurance                        | Test_Wi_02_zip | ● ● ● ● | Freitag, 07. Juli 2017, 10:55 Uhr | 00:00:06 | anzeigen<br>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 'Workflows' page in the 'docuteam feeder' application. The page has a navigation bar with 'Organisation\_A', 'Ausführungen', 'Workflows', 'Cockpit', 'Benutzer', and 'Admin'. The main content area is titled 'Workflows' and contains a table with 9 rows. Each row has an 'ID' column, a 'Name' column, and two action buttons: 'starten' and 'verwalten'.

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

Mit „Verwalten“ und „Anzeigen“ werden die einzelnen Schritte des jeweiligen Workflows angezeigt.

The screenshot displays the Docuteam Feeder web application. The browser address bar shows `localhost/organizations/1/workflows/5`. The application header includes the logo "docuteam feeder" and navigation links: "Organisation\_A", "Ausführungen", "Workflows", "Cockpit", "Benutzer", and "Admin". The main content area is titled "Workflow: Quality Assurance, Migration and Storage". Below the title, it specifies the "Schritt-Präfix" as `cd %FEEDER_JAVA%`. A list of 12 Java commands follows, detailing the workflow steps from SIP extraction to cleanup. Metadata at the bottom indicates the workflow was created by admin on May 25, 07:01 Uhr, and last modified on July 6, 16:11 Uhr. Action links "bearbeiten" and "zurück" are provided. The footer shows the user is logged in as admin with links for "Logout", "Dokumentation", and "Hilfe".

Workflow: Quality Assurance, Migration and Storage

Schritt-Präfix `cd %FEEDER_JAVA%`

1. `java ch.docuteam.feeder.qualityassurance.SIPExtractor "${SIP}"`
2. `java ch.docuteam.feeder.qualityassurance.SIPSubmissionAgreementCheck "${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 „Verwalten“ und „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: X

localhost/organizations/1/workflows/3/edit

docuteam feeder Organisation\_A Ausführungen Workflows Cockpit Benutzer Admin

### Workflow bearbeiten

|                                        |                |                   |           |                 |
|----------------------------------------|----------------|-------------------|-----------|-----------------|
| Quality Assurance: extract SIP into wo | Name *         | Quality Assurance | Parameter | "\${SIP}"       |
| Quality Assurance: check formats aga   | Schritt-Präfix | cd %FEEDER_JAVA%  | Parameter | "\${SIP}" false |
| Quality Assurance: fixity check (md5)  |                |                   | Parameter | "\${SIP}"       |
| Quality Assurance: virus check         |                |                   | Parameter | "\${SIP}" false |
| Ingest: cleanup working copies         |                |                   | Parameter | "\${SIP}"       |
|                                        |                |                   | Parameter |                 |
|                                        |                |                   | Parameter |                 |
|                                        |                |                   | Parameter |                 |
|                                        |                |                   | Parameter |                 |
|                                        |                |                   | Parameter |                 |
|                                        |                |                   | Parameter |                 |
|                                        |                |                   | Parameter |                 |
|                                        |                |                   | Parameter |                 |
|                                        |                |                   | Parameter |                 |
|                                        |                |                   | Parameter |                 |
|                                        |                |                   | Parameter |                 |
|                                        |                |                   | Parameter |                 |

Workflow aktualisieren

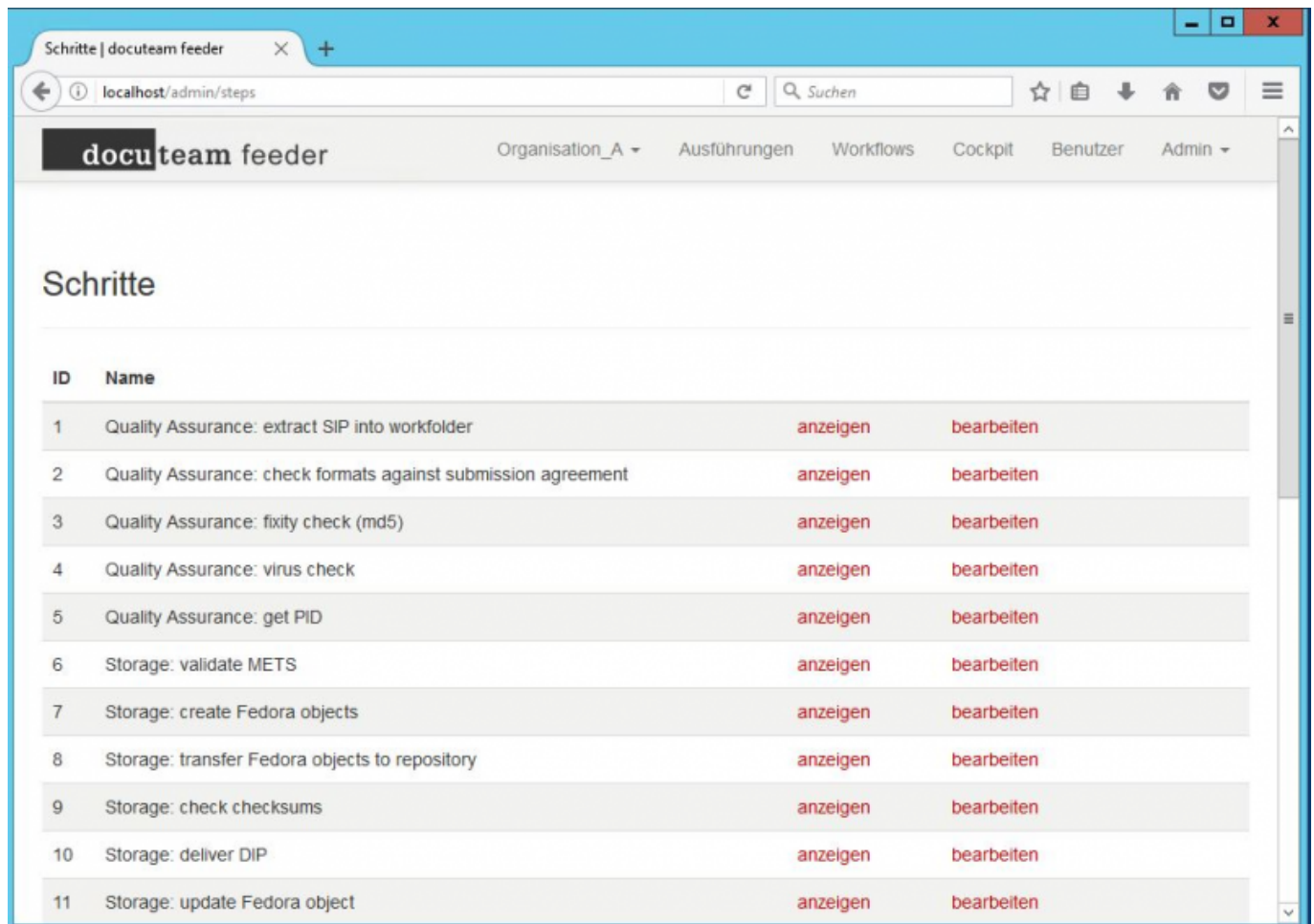
Für den Testbetrieb kann der Ingest einzelner SIPs auch manuell ausgelöst werden. Diesem Zweck dient der Befehl `starten`. Im Formularfeld „SIP“, das nach der Aktivierung des Befehls erscheint, wird der Name des SIP angegeben. **Wichtig:** der Name sollte normalerweise ohne die Endung „.zip“ angegeben 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 navigation bar includes "Organisation\_A", "Ausführungen", "Workflows", "Cockpit", "Benutzer", and "Admin". The main form has two input fields: "Objekt-ID" with a "SIP" dropdown menu, and "Manuelle Eingabe". A button labeled "Ausführung erstellen" is at the bottom. The footer shows "Angemeldet als admin | Logout | Dokumentation | 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.

Mögliche Schritte: [Feeder Steps](#)

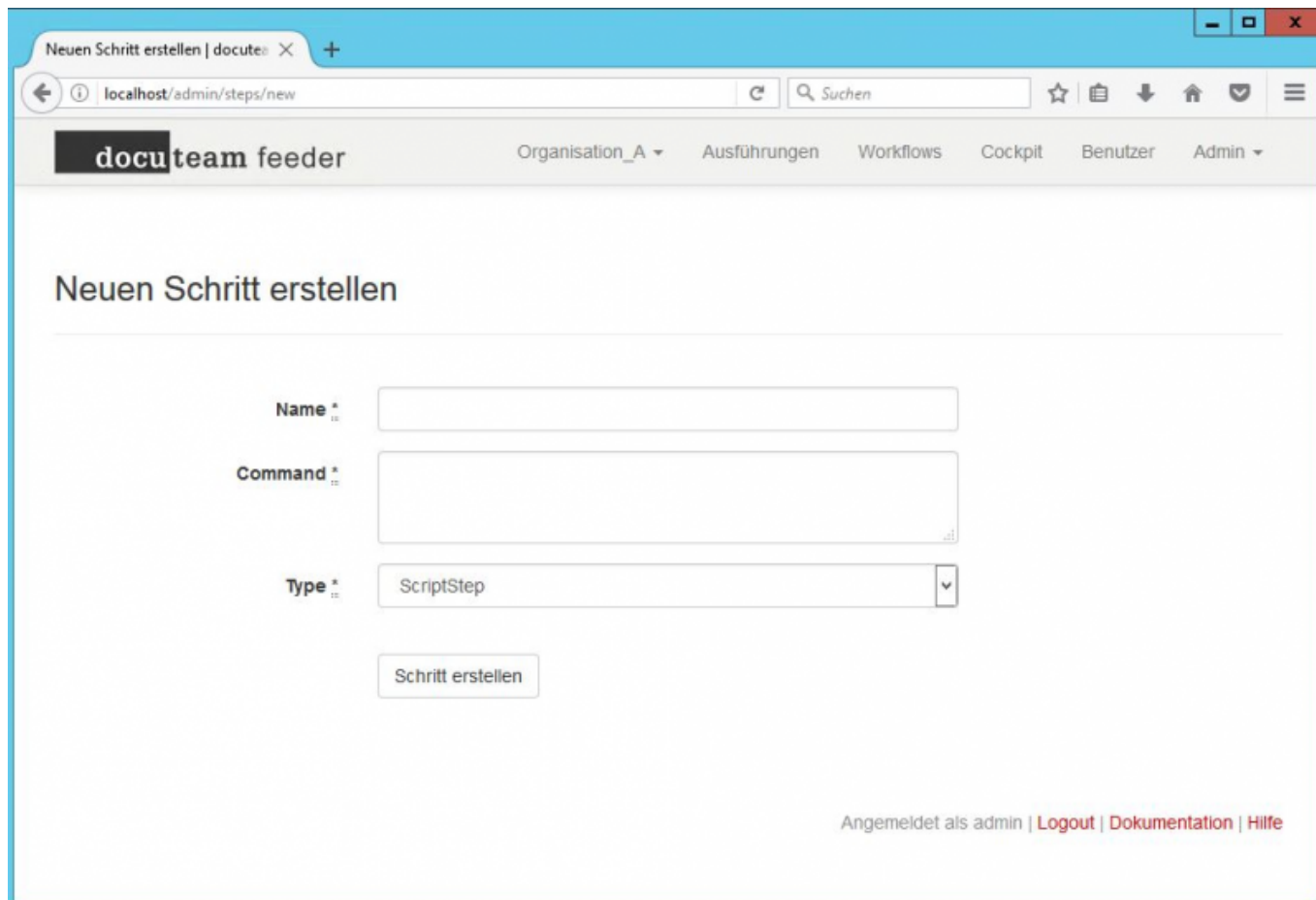


The screenshot shows the 'Schritte' (Steps) page in the Docuteam Feeder application. The interface includes a top navigation bar with the application name 'docuteam feeder' and several menu items: 'Organisation\_A', 'Ausführungen', 'Workflows', 'Cockpit', 'Benutzer', and 'Admin'. Below the navigation bar, the title 'Schritte' is displayed. The main content area contains a table with 11 rows, each representing a step in the migration process. Each row has an 'ID' column, a 'Name' column, and two action columns labeled 'anzeigen' (show) and 'bearbeiten' (edit).

| 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.





The screenshot shows the 'Neuen Schritt erstellen' page in the docuteam feeder application. The browser address bar shows 'localhost/admin/steps/new'. The application header includes the 'docuteam feeder' logo and navigation links: 'Organisation\_A', 'Ausführungen', 'Workflows', 'Cockpit', 'Benutzer', and 'Admin'. The main heading is 'Neuen Schritt erstellen'. Below it, there are three form fields: 'Name \*', 'Command \*', and 'Type \*'. The 'Type' dropdown is set to 'ScriptStep'. A 'Schritt erstellen' button is at the bottom. The footer shows 'Angemeldet als admin' with links for 'Logout', 'Dokumentation', and 'Hilfe'.

Neuen Schritt erstellen

Name \*

Command \*

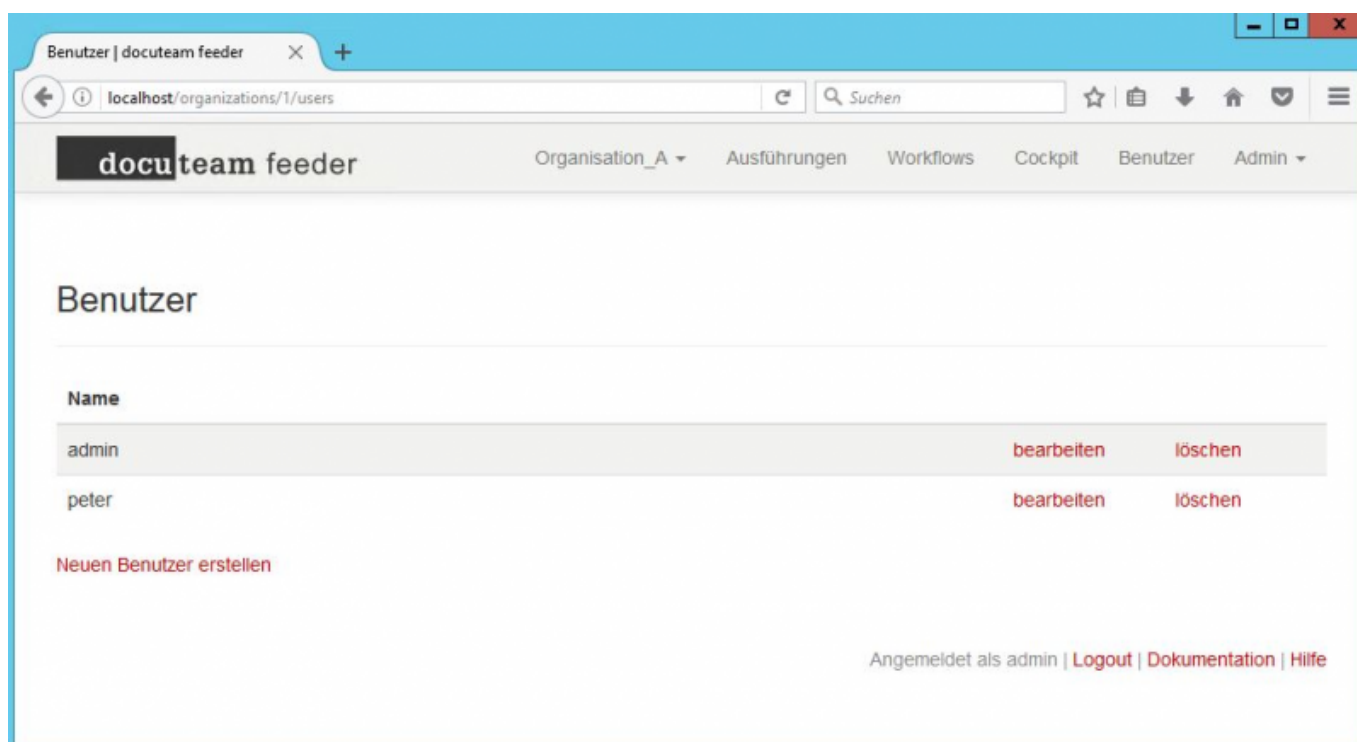
Type \* ScriptStep

Schritt erstellen

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



## Menüpunkt "Benutzer"



The screenshot shows the 'Benutzer' page in the docuteam feeder application. The browser address bar shows 'localhost/organizations/1/users'. The application header is the same as the previous screenshot. The main heading is 'Benutzer'. Below it, there is a table with two columns: 'Name' and actions. The table lists two users: 'admin' and 'peter'. Each user has 'bearbeiten' (edit) and 'löschen' (delete) actions. A link 'Neuen Benutzer erstellen' (Create New User) is at the bottom left. The footer shows 'Angemeldet als admin' with links for 'Logout', 'Dokumentation', and 'Hilfe'.

Benutzer

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

Benutzer bearbeiten | docuteam

localhost/organizations/1/users/2/edit

docuteam feeder

Organisation\_A ▾ Ausführungen Workflows Cockpit Benutzer Admin ▾

## Benutzer bearbeiten

Name \*

E-Mail \*

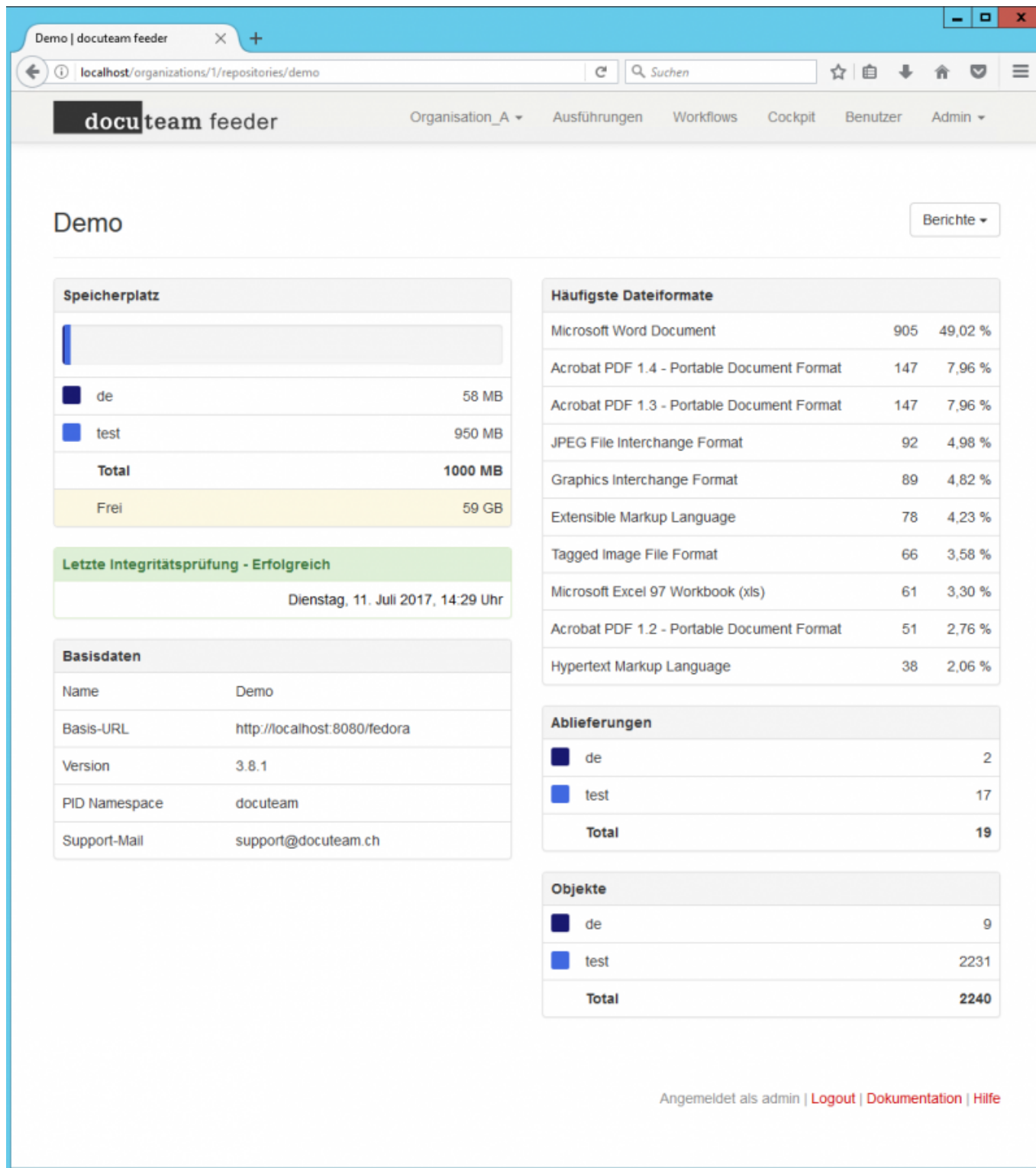
Passwort

Bestätigung

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 +

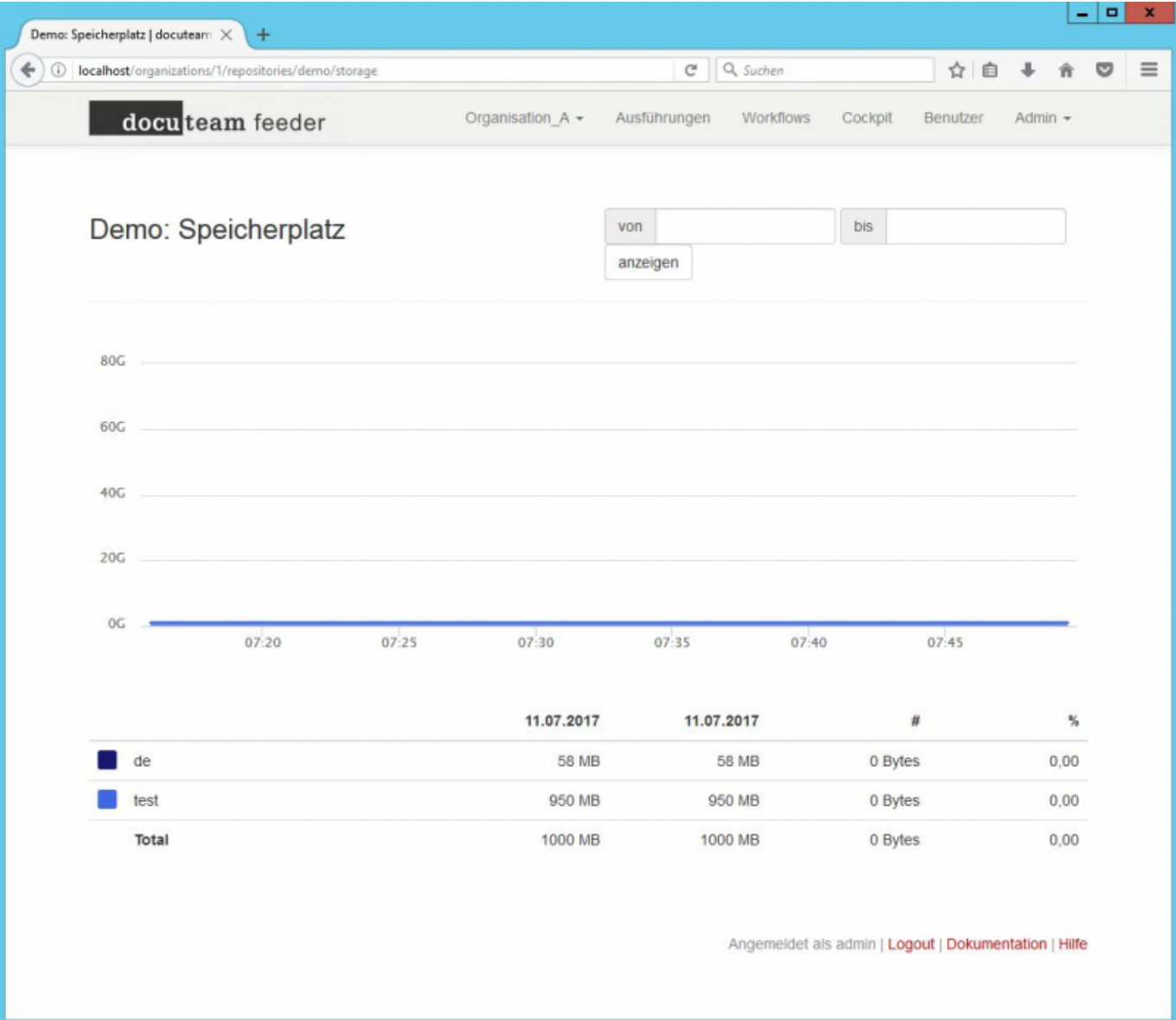
localhost/organizations/1/repositories/demo/file\_formats Suchen

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\\_340&rev=1506073763](https://wiki.docuteam.ch/doku.php?id=docuteam:feeder_340&rev=1506073763)

Last update: **2019/01/07 11:14**

