User Manual „docuteam packer“

docuteam packer 2.1.2d ETH

Contact for questions:
Martin Bärlocher,
ETH-Bibliothek, Library-IT-Services
Rämistrasse 101, 8092 Zurich
Tel. +41 44 632 0405

Author: Digital Curation Office (Martin Bärlocher)
http://www.library.ethz.ch/Digital-Curation
Version: 0.2
Date: 24.10.2014
Location: This or an updated version of the document can be found here:
http://download.library.ethz.ch/docuteam-packer/

ETH-Bibliothek
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Directory of edits

Date	Version	Description of edit	Name
24.10.2014	0.2	Adaptions to new version „docuteam packer“ 2.1.2d ETH	M. Bärlocher

Contents

Directory of edits II

1 Basis for the workshop 1

1.1 Download and installation 1

1.2 Configuration 2

1.3 Starting the application 2

2 Open Archival Packages 3

3 Searching in Archival Packages 3

4 Creating Archival Packages 5

4.1 Create a new Archival Package 5

4.2 Create a new Archival Package from template 6

5 Working areas in „docuteam packer“ 7

6 Editing Archival Packages 9

6.1 Edit the Archival Package’s structure 9

6.2 Add or delete files 9

7 Add metadata within Archival Package 11

7.1 Assign element levels 11

7.2 Add metadata 12

8 Metadata fields Standard DOI (Standard configuration) 14

8.1 Level Project 14

8.2 Level Folder 15

8.3 Level Complex_Object 16

8.4 Level Single_Object 17

8.5 Level Undefined 18

9 Save Archival Package 19

10 „docuteam packer“ Standard-Documentation 19

11 Long term preservation solution Rosetta 19

User Manual „docuteam packer“ ETH

1 Basis for the workshop

This manual describes, how an archival package („package“) is manually produced from a file folder using „docuteam packer“. „docuteam packer“ is generally used for visualising and editing information packages which comprise digital objects and their metadata in an XML-structure according to the METS-standard (Metadata Encoding and Transmission Standard[1]). In the terminology of the OAIS Reference Model (Open Archival Information System) these packages can be SIP (Submission Information Packages), AIP (Archival Information Packages) or DIP (Dissemination Information Packages). For „docuteam packer“ as a tool these distinctions are not relevant. In the following manual the term “Archival Package“ is used because mainly the formation of submissions to a digital archive is to be shown. In the application’s menu, the term „Package“ is used..

Purpose:
The „docuteam packer“ is a configurable tool to locally collect, structure and describe the digital objects within an Archival Package. It can be configured which metadata fields can be used for the description. Following the generation of the Archival Package the workflow for submissionn to the long term archive will be triggered from within „docuteam packer” in the future. This functionality is currently being tested and will be available for production early in 2015.
The „docuteam packer“ also facilitates a search within the locally stored Archival Packages. However, the tool is no substitute for collaborative applications or for document management systems (e.g. MS Sharepoint) or for comprehensive repository-software (e.g. DSpace, Fedora etc.). Its use for such purposes is not recommended. „Docuteam packer“ should be regarded as editor to prepare for digital preservation.

Version 2.1.2d of „docuteam packer“ should be considered as a demo version.

The screenshots were produced in a Windows-environment.

1.1 Download and installation

„docuteam packer“ (ETH version) is available under the following link:
http://download.library.ethz.ch/docuteam-packer/ .

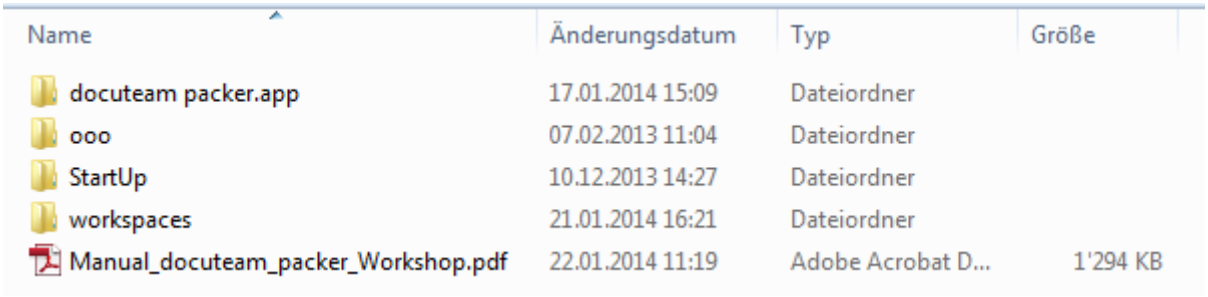
- Manual_docuteam_packer_Workshop.pdf (1.3 MB) - Short instructions
- packer_portable.zip (38 MB) - Startscripts for Windows und Mac
  - Docuteam packer application
  - Example Archival Packages
- ooo.zip (333 MB) - Open-Office for internal Previews

The Open Office-package ooo.zip is only required if you want to make use of the internal preview for office-files in „docuteam packer”. Open Office is otherwise not part of „docuteam packer“.

[1] http://www.loc.gov/standards/mets/

Installation: No installation is required. Copy the Zip-files into a directory on your computer (e.g. C:\APPS\) and unpack it via the context menu („Extract to here“).

Important:
„docuteam packer“ requires an up-to-date version of the Java Runtime Environment. It can be found here for download: http://www.java.com/de/download/manual.jsp .



Following download and unpacking of all files, the folders should show this structure.

1.2 Configuration

„docuteam packer“ works „out of the box“, but can be configured very flexibly. The configuration will usually be set up and maintained by the Digital Curation Office. Changes in the configuration can result in data not being able to be submitted to the long term archive!

Standard configurations

Standard_DOI
If you plan to make all or part of your archival data publicly available and reusable, you best use the DOI-standard-configuration. This configuration causes all archival elements (folders and files) to be assigned a unique DOI-name (e.g. 10.5405/ethz-0023-12345678) immediately upon import. Digital Object Identifiers (DOI)[2] make the archival object citable in the long term, comparable to journal articles. In addition, the DOI facilitates direct linking to an archival object (e.g. http://doi.org/10.5405/ethz-0023-12345678). Only upon definite archiving in ETH Data Archive, the DOI-name will be registered with the global DOI-system and become actionable as a link. At the same time, descriptive metadata is published to the DataCite Metadata Store.[3]

Standard_Simple
If you only plan to safeguard your archival data without making them accessible to a larger audience, you may use this simpler configuration. In this case, less descriptive metadata are mandatory.

1.3 Starting the application

„docuteam packer“ does not have to be installed locally, but can be launched directly from the folder „StartUp“. Developers test „docuteam packer“ under both Windows and Mac OS X and guarantee that die application runs in both environments.

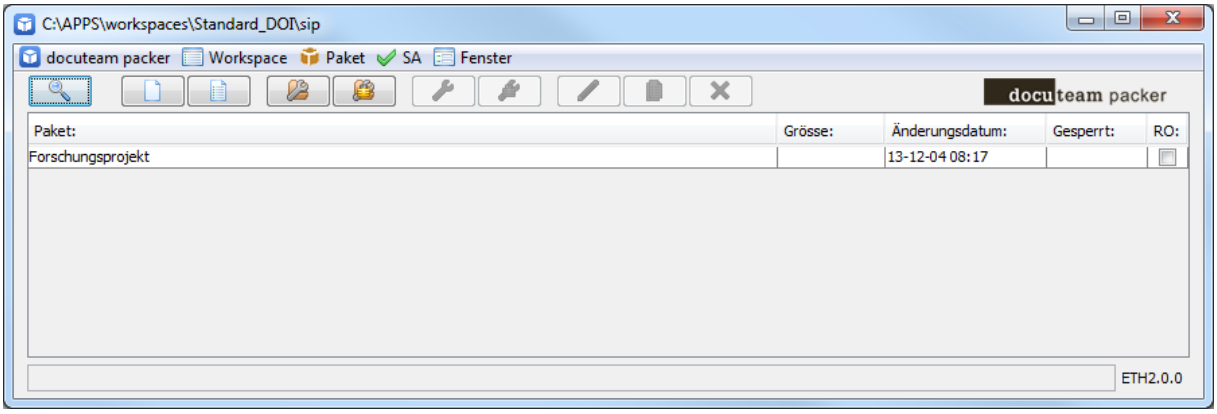
[2] Siehe www.doi.org
[3] Siehe www.datacite.org und für die Suche http://search.datacite.org/ui

Start under Windows: The folder StartUp contains two Command-Files „Standard_DOI-Win.cmd“ and „Standard_simple-Win.cmd“. By executing either of these files, „docuteam-packer“ is launched with the respective configuration.

Start under Mac OS: In the folder StartUp there are two Script-Files, „Standard_DOI-OSX.command“ and „Standard_simple-OSX.command“. By executing either of these scripts, „docuteam-packer“ is launched with the respective configuration.

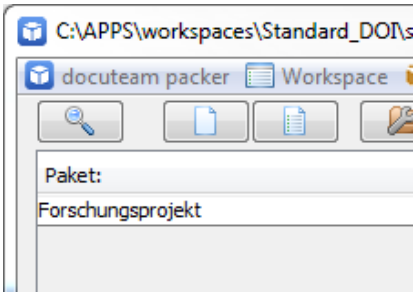
2 Open Archival Packages

After executing the StartUp file (e.g. Standard_DOI) the „docuteam launcher“ is started. It displays all available packages. Double clicking the respective line opens „docuteam packer“ with the selected package.

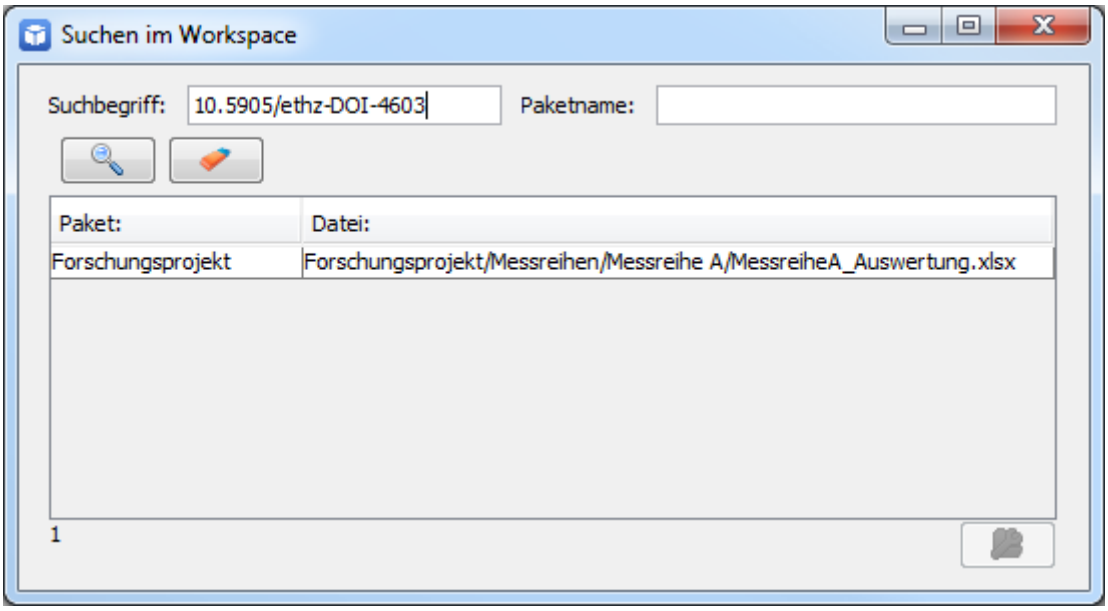


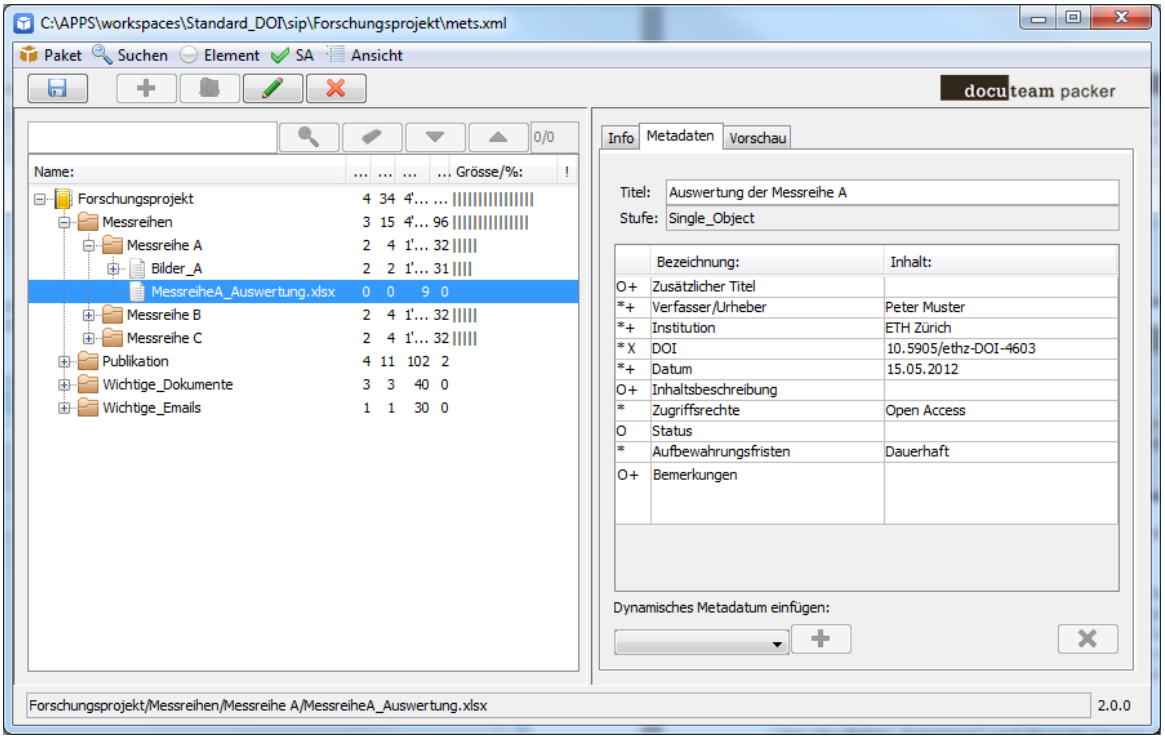
3 Searching in Archival Packages

From the "docuteam launcher" a search can be performed within the metadata of the Archival Packages of a workspace, usually in the Archival Packages of a working group.



Clicking the “magnifying glass” opens the search form. After entering a search term and executing the search, the results list presents the packages containing the hits. Double clicking opens the respective Archival Package at the position of the element containing the hit.

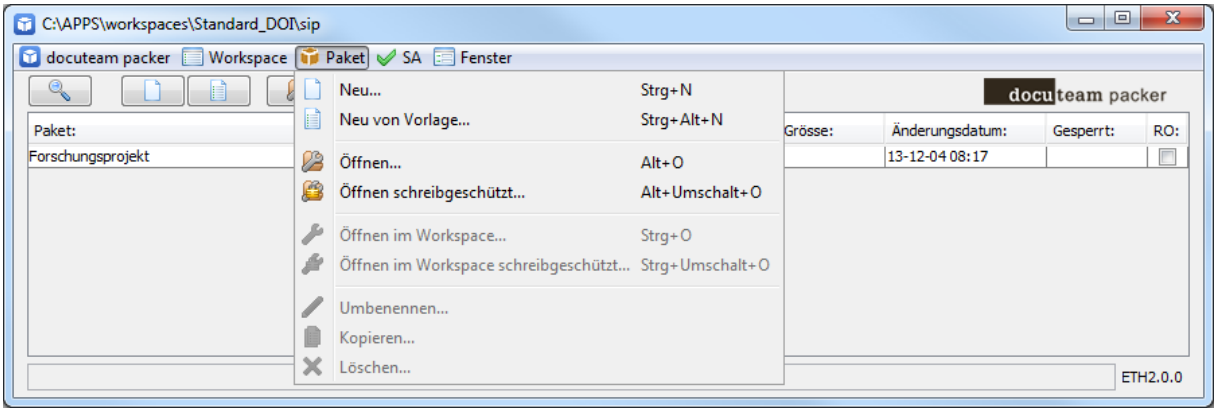

Suchen im Workspace
Suchbegriff:
10.5905/ethz-DOI-4603
Paketname:
Paket:
Datei:
Forschungsprojekt
Forschungsprojekt/Messreihen/Messreihe A/MessreiheA_Auswertung.xlsx
1

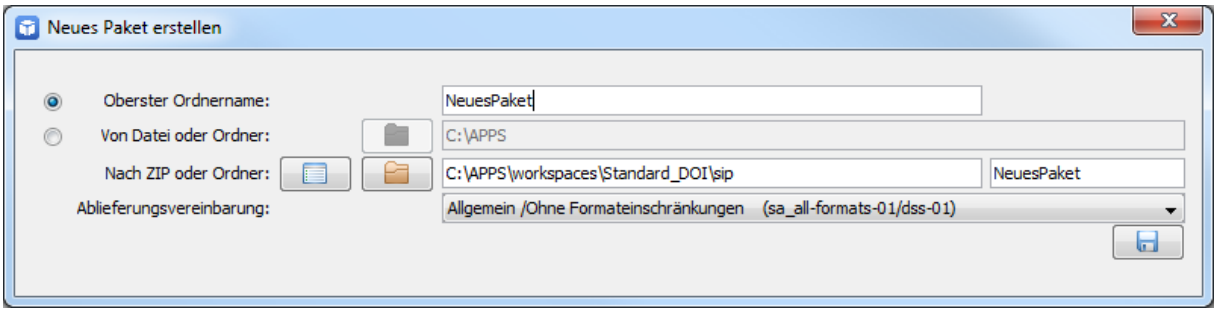

C:\APPS\workspaces\Standard_DOI\sip\Forschungsprojekt\mets.xml
Paket
Suchen
Element
SA
Ansicht
docuteam packer
Name:
Grösse/%:
Forschungsprojekt
Messreihen
Messreihe A
Bilder_A
MessreiheA_Auswertung.xlsx
Messreihe B
Messreihe C
Publikation
Wichtige_Dokumente
Wichtige_Emails
Info
Metadaten
Vorschau
Titel:
Auswertung der Messreihe A
Stufe:
Single_Object
Bezeichnung:
Inhalt:
Zusätzlicher Titel
Verfasser/Urheber
Peter Muster
Institution
ETH Zürich
DOI
10.5905/ethz-DOI-4603
Datum
15.05.2012
Inhaltsbeschreibung
Zugriffsrechte
Open Access
Status
Aufbewahrungsfristen
Dauerhaft
Bemerkungen
Dynamisches Metadatum einfügen:
Forschungsprojekt/Messreihen/Messreihe A/MessreiheA_Auswertung.xlsx
2.0.0

4 Creating Archival Packages

4.1 Create a new Archival Package

Create a new Archival Package in „docuteam launcher“: „Package“ –> „New...“

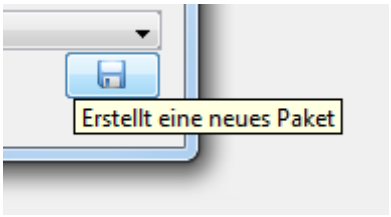




By selecting the radio button a „Top folder name” a new, empty package can be created.

By selecting the radio button „From file or folder“ a folder with the files to be imported and the respective Submission Agreement (SA)[4] can be chosen. An Archival Package does not have to be based on a folder, but may also contain one file only. Archival Packages with individual files might be useful e.g. for digital libraries or photo collections where no hierarchical structure is used.

The command „Save“ creates the first version of the Archival Package:



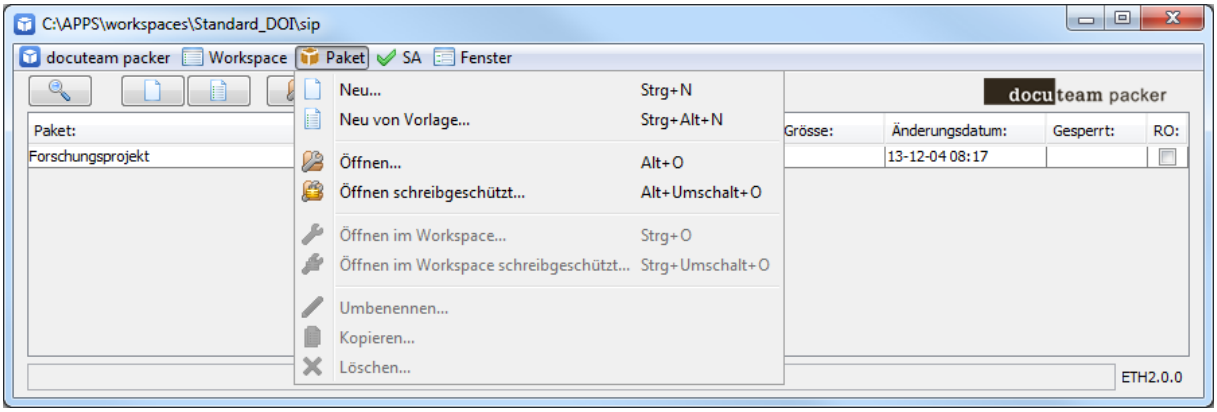
The basic structure of the Archival Package is now created automatically (copy of the data, creation of metadata). With large data volumes, this step might take some time to complete be-

[4] Such a previously agreed upon Submission Agreement can for example be used to ensure that only certain file formats may be submitted to the ETH Data Archive.

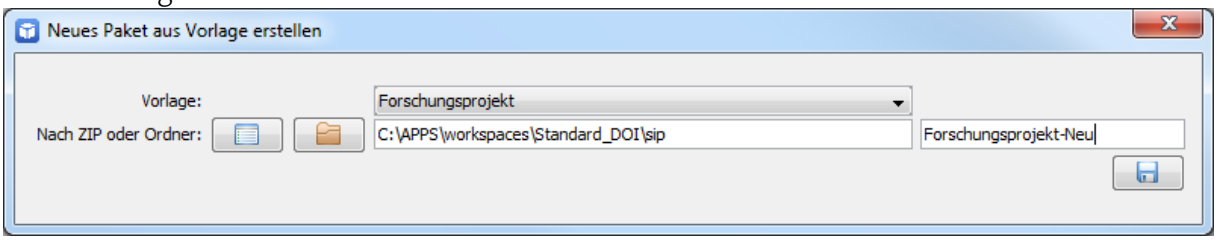
cause among other activities, checksums are calculated for all files which is a computation intensive task.
As soon as data has been imported, the newly created Archival Package is displayed and ready for further editing.

4.2 Create a new Archival Package from template

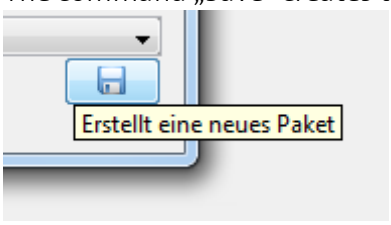
A new Archival Package based on an existing template can be opened by the command „Package“ –> „New from template...“.



In the dialogue, the respective template can be chosen and a name can be entered for the Archival Package.



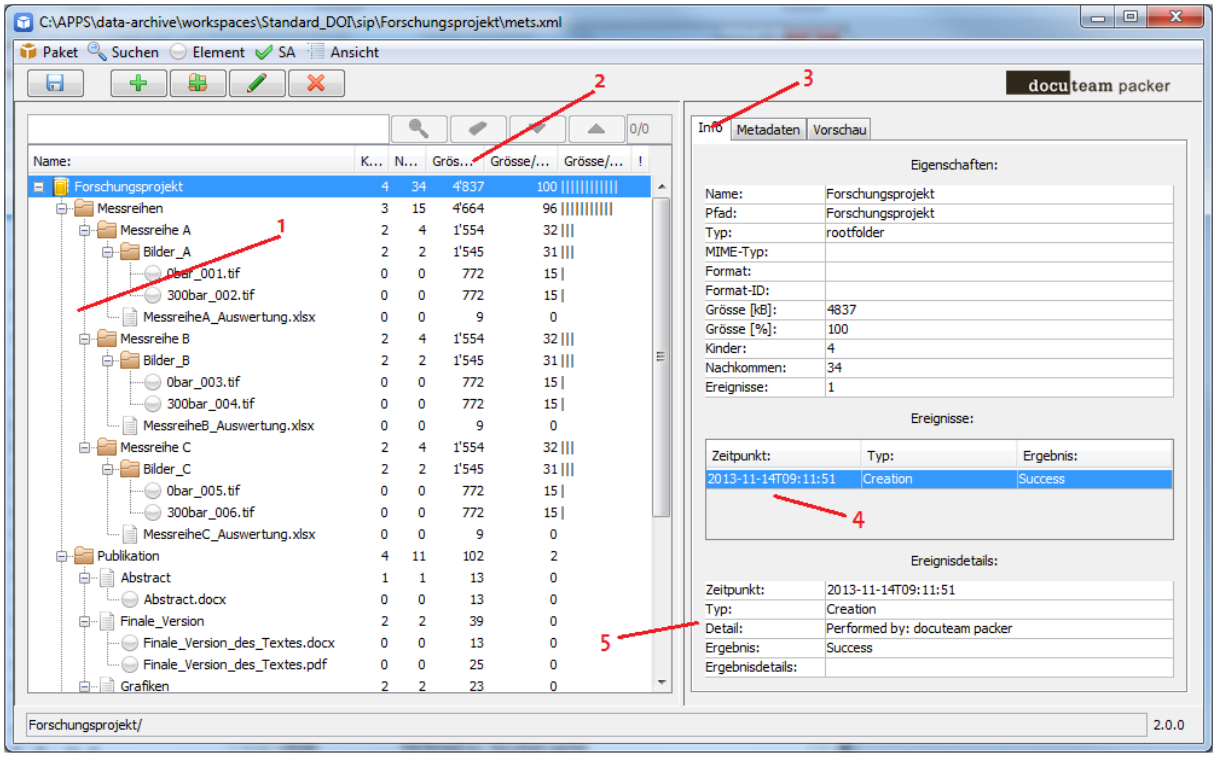
The command „Save“ creates the first version of the Archival Package:



The structure is then created according to the template without data and metadata.
As soon as the empty Archival Package has been created, it is displayed in „docuteam packer“.

5 Working areas in „docuteam packer“

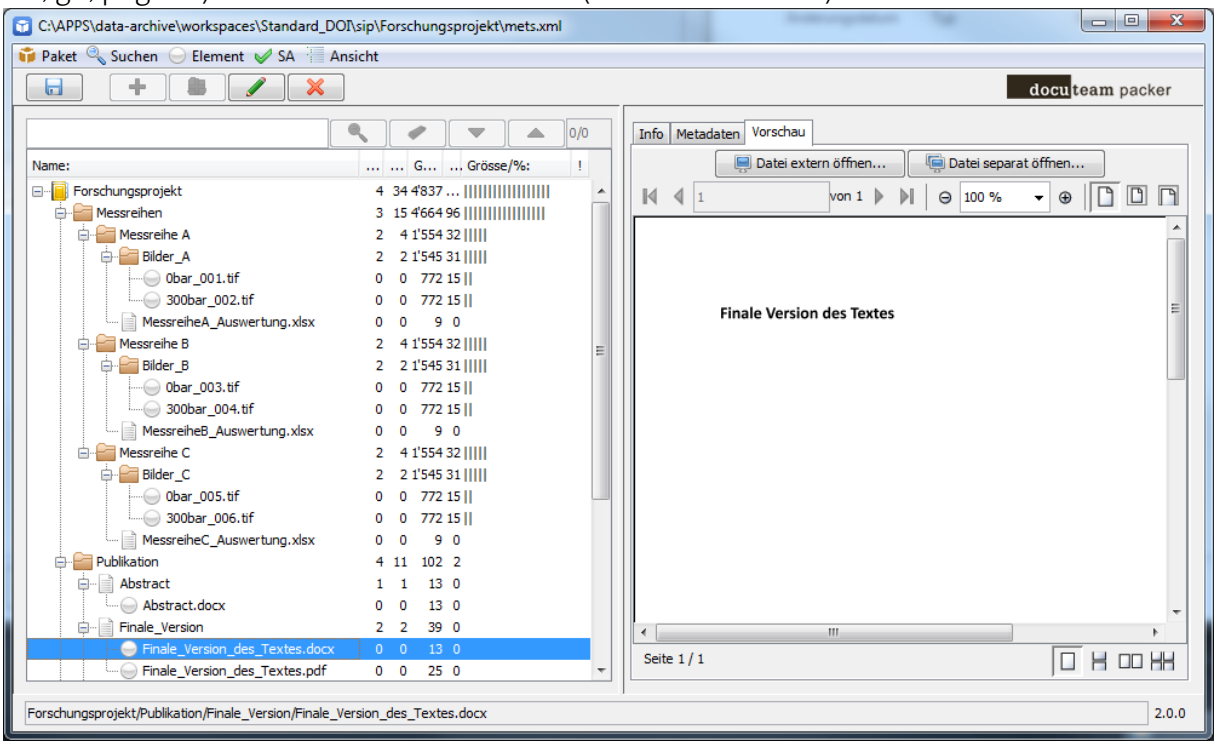
In the overview, „docuteam packer“ presents itself as follows:



1. Display of folders and files within the Archival Package
2. Statistical data on the individual files of the Archival Package. It is indicated, how many elements are directly attached to a folder (=children) and how many sub-elements it has altogether (=descendants). The absolute size in kilobyte is displayed as well as the relative size of a file in relation to the whole Archival Package. The relative size is also visualized with a bar as an immediate alert to particularly large files.
3. On the right side of the window, the tab „properties“ displays information on each individual object (folder, file). In particular, there are also indications on file format and file format version where available.
4. The section „Events“ contains the list of operations which have been carried out on the Archival Package and its parts. The full „Audit Trail“ is displayed to allow a tracking of how the Archival Package was created and prepared for long term preservation (migration steps etc.).
5. “Event Details” contain additional information on the measures taken, on the agent and the point of time.

The tab „Metadata“ is explained under „Metadaten im Archivpaket vergeben“.

When switching the right part of the window from „Info“ to „Preview“, the internal viewer of „docuteam packer“ displays a file preview. The preview works for PDF-files, image formats (jpg, tiff, gif, png etc.) and common office-formats (via creation of PDF).



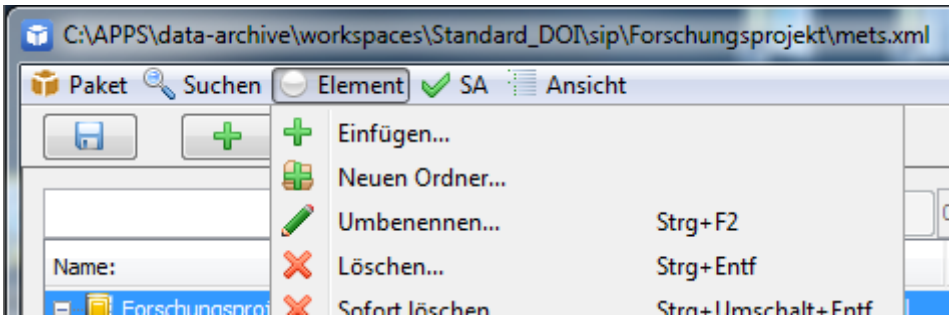
Using the function „Open file separately...“ opens the file preview of „docuteam packer“ in a separate window to enable metadata editing while viewing the file. The file preview is updated automatically when another file is selected in the tree on the left.

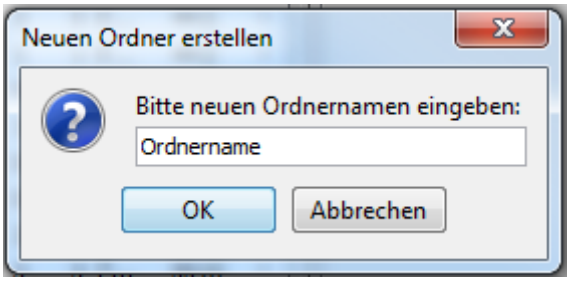
With the function „Open file externally...“ a file can be opened in the appropriate, locally installed application outside of „docuteam packer“.

6 Editing Archival Packages

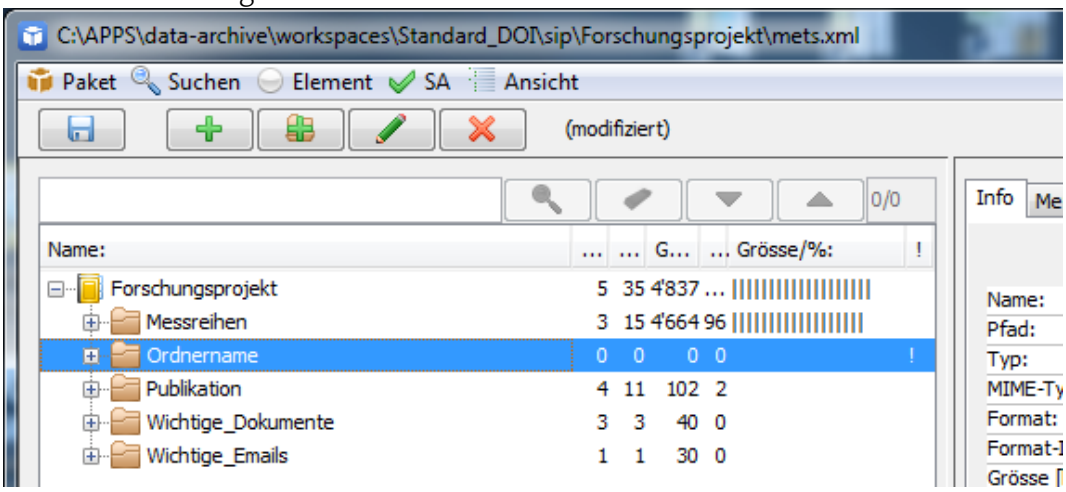
6.1 Edit the Archival Package's structure

To create an additional subfolder within an Archival Package, select the higher level folder and via the menu item „Element" –> „Create Folder..." create and name the subfolder.





The Archival Package then contains the additional substructure:



Likewise folders can be removed via the menu item „Element" –> „Delete...". Caution: This is a powerful command which will delete all of the folder's content and any substructure.

Files and folders may also be shifted within the Archival Package via Drag & Drop.

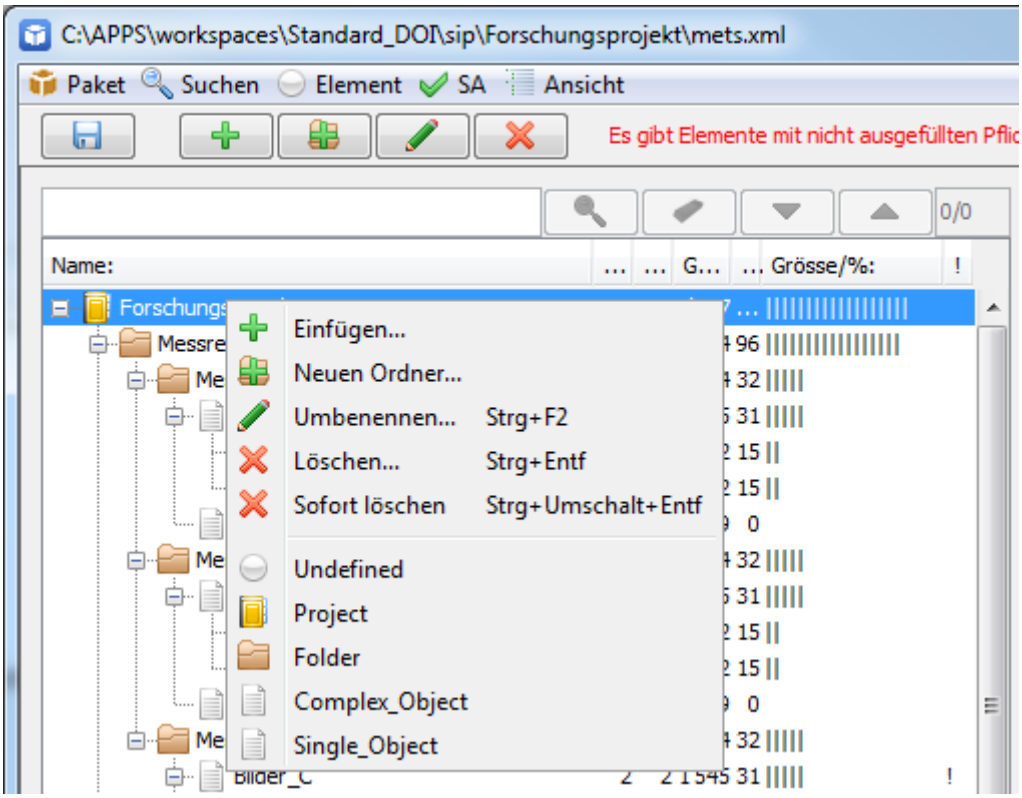
6.2 Add or delete files

Using the command „Element" –> „Insert..." files can be added to the Archival Package. This can be done for

1. an individual file which can be selected in the file dialogue,

2. several files which can be selected individually while the Ctrl-key is pressed,
3. a whole folder by selecting a folder instead of a file in the file dialogue.

Using the context menu (right mouse key) a file or folder can be inserted at a certain position within the Archival Package. Via the same menu, files or whole folders can be deleted.



As an alternative to adding elements via the file dialogue, files and folders may be copied from the file system to the desired folder in the Archival Package using Drag & Drop.

7 Add metadata within Archival Package

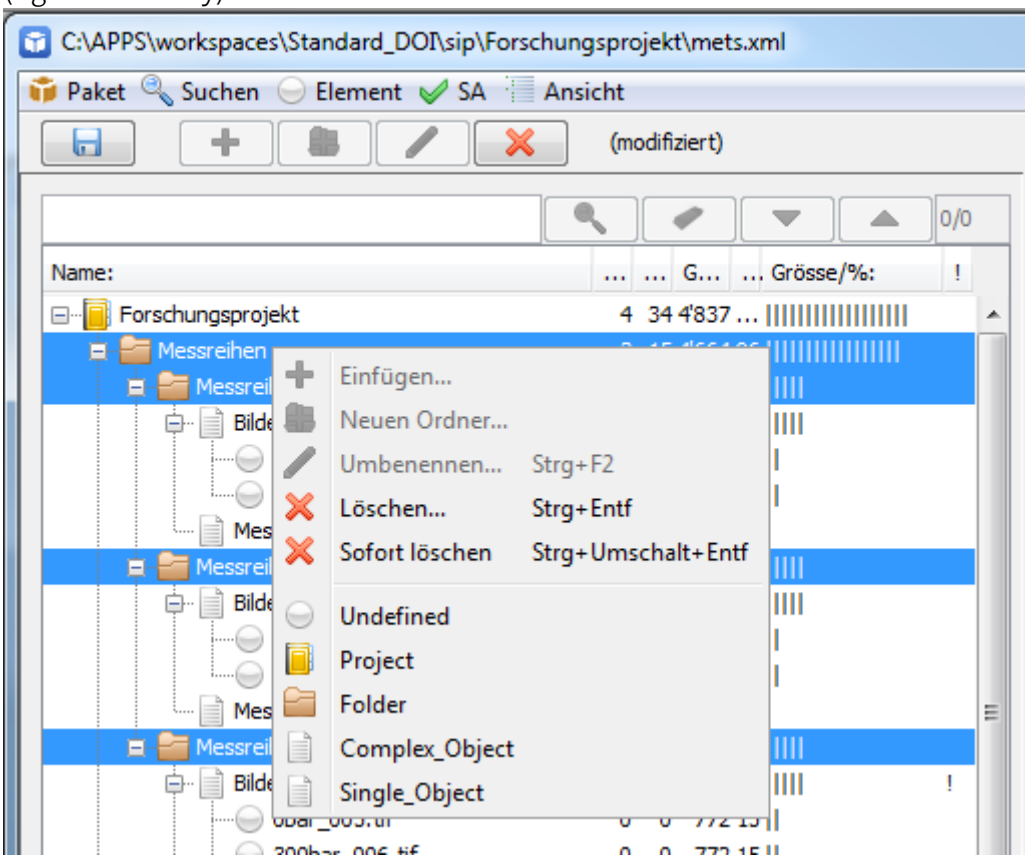
„docuteam packer“ facilitates building hierarchical structures from pre-defined elements which are assigned to levels of the hierarchy. Depending on the configuration, certain metadata fields with different properties are available per element.

7.1 Assign element levels

There are three ways to assign a certain level to an element: individually, by layer or by element-name and type.

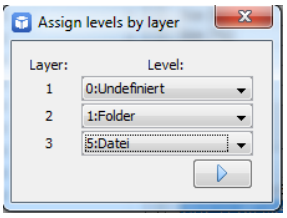
Individually

Select individual folders or files by mouse click. Use “Shift” and “Ctrl” keys to select several folders or files at a time. Using the menu command „Element“ → „Assign level“ all marked elements will be assigned the chosen level. The level can also be assigned using the context menu (right mouse key).



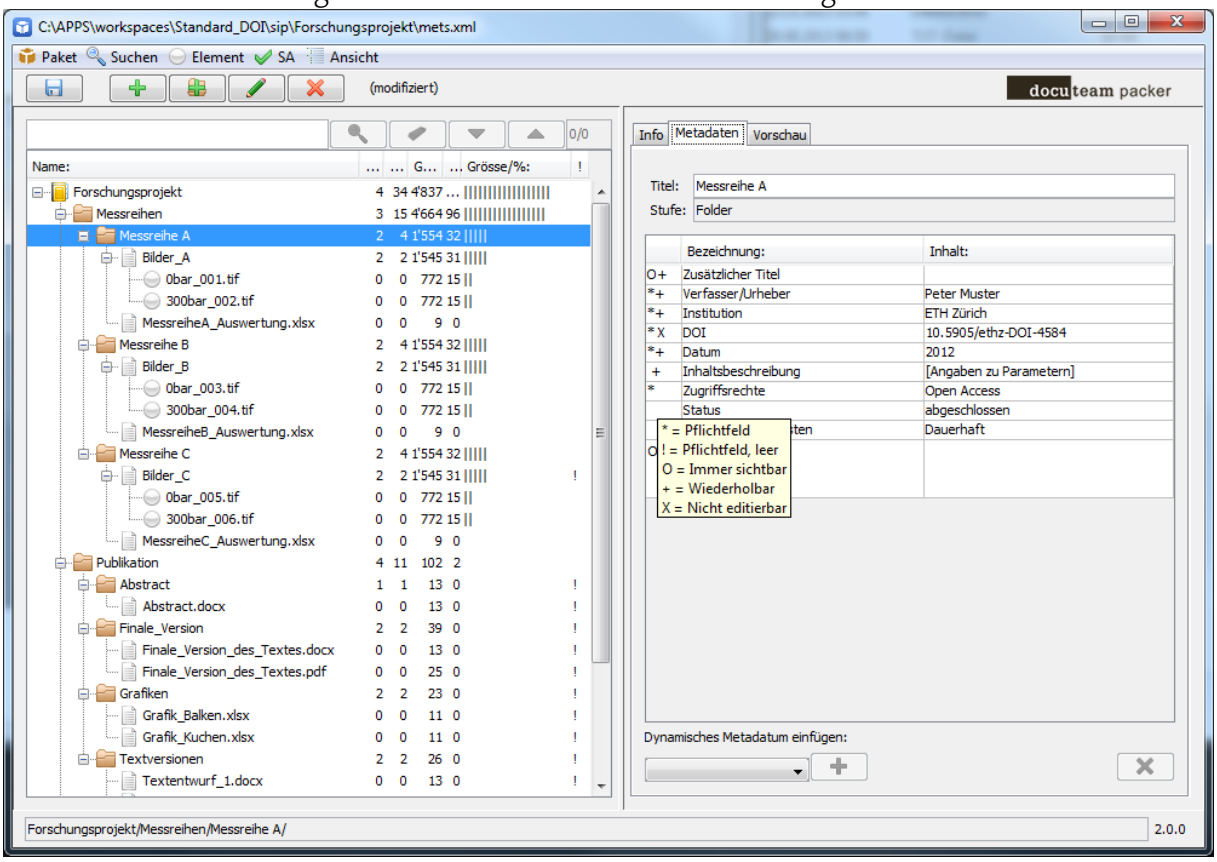
Assign per layer

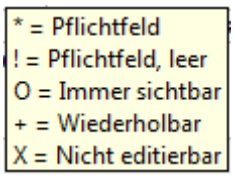
If a folder is selected, the menu command „Element“ → „Assign levels by layer...“ opens a window which allows to assign a level to all layers below (including) the selected folder. The first level is the one of the selected folder. If a level is left empty, elements on this level will not be changed.



7.2 Add metadata

In „docuteam packer“ metadata can be added on each level (i.e. folders and files). Selecting the tab “Metadata” on the right side of the client shows the following view:

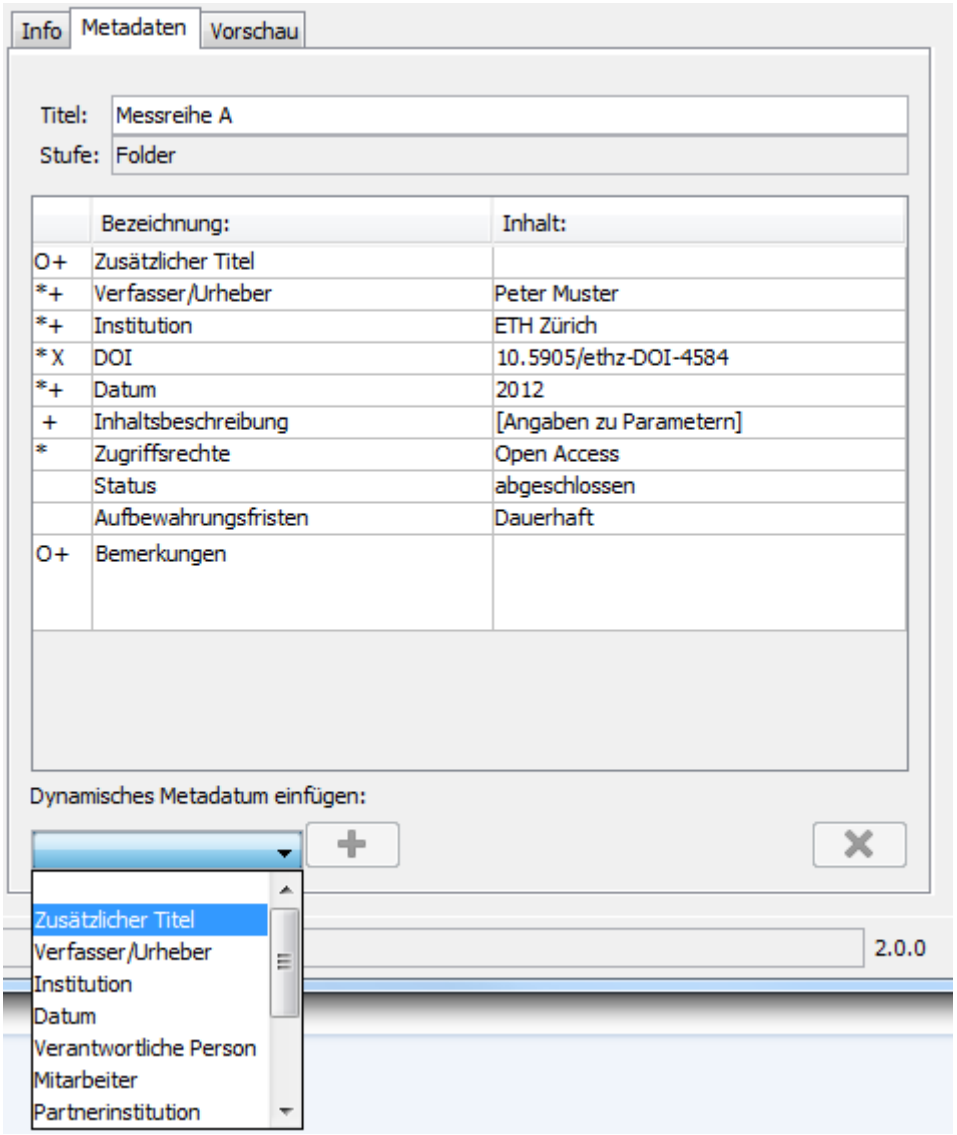




- * Mandatory metadata fields are marked with asterics *, if they are completed
- ! Mandatory metadata fields are marked with exclamation mark ! as long as no value has been entered. The Archival Package can nevertheless be saved at any time.
- O Optional fields which are always displayed are marked with O.
- \+ Repeatable fields are marked with +.
- X Read only fields such as the DOI are marked with X.

While some metadata fields must be completed as a prerequisite for registration of a DOI (Digital Object Identifier) gefüllt werden müssen, further mandatory fields may be defined according to customer demand.

Additional metadata fields for an element may be added via the dropdown menu at the bottom of the window. Which fields are available here can be defined in the configuration.



8 Metadata fields Standard DOI (Standard configuration)

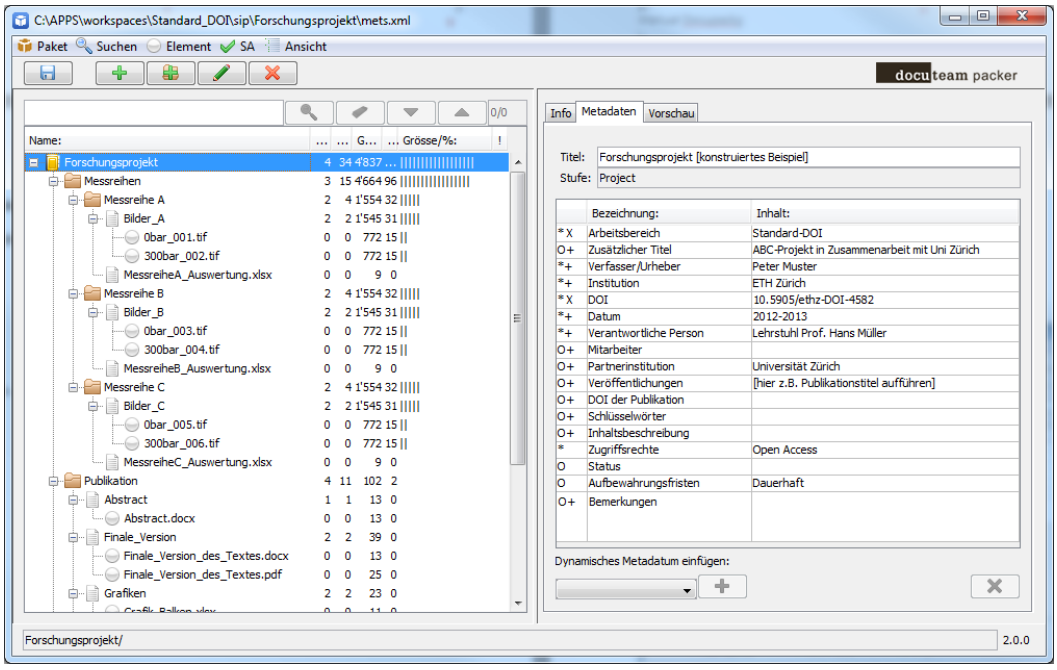
8.1 Level Project

In this use case, members of a research group collect their data in archival packages. DOI are generated for all objects in the Packages. These DOI are registered and published together with part of the metadata after the package has been submitted to ETH Data Archive.

Metadata fields:

Values can be handed down to lower hierarchical levels from this topmost level only.

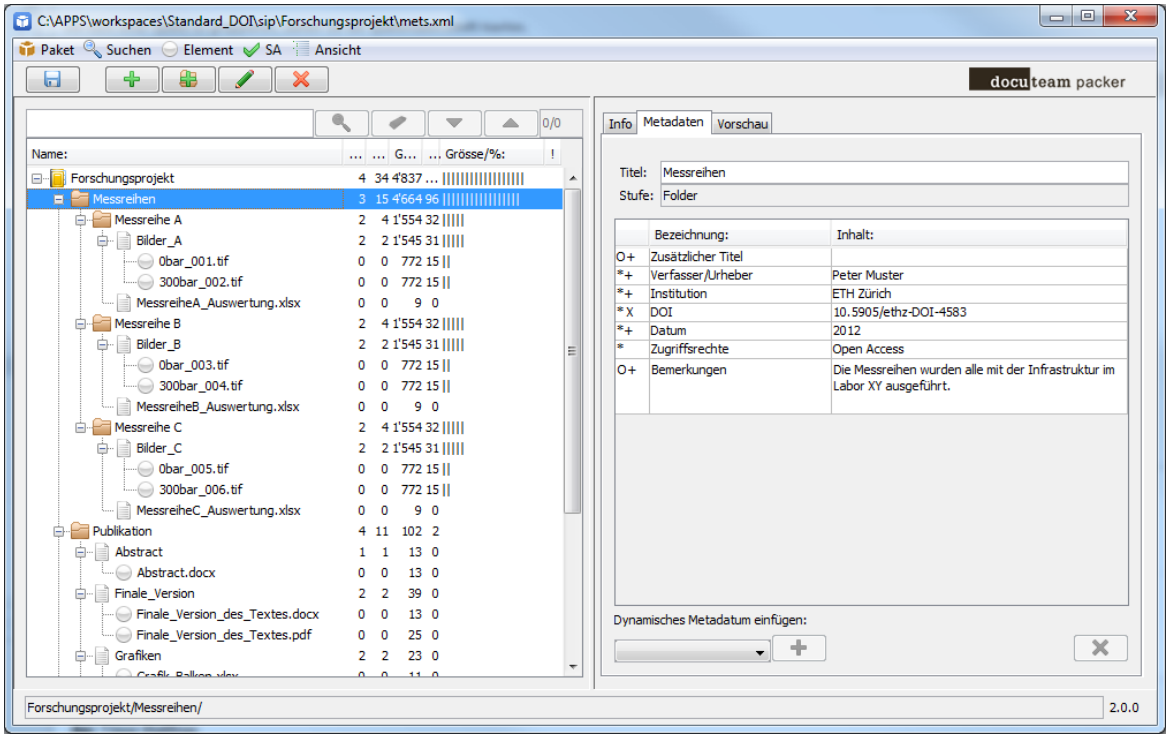
- Workspace (unique identifier for ETH Data Archive / Rosetta, Read-Only)
- Additional title
- Author / Creator
- Institution
- DOI (automatic creation, Read-Only)
- Date
- Responsible person
- Involved
- Cooperation Partner
- Publications
- DOI of the publications
- Keywords
- Description of content
- Access Rights (to be enforced in ETH Data Archive)
- Status
- Retention Period
- Comment



8.2 Level Folder

Metadata fields:

- Additional title
- Author / Creator
- Institution
- DOI (automatic creation, Read-Only)
- Date
- Access Rights (to be enforced in ETH Data Archive)
- Comment

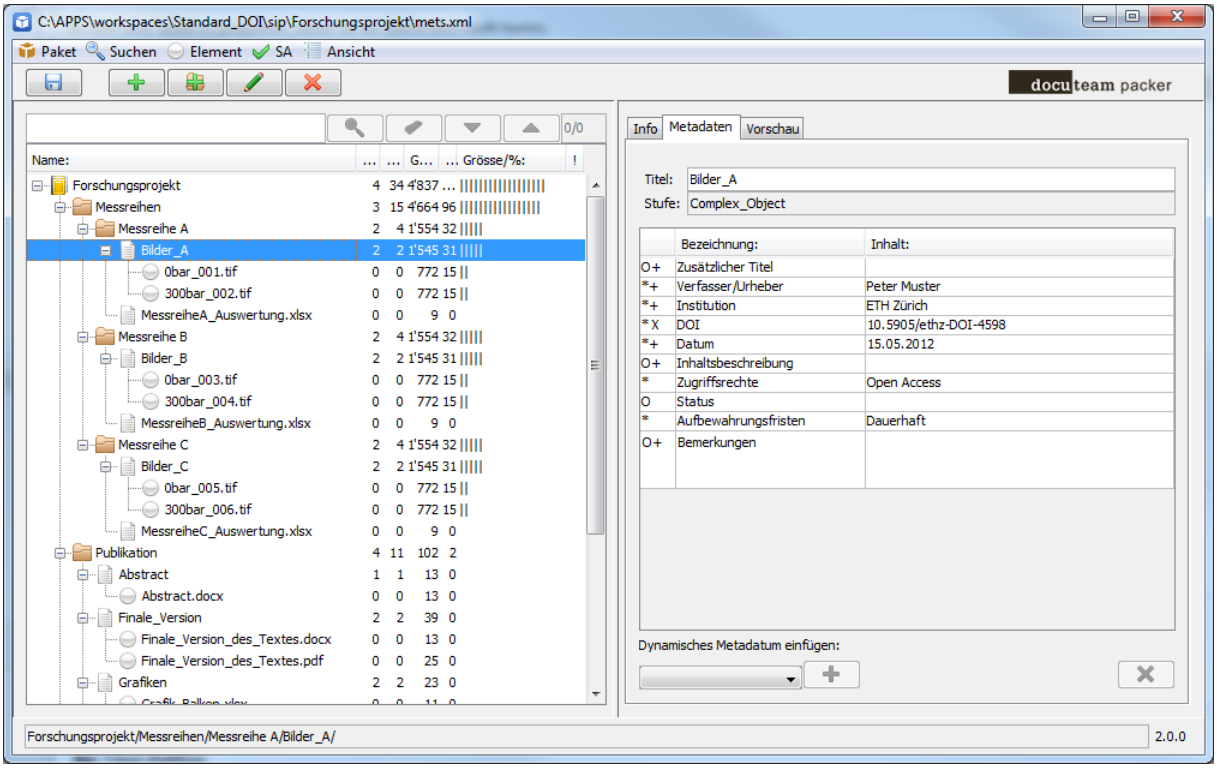


8.3 Level Complex_Object

This level differs from a folder in the following way: objects which are part of a "Complex_Object" do not need to and cannot be described with metadata themselves. A “Complex_Object” therefore is a container and it only the container is assigned metadata.

Metadata fields:

- Additional title
- Author / Creator
- Institution
- DOI (automatic creation, Read-Only)
- Date
- Description of content
- Access Rights (to be enforced in ETH Data Archive)
- Status
- Retention Period
- Comment

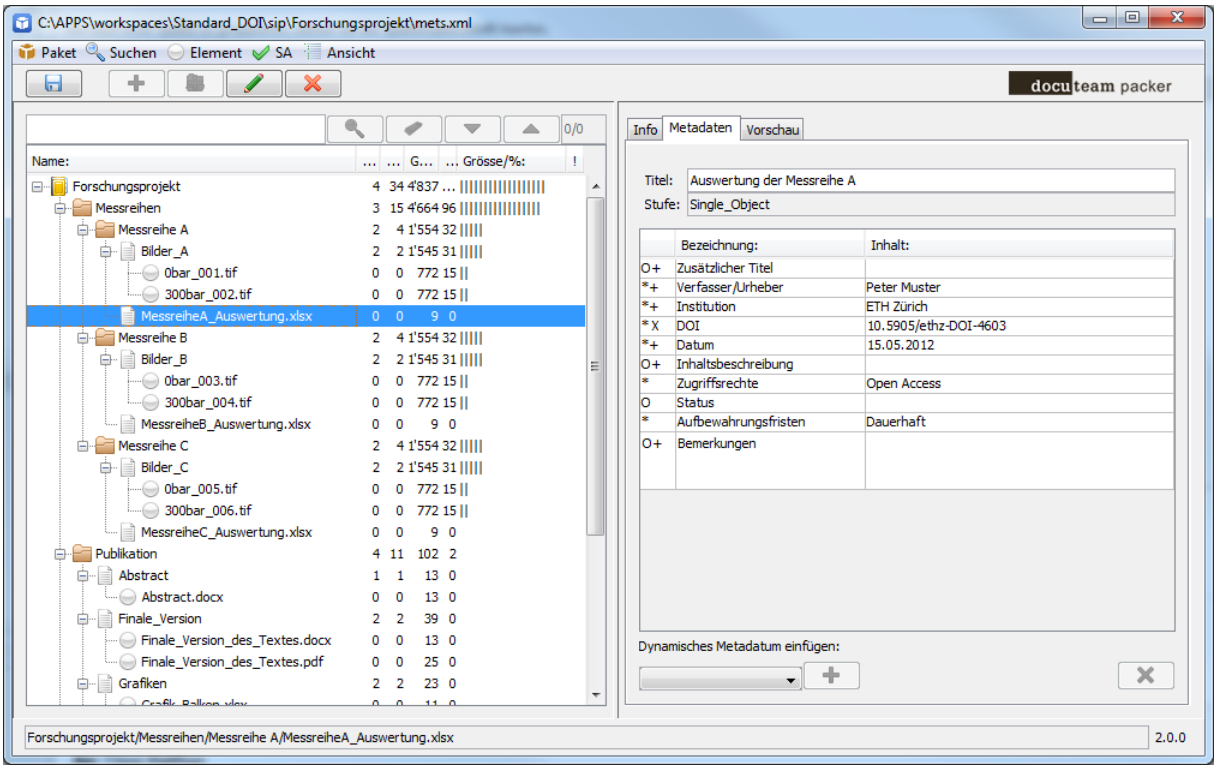


8.4 Level Single_Object

This level is assigned to files which are to be described with their own metadata.

Metadata fields:

- Additional title
- Author / Creator
- Institution
- DOI (automatic creation, Read-Only)
- Date
- Description of content
- Access Rights (to be enforced in ETH Data Archive)
- Status
- Retention Period
- Comment

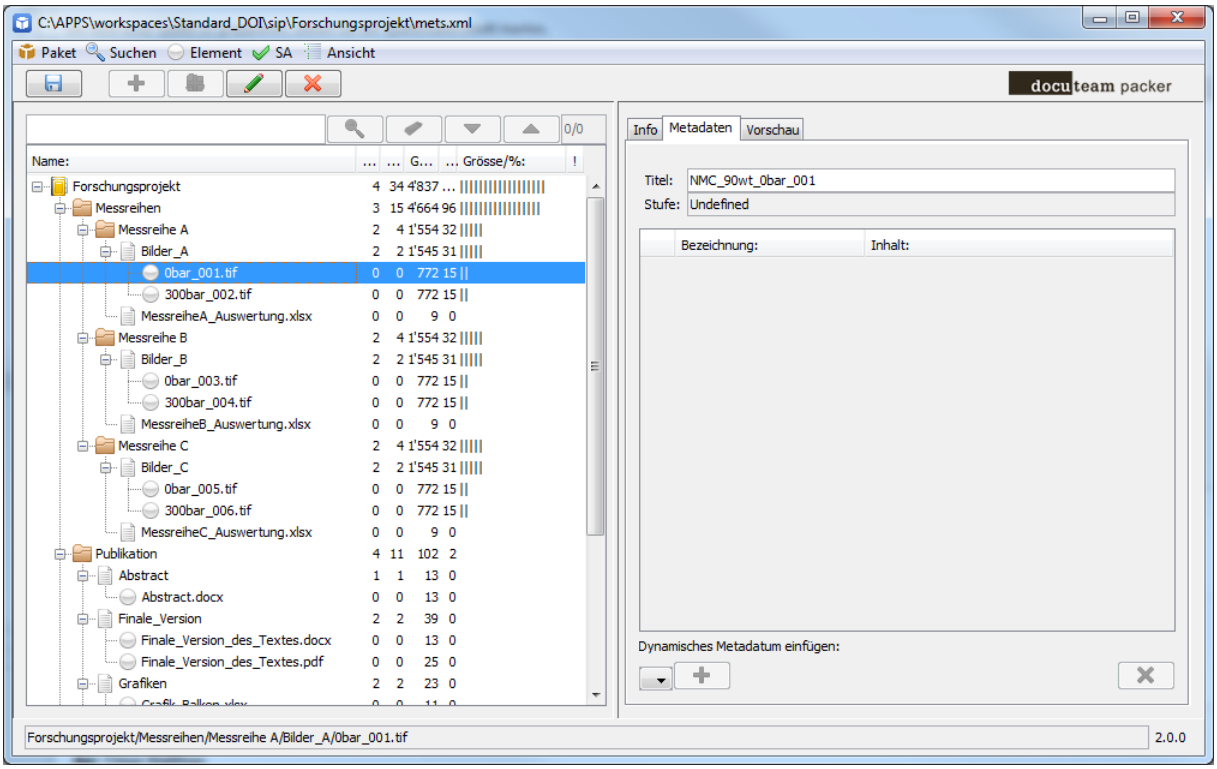


8.5 Level Undefined

The elements within a "Complex_Object" are assigned this level.

Metadata field:

- No metadata fields are defined.



9 Save Archival Package

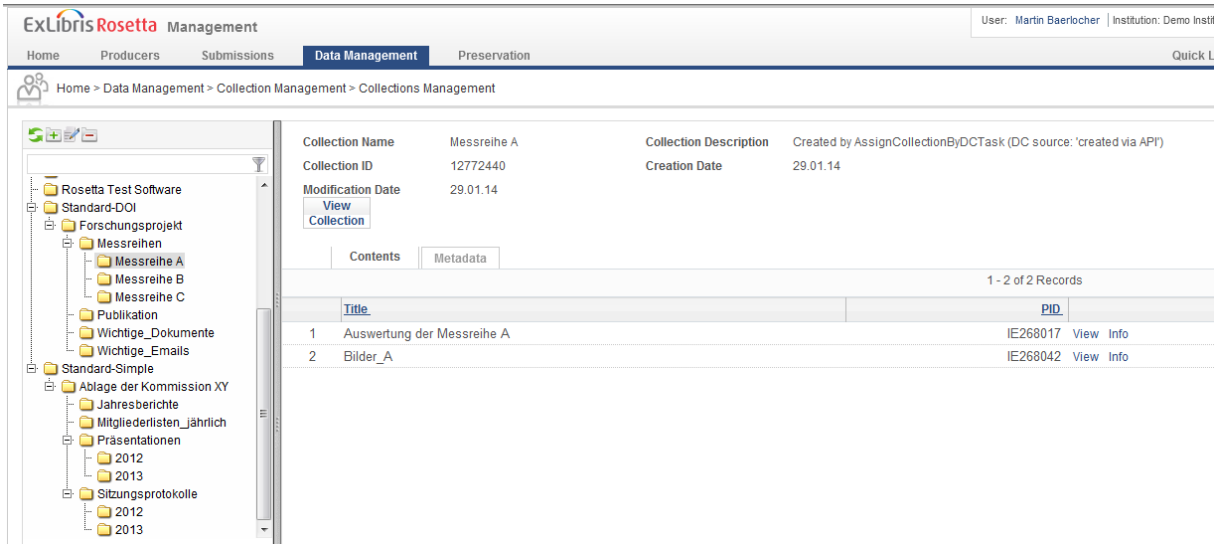
The modified Archival Package can be saved via the menu item „Package“ –> „Save“. „docuteam packer“ is exclusively working on a copy of the Archival Package. In the case of errors, the original data remains available. The original files are not touched by saving the Package, but the original Archival Package is renamed with the addition ORIGINAL in its file name to distinguish it from the current version.
The configuration allows to define how many versions of a package should be kept before the oldest one is overwritten.

10 „docuteam packer“ Standard-Documentation

A current version of the documentation for „docuteam packer“ kann be found with the supplier docuteam (currently in German only). Please note that the general release of the software differs from ETH Zurich’s version in some aspects.
http://wiki.docuteam.ch/doku.php?id=docuteam:packer_210

11 Long term preservation solution Rosetta

View of the content of an Archival Package from “docuteam packer” in Rosetta after submission.



http://rossa-test-ap11.ethz.ch:1801/delivery/action/collectionViewer.do?collectionId=14463320