

eContent Programme reUSE digital master files of printed publications

(eContent Project No. 11173)

Deliverable D1.2 Digital Repository - Austria



Date of delivery: 13. July 2005

Responsible author: Marco Köttstorfer

Author's organisation: UBI

Version: 1.0

Contributors: i3s3, UBG

Confidentiality: PP

Executive Summary

The Deliverable 1.2, Digital Repositories in Austria, is divided into four main chapters:

- Chapter 1 – ALO-Client Documentation
- Chapter 2 – ALO-Server Description
- Chapter 3 – Metadata Usage
- Chapter 4 - ALO Guide

Chapter 1 is called “ALO-Client Documentation” and includes basically an overview about the ALO-Client. This part illuminates the six categories of which the main menu of the Client is composed. Further, the forms and dialogs of the Client are exemplified.

Chapter 2 the “ALO-Server Description” includes the server side of the system as well as some features of the webpage www.literature.at. Also the content of the system is described. Moreover a brief explanation is given of how communication between the server and the client is realised.

The “Metadata Usage” is described in chapter 3. This covers the annotation of the 5 standards on which the ALO-System is based on, namely Dublin Core, RDF, Dig35 Metadata Schema, Metadata encoding and transmission standard (METS) Metadata Schema and Metadata in the TIFF File.

Chapter 4 contains a guide for working with the ALO-System. Therefore it is mentioned how to upload metadata and imagedata to the system and how the Dublin Core for books and journals should look like. Additionally, various examples are given.



Deliverable 1.2

Digital Repositories – Austria

Chapter 1

ALO-Client Documentation

ALO-Client Documentation



Table of Contents

1. GENERAL	5
1.1 ALO.....	5
1.2 About the ALO Client.....	5
1.3 Conventions	5
1.3.1 General conventions.....	5
1.3.2 Typographic conventions	6
2. MAIN MENU.....	7
2.1 File	7
2.1.1 New	7
2.1.2 Exit.....	8
2.2 Edit	8
2.2.1 Change.....	8
2.2.2 Delete	9
2.2.3 Save to file.....	9
2.2.4 Load from file.....	9
2.2.5 Send to server	9
2.2.6 Export to file.....	9
2.2.7 R ↔ V	9
2.3 Database.....	9
2.3.1 Connect	9
2.3.2 Disconnect.....	9
2.3.3 INI-File.....	10
2.4 Options	10
2.4.1 Show Objects in Collections	10
2.4.2 Scanner.....	10
2.4.3 Meta-data.....	10
2.4.4 Sort files when added	10
2.5 Window	10
2.5.1 Tile Windows	10
2.5.2 Cascade	11
2.5.3 Minimize all	12
2.5.4 Maximize all.....	12
2.5.5 Close all.....	12
2.6 Help	12
2.6.1 Contents.....	12
2.6.2 Search for Help On	12
2.6.3 How to Use Help.....	12
2.6.4 About.....	12
3. FORMS AND DIALOGS.....	13
3.1 Austrian Literature Online – Main form.....	13

3.2 Objects	14
3.2.1 Double click	14
3.2.2 Delete	14
3.3 Collection	15
3.4 Book.....	16
3.4.1 Tabsheet <Meta-data electronic>	17
3.4.2 Tabsheet <Meta-data original>.....	18
3.4.3 Tabsheet <Files>	19
3.4.4 Button [Save to file]	23
3.4.5 Button [Load from File]	23
3.4.6 Button [Send to Server]	23
3.4.7 Button [Export to File]	23
3.5 Postcard	23
3.6 PDF and RTF	24
3.7 Database Options	24
3.7.1 Menu-Items	25
3.7.2 Connection details	25
3.8 Scan Options	26
3.8.1 Menu-Items	26
3.8.2 Scanner details.....	27
3.9 Meta-data Options	28
3.9.1 Menu-Items	28
3.9.2 Meta-data details	29
3.10 INI-File.....	29
3.10.1 General information	29
3.10.2 Database information	30
3.10.3 Scan information	30
3.10.4 Meta-data information	31

1. General

1.1 ALO

The digitisation of printed material such as books, journals, magazines and newspapers is still faced with the prejudice that it is a complicated and expensive process. Setting up a digitisation line requires a patchwork of software programs for scanning, image pre-processing, OCRing, structural mark-up and generating XML output files.

Many libraries therefore shrink from taking steps to introduce the digitisation of printed materials as a standard service activity – even though they know very well that their users want to have all the library's holdings on the spot and right on their desktop in digital form.

It is exactly that gap between a growing demand and the problem how to fill it in an effective way that the ALO (Austrian Literature Online) working group intends to address with its projects. The working group consists of:

- University Library Graz
- University Library Innsbruck
- “integrated study” - Austrian-wide Institute for Information Systems Supporting Print Disabled People (University of Linz, University of Graz, Vienna University of Technology, University of Vienna)

The goal was to digitise the 1000 most important books of the Austrian literature and make them available via the Internet. The amount of 1000 books has already been reached and is still going to be extended. The digitised objects can be viewed via the ALO web interface at <http://www.literature.at>.

1.2 About the ALO Client

Because of the successfulness of ALO, some other projects followed to extend the functionality of ALO. On of these extensions is the ALO Client, which is a tool for uploading digitised material (e.g. books, postcards) and the associated meta-data to the ALO server easily.

The ALO Client should support libraries, organisations and private individuals in their work, simplify their workflow and therefore push the digitisation of antique books, postcards etc.

1.3 Conventions

1.3.1 General conventions

- Some of the menu items can also be selected via a context menu. Context menus can be reached by right-clicking into the area / onto the object that is mentioned in the description.
- Shortcuts are available for most of the menu items and edit fields. They can be used by holding the “Alt” key and pressing the according abbreviation key (the underlined

letter of the label corresponding to the edit field or of the menu item respectively) at the same time.

1.3.2 Typographic conventions

- Strings that appear on the graphical user interface of the client, such as names of panels, tabsheets etc., are written in the following form: <Panel>, <Tabsheet>, etc.
- Strings that are names of buttons, menu items, etc. (everything you can execute by clicking) are written in the following form: [Button], [Menu item], etc.
- Strings that have to be entered by the user, such as meta-data, are enclosed into double quotation marks, for example “Author”, “Title”, etc.
- Titles of windows (forms, dialogs) always appear bold face and italic, for example ***Book, Postcard, PDF, RTF, etc.***

2. Main Menu

As can be seen in Figure 0.1, the main menu of the ALO Client consists of six categories: <File>, <Edit>, <Database>, <Options>, <Window> and <Help>.

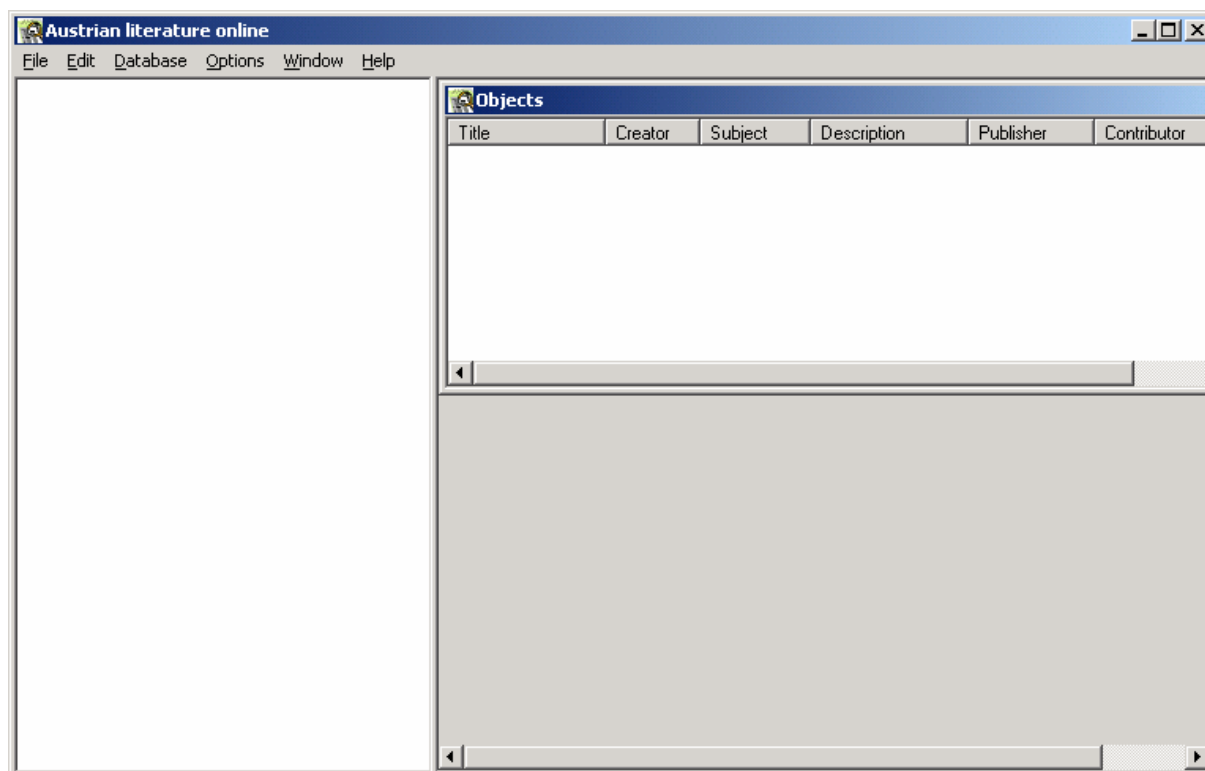


Figure 0.1: ALO Client with main menu

These six main categories of the menu always stay the same, but the entries within these categories can vary according to which window currently has the focus.

In the following, all possible entries of the main menu are described.

2.1 File

2.1.1 New

Collection:

Opens a new Collection window where you can fill in the meta-data about a new collection. This operation is only possible if there is a connection to the database and if a collection (the root collection or any sub-collection) has been selected from the list. The new collection will be a sub-collection of the selected collection.

This menu item is also available via the context menu of the collection-tree-area on the left. For further information about the Collection window see chapter 0.

Book:

Opens a new Book window to fill in the meta-data about a new book-object and specify the image files that belong to the book-object.

If there is no connection to the database, the object can only be stored locally.

If there is a connection to the database, the execution of the menu item [New→Book] will work only if a collection has been selected. After sending it to the server, the newly-created book-object will be part of this collection.

This menu item is also available via the context menu of the collection-tree-area on the left. For further information about the Book window see chapter 0.

Postcard:

Opens a new Postcard window to fill in the meta-data about a new postcard-object and specify the image file that belongs to the postcard-object.

If there is a connection to the database, the execution of the menu item [New→Postcard] will work only if a collection has been selected. After sending it to the server, the newly-created postcard-object will be part of this collection.

This menu item is also available via the context menu of the collection-tree-area on the left. For further information about the Postcard window see chapter 0.

PDF:

Opens a new PDF window to fill in the meta-data about a new PDF-object and specify the PDF-file that belongs to the PDF-object.

If there is a connection to the database, the execution of the menu item [New→PDF] will work only if a collection has been selected. After sending it to the server, the newly-created PDF-object will be part of this collection.

This menu item is also available via the context menu of the collection-tree-area on the left. For further information about the PDF window see chapter 0.

RTF:

Opens a new RTF window to fill in the meta-data about a new RTF-object and specify the RTF-file that belongs to the RTF-object.

If there is a connection to the database, the execution of the menu item [New→RTF] will work only if a collection has been selected. After sending it to the server, the newly-created RTF-object will be part of this collection.

This menu item is also available via the context menu of the collection-tree-area on the left. For further information about the RTF window see chapter 0.

2.1.2 Exit

Closes the application.

2.2 Edit

2.2.1 Change

Changes the details about the item that is currently selected.

If the selected item is a collection, a dialog for changing the meta-data of the collection will be opened.

If the selected item is of the type Book, Postcard, PDF or RTF, the corresponding window will be opened and gives the possibility for changes. The changes can be saved by sending the information to the server again.

This menu item is also available via the context menu of the collection-tree-area on the left.

2.2.2 Delete

This menu item is only visible if the active window is the Objects window, it is not visible if the Book, Postcard, PDF or RTF window has the focus.

A click on the menu item [Delete] deletes the selected item. If there is no item selected, a click on [Delete] simply has no effect.

This menu item is also available via the context menu of the collection-tree-area on the left.

2.2.3 Save to file

This menu item is visible only if one of the following windows is focused: Book, Postcard, PDF or RTF. Since this menu item is also realised as a button on each of these windows, we refer to 0 for details.

2.2.4 Load from file

This menu item is visible only if one of the following windows is focused: Book, Postcard, PDF or RTF. Since this menu item is also realised as a button on each of these windows, we refer to 0 for details.

2.2.5 Send to server

This menu item is visible only if one of the following windows is focused: Book, Postcard, PDF or RTF. Since this menu item is also realised as a button on each of these windows, we refer to 0 for details.

2.2.6 Export to file

This menu item is visible only if one of the following windows is focused: Book, Postcard, PDF or RTF. Since this menu item is also realised as a button on each of these windows, we refer to 0 for details.

2.2.7 R ↔ V

This menu item is visible only if a Book window is focused. If (at least two) image files are selected, a click on [R↔V] swaps two consecutive items, otherwise an error message will be displayed.

The basic idea of this menu item is to quickly change the order of recto and verso pages if the order is wrong because of alphabetical file name conventions.

2.3 Database

2.3.1 Connect

This menu item is enabled only if there is currently no connection to the database. When it is clicked, an option dialog is opened which lets you choose a database to connect to (see also chapter 0).

2.3.2 Disconnect

This menu item is enabled only if there is currently a connection to the database. When it is clicked, the current connection is closed.

2.3.3 INI-File

A click on the menu item [INI-File] opens a dialog that lets you choose (the path of) an INI-File (see also chapter 0). The INI-File is a text file that contains details about the basic settings and predefined values.)

2.4 Options

2.4.1 Show Objects in Collections

The menu item [Show Objects in Collections] can be enabled / disabled by a single click. The current status is indicated by a check mark in front of the menu item name.

If the menu item [Show Objects in Collections] is active, the collections are expanded and the objects within the collections are shown immediately. Otherwise, only the collections themselves are displayed.

For a detailed description see chapter 0.

2.4.2 Scanner

A click on the menu item [Scanner] opens a new dialog that lets you choose existing scanner information or insert the details about a new scanner. For details see chapter 0.

2.4.3 Meta-data

A click on the menu item [Meta-data] opens a new dialog that lets you choose existing meta-data information or insert the details about new meta-data. For details see chapter 0.

2.4.4 Sort files when added

The menu item [Sort files when added] can be enabled / disabled by a single click. The current status is indicated by a check mark in front of the menu item name.

If the menu item [Sort files when added] is active, a click on the button [Add files] implies that the files will be displayed alphabetically, no matter in which order they have been selected.

2.5 Window

2.5.1 Tile Windows

A click on the menu item [Tile Windows] tiles all open windows as can be seen in Figure 0.2.

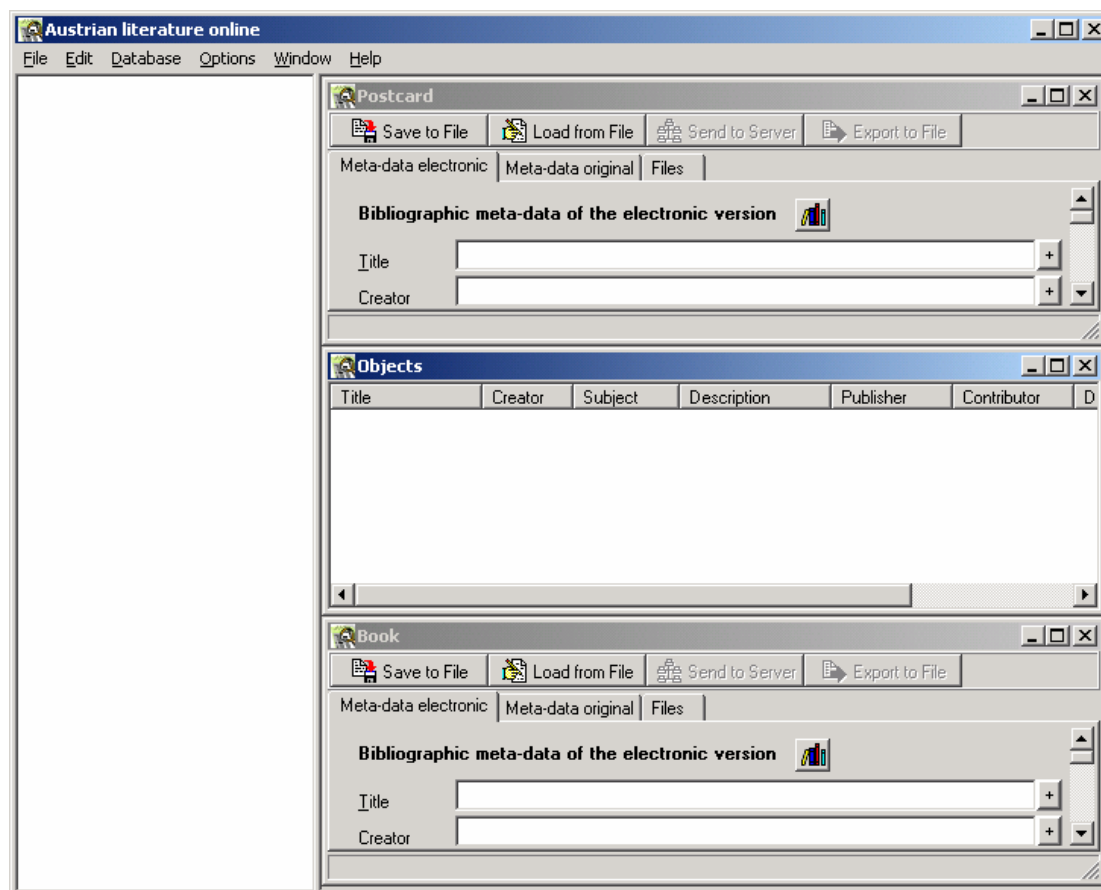


Figure 0.2: Tiled windows

2.5.2 Cascade

A click on the menu item [Cascade] cascades all open windows as can be seen in Figure 0.3.

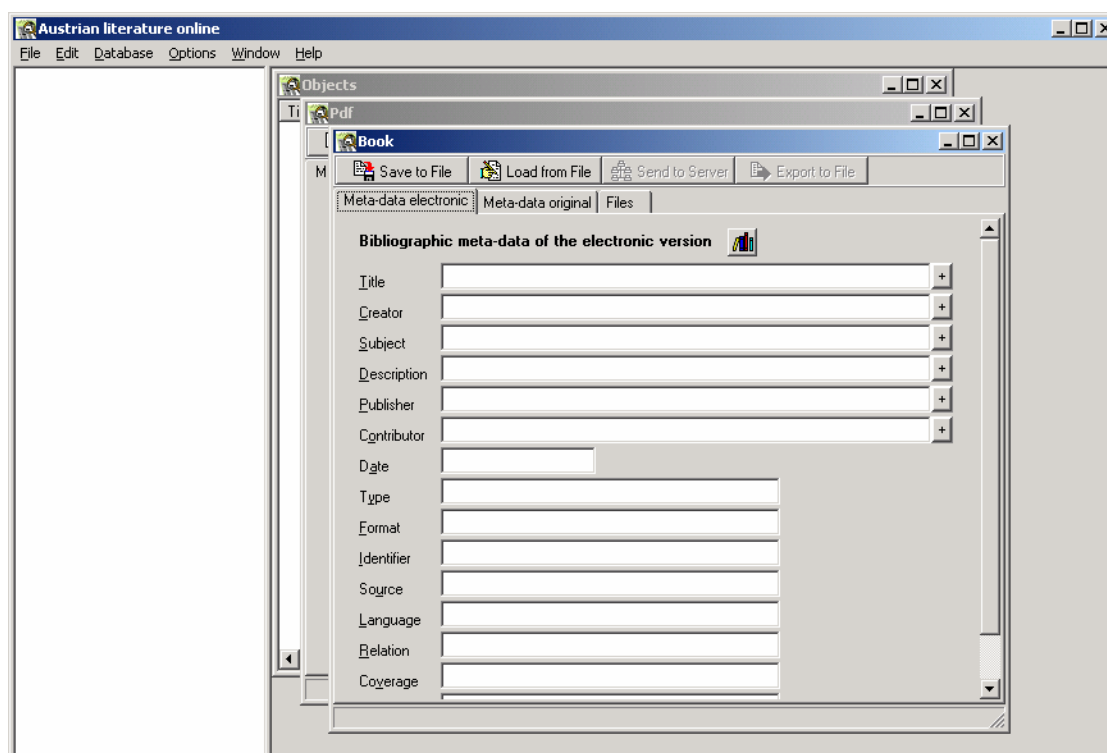


Figure 0.3: Cascaded windows

2.5.3 Minimize all

A click on the menu item [Minimize all] minimizes all open windows as can be seen in Figure 0.4.

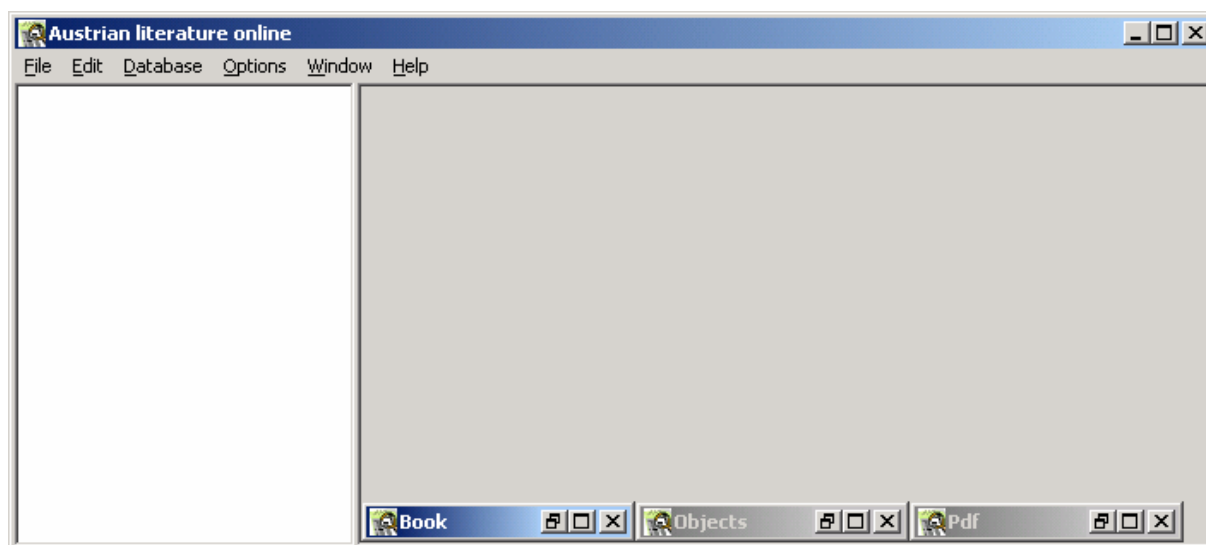


Figure 0.4: Minimized windows

2.5.4 Maximize all

A click on the menu item [Maximize all] maximizes all open windows.

2.5.5 Close all

A click on the menu item [Close all] closes all open windows except the book window (which can never be closed).

2.6 Help

2.6.1 Contents

A click on the menu item [Contents] opens the table of contents of the help file.

As an alternative to clicking the menu item [Contents] you can also press the key [F1] as a shortcut.

2.6.2 Search for Help On...

A click on the menu item [Search for Help On ...] opens a window that let's you search for a keyword or let's you choose an entry from a list of alphabetically ordered keywords.

2.6.3 How to Use Help

A click on the menu item [How to Use Help] gives an introduction in how to use the help provided.

2.6.4 About

A click on the menu item [About] opens a new window with information about the ALO-Client such as version number and copyright information.

3. Forms and Dialogs

3.1 Austrian Literature Online – Main form

The main window of the ALO client is separated into two regions (see Figure 0.5):

- A tree view containing information about collections on the left hand side (also denoted as collection-tree-area in the following)
- A region to display different windows on the right hand side

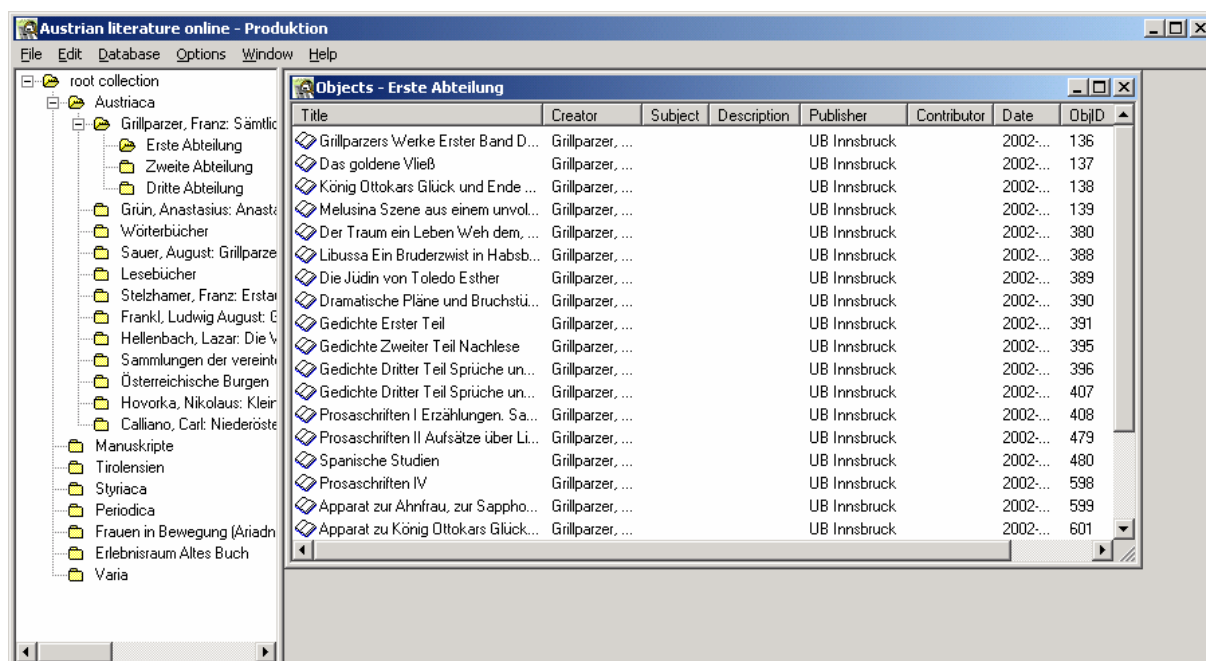


Figure 0.5: Austrian Literature Online – The main form

The left hand side contains a list of all collections, represented by a tree. When connecting to the database, only the root collection will be displayed. It contains no objects, but some other (sub)collections. These subcollections on the first level can be viewed by clicking on the plus symbol in front of the root collection.

If and how subcollections on deeper levels and objects within collections are displayed depends on whether the item [Show Objects in Collection] in the main menu is checked.

[Show Objects in Collection] is checked:

By clicking (single click) on one of the collections, the objects contained in this collection will be displayed in the Objects-window on the right hand side.

If the collection contains subcollections, a plus sign will appear in front of the collection name, which can again be opened by a single click.

[Show Objects in Collection] is not checked:

When doing a single click on one of the collections, nothing will happen this time.

But by doing a double click on one of the collections, the collection will be searched for subcollections. If there could be found any, they will again be displayed beneath the collection itself in the tree view.

Objects themselves will never be displayed in this mode, the Objects window on the right hand side remains empty.

3.2 Objects

Figure 0.6 shows the Objects window. It displays a list of objects that are part of the collection that is currently selected.

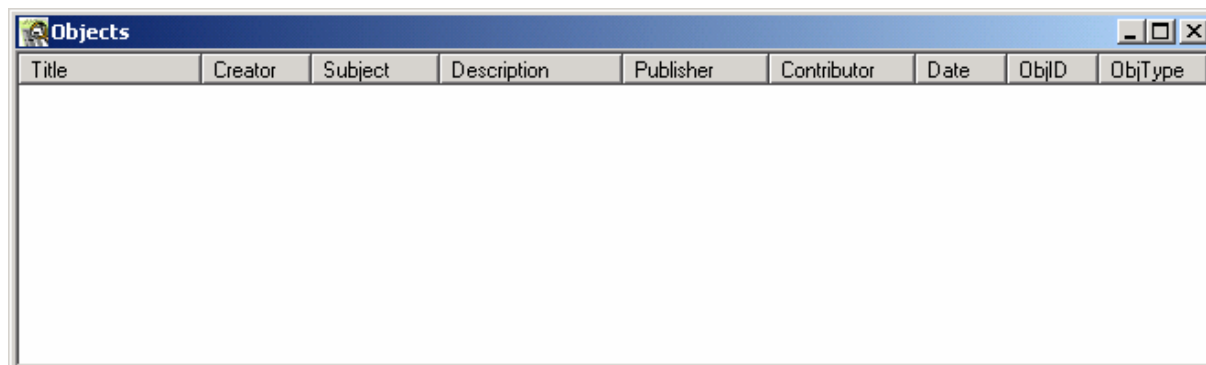


Figure 0.6: The *Objects* window

The first seven columns show the values that have been stated by the user of the client when uploading the object to the server. Additionally, the values for the ObjID and the ObjType, which are generated automatically, are displayed.

The columns have the following meaning:

Title	Title of the object
Creator	Creator of the object
Subject	Keywords or Classification-number
Description	Additional description of the object
Publisher	Publisher of the object
Contributor	Contributor of the object
Date	Date of creation of the electronic object
ObjID	ID of the object (has been uniquely assigned by the server after uploading)
ObjType	Type of the object (Book, PDF, RTF, Postcard; has been assigned automatically when uploading)

3.2.1 Double click

A double click on one of the objects in the list opens the corresponding Book, PDF, RTF or Postcard dialog which displays detailed information about the object.

For further information about the single dialogs we refer to the chapter 0, chapter 0 and chapter 0.

3.2.2 Delete

A click on the menu item [Edit→Delete] in the main menu deletes the object(s) currently selected in the list.

Alternatively, you can also press the Delete-Button on the keyboard or use the context menu of the Objects window.

3.3 Collection

When creating a new collection, the client asks you to state some meta-data which you have to insert into the form which can be seen in Figure 0.7.

Figure 0.7: The *Collection* window

The <Collection Type> can be selected from a choice box. By now, two different collection types are supported: ordinary (for Books, PDFs and RTFs) and Postcard (for Postcards).

The other entries have the following meaning:

Title	Title of the collection (obligatory)
Creator	Name of the person that created this collection (optional)
Subject	Keywords or Classification-number (optional)
Description	Additional description of the collection (optional)
Publisher	Not assigned
Contributor	Not assigned
Date	Creation date of the collection (by default the current date will be set)
Type	Not assigned
Format	Not assigned

Identifier	Not assigned
Source	Not assigned
Language	If all objects that will be part of this collection will have the same language, the ISO Language code (for example “ISO 639: 1988 de”) can be stated here. The language codes are available at http://www.oasis-open.org/cover/iso639a.html
Relation	Not assigned
Coverage	Not assigned
Rights	Not assigned

3.4 Book

The Book window (see Figure 0.8) consists of three different tab sheets: <Meta-data electronic>, <Meta-data original> and <Files>. Additionally, there are four buttons: [Save to file], [Send to Server], [Export to File], [Load from File], whereas the button [Send to server] is enabled only if there is a valid connection to the server when the Book window is opened, and the button [Export to File] is enabled only when the Book window displays an object that is already on the server.

Book

Save to File Load from File Send to Server Export to File

Meta-data electronic Meta-data original Files

Bibliographic meta-data of the electronic version

Title

Creator

Subject

Description

Publisher

Contributor

Date

Type

Format

Identifier

Source

Language

Relation

Coverage

Rights

Copy to original

Figure 0.8: The *Book* window

3.4.1 Tabsheet <Meta-data electronic>

Speedbutton 'Meta-data':

The small button showing some coloured books (next to the caption <Bibliographic meta-data of the electronic version>) gives you the possibility to select predefined meta-data and to edit them (change existing meta-data and add new meta-data).

A click on the button pops up a context menu containing a list of all predefined meta-data that are currently available and the entry [Edit]. If you select one of the names in the list of predefined meta-data, the corresponding entries will be read from the INI-File and inserted into the fields of the Book window. If you select [Edit], the Meta-data Options dialog will be opened. For details about this dialog we refer to chapter 0.

Bibliographic meta-data of the electronic version:

Title	<ul style="list-style-type: none"> - Title, subtitles, add-ons to titles etc. For example: "Second edition", "Containing 3 illustrations" - If available, also translations of foreign-language titles can be listed.
Creator	<ul style="list-style-type: none"> - Author or creator of the book; the person or institution responsible for the content. - If working on journals, magazines or anthologies: Editor/Publisher - Exception: Scientific editions (for example "Goethe's Collected Works": Creator = "Goethe", Contributor = name of the scientific editor) - Names have to be entered in the commonly used form. If needed, some research within the catalogues Aleph and KVK can be done. Aleph: http://www.bibvb.ac.at/verbund-opac.htm KVK: http://kvk.uni-karlsruhe.de - European names: Last name – comma blank – First name, for example "Stifter, Adalbert" - Non-European names: For example "Mao Tse Tung" - Abbreviations: "Georg K. Feuerbach" changes to „Feuerbach, Georg K." - Pseudonyms: Insert the most popular form (For example "Jean Paul", not "Richter, Jean Paul"). Besides, if it is also known, the real name can be inserted into a second field. - Books without authors: keep the field empty - Multiple authors: insert the names in separate lines, without "and" etc.
Subject	<ul style="list-style-type: none"> - Keywords or Classification-number (optional)
Description	<ul style="list-style-type: none"> - Additional description of the object (optional)
Publisher	<ul style="list-style-type: none"> - The library or organisation that has done the scanning and uploading to the ALO system. The name of the library or organisation has to be written without abbreviations (for example "Universitätsbibliothek Innsbruck" instead of "UBI")
Contributor	<ul style="list-style-type: none"> - Persons who contributed the book (for example illustrators, authors of introductions, publisher/editor of scientific editions, etc.)
Date	<ul style="list-style-type: none"> - Date of the upload - Format: YYYY-MM-DD
Type	<ul style="list-style-type: none"> - Not assigned
Format	<ul style="list-style-type: none"> - Not assigned

Identifier	- Not assigned
Source	- Not assigned
Language	- ISO Language code (for example “ISO 639: 1988 de”) - The language codes are available at http://www.oasis-open.org/cover/iso639a.html
Relation	- Not assigned
Coverage	- Not assigned
Rights	- Not assigned

Button [Copy to original]:

The Button [Copy to original] copies all entries made on the tabsheet <Meta-data electronic> to the tabsheet <Meta-data original>.

3.4.2 Tabsheet <Meta-data original>

Most of the meta-data entries of the electronic object are similar or equal to the meta-data entries of the original object. Therefore, all entries that are marked with ‘See above’ in the following table refer to the table captioned ‘Bibliographic meta-data of the electronic version’ in chapter 0.

Bibliographic meta-data of the original version:

Title	- See above
Creator	- See above
Subject	- See above
Description	- See above
Publisher	- First line: publisher of the book - Second line: publishing location - If there are more publishers, only the most important one should be stated. If that is not obvious, insert the first one. (The same rule applies to the location.)
Contributor	- See above
Date	- Year of publication of the book in the format: YYYY - If there is an area of years, separate them with a dash without blanks (for example “YYYY-YYYY”) - Insert the year if it is known, even if it is not stated in the book. - Leave the field empty, if the year is not known.
Type	- Not assigned
Format	- Not assigned
Identifier	- Abbreviation of the library and signature of the original book - For example: “UBG-HB I 45958” or “UBI-GER II 846”
Source	- Not assigned
Language	- See above
Relation	- Not assigned
Coverage	- Not assigned
Rights	- Not assigned

Button [Copy to electronic]:

The Button [Copy to electronic] copies all entries made on the <Meta-data original>-Tabsheet to the <Meta-data electronic>-Tabsheet.

3.4.3 Tabsheet <Files>

The tabsheet <Files> is separated into a left and a right area as can be seen in Figure 0.9.

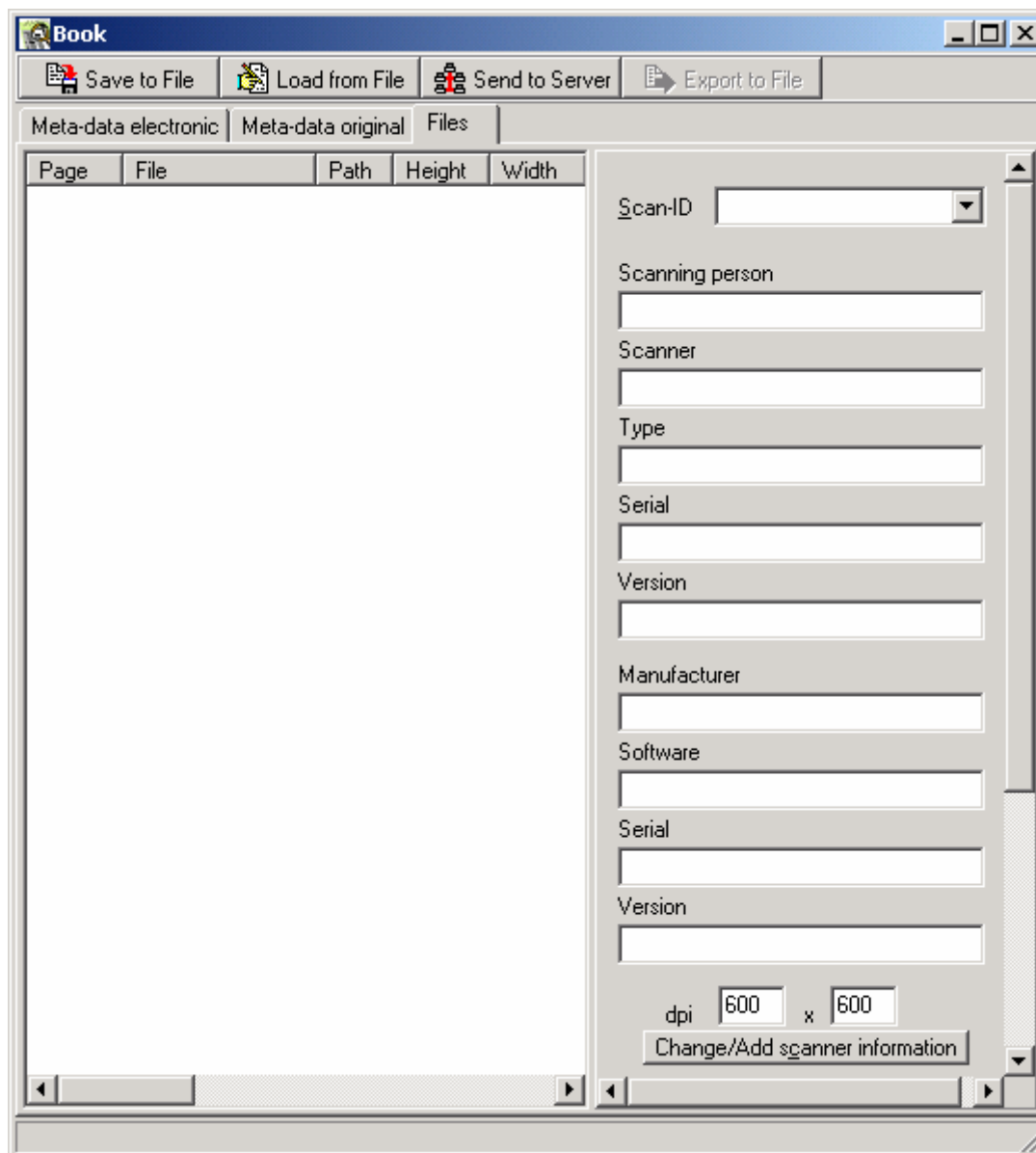


Figure 0.9: Tabsheet <Files>

On the left hand side the list of files that belong to the current object (Book, PDF, RTF, Postcard) will be displayed. The list contains the following columns:

Page	Page number
File	File name
Path	Local path, if the file has just been added from the local disk Empty, if the data has been loaded from the server or from an XML-file
Height	Scanner resolution – height

Width	Scanner resolution – width
Scanning person	Last name of the person who created the files / scanned the images
Scanner	Brand name of the scanner (e.g. “Minolta”)
Type	Type of the scanner (e.g. “PS7000”)
Serial scanner	Serial number of the scanner
Version scanner	Version number of the scanner
Manufacturer	Manufacturer of the software used for scanning
Software	Name of the software used for scanning
Serial software	Serial number of the software used for scanning
Version software	Version number of the software used for scanning
Created	Date of the creation of the file
Format type	Type of the files (“JFIF”, “GIF”, “PNG”, “BMP”, “TIFF”, “PDF”, “RTF”)
Image width	Width of the image, if format type is “JFIF”, “PNG”, “BMP”, “TIFF”; not assigned otherwise.
Image height	Height of the image, if format type is “JFIF”, “PNG”, “BMP”, “TIFF”; not assigned otherwise.
Compression	Type of compression of image files (e.g. “BI_RGB” for Bitmap, “CCITT 64” for Tif, “deflate/inflate” for Png)
URL	Empty, if the file has just been added from the local disk URL of the file, if the file has been loaded from the server (“http://...”) or from a file (“file://...”)

When looking at an object that is already on the server, the server sends the information displayed in this list. When creating a new object, the list will be empty at the beginning. After adding files to the object (by using the button [Add files] on the right hand side), the columns about the scanner information (Scanning person, Scanner, Type, etc.) will contain the information given on the right hand side of the tabsheet <Files>, the missing information (such as Page, File, Path etc.) will be extracted directly from the files themselves or generated automatically.

A double click on one of the elements in the list opens a new window that displays the file according to the list element. As illustrated in Figure 0.10, this new window offers some zooming possibilities for the image displayed:

- Button [+]: Zoom in
- Button [-]: Zoom out
- Button [full page]: Changes the size of the image such that it fits to the windows
- Button [original size]: Displays the image in its original size



Figure 0.10: Zooming possibilities

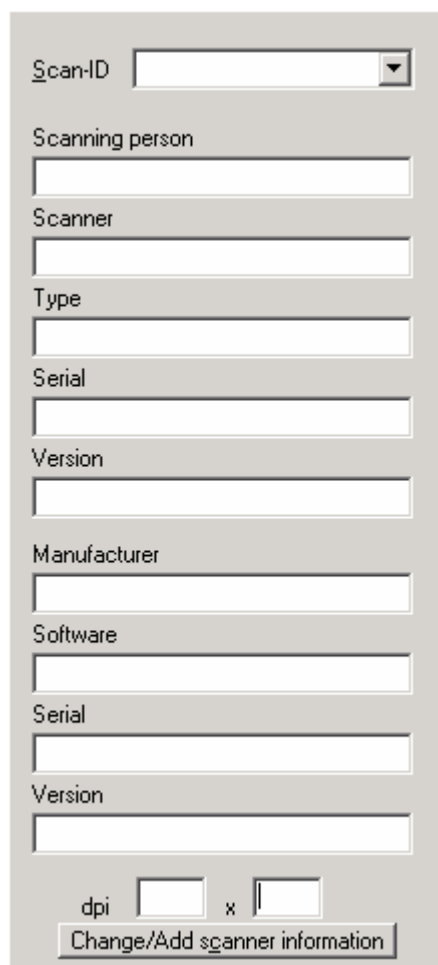
The right hand side of the tabsheet <Files> is separated into three parts that will be described in the following.

Part 1: Scan information

Part 1, which is illustrated in Figure 0.11, prepares some fields for displaying or entering information about the scanner.

You can either choose an existing scan information from the <Scan-ID> list or enter some new information yourself.

If you enter some new information here, be aware that the information is only valid for the single object you are currently working on. So if you want to add scan information for permanent use, use the item [Options→Scanner] in the main menu (see chapter 0). For a detailed description of the scan options take a look at chapter 0.



A vertical form titled 'Scan Information' with the following fields: Scan-ID (dropdown), Scanning person (text), Scanner (text), Type (text), Serial (text), Version (text), Manufacturer (text), Software (text), Serial (text), Version (text), and a dpi field with two input boxes and an 'x' separator. At the bottom is a button labeled 'Change/Add scanner information'.

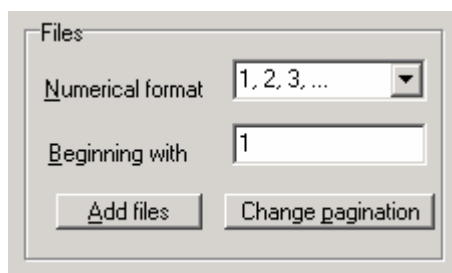
Figure 0.11: Scan Information

Part 2: Files

Part 2, which is illustrated in Figure 0.12, provides functionalities for adding files to the current object and changing the pagination of these files.

When clicking on the button [Add files], a standard Windows dialog for selecting files opens. There, you can select the (image) files you want to add to the object. The list of selected files will be displayed on the left, the page numbers will be displayed as Arabic or as Roman numbers, in dependence on which format has been chosen from the choice box <Numerical format>. You can also change the number the numbering should start with by changing the value of the <Beginning with> text field.

If you found that there was something wrong with the pagination, you can easily change the page numbers by selecting the concerned files from the list on the left, making the desired modifications and clicking on the button [Change pagination] afterwards.



A dialog box titled 'Files' with the following elements: 'Numerical format' dropdown menu (showing '1, 2, 3, ...'), 'Beginning with' text field (showing '1'), and two buttons: 'Add files' and 'Change pagination'.

Figure 0.12: Files

Part 3: Special

Part 3, which is illustrated in Figure 0.13, offers two special features.

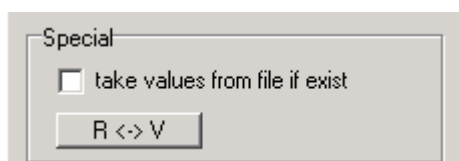


Figure 0.13: Special features

Checkbox [take values from file if exist]:

Some image files can already contain some scan information. By default, these values are ignored.

If the checkbox [take values from file if exist] is marked, the information in the files is preferred. If some information is missing, the values of the appropriate user input are taken instead.

Button [R<->V]:

The button [R<->V], has the same functionality as the item [R<->V] in the main menu. For details see chapter 0.

3.4.4 Button [Save to file]

A click on [Save to file] saves the information about the active object onto the local disk as an XML-File. This function is useful if you want to create the files now, but send them to the server later. This can be done either by loading the file into the client again later (with the help of the button [Load from File]) or by using the ACLnight-client for sending files during the night.

3.4.5 Button [Load from File]

A click on [Load from File] loads the information about a new object from a file that has previously been saved by clicking the button [Save to file] or [Export to file]. The XML-file is extracted again and the information about meta-data and files is displayed.

3.4.6 Button [Send to Server]

This button is enabled only if there has been a connection to the database at the time the Book window has been opened. Otherwise, it will be grayed.

A click on [Send to server] sends the active object (data and files) to the server immediately.

3.4.7 Button [Export to File]

A click on [Export to file] saves the selected object (XML file and image-, PDF- or RTF-files) onto the local disk. All files will be saved in the same directory.

The button [Export to file] is only useful and for objects viewed from the server. Therefore, it is enabled for objects from the server and grayed for newly-created objects.

3.5 Postcard

The Postcard window is very similar to the Book window (see chapter 0).

There are only a few differences:

- You can add only a single file to the current object.

- As a consequence, the button [r→v] of part 3 (<Special>) is missing on the tabsheet <Files> because there cannot be a recto-verso-problem with only one file.

3.6 PDF and RTF

The PDF and the RTF window are also very similar to the Book window (see chapter 0).
The differences are:

- You can add only a single file to the current object.
- Because of the previous point (there can be no recto-verso-problem with only one file) and the fact that there is no image information within PDF- and RTF-files, part 3 (<Special>) is missing on the tabsheet <Files>.
- Instead of the scan information, the user can enter information about the software that has been used to create the PDF-File or the RTF-File respectively, as can be seen in Figure 0.14.

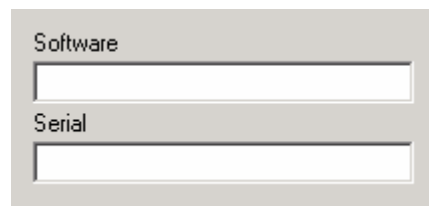


Figure 0.14: Software information

The two fields have the following meaning:

Software:

Name of the software product that has been used to create the files (e.g. “Adobe Acrobat Writer”, “Microsoft Word 2000”)

Serial:

Serial number of the software that has been used to create the files

3.7 Database Options

The dialog Database Options, which can be seen in Figure 0.15, lets you choose which database to connect to and allows you to add, change or remove information about databases and save the changes in the INI-File.

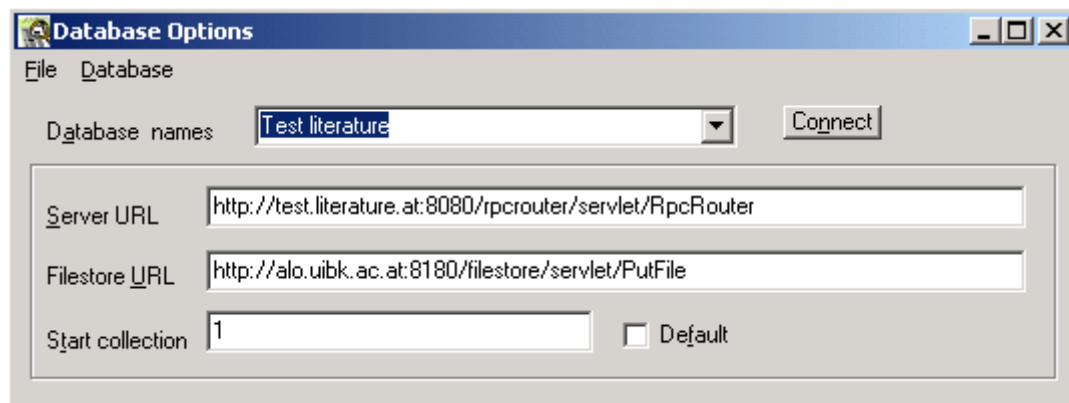


Figure 0.15: The Database Options dialog

3.7.1 Menu-Items

[File→Exit]:

A click on the menu item [Exit] closes the dialog Database Options.

[Database→Add]:

A click on the menu item [Add] gives you the possibility to add the details about a new database connection to the list. The changes can be confirmed by clicking the [Ok] button or discarded by clicking the [Cancel] button, which were both made visible with the click on [Add].

[Database→Change]:

A click on the menu item [Change] gives you the possibility to change the details about the database currently selected. The changes can be confirmed by clicking the [Ok] button or discarded by clicking the [Cancel] button, which were both made visible with the click on [Change].

[Database→Remove]:

A click on the menu item [Remove] removes the information about the current database details from the INI-File.

[Database→Connect]:

A click on the menu item [Connect] connects you to the chosen database.

[Database→Details]:

If the menu item [Details] is not checked, a click on [Details] opens a panel with additional information about the connection to the database. Otherwise, if the menu item is already checked, a click on [Details] implies that the panel containing the details is closed.

3.7.2 Connection details

Database names:

A name for the connection that can be chosen by the user.

Server URL:

URL of the server you want to connect to, for example:

“http://test.literature.at:8080/rpcrouter/servlet/RpcRouter” for the test database

“http://www.literature.at:8180/rpcrouter/servlet/RpcRouter” for the production database

Filestore URL:

URL of the filestore you want to connect to, for example:

“http://alo.uibk.ac.at:8180/filestore/servlet/PutFile”

Start collection:

Indicates which collection to start with, for example:

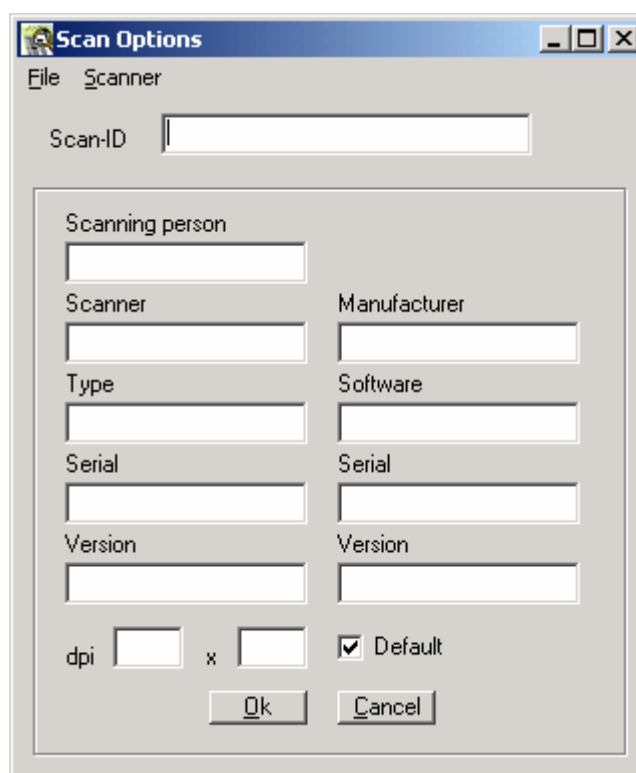
“1” for starting with the root collection

Default checkbox:

If the checkbox is marked, the database connection that is currently chosen is set as default connection.

3.8 Scan Options

The dialog Scan Options, which can be seen in Figure 0.16, lets you choose existing scanner information or insert the details about a new scanner.



The 'Scan Options' dialog box is a standard Windows-style window. It features a title bar with the text 'Scan Options' and standard window control buttons (minimize, maximize, close). Below the title bar is a menu bar with two items: 'File' and 'Scanner'. The main area of the dialog contains several input fields arranged in a grid-like fashion. At the top is a single-line text field labeled 'Scan-ID'. Below this is a section titled 'Scanning person' followed by a text field. To the right of this section are two columns of fields: 'Scanner' and 'Manufacturer' in the first row, 'Type' and 'Software' in the second, 'Serial' and 'Serial' in the third, and 'Version' and 'Version' in the fourth. At the bottom left, there is a 'dpi' label followed by a text field, then an 'x' label, then another text field. To the right of these is a checked checkbox labeled 'Default'. At the very bottom are two buttons: 'Ok' and 'Cancel'.

Figure 0.16: Scan Options

3.8.1 Menu-Items

[File→Exit]:

A click on the menu item [Exit] closes the dialog Scan Options.

[Scanner→Add]:

A click on the menu item [Add] gives you the possibility to add information about a new scanner to the list. The changes can be confirmed by clicking the [Ok] button or discarded by clicking the [Cancel] button, which were both made visible with the click on [Add].

[Scanner→Change]:

A click on the menu item [Change] gives you the possibility to change the details about the scanner currently selected. The changes can be confirmed by clicking the [Ok] button or discarded by clicking the [Cancel] button, which were both made visible with the click on [Change].

[Scanner→Remove]:

A click on the menu item [Remove] removes the information about the current database details from the INI-File.

3.8.2 Scanner details

Scanning person:

Last name of the person who scanned the files

Scanner:

Brand name of the scanner (e.g. “Minolta”, “BookEye”)

Type:

Type of the scanner (e.g. “PS7000”)

Serial:

Serial number of the scanner

Version:

Version number of the scanner

dpi:

Height and width of the scanner resolution (e.g. “300”x”300”)

Manufacturer:

Manufacturer of the software used for scanning (e.g. “SRZ” (Satz-Rechen-Zentrum) or “ImageWare”)

Software:

Name of the software used for scanning (e.g. “ProScan Book” or “BCS-2”)

Serial:

Serial number of the software used for scanning

Version:

Version number of the software used for scanning

Default checkbox:

If the checkbox is marked, the scanner details that are currently chosen are set as default scanner details.

3.9 Meta-data Options

The dialog Meta-data Options, which can be seen in Figure 0.17, lets you choose existing meta-data information or insert the details about new meta-data.

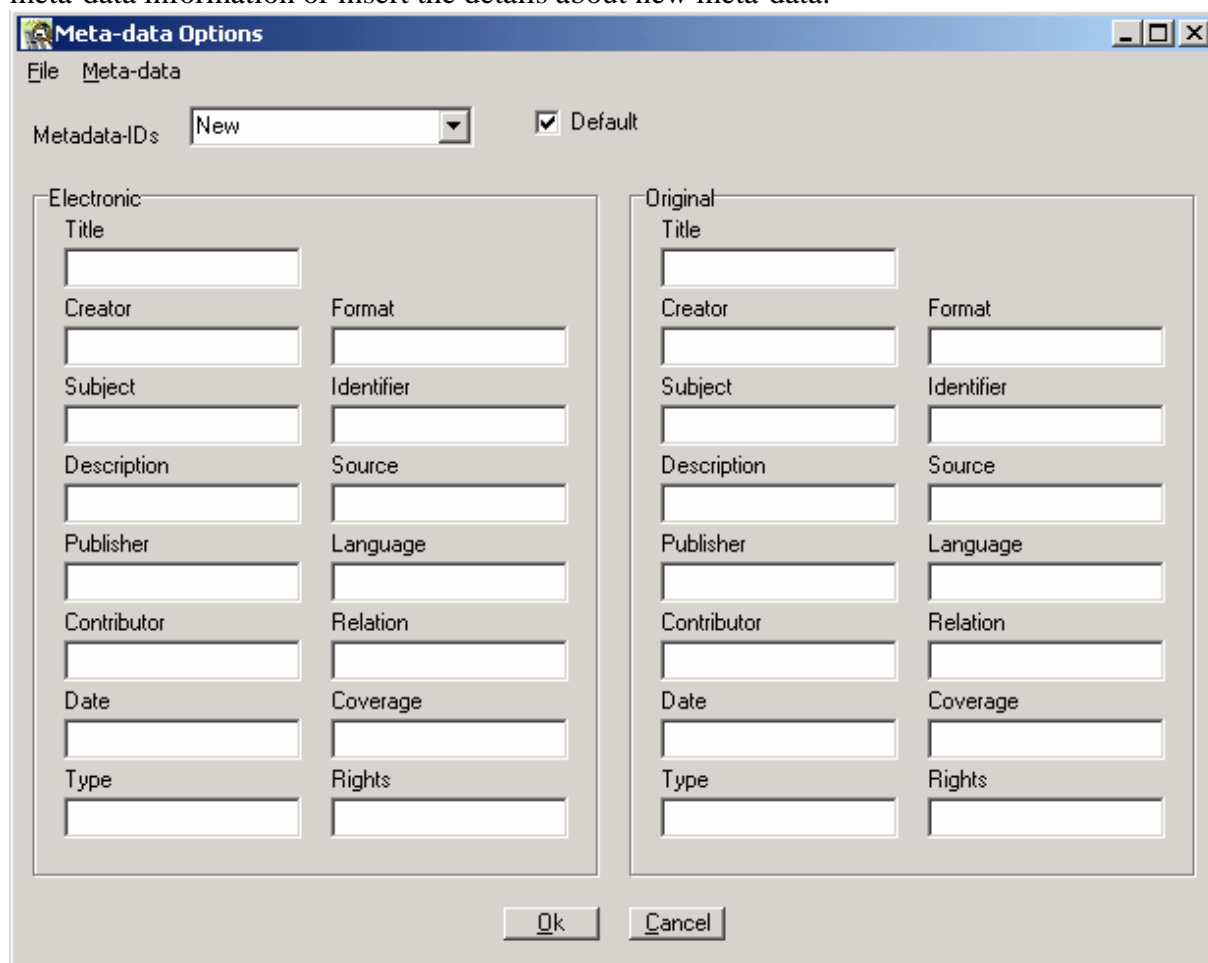


Figure 0.17: Meta-data Options

3.9.1 Menu-Items

[File→Exit]:

A click on the menu item [Exit] closes the dialog Meta-data Options.

[Meta-data→Add]:

A click on the menu item [Add] gives you the possibility to add information about new meta-data to the list. The changes can be confirmed by clicking the [Ok] button or discarded by clicking the [Cancel] button, which were both made visible with the click on [Add].

[Meta-data→Change]:

A click on the menu item [Change] gives you the possibility to change the details about the meta-data currently selected. The changes can be confirmed by clicking the [Ok] button or discarded by clicking the [Cancel] button, which were both made visible with the click on [Add].

[Meta-data→Remove]:

A click on the menu item [Remove] removes the information about the currently selected meta-data details from the INI-File.

3.9.2 Meta-data details

MetaDataIDs: Name of the predefined meta-data, that can be chosen by the user.

GroupBox Electronic/GroupBox Original: For a detailed information regarding the entries for “Title”, “Creator”, “Subject”, etc. we refer to chapter 0 and chapter 0.

Default checkbox: If the checkbox is marked, the meta-data details that are currently chosen are set as default values.

3.10 INI-File

The INI-File contains details about database connections, scan information, meta-data information and other basic settings. It is a simple text file and has the extension “.ini”. Within the ALO Client, the path of the INI-File has to be specified in the dialog shown in Figure 0.18. Below the figure, some information that can be contained in the INI-File is listed.

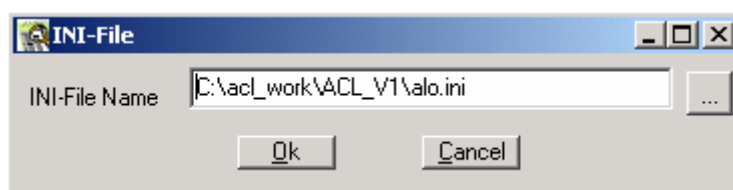


Figure 0.18: The *INI-File* window

3.10.1 General information

Figure 0.19 shows how general information, such as for example the date of the creation of the INI-File, is stated.

```
[Austrian literature online]  
Created=19.08.2002
```

Figure 0.19: General information in the INI-File

3.10.2 Database information

Figure 0.20 shows an example of how the information about two different databases, the “Production” database and the “Test literature” database, is stated in the INI-File.

```
[Connect]
Produktion=
Test literature=

[Produktion]
Rpcrouter=http://www.literature.at:8180/rpcrouter/servlet/RpcRouter
Filestore=http://filestore.literature.at:8180/filestore/servlet/PutFile
StartCollection=1

[Test literature]
Rpcrouter=http://test.literature.at:8080/rpcrouter/servlet/RpcRouter
Filestore=http://alo.uibk.ac.at:8180/filestore/servlet/PutFile
StartCollection=1

[Default]
Connect=Test literature
```

Figure 0.20: Database information in the INI-File

3.10.3 Scan information

Figure 0.21 shows an example of how the scan information is stated in the INI-File:

```
[Scan]
Minolta default III=

[Scan_Minolta default III]
ScanID=Minolta default III
ScanPerson=Kulmer
Scanner=Minolta
ScannerType=PS7000
ScannerSerial=scanSerial
ScannerVersion=scanVersion
Manufacturer=SRZ
Software=ProScan Book
SoftwareSerial=manSerial
SoftwareVersion=manVersion
DpiWidth=600
DpiHeight=700
```

Figure 0.21: Scan information in the INI-File

3.10.4 Meta-data information

Figure 0.22 shows an example of how the meta-data information is stated in the INI-File:

```
[Meta]
i3s3=

[Meta_i3s3]
MetaElecTitle=
MetaElecCreator=
MetaElecSubject=
MetaElecDescription=
MetaElecPublisher=i3s3
MetaElecContributor=
MetaElecDate=
MetaElecType=
MetaElecFormat=
MetaElecIdentifier=
MetaElecSource=
MetaElecLanguage=DE
MetaElecRelation=
MetaElecCoverage=
MetaElecRights=
MetaOrigTitle=
MetaOrigCreator=
MetaOrigSubject=
MetaOrigDescription=
MetaOrigPublisher=
MetaOrigContributor=
MetaOrigDate=
MetaOrigType=
MetaOrigFormat=
MetaOrigIdentifier=
MetaOrigSource=
MetaOrigLanguage=DE
MetaOrigRelation=
MetaOrigCoverage=
MetaOrigRights=

[Meta_Default]
MetaID=i3s3
```

Figure 0.22: Meta-data information in the INI-File

Deliverable 1.2

Digital Repositories – Austria

Chapter 2

ALO-Server Description

Table of Contents

1. ALO – AUSTRIAN LITERATURE ONLINE	3
2. FACTS ABOUT THE RUNNING SYSTEM WWW.LITERATURE.AT	3
3. SOFTWARE/TECHNOLOGY USED FOR THE SERVER SIDE	3
4. FEATURES OF THE WEBSITE WWW.LITERATURE.AT	4
4.1 Hierarchy/browsing	4
4.2 Metadata Searching	4
4.3 Full text	4
4.4 PDF on the fly	4
4.5 Skin ability	5
5. ALO ARCHITECTURE /COMPONENTS	5
5.1. Metsserver	5
5.2 Metsobject.....	6
5.3 Webinterface.....	6
5.4 Filestore.....	6
5.5 Rpcrouter.....	6
6. STORAGE OF THE OBJECT TYPES.....	7
7. ALO TEST-SERVER	8

1. ALO – Austrian literature online

This part of the ALO-description covers the server side of the system as well as some features of the webpage www.literature.at. Also some facts about the content of the system are given. The description of the server also includes a brief explanation of how communication between the server and the client is realised.

2. Facts about the running system www.literature.at

The idea behind ALO was to bring the 1000 most important Austrian books online. Right now, the system contains about 4000 digital objects (which means about 650 000 pages). The space occupied by these documents on the NFS-Server of the central information technology service of the University of Innsbruck, where all files are hosted, is about 100 GB. These files and the files on the server machine are backed up daily on file level, which means also the database files are backed up.

3. Software/Technology used for the server side

The whole server application is written in Java. The currently used version is JDK 1.4.0. To process the parameters of a client request and for communicating to the client, Servlets are used. As Servlet Engine, Tomcat 4.0.6. is applied.

The Metsserver, which is the core element of the whole system, uses RMI technology. So every component which needs the functionality of this application server has to connect via RMI. Because of this, it is theoretically possible that the Metsserver runs on a different machine than the rest of the system (which means the database and the Webinterface).

For the presentation of the objects on the website, XML/XSL-transformation is used. All necessary content information for displaying is built within Servlets.

As database, an older version of mySQL(3.23) is used. Unfortunately, the ALO-system uses mySQL-specific SQL-commands, which means that replacing the database would cause some programming effort.

Some additional libraries of the Apache Jakarta Group are used for different purposes (e.g. logging, XML parsing and XSL-transformation). For on the fly scaling and converting of images (this is done by the Filestore Servlet), Java Advanced Imaging from SUN is used. PDF-generation on the fly is done by iText, a very easy to use and stable library for PDF reading and creating. One problem is that this package is a bit slow. For full text generation, we use ABBYY Finereader 6.0 scripting edition.

The server, where the system is hosted, is a dell Poweredge server 2500, and as operating system, Redhat Linux 7.2 is used.

4. Features of the website www.literature.at

4.1 Hierarchy/browsing

All objects are hierarchically organised within collections. Each object has to be a child node of exactly one collection. With the navigation bar, one can flick through the document list similar to the hit list presentation in google. To get a summary about the content of a collection, a bibliography can be created. Because of server-sided caching technology, browsing works very fast. Each collection and object is displayed depending on its object type.

4.2 Metadata Searching

Metadata search can be performed on either all Dublin core elements- of a book or only on specific fields. The results can be sorted by a certain criteria (for example by author or date).

4.3 Full text

There are also about 200 books with full text available. The recognition was done after uploading to the system. The text for a single page and also an html-file of the whole book is available and can easily be downloaded.

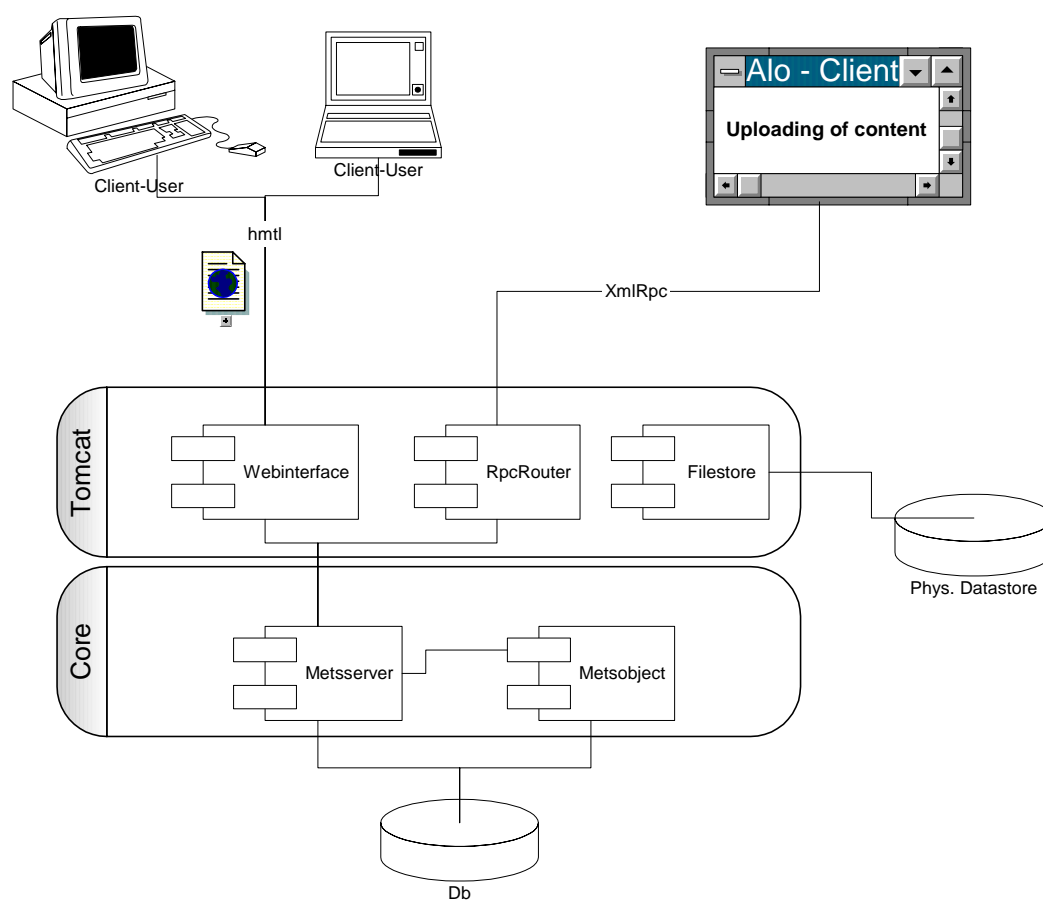
4.4 PDF on the fly

It is possible to get a PDF of a part of a book (limited to 10 pages per request) on the fly.

4.5 Skin ability

XML/XSLT-methodology enables user interface language independency (languages German and English are already implemented), and also different skinning independent from content itself is possible.

5. ALO Architecture /Components



5.1. Metsserver

The Metsserver is the core element of the system. Here, all tasks such as searching within the metadata, preview generation, adding objects or collections, deleting objects or collections, unique id generation, and so on, are implemented. Also a rudimentary administration client exists for tasks like re-creation of previews, deleting objects or collections and so on, exists.

5.2 Metsobject

This package is responsible for mapping a certain METS-xml file to a java object, to a relational database and vice versa. Currently, only mapping of METS Version 1.1 is implemented.

5.3 Webinterface

The Webinterface is realized with Servlets and XSL/XSLT. Creation of the necessary XML data happens in the Servlets (a Servlet exists for each object type and each task (such as searching)).

5.4 Filestore

The Filestore contains 2 Servlets, one for uploading files and the other one for accessing them. When uploading, a unique id is assigned to each file and returned to the client system who requested the upload. These ids are written to database afterwards to assign these files to the object where they belong to. Most of the images of ALO are encoded as TIF group 4 images, and files with TIF format cannot be displayed inside a browser without using a plug-in. ALO solves this problem in offering on the fly conversion, where the desired output format can be specified. Also tasks like resizing the file on the fly are possible.

One big advantage of the Filestore is that this component can be used without the rest of the system, so it can be integrated in projects with a totally different focus, just for storing files externally.

5.5 Rpcrouter

The Rpcrouter is the interface to the ALO-Client, which is written in Delphi. It enables platform and technology independent communication. It's similar to the way web services work, but the format or protocol used here, is proprietary.

Here some examples how communication between the server and the client works. A Request for a description of the root collection of the whole system, e.g., looks like this:

```
<Request>  
  <Method>getRootCollection</Method>  
</Request>
```

As response, a METS encoded description of the root collection (containing a list of all subcollections, is delivered)

Or, as another example, this is the way a collection with a certain id can be requested:

```
<Request>
  <Method>getCollection</Method>
  <Parameter>
    <Name>collection_id</Name>
    <Type>String</Type>
    <Value>2</Value>
  </Parameter>
</Request>
```

An object can be retrieved almost the same way:

```
<Request>
  <Method>getObject</Method>
  <Parameter>
    <Name>object_id</Name>
    <Type>String</Type>
    <Value>1174</Value>
  </Parameter>
</Request>
```

For further description of how such an ALO-METS-object (which would be the result of this example) is encoded within the ALO-system, please take a look at the description of the client.

6. Storage of the object types

As supported objects, there is an object-type simply called ALO_OBJECT (the clear majority of all objects within the system), which consists of a scan per page, administrative metadata for each scan and metadata for the whole object. The whole object is wrapped by METS. In ALO, METS is used for storing almost everything. Also a few PDFs are already in the system (about 40, which means about 1 percent of all objects), which are stored as a whole (no structure added at all). Right now, no RTF file has been uploaded, but the client can deal with this format. For the server, just a few extensions would be necessary when this type should also be used. Both, PDF and RTF- objects are also stored within METS.

To add any new object type to the system, just a few configuration settings and the implementation of style sheets for this type are necessary. There is only one constraint: METS has to be used as encoding language.

7. ALO Test-Server

At the moment the ALO Test-Server is used for uploading the digital documents, which are collected from various public institutions in Austria. The screenshots below show examples from the collection of reUSE styria.

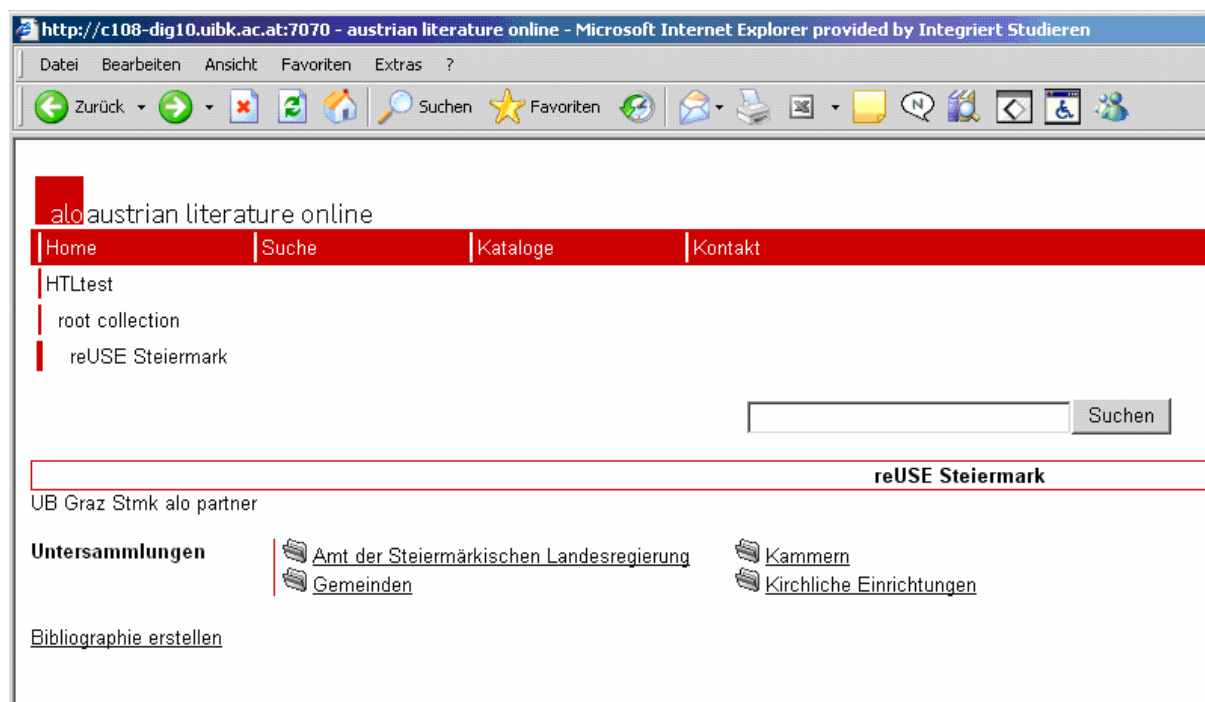


Figure 1: list of public institutions that are using the reUSE-Service so far.

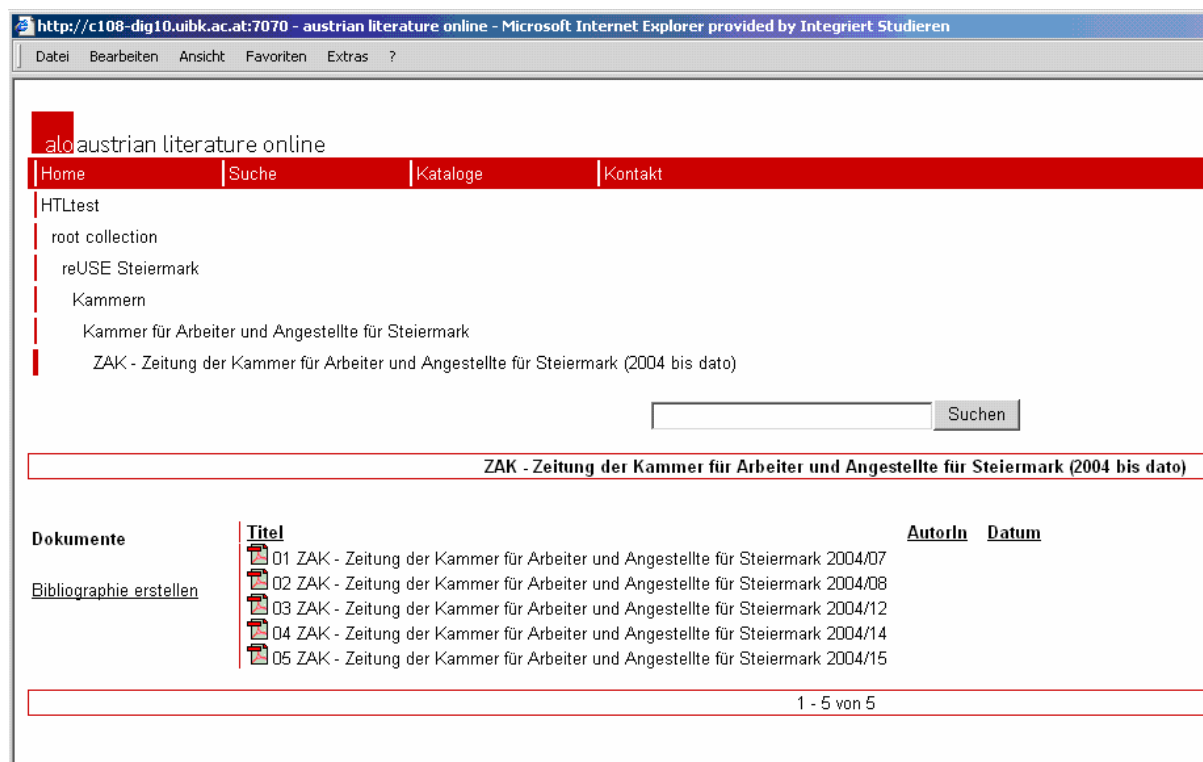


Figure 2: list of available digital documents from one institution

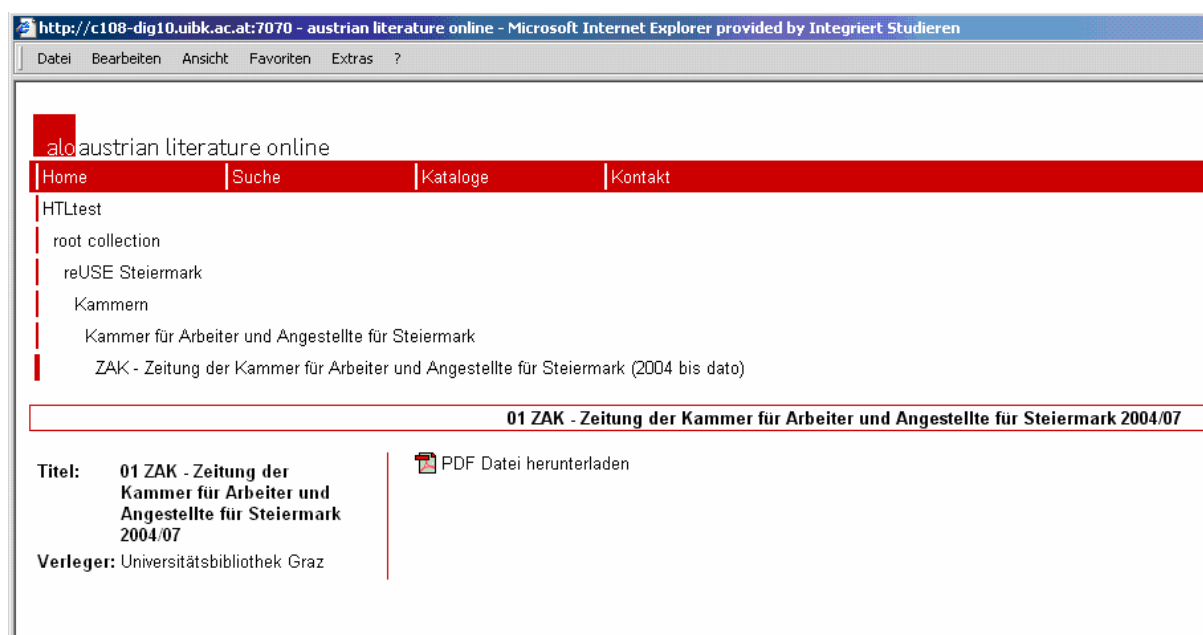


Figure 3: download of a document in PDF-Format

Deliverable 1.2

Digital Repositories – Austria

Chapter 3

Metadata Usage

Table of Contents

1. METADATA STANDARDS, ON WHICH THE ALO SYSTEM IS BASED	3
2. DUBLIN CORE.....	3
3. RDF	4
3.1 Example of RDF and DC.....	6
4. DIG35 METADATA SCHEMA.....	7
4.1 Basic Image Parameter Metadata	7
4.2 Image Creation Metadata.....	8
4.3 Content Description Metadata.....	10
4.5 History Metadata	10
4.6 Intellectual Property Rights (IPR) Metadata	11
5. METADATA ENCODING AND TRANSMISSION STANDARD (METS) METADATA SCHEMA	13
5.1 Descriptive Metadata.....	13
5.2 Administrative Metadata.....	14
5.3 File Groups	15
5.4 Structural Map.....	15
6. METADATA IN THE TIFF FILE	20
7. REFERENCES	24

1. Metadata standards, on which the ALO system is based

There are 5 standards, the metadata format for the ALO system is based on:

- eXtensible Mark-up Language (XML). XML has the advantage, that it is not only readable by machines, but can also be read by humans. By supporting Unicode as the standard character encoding most encoding problems can be overcome. For further information about XML, have a look at <http://www.w3.org/XML/> or for a lot of links e.g. at <http://www.oasis-open.org/cover/xml.html>
- Resource Description Framework (RDF) is used together with XML to transmit structure information and metadata. For further information on RDF please have a look at <http://www.w3.org/RDF/Overview.html>. or for links at <http://www.cs.ukc.ac.uk/people/staff/djb1/research/metadata/rdf.shtml>.
- Dublin Core (DC) utilise a formal standard for the description of metadata elements. The formalisation helps to improve consistency with other metadata communities and enhances the clarity, scope, and internal consistency of the Dublin Core. For further information about DC have a look at <http://dublincore.org/documents/dces/>.
- Metadata Encoding & Transmission Standard (METS) schema is a standard for encoding descriptive, administrative, and structural metadata regarding objects within a digital library. For further information about METS have a look at <http://www.loc.gov/standards/mets/>.
- dig35 is a definition of a set of public metadata for digital still images. These images may encompass a single image or a collection of images that could be supported and exchanged by current and future devices, software, and services in an open environment. For further information about METS have a look at <http://members.digitalimaging.org/shareddocs/downloads/dig35v1.0-Sept00.pdf>.

2. Dublin Core

Dublin Core is a metadata element set intended to facilitate discovery of electronic resources. Originally conceived for author-generated description of Web resources, it has attracted the attention of formal resource description communities such as museums, libraries, government agencies, and commercial organisations.

The characteristics of the Dublin Core that distinguish it as a prominent candidate for description of electronic resources fall into several categories:

- **Simplicity**
The Dublin Core is intended to be usable by non-cataloguers as well as resource description specialists. Most of the elements have a commonly understood semantics of roughly the complexity of a library catalogue card.

- **Semantic Interoperability**

In the Internet Commons, disparate description models interfere with the ability to search across discipline boundaries. Promoting a commonly understood set of descriptors that helps to unify other data content standards increases the possibility of semantic interoperability across disciplines.

- **International Consensus**

Recognition of the international scope of resource discovery on the Web is critical to the development of effective discovery infrastructure. The Dublin Core benefits from active participation and promotion in some 20 countries in North America, Europe, Australia, and Asia.

- **Extensibility**

The Dublin Core provides an economical alternative to more elaborate description models such as the full MARC cataloguing of the library world. Additionally, it includes sufficient flexibility and extensibility to encode the structure and more elaborate semantics inherent in richer description standards

- **Metadata Modularity on the Web**

The diversity of metadata needs on the Web requires an infrastructure that supports the coexistence of complementary, independently maintained metadata packages. The World Wide Web Consortium (W3C) has begun implementing an architecture for metadata for the Web. The Resource Description Framework, or RDF, is designed to support the many different metadata needs of vendors and information providers.

For further information see the Dublin Core Home Page [1].

Targets of the DC:

- to provide a small set of Elements and keep this set as small as possible.
- to be syntax independent
- all elements are optional
- all elements are repeatable
- each element can be extended by an optional qualifier

About disadvantages of the Dublin Core and how to get around them, please read [2].

3. RDF

The *Resource Description Framework* (RDF) [3] is a foundation for processing metadata; it provides interoperability between applications that exchange machine-understandable information on the Web. RDF emphasises facilities to enable automated processing of Web resources. RDF can be used in a variety of application areas; for example: in resource discovery to provide better search engine capabilities, in cataloguing for describing the content and content relationships available at a particular Web site, page, or digital library, by intelligent software agents to facilitate knowledge sharing and exchange, in content rating, in describing collections of pages that represent a single logical "document", for describing intellectual property rights of Web pages, and for expressing the privacy preferences of a user

as well as the privacy policies of a Web site. RDF with digital signatures will be key to building the "Web of Trust" for electronic commerce, collaboration, and other applications.

One of the goals of RDF is to make it possible to specify semantics for data based on XML in a standardised, interoperable manner. RDF and XML are complementary: RDF is a model of metadata and only addresses by reference many of the encoding issues that transportation and file storage require (such as internationalisation, character sets, etc.). For these issues, RDF relies on the support of XML. It is also important to understand that this XML syntax is only one possible syntax for RDF and that alternate ways to represent the same RDF data model may emerge.

The broad goal of RDF is to define a mechanism for describing resources that makes no assumptions about a particular application domain, nor defines (a priori) the semantics of any application domain. The definition of the mechanism should be domain neutral, yet the mechanism should be suitable for describing information about any domain.

This specification will be followed by other documents that will complete the framework. Most importantly, to facilitate the definition of metadata RDF will have a class system much like many object-oriented programming and modelling systems. A collection of classes (typically authored for a specific purpose or domain) is called a schema. Classes are organised in a hierarchy, and offer extensibility through subclass refinement. This way, in order to create a schema slightly different from an existing one it is not necessary to "reinvent the wheel" but one can just provide incremental modifications to the base schema. Through the sharability of schemas RDF will support the reusability of metadata definitions. Due to RDF's incremental extensibility, agents processing metadata will be able to trace the origins of schemata they are unfamiliar with back to known schemata and perform meaningful actions on metadata they weren't originally designed to process. The sharability and extensibility of RDF also allows metadata authors to use multiple inheritance to "mix" definitions, to provide multiple views to their data, leveraging work done by others. In addition, it is possible to create RDF instance data based on multiple schemata from multiple sources (i.e., "interleaving" different types of metadata). Schemas may themselves be written in RDF; a companion document to this specification, describes one set of properties and classes for describing RDF schemas.

The foundation of RDF is a model for representing named properties and property values. The RDF model draws on well-established principles from various data representation communities. RDF properties may be thought of as attributes of resources and in this sense correspond to traditional attribute-value pairs. RDF properties also represent relationships between resources and a RDF model can therefore resemble an entity-relationship diagram. (More precisely, RDF Schemas — which are themselves instances of RDF data models — are ER diagrams.) In object-oriented design terminology, resources correspond to objects and properties correspond to instance variables.

The RDF data model is a syntax-neutral way of representing RDF expressions. The data model representation is used to evaluate equivalence in meaning. Two RDF expressions are equivalent if and only if their data model representations are the same. This definition of equivalence permits some syntactic variation in expression without altering the meaning.

3.1 Example of RDF and DC

```
<?xml version="1.0" encoding="iso-8859-1"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:dc="http://purl.org/dc/elements/1.0/"
xmlns:dcq="http://purl.org/dc/qualifiers/1.0"
xmlns:METAE="http://metae.aib.uni-linz.ac.at/2001/syntax">

  <rdf:Description>
    <dc:identifier>
      <rdf:Description>
        <rdf:value>urn:issn:1432-1823</rdf:value>
        <dcq:identifierType>
          <METAE:URN/>
        </dcq:identifierType>
        <dcq:identifierScheme
rdf:resource="http://purl.org/dc/qualifiers/1.0/identifier/scheme/ISSN/">
          </dcq:identifierScheme>

      </rdf:Description>
    </dc:identifier>
    <!-- this is the main title -->
    <dc:title>
      <rdf:Description>
        <rdf:value>Mathematische Zeitschrift</rdf:value>
        <dcq:titleType>
          <METAE:MainTitle/>
        </dcq:titleType>
      </rdf:Description>
    </dc:title>
    <dc:contributor>
      <rdf:Description>
        <rdf:value>Lichtenstein, L.</rdf:value>
        <dcq:contributorType>
          <METAE:Editor/>
        </dcq:contributorType>
      </rdf:Description>
    </dc:contributor>
    <dc:publisher>
      <rdf:Description>
        <rdf:value>Springer</rdf:value>
      </rdf:Description>
    </dc:publisher>
    <dc:subject>
      <rdf:Description>
        <rdf:value>510</rdf:value>
        <!-- 510 is the DDC class number for mathematics. -->
        <dcq:subjectScheme>DDC</dcq:subjectScheme>
      </rdf:Description>
    </dc:subject>
  </rdf:Description>
</rdf:RDF>
```


4. dig35 Metadata Schema

The dig35 specification defines a set of public metadata for digital still images. In addition to the metadata definition, a recommendation for implementation is defined to enable exchange of such metadata. One big advantage of dig35 is its Image File Format Independence, it does not rely on any specific file format and can therefore support many file formats and compression mechanisms. Another advantage is the allowance of metadata redundancy. If certain values exist that can be calculated from other fields, redundant metadata do exist and need to be managed. Where this is applicable the specification makes recommendations how to manage such redundancy.

DIG35's Metadata definition consists of five logical sections with a separate common definition that is referred to by other sections. While each section is logically partitioned, they may be linked to each other to form additional semantics.

These 5 sections are:

- Basic Image Parameter
- Image Creation
- Content Description
- History
- IPR (Intellectual Property Rights Metadata) & Exploitation

4.1 Basic Image Parameter Metadata

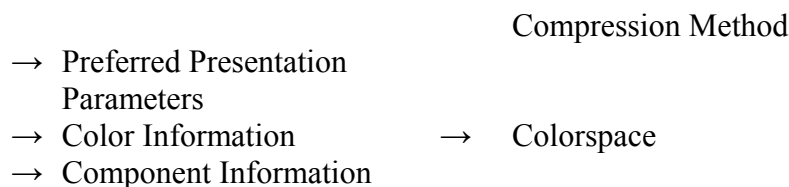
This section defines basic image parameter metadata that contains generic information about the image, such as the image size and number of components. These metadata are commonly referred as the "Header Data" of an image file and are file format independent. These metadata should be considered informational and not be used to decode the image data stored in the associated file.

This section may contain information similar or identical in use to the file header metadata. The dig35 manual [4] says: "There should never be conflicts between this block and the file header metadata as this block is intended to be used, as stated above, only when there is no file header metadata. However, if there is a conflict between the file-format header information and the Basic Image Parameter Metadata, the file header should always take precedence."

We do not agree with that in the METAe project, because the metadata in this section are mainly extracted from the TIFF files we scan. We found out that sometimes the operators who manage the scanning are sleazy entering new metadata when scanning a new book, so that the metadata in the XML Files gets corrected and are more precise than the one in the TIFF header. So the metadata in the XML file is preferred except the Timestamp in the TIFF file is newer than the one in the XML file. (But this indicates that the image has been transformed or is newly created)

Layout

Basic Image Parameter → Basic Image Information → File and Format
Image Identifier
Image Size

**Example:**

```

<dig35:METADATA>
  <dig35:BASIC_IMAGE_PARAM>
    <dig35:BASIC_IMAGE_INFO>
      <dig35:FILE_FORMAT>
        <dig35:FILE_NAME>00000003.tif</dig35:FILE_NAME>
        <dig35:FILE_FORMAT>TIFF</dig35:FILE_FORMAT>
      </dig35:FILE_FORMAT>
      <dig35:IMAGE_SIZE>
        <dig35:WIDTH>2400</dig35:WIDTH>
        <dig35:HEIGHT>3768</dig35:HEIGHT>
      </dig35:IMAGE_SIZE>
      <dig35:COMPRESSION>CCITT G4</dig35:COMPRESSION>
    </dig35:BASIC_IMAGE_INFO>
  </dig35:BASIC_IMAGE_PARAM>
</dig35:METADATA>
  
```

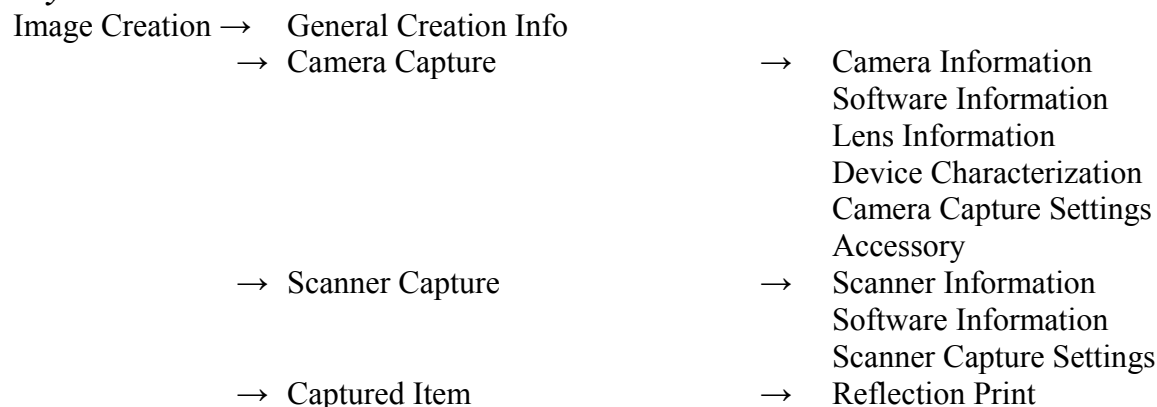
Note:

Annex A of the dig35 Specification suggests only 6 compression methods. Since we are using TIFF files b/w, CCITT G4 is the most used compression method. So we extend the Table with the following values:

Values	Meaning
CCITT 1D	CCITT 1D compression used
CCITT G3	CCITT G3 compression used
CCITT G4	CCITT G4 compression used

4.2 Image Creation Metadata

The Image Creation Metadata defines the "how" metadata that specifies the source of which the image was created. For example, the camera and lens information and capture condition are useful technical information for professional and serious amateur photographers as well as advanced imaging applications. In general, creation metadata is considered to be "read-only" being persistent and associated with the image through the workflow.

Layout

Film

Example:

```

- <amdSec ID="TMD_CREA0">
  - <techMD ID="AMD_CREA0_1">
    - <mdWrap MDTYPE="OTHER">
      - <dig35:METADATA>
        - <dig35:IMAGE_CREATION>
          - <dig35:GENERAL_CREATION_INFO>
            <dig35:CREATION_TIME>2000:04:16 13:05:01</dig35:CREATION_TIME>
            <dig35:IMAGE_SOURCE>Reflection Print Scanner</dig35:IMAGE_SOURCE>
          - <OPERATOR_ORG>
            <ORG_NAME>Centrale Bibliotheek, Katholieke Universiteit
              Leuven</ORG_NAME>
            </OPERATOR_ORG>
          - <dig35:IMAGE_CREATOR>
            - <dig35:PERSON_NAME>
              <dig35:NAME_COMP TYPE="Family">Kulmer</dig35:NAME_COMP>
              </dig35:PERSON_NAME>
            </dig35:IMAGE_CREATOR>
          </dig35:GENERAL_CREATION_INFO>
        - <dig35:SCANNER_CAPTURE>
          - <dig35:SCANNER_INFO>
            - <dig35:MANUFACTURER>
              <dig35:ORG_NAME>Minolta</dig35:ORG_NAME>
              </dig35:MANUFACTURER>
              <dig35:MODEL>PS7000</dig35:MODEL>
              <dig35:SERIAL>1</dig35:SERIAL>
              <dig35:VERSION>2</dig35:VERSION>
            </dig35:SCANNER_INFO>
          - <dig35:SOFTWARE_INFO>
            - <dig35:MANUFACTURER>
              <dig35:ORG_NAME>SRZ</dig35:ORG_NAME>
              </dig35:MANUFACTURER>
              <dig35:MODEL>ProScan Book</dig35:MODEL>
              <dig35:SERIAL>3</dig35:SERIAL>
              <dig35:VERSION>4</dig35:VERSION>
            </dig35:SOFTWARE_INFO>
          - <dig35:SCANNER_SETTINGS>
            <dig35:PIXEL_SIZE />
            - <dig35:PHYSICAL_SCAN_RES>
              <dig35:WIDTH>600</dig35:WIDTH>
              <dig35:HEIGHT>600</dig35:HEIGHT>
            </dig35:PHYSICAL_SCAN_RES>
            </dig35:SCANNER_SETTINGS>
          </dig35:SCANNER_CAPTURE>
        </dig35:IMAGE_CREATION>
      </dig35:METADATA>
    </mdWrap>
  </techMD>
</amdSec>

```

Note:

Annex B of the dig35 Specification suggests only 5 different Image sources. When digitising books the most used Scanner is a book scanner, not listed. So we extend the Table with the following values:

Values	Meaning
Flatbed Scanner	Image created by a flatbed scanner
Book Scanner	Image created by a book scanner

For the scene type value we suggest to use “Second Generation Scene” because our images are captured from an existing hardcopy image.

4.3 Content Description Metadata

The Content Description Metadata defines the descriptive information of "who", "what", "when" and "where" aspect of the image. Often this metadata takes the form of extensive words, phrases, or sentences to describe a particular event or location that the image illustrates. Typically, this metadata consists of text that the user enters, either when the images are taken or scanned or later in the process during manipulation or use of the images. Since Content Description is only good for photographs we do not support this section.

4.5 History Metadata

The History Metadata is used to provide partial information about how the image got to the present state. This metadata is only approximate because;

- some of the data (both metadata and image data) is collapsed, thus providing only a summary
- some of the data may not have been properly entered because applications used were not capable of updating the history metadata.

The History metadata contains a summary of basic image editing operations that have already been applied to the image and previous version of the image metadata. The History metadata is not designed to be used to reverse image editing operations.

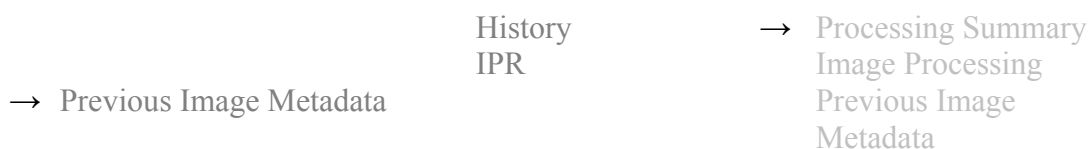
To permit flexibility in construction of the image history metadata, two alternate representations of the history are permitted. In the first, the history metadata is embedded in the image metadata. In the second, the previous versions of the image are included in the history metadata as pointers to the location of the actual history. While this specification does not define the "how" or "how much" part of the processing aspect, it does enable logging of certain processing steps by adding textual descriptions of operations applied to an image as hints for future use.

There are two basic problems for the METAe project that it will not use the History Metadata Section. The first thing is that these Metadata have to be applied when the images are scanned. But at the moment there exists no software that supports this step. Because of the amount of the images the history metadata can not be applied manually. The second problem is that these metadata are not precise. It is just specified what was done, but not how. Some metadata are entered in a textual way, like "Image rotated by 10 degrees" which depends on a language and on the person who writes the text.

The METAe software will not use this section if it is applied, it will ignore it and leave it untouched.

Layout

History → Processing Summary
→ Image Processing Hints
→ Previous History Metadata → Basic Image Parameters
Image Creation
Content Description



The following assumption and definitions govern the proposed logical structure for Change History:

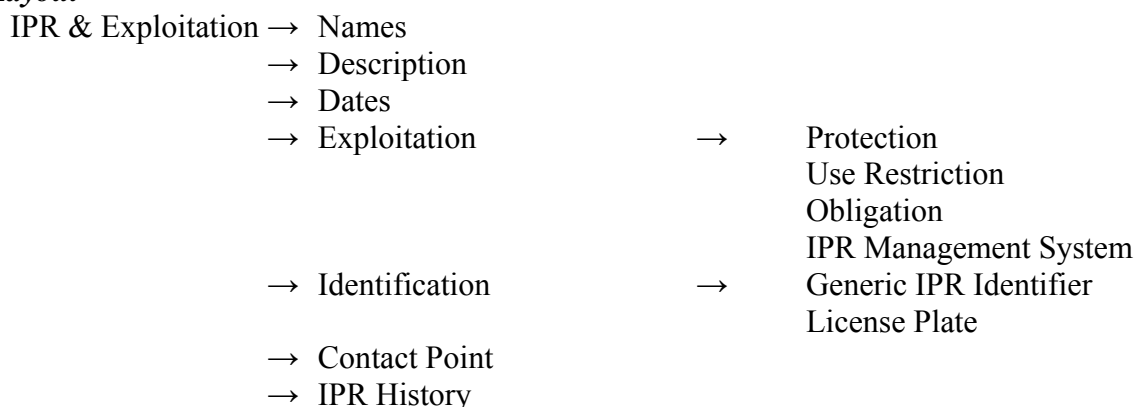
- image processing may occur multiple times throughout the life cycle of an image
- the image life cycle may consist of multiple generations of the image
- changes to any of the values in Format create a new generation of the image
- image transformation refers to any processing that produces a new generation image
- all other processes (i.e., those that do not create new values) are classified as image editing

4.6 Intellectual Property Rights (IPR) Metadata

The Intellectual Property Rights Metadata (IPR) defines metadata to either protect the rights of the owner of the image or provide further information to request permission to use it. IPR metadata are designed to protect the contents of an image file against misuse and must preserve both moral rights and copyrights. Moral rights are those rights attached to the creation process; therefore, moral rights persistently pertain to the author or creator of the artwork, whereas copyrights can be repeatedly transferred to different owners, under exploitation conditions that are also part of the IPR and exploitation metadata.

Because of international agreements (e.g., the WIPO Treaty), this particular part of the metadata set is subject to specific conditions in terms of persistence. According to the WIPO Treaty, it is strictly forbidden to "remove" or "alter" IPR metadata that has been inserted inside a file; therefore, the only allowable operation is the addition of IPR metadata. The IPR History metadata stores all IPR metadata-related modifications. IPR modifications may also be stored as part of the Image History metadata.

Layout



Example:

```

<IPR>
  <IPR_NAMES>
    <IPR_PERSON ID="1" DESCRIPTION="Original Work Author">
      <PERSON_NAME>
        <NAME_COMP TYPE="Given">Marco</NAME_COMP>
        <NAME_COMP TYPE="Family">Koettstorfer</NAME_COMP>
      </PERSON_NAME>
    </IPR_PERSON>
  </IPR_NAMES>
</IPR>
  
```

```

</IPR_PERSON>
= <IPR_PERSON ID="2" DESCRIPTION="Image Creator">
  = <PERSON_NAME>
    <NAME_COMP TYPE="Given">John</NAME_COMP>
    <NAME_COMP TYPE="Family">Kulmer</NAME_COMP>
    <COMMENT />
  </PERSON_NAME>
</IPR_PERSON>
= <IPR_ORG ID="3" DESCRIPTION="Right Holder">
  <ORG_NAME>Uni Graz</ORG_NAME>
  <WEB>www.uni-graz.at</WEB>
</IPR_ORG>
</IPR_NAMES>
= <IPR_DESCRIPTION>
  <COPYRIGHT>Copyright 2001 UBG</COPYRIGHT>
</IPR_DESCRIPTION>
= <IPR_DATE DESCRIPTION="Scanned">
  <DATE>2001-06-02</DATE>
</IPR_DATE>
= <IPR_DATE DESCRIPTION="Original Work Created">
  <DATE>2000-06-01</DATE>
</IPR_DATE>
= <IPR_EXPLOITATION>
  <IPR_USE_RESTRICTION>No copy</IPR_USE_RESTRICTION>
  <IPR_USE_RESTRICTION>No print</IPR_USE_RESTRICTION>
</IPR_EXPLOITATION>
  <IPR_OBLIGATION>Copyright mention mandatory</IPR_OBLIGATION>
</IPR>

```

5. Metadata Encoding and Transmission Standard (METS) Metadata Schema

METS, a Digital Library Federation initiative, attempts to build upon the work of MOA2 and provide an XML document format for encoding metadata necessary for both management of digital library objects within a repository and exchange of such objects between repositories. In Metae the METS schema is used as the outer wrapper to transport all metadata.

A METS document consists of four major sections:

- **Descriptive Metadata**
- **Administrative Metadata**
- **File Groups**
- **Structural Map**

5.1 Descriptive Metadata

The descriptive metadata section may point to descriptive metadata external to the METS document (e.g., a MARC record in an OPAC ...) or contain internally embedded descriptive metadata, or both. The descriptive metadata section of a METS document consists of one or more <dmdSec> (Descriptive Metadata Section) elements

Metae only uses the Internal Descriptive Metadata with is wrapped by the <mdWrap> element. An mdWrap element provides a wrapper around metadata embedded within a METS document. In it DC is used to describe the data.

Example:

```
<dmdSec ID="DCMD1">
  <mdWrap MIMETYPE="text/xml" MDTYPE="DC" LABEL="Bibliographic meta-data">
    <rdf:RDF>
      <rdf:Description>
        <dc:title>
          <rdf:Description>
            <rdf:value>IRIS</rdf:value>
          </rdf:Description>
        </dc:title>
        <dc:creator>
          <rdf:Description>
            <rdf:value>Koettstorfer</rdf:value>
          </rdf:Description>
        </dc:creator>
        <dc:subject>
          <rdf:Description>
            <rdf:value>0862</rdf:value>
          </rdf:Description>
        </dc:subject>
        <dc:identifier>
          <rdf:Description>
            <rdf:value>URN:ISSN:1234-5679</rdf:value>
          </rdf:Description>
        </dc:identifier>
        <dc:language>
          <rdf:Description>
            <rdf:value>DE</rdf:value>
          </rdf:Description>
        </dc:language>
      </rdf:Description>
    </rdf:RDF>
  </mdWrap>
</dmdSec>
```

```

        </rdf:Description>
      </dc:language>
    </rdf:Description>
  </rdf:RDF>
</mdWrap>
</dmdSec>
< dmdSec ID="DCMD2">
  < mdWrap MIMETYPE="text/xml" MDTYPE="DC" LABEL="Electronic meta-data">
    < rdf:RDF>
      < rdf:Description>
        < dc:title>
          < rdf:Description>
            < rdf:value>IRIS</rdf:value>
          </rdf:Description>
        </dc:title>
        < dc:subject>
          < rdf:Description>
            < rdf:value>0862</rdf:value>
          </rdf:Description>
        </dc:subject>
        < dc:language>
          < rdf:Description>
            < rdf:value>DE</rdf:value>
          </rdf:Description>
        </dc:language>
      </rdf:Description>
    </rdf:RDF>
  </mdWrap>
</dmdSec>

```

It is important that all <dmdSec> elements must possess an ID attribute. This attribute provides a unique, internal name for each <dmdSec> element which can be used in the structural map to link a particular division of the document hierarchy to a particular <dmdSec> element. This allows specific sections of descriptive metadata to be linked to specific parts of the digital object.

5.2 Administrative Metadata

The administrative metadata section provides information regarding how the files were created and stored, intellectual property rights, metadata regarding the original source object from which the digital library object derives. As with descriptive metadata, administrative metadata may be either external to the METS document, or encoded internally.

The <amdSec> elements contain the administrative metadata.

There are four main forms of administrative metadata provided for in a METS document:

1. Technical Metadata <techMD> (information regarding files' creation, format, and use characteristics)
2. Intellectual Property Rights Metadata <rightsMD> (copyright and license information),
3. Source Metadata <sourceMD> (descriptive and administrative metadata regarding the analog source from which a digital library object derives)
4. Digital Provenance Metadata <digiprovMD> (information regarding source/destination relationships between files).

The <amdSec> element employs the same content model as <dmdSec>: it may contain an <mdRef> element to point to external administrative metadata, an <mdWrap> element to use when embedding administrative metadata within a METS document, or both.

Metae again uses only the internal metadata and these data are wrapped and described via the dig35 Metadata schema.

Example:

```

<amdSec ID="TMD1">
  <techMD ID="TMD1_1">
    <mdWrap MDTYPE="OTHER" OTHERMDTYPE="dig35">
      <dig35:METADATA>
        <dig35:BASIC_IMAGE_PARAM>
          <dig35:BASIC_IMAGE_INFO>
            <dig35:FILE_FORMAT>

              <dig35:FILE_NAME>00000003.tif</dig35:FILE_
                NAME>
              <dig35:FILE_FORMAT>TIFF</dig35:FILE_FORMAT>
            </dig35:FILE_FORMAT>
          <dig35:IMAGE_SIZE>
            <dig35:WIDTH>2400</dig35:WIDTH>
            <dig35:HEIGHT>3768</dig35:HEIGHT>
          </dig35:IMAGE_SIZE>
          <dig35:COMPRESSION>CCITT G4</dig35:COMPRESSION>
        </dig35:BASIC_IMAGE_INFO>
      </dig35:BASIC_IMAGE_PARAM>
    </dig35:METADATA>
  </mdWrap>
</techMD>
</amdSec>

```

5.3 File Groups

The file group section lists all files comprising all electronic versions of the digital object. File group elements may nest, to provide for subdividing the files by object version.

The file group section consists of one or more <fileGrp> elements used to group together related files. A <fileGrp> lists all of the files which comprise a single electronic version of the digital library object.

Example:

```

<fileGrp ID="ORIGIMG">
  <file ID="ORIGIMG1" CREATED="2000:04:16 13:05:19" MIMETYPE="image/tif"
    SEQ="1" ADMID="TMD1 TMD_CREA1">
    <FLocat>./00000003.tif</FLocat>
  </file>
  <file ID="ORIGIMG2" CREATED="2000:04:16 13:05:19" MIMETYPE="image/tif"
    SEQ="2" ADMID="TMD2 TMD_CREA2">
    <FLocat>./00000002.tif</FLocat>
  </file>
  <file ID="ORIGIMG3" CREATED="2000:04:16 13:05:01" MIMETYPE="image/tif"
    SEQ="3" ADMID="TMD3 TMD_CREA3">
    <FLocat>./00000001.tif</FLocat>
  </file>
</fileGrp>

```

Note that all <file> elements must have a unique ID attribute. This attribute provides a unique, internal name for this file which can be referenced by other portions of the document.

5.4 Structural Map

The structural map is the heart of a METS document. It outlines a hierarchical structure for the digital library object, and links the elements of that structure to content files and metadata that pertain to each element.

The <structMap> element encodes this hierarchy as a nested series of <div> elements. Each <div> carries attribute information specifying what kind of division it is. It may contain multiple file pointer (<fptr>) elements. File pointers specify files (or locations within files) within the current METS document's <fileGrp> section that correspond to the portion in the hierarchy represented by the current <div>.

In Metae the structMap is used to bind together the textual pages with the scanned images

Example:

```
<structMap LABEL="Physical structure" TYPE="PHYSICAL">
  <div DMD="DCMD1" LABEL="">
    <div ORDER="1" TYPE="PAGE">
      <fptr>
        <par>
          <area FILEID="ORIGIMIG1" />
          <area FILEID="XMLTEXT1" BETYPE="IDREF" BEGIN="PAGE1" />
        </par>
      </fptr>
    </div>
    <div ORDER="2" TYPE="PAGE">
      <fptr>
        <par>
          <area FILEID="ORIGIMIG2" />
          <area FILEID="XMLTEXT2" BETYPE="IDREF" BEGIN="PAGE2" />
        </par>
      </fptr>
    </div>
    <div ORDER="3" TYPE="PAGE">
      <fptr>
        <par>
          <area FILEID="ORIGIMIG3" />
          <area FILEID="XMLTEXT3" BETYPE="IDREF" BEGIN="PAGE3" />
        </par>
      </fptr>
    </div>
  </div>
</structMap>
```

Complete Example:

Find here the code of an complete example as it is used in the Metae project.

```
<?xml version="1.0" encoding="UTF-8" ?>
<mets xmlns:xsi="http://www.w3.org/2000/10/XMLSchema-instance"
  xmlns="http://www.loc.gov/METS/"
  xsi:schemaLocation="http://www.loc.gov/standards/mets/mets.xsd"
  xmlns:dc="http://purl.org/dc/elements/1.0/"
  xmlns:dcq="http://purl.org/dc/qualifiers/1.0/"
  xmlns:dig35="http://www.digitalimaging.org/dig35/1.0/xml"
  xmlns:rdp="http://www.w3.org/1999/02/22-rdf-syntax-ns#" TYPE="" LABEL=""
  CREATEDATE="14.09.2001 13:10:07">
  <dmdSec ID="DCMD1">
    <mdWrap MIMETYPE="text/xml" MDTYPE="DC" LABEL="Bibliographic meta-data">
      <rdp:RDF>
        <rdp:Description />
      </rdp:RDF>
    </mdWrap>
  </dmdSec>
  <dmdSec ID="DCMD2">
    <mdWrap MIMETYPE="text/xml" MDTYPE="DC" LABEL="Electronic meta-data">
      <rdp:RDF>
        <rdp:Description />
      </rdp:RDF>
    </mdWrap>
  </dmdSec>
  <amdSec ID="TMD1">
```

```

<techMD ID="TMD1_1">
  <mdWrap MDTYPE="OTHER" OTHERMDTYPE="dig35">
    <dig35:METADATA>
      <dig35:BASIC_IMAGE_PARAM>
        <dig35:BASIC_IMAGE_INFO>
          <dig35:FILE_FORMAT>

            <dig35:FILE_NAME>00000003.tif</dig35:FILE_
              NAME>
            <dig35:FILE_FORMAT>TIFF</dig35:FILE_FORMAT>
          </dig35:FILE_FORMAT>
        <dig35:IMAGE_SIZE>
          <dig35:WIDTH>2400</dig35:WIDTH>
          <dig35:HEIGHT>3768</dig35:HEIGHT>
        </dig35:IMAGE_SIZE>
        <dig35:COMPRESSION>CCITT G4</dig35:COMPRESSION>
      </dig35:BASIC_IMAGE_INFO>
    </dig35:BASIC_IMAGE_PARAM>
  </dig35:METADATA>
</mdWrap>
</techMD>
</amdSec>
<amdSec ID="TMD2">
  <techMD ID="TMD2_1">
    <mdWrap MDTYPE="OTHER" OTHERMDTYPE="dig35">
      <dig35:METADATA>
        <dig35:BASIC_IMAGE_PARAM>
          <dig35:BASIC_IMAGE_INFO>
            <dig35:FILE_FORMAT>

              <dig35:FILE_NAME>00000002.tif</dig35:FILE_
                NAME>
              <dig35:FILE_FORMAT>TIFF</dig35:FILE_FORMAT>
            </dig35:FILE_FORMAT>
          <dig35:IMAGE_SIZE>
            <dig35:WIDTH>2173</dig35:WIDTH>
            <dig35:HEIGHT>3614</dig35:HEIGHT>
          </dig35:IMAGE_SIZE>
          <dig35:COMPRESSION>CCITT G4</dig35:COMPRESSION>
        </dig35:BASIC_IMAGE_INFO>
      </dig35:BASIC_IMAGE_PARAM>
    </dig35:METADATA>
  </mdWrap>
</techMD>
</amdSec>
<amdSec ID="TMD3">
<amdSec ID="TMD_CREAO">
  <techMD ID="AMD_CREAO_1">
    <mdWrap MDTYPE="OTHER" OTHERMDTYPE="dig35">
      <dig35:METADATA>
        <dig35:IMAGE_CREATION>
          <dig35:GENERAL_CREATION_INFO>
            <dig35:CREATION_TIME>2000:04:16
              13:05:19</dig35:CREATION_TIME>
            <dig35:IMAGE_SOURCE>Reflection Print
              Scanner</dig35:IMAGE_SOURCE>
          <dig35:OPERATOR_ORG>
            <dig35:ORG_NAME>Centrale Bibliotheek, Katholieke
              Universiteit Leuven</dig35:ORG_NAME>
          </dig35:OPERATOR_ORG>
          <dig35:IMAGE_CREATOR>
            <dig35:PERSON_NAME>
              <dig35:NAME_COMP
                TYPE="Family">Kulmer</dig35:NAME_CO
                MP>
            </dig35:PERSON_NAME>
          </dig35:IMAGE_CREATOR>
        </dig35:GENERAL_CREATION_INFO>
      <dig35:SCANNER_CAPTURE>
        <dig35:SCANNER_INFO>
          <dig35:MANUFACTURER>

```

```

        <dig35:ORG_NAME>Minolta</dig35:ORG_
        NAME>
    </dig35:MANUFACTURER>
    <dig35:MODEL>PS7000</dig35:MODEL>
    <dig35:SERIAL />
    <dig35:VERSION />
</dig35:SCANNER_INFO>
± <dig35:SOFTWARE_INFO>
    <dig35:MANUFACTURER />
    <dig35:MODEL>SRZ Proscan
    V2.0.15</dig35:MODEL>
</dig35:SOFTWARE_INFO>
± <dig35:SCANNER_SETTINGS>
    <dig35:PIXEL_SIZE />
    ± <dig35:PHYSICAL_SCAN_RES>
        <dig35:WIDTH />
        <dig35:HEIGHT />
    </dig35:PHYSICAL_SCAN_RES>
</dig35:SCANNER_SETTINGS>
</dig35:SCANNER_CAPTURE>
</dig35:IMAGE_CREATION>
</dig35:METADATA>
</mdWrap>
</techMD>
</amdSec>
± <amdSec ID="TMD_CREA1">
± <amdSec ID="TMD_CREA2">
± <fileGrp ID="ORIGIMG">
    ± <file ID="ORIGIMG1" CREATED="2000:04:16 13:05:19" MIMETYPE="image/tif"
        SEQ="1" ADMID="TMD1 TMD_CREA1">
        <FLocat>./00000003.tif</FLocat>
    </file>
    ± <file ID="ORIGIMG2" CREATED="2000:04:16 13:05:19" MIMETYPE="image/tif"
        SEQ="2" ADMID="TMD2 TMD_CREA2">
        <FLocat>./00000002.tif</FLocat>
    </file>
    ± <file ID="ORIGIMG3" CREATED="2000:04:16 13:05:01" MIMETYPE="image/tif"
        SEQ="3" ADMID="TMD3 TMD_CREA3">
        <FLocat>./00000001.tif</FLocat>
    </file>
</fileGrp>
± <structMap LABEL="Physical structure" TYPE="PHYSICAL">
    ± <div DMD="DCMD1" LABEL="">
        ± <div ORDER="1" TYPE="PAGE">
            ± <fptr>
                ± <par>
                    <area FILEID="ORIGIMG1" />
                    <area FILEID="XMLTEXT1" BETYPE="IDREF" BEGIN="PAGE1"
                    />
                </par>
            </fptr>
        </div>
        ± <div ORDER="2" TYPE="PAGE">
            ± <fptr>
                ± <par>
                    <area FILEID="ORIGIMG2" />
                    <area FILEID="XMLTEXT2" BETYPE="IDREF" BEGIN="PAGE2"
                    />
                </par>
            </fptr>
        </div>
        ± <div ORDER="3" TYPE="PAGE">
            ± <fptr>
                ± <par>
                    <area FILEID="ORIGIMG3" />
                    <area FILEID="XMLTEXT3" BETYPE="IDREF" BEGIN="PAGE3"
                    />
                </par>
            </fptr>
        </div>
    </div>
</structMap>

```

</mets>

6. Metadata in the TIFF file

When dealing with images the most used file format is tiff because it can, as its name “Tagged Interchange File Format” says, store metadata in its header. Most of these parameter are also stored in the dig35 Metadata Schema (Basic Image Parameters)

Compression

Designates the compression scheme used to store the image data.

Values 1 = Uncompressed
 2 = CCITT 1D
 3 = CCITT Group 3
 4 = CCITT Group 4
 5 = LZW
 6 = JPEG
 32773 = PackBits (simple byte-oriented run-length scheme)

PhotometricInterpretation

Designates the colour space of the decompressed image data.

Values 0 = WhiteIsZero. For bilevel and grayscale images: 0 is imaged as white.
 1 = BlackIsZero. For bilevel and grayscale images: 0 is imaged as black.
 2 = RGB
 3 = Palette color
 4 = Transparency Mask
 5 = CMYK
 6 = YCbCr
 8 = CIELab

SegmentType

Specifies whether image data is stored in strips or tiles.

Values 0 = strips
 1 = tiles

StripOffsets

For each strip, the byte offset of that strip.

RowsPerStrip

The number of rows per strip.

StripByteCounts

The number of image data bytes stored within each strip after compression.

TileWidth

The tile width in pixels. This is the number of columns in each tile.

TileLength

The tile length (height) in pixels. This is the number of rows in each tile.

TileOffsets

For each tile, the byte offset of that tile, as compressed and stored on disk.

TileByteCounts

For each tile, the byte offset of that tile, as compressed and stored on disk.

Values N = TilesPerImage for PlanarConfiguration = 1
 = SamplesPerPixel * TilesPerImage for PlanarConfiguration = 2

PlanarConfiguration

Designates how the components of each pixel are stored.

Values 1 = chunky format
 2 = planar format

Orientation

Designates the orientation of the image, with respect to the placement of its rows (ImageWidth) and columns (ImageLength), as it was saved to disk.

Values 1 = normal¹
 3 = normal rotated 180°
 6 = normal rotated cw 90°
 8 = normal rotated ccw 90°
 9 = unknown

HostComputer

Computer and/or operating system in use at the time of image creation.

ScanningSystem

The scanner manufacturer and model name or number, and the name and version of the capture software used to create the image.

DateTimeCreated

Date or DateTime image was created.

ImageWidth

This specifies the width of the digital image, i.e. horizontal or X dimension, in pixels.

ImageLength

This specifies the length of the digital image, i.e. vertical or Y dimension, in pixels.

XSamplingFrequency

¹ "normal" is defined as follows: when opened, the top (0th) row of pixels corresponds to the visual top of the image, and the first (0th) column of pixels on left corresponds to the visual left-hand side of the image.

This field specifies the number of pixels per SamplingFrequencyUnit in the image width.

YSamplingFrequency

This field specifies the number of pixels per SamplingFrequencyUnit in the image length.

SamplingFrequencyUnit

The unit of measurement for XSamplingFrequency and YSamplingFrequency.

Values 1 = no absolute unit of measurement.
 2 = inch
 3 = centimeter

BitsPerSample

The number of bits per component for each pixel. This field provides N values depending upon SamplesPerPixel present.

Values 1 = 1-bit (bitonal)
 4 = 4-bit grayscale
 8 = 8-bit grayscale or palletizedcolor
 8,8,8 = RGB
 16,16,16 = TIFF, HDR (high dynamic range)
 8,8,8,8 = CMYK

SamplesPerPixel

Designates the number of colour components per pixel.

Values 1 = when PhotometricInterpretation = 0 or 1
 3 = when PhotometricInterpretation = 2 (RGB), 6 (YcbCr)
 4 = when PhotometricInterpretation = 5 (CMYK)

Extrasamples

Specifies that each pixel has M extra components whose interpretation is defined by one of the values listed below.

Values 0 = unspecified data
 1 = associated alpha data (with pre-multiplied colour)
 2 = unassociated alpha data
 3 = range or depth data

Colormap

This field defines a Red-Green-Blue colour map (often called a lookup table) for palette-colour images.

Values [n bit code value] [red value] [green value] [blue value]
 (Examples)
 0 0 4 5
 1 1 5 8
 2 3 7 10
 255 129 250 150

GrayResponseCurve

For grayscale data, the optical density of each possible pixel value.

Values $N = 2^{**} \text{BitsPerSample}$

GrayResponseUnit

The precision of the information contained in the GrayResponseCurve.

Values 1 = Number represents tenths of a unit.
 2 = Number represents hundredths of a unit.
 3 = Number represents thousandths of a unit.
 4 = Number represents ten-thousandths of a unit.
 5 = Number represents hundred-thousandths of a unit.

WhitePoint

The white point chromaticity of the effective illumination source of the capture process.

Values 3127/10000, 3290/10000, 1 = [source-to-intermediate capture]
 3127/10000, 3290/10000, 2 = [source-to-digital capture]

PrimaryChromaticities

The chromaticities of the primary colors of the imaging process

Values 640/1000, 330/1000, 300/1000, 600/1000, 150/1000, 60/1000, 1 = [source-to-intermediate capture]
 (Examples) 640/1000, 330/1000, 300/1000, 600/1000, 150/1000, 60/1000, 2 = [source-to-digital capture]

7. References

- [1] Dublin Core Homepage
<http://purl.org/dc/>
- [2] Beyond the Dublin Core: Rich Meta-Data and Convenience-of-Use Are Compatible After All, Roger Clarke
<http://www.anu.edu.au/people/Roger.Clarke/II/DublinCore.html>
- [3] Resource Description Framework (RDF) Model and Syntax Specification,
W3C Proposed Recommendation 05 January 1999
Ora Lassila (ora.lassila@research.nokia.com), Nokia Research Center
Ralph R. Swick (swick@w3.org), World Wide Web Consortium
<http://www.w3.org/TR/PR-rdf-syntax/>
- [4] DIG35 Specification, Metadata for Digital Images, Version 1.0, August 30,
2000 Digital Imaging Group, Inc
<http://www.digitalimaging.org>

Deliverable 1.2

Digital Repositories – Austria

Chapter 4

ALO Guide

Table of Contents

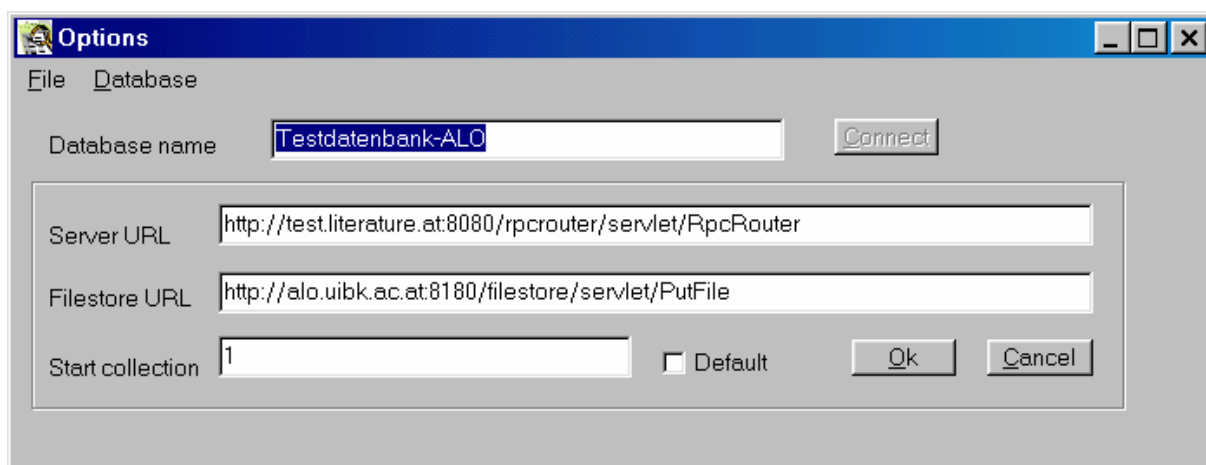
1. ALO CLIENT	3
1.1 Options	3
1.2. Sammlungen	3
1.3 Periodica	4
1.4 Mehrbändige Werke	4
1.5 Gesammelte Werke, Wissenschaftliche Editionen	5
2. METADATENEINGABE	5
2.1 Allgemeines	5
2.2 Grundsätzliches zur Titel und Autoraufnahme	6
2.3 Mehrbändige Werke	6
2.4 Dublin Core für Bücher und Zeitschriften	7
2.4.1 Allgemeines	7
2.4.2 DC: Title	7
2.4.3 DC: Creator	7
2.4.4 DC: Subject	8
2.4.5 DC: Description	8
2.4.6 DC: Publisher	8
2.4.7 DC: Contributor	8
2.4.8 DC: Date	8
2.4.9 DC: Type	8
2.4.10 DC: Format	8
2.4.11 DC: Identifier	9
2.4.12 DC: Source	9
2.4.13 DC: Language	9
2.4.14 DC: Relation	9
2.4.15 DC: Coverage	9
2.4.16 DC: Rights	9
2.5 Technische Metadaten	10
2.6 Kontrollblick in ALO	10
3. STANDARDAUFNAHME	15
3.1 Mehrbändige Werke	15
3.2 Zeitschriften	19

1. ALO Client

Mit dem ALO Client werden die Imagedaten und Metadaten auf den Server hochgeladen.

1.1 Options

- Das Hochladen erfolgt über das Internet, deshalb müssen folgende Daten eingegeben werden („Datenbank“):
- Produktionsdatenbank:
Server: `http://www.literature.at:8180/rpcrouter/servlet/RpcRouter`
Filestore: `http://alo.uibk.ac.at:8180/filestore/servlet/PutFile`
- Testdatenbank:
Server: `http://test.literature.at:8080/rpcrouter/servlet/RpcRouter`
Filestore: `http://alo.uibk.ac.at:8180/filestore/servlet/PutFile`
- Dieper-Server ...
Server URL: `http://dieper.aib.uni-linz.ac.at:8080/rpcrouter/servlet/RpcRouter`
Filestore URL: `http://alo.uibk.ac.at:8180/filestore/servlet/PutFile`



1.2. Sammlungen

- reUSE - Aufbau der Sammlungen
 - Jeder reUSE Node hat eigenes Skin
 - Oberste Ebene: Fokusgruppen
 - Darunter nach Institutionen
 - Innerhalb der Institutionen nach Publikationsarten und Wünschen der Partner

- Rein mechanische Kriterien zählen:
Austriaca sind:
 - Bücher aus Österreich (Verlag, Verlagsort, etc.)
 - Bücher von Österreichern (geboren, wohnhaft, sonstwie mit Österreichbezug)
 - Bücher über Österreich (auch fremdsprachige, etc.)Das gleiche gilt sinngemäß für Styriaca (Steiermark) und Tirolensien.
- Die Auswahl der richtigen Collection ist der schwierigste Bereich der Erfassung.
Achtung: Trotzdem nicht zuviel Zeit darauf verschwenden und in keine ideologischen Diskussionen einlassen.
- Am besten vorher prüfen, ob es sich um ein Styriaca, ein Austriaca, etc. handelt
- Folgende Collections werden derzeit unterstützt:
 - Austriaca
 - Tirolensien: Wie oben, aber auf Tirol bezogen
 - Styriaca: Wie oben, aber auf die Steiermark bezogen
 - Periodica: Unabhängig ob Tyrolensien, etc. werden hier alle Serien aufgenommen (Zeitschriften, Jahrbücher, etc.), nicht aber mehrbändige Werke
 - Varia: Alles was oben nicht passt, also z.B. nichts mit Österreich zu tun hat, aber trotzdem aus irgendeinem Grund gescannt wird (z.B. Fernleiheauftrag, Lehrveranstaltung oder ähnliches)
 - Manuskripte: Handschriftliches
 - Frauen in Bewegung: ÖNB-Projekt
 - Erlebnisraum Altes Buch: SOSA UB Graz

1.3 Periodica

- reUSE: Periodica liegen ebenfalls unter den Institutionen
 - Ansonsten gleiche Vorgangsweise: Collection mit Haupttitel der alles in Kurzform enthält
- Eine eigene (Sub-)Collection unter Periodica (derzeit Serial) wird angelegt
- Diese Collection wird allerdings nur mit einem einzigen DC Feld, dem Titelfeld, erfasst. Es wird der Titel und das Jahr bzw. die Jahre von bis eingetragen.
Bsp: Jahrbücher der Grillparzer Gesellschaft (18xy-19xy)
- Unter diese Sub-Collection werden dann die einzelnen Bände der Jahrgänge mit voller Angabe eingegeben.

1.4 Mehrbändige Werke

- Achtung: Werden zwar ähnlich behandelt, nicht aber unter Periodica eingespielt sondern inhaltlich (Tirolensien, Styriaca, etc.) zugeordnet.
- Autor oder Herausgeber werden genannt
- Typisches Zeichen für mehrbändiges Werk: Paginierung beginnt wieder bei 1. (Deshalb manchmal mehrbändige Werke auch innerhalb eines Bandes)
- Bsp: DC Title: Pichler, Adolf: Gesammelte Werke in 17 Bänden. (18xy-18xy)

1.5 Gesammelte Werke, Wissenschaftliche Editionen

- Ganz wie mehrbändige Werke, allerdings wird der Autor der Schriften, nicht aber der Herausgeber als DC: Creator genannt.
- Der Herausgeber als DC: Contributor
- Manchmal wird die Anlage von Untersammlungen (Abteilungen) sinnvoll sein.

2. Metadateneingabe

siehe auch Anhang: Beispiele für Metadatenaufnahmen

2.1 Allgemeines

- Im ALO Client erscheinen alle DC Felder, aber nur die unten beschriebenen werden auch verwendet.
- Die Metadaten können in der neuesten Client Version auch vordefiniert werden. Dies ist besonders praktisch, wenn sich die Metadaten über mehrere Bücher hinweg nicht oder nur kaum unterscheiden (zB Haupttitel, Autor, Publisher, ...) Das Symbol zur Vordefinition der Metadaten befindet sich in der Maske der elektronischen Version rechts von dem Register „Files“.
- Genau unterscheiden zwischen den Metadaten für elektronische Version und Original!
- Wenn mehrere Informationen zu einem Feld vorliegen, z.B. mehrere Autoren, oder Titel, Untertitel, Titelzusatz etc. dann jeweils eine eigene Zeile dafür öffnen (im Klient durch + gekennzeichnet)
- Grundsätzlich keine Satzzeichen am Ende einer Eingabe schreiben, auch dann nicht, wenn z.B. ein Titel ein Rufzeichen oder Fragezeichen enthält!
- Damit man sich das doppelt schreiben für Titel, Autor, Contributor erspart: copy-Taste beachten.
- Bei Unklarheiten kurze Research nach dem Buch in online-Datenbanken durchführen (zB www.zvab.com, <http://www.ubka.uni-karlsruhe.de/kvk.html>, oder im ALEPH)

2.2 Grundsätzliches zur Titel und Autoraufnahme

Das Wesentliche soll verzeichnet werden.

- Deshalb bei Autoren werden jene Infos weggelassen wie „Dr.“ „Prof.“, „Herausgeber“, „Mag. Phil.“ etc, genauso Ehren- und Adelstitel
- Hingegen sollte bei Abkürzungen mit einfachen Mitteln versucht werden, die richtige Ergänzung einzufügen: „L. Mühlbach“ wird daher zu „Mühlbach, Louise“
- Auch ein fehlendes Jahr kann ergänzt werden, sofern aus einem Bibliothekskatalog eine eindeutige Angabe hervorgeht.

Wiederholungen und Selbstverständlichkeiten vermeiden

- „Druck und Verlag von Hermann Schoenfeld“ steht unter DC: Publisher und der Name reicht daher: „Hermann Schoenfeld“.

Für mehrere Informationen innerhalb einer Kategorie wird jeweils ein eigenes Feld angelegt, dafür die Füllwörter weggelassen

- „Johann Graf Mailáth und Dr. S. Saphir“ wird zu „Mailáth, Johann“, „Saphir, Sigmund“

2.3 Mehrbändige Werke

Die Unterscheidung zwischen mehrbändigen Werken und Periodica: Periodica waren grundsätzlich auf eine unbegrenzte Erscheinungsdauer angelegt, also Zeitschriften, Jahrbücher, etc. Beide Arten werden in afo derzeit noch ganz ähnlich erfasst, d.h. als „Sammlung/Collection“. Teile dieser Sammlungen sind dann die einzelnen Objekte, also Bücher. Bei Zeitschriften mit vielen Unterheften werden auch noch Jahrgänge zu eigenen Sammlungen. Die Sammlungen erhalten einen kurzen prägnanten Titel, der die wichtigsten Informationen enthalten soll und dem Benutzer gleich verrät worum es sich handelt.

2.4 Dublin Core für Bücher und Zeitschriften

2.4.1 Allgemeines

- Die Informationen finden sich grundsätzlich auf dem Haupttitelblatt des Werks, das in der Regel rechts abgebildet ist.
- Titelblätter, die links erscheinen, beziehen sich meist auf ein übergeordnetes Werk oder eine Reihe.
- Druckvorlagen neuer Bücher werden genauso behandelt wie gescannte Bücher, d.h. es gibt Metadaten für Original und Metadaten für die elektronische Version.
- Wichtig: Bei laufenden Veröffentlichungen kann der Titel und die Information zu einem Band auf jene Daten verkürzt werden, die nur für diesen Band spezifisch sind. Z.b. „Heft 04“ oder „Ausgabe 10/2000“ oder ähnliches. Vorangestellte „0“ nur wenn mehr als 10 Ausgaben

2.4.2 DC: Title

- Metadaten für electronic und original sind dieselben.
- Alle Titel werden hier aufgenommen, aber in jeweils eigener Zeile (Return Taste): Untertitel etc. und Titelzusätze.
Beispiel: Vierte verbesserte Auflage. Mit 3 Illustrationen.
- Auch Übersetzungen fremdsprachiger Titel – sofern sie vorliegen - können hier eingetragen werden.
- Bei laufenden Veröffentlichungen kann der ti

2.4.3 DC: Creator

- Metadaten für electronic und original sind dieselben.
- Verfasser, Schöpfer eines Werkes; diejenige Person, oder Institution die für den Inhalt verantwortlich ist.
- Bei Zeitschriften, Sammelbänden etc. der Herausgeber
- Ausnahme: Wissenschaftliche Editionen (Bsp.: Goethes Gesammelte Werke.
DC:Creator = Goethe. Name des wissenschaftlichen Herausgebers ist bei Contributor einzufügen)
- Name wird in vollständiger und gängiger Form eingetragen, nicht quellentreu. Falls unvollständige Namen im Buch genannt werden, Recherche in Aleph und KVK.
- Europäische Namen: Familienname – Beistrich – Vorname.
- Außereuropäische Namen: Mao Tse Tung
- Abkürzungen: Georg K. Feuerbach wird zu: Feuerbach, Georg K.
- Pseudonyme: Die gängige Form wird eingegeben: Jean Paul (nicht: Richter, Jean Paul)
- Falls bekannt kann auch ein zweites Feld aufgemacht werden und der tatsächliche Name eingegeben werden
- Autoren die praktisch nur unter ihrem Pseudonym bekannt sind, werden nicht aufgelöst
- Bücher ohne Autoren: Autorenfeld bleibt leer.
- Mehrere Autoren werden ohne „und“ etc. in mehrere Autorenfelder eingetragen.

2.4.4 DC: Subject

Derzeit nicht vergeben.

2.4.5 DC: Description

- Sowohl in elektronischer als auch Originalversion
- Kann man verwenden um kurz etwas über den Inhalt des Buches zu berichten.
- Auch Links können hier eingefügt werden: Eigene Zeile und <http://...> führt zur Anzeige des Links.

2.4.6 DC: Publisher

Elektronische Version

- Immer die Bibliothek, aus der das Buch stammt. Unbedingt Namen der UB ausschreiben:
- Bsp.: Universitätsbibliothek Innsbruck, Universitätsbibliothek Graz, Brenner Archiv

Metadata original

- Hier wird der Verlag oder Verleger der im Buch steht eingetragen.
- Verlagsort wird ebenfalls in einem Zusatzfeld Publisher verzeichnet.
- Reihenfolge lautet immer: Verlag – Ort
- Mehrere Verlage und Orte können ignoriert werden. Der wichtigste Verlag sollte angegeben werden oder – falls nicht offensichtlich – der erstgenannte. Gleiches gilt für den Ort

2.4.7 DC: Contributor

- Metadaten für electronic und original sind dieselben.
- Personen, die einen wesentlichen Beitrag zum Inhalt des Werks geliefert haben.
- Bsp: Illustratoren, Beiträger in Sammelbänden, Schreiber von Einleitungen, Herausgeber bei wissenschaftlichen Editionen, u.ä.

2.4.8 DC: Date

Für die elektronische Version

- Datum des Hochladens: diese wird automatisch vergeben, es ist nicht nötig etwas einzugeben!
- Eingabe: 2002-01-24

Für das Original:

- Publikationsdatum des Buches
- Eingabe: Jahreszahl.
- Wenn mehrere Jahre: 1906-1911 (Bindestrich ohne Leerzeichen)
- Wenn keine Jahreszahl im Buch abgedruckt ist, aber das Jahr bekannt ist: 1906
- Wenn keine Jahreszahl abgedruckt und nicht eruierbar ist Feld frei lassen

2.4.9 DC: Type

- Für das Original: Auswahl aus Liste: Texttypen (derzeit in Ausarbeitung)
- Hier wird eingetragen: Zeitschrift, Gemeindebote, Adressbuch,...

2.4.10 DC: Format

- Derzeit nicht vergeben.

2.4.11 DC: Identifier

Bei gescannten Büchern für das Original:

- Signatur
- Die Bibliothek in Abkürzung und die Signatur
- Derzeitige Abkürzungen (nur für Innsbruck):
- Hier muss die Bibliothek und die Katalognummer des Buches eingetragen werden.
- Falls nicht bekannt ist ob das Buch Magazin oder Sondersammlung ist, dann einfach UBI-HB und die Signatur.
- UBI-HB-M = Hauptbibliothek, Magazin
- UBI-HB-S = Abt. für Sondersammlungen (hier auch Bestände des Handschriftenleseraums)
- UBI-GER = Fachbibliothek für Germanistik
- UBI-FBG = Fachbibliothek für Geschichte
- UBG-???

Wichtig: Bei allen Büchern (v.a. aber reUSE Büchern) wo vorhanden, zusätzlich:

- ISBN oder ISSN
- ...

Für die elektronische Version: Bei allen Büchern, v.a. aber bei reUSE Büchern und Zeitschriften die ALEPH ID (AC Nummer) eintragen

- AC
- Bei jedem Objekt eintragen!

2.4.12 DC: Source

Derzeit nicht vergeben.

2.4.13 DC: Language

- ISO Language code 639 (1988)
- <http://www.oasis-open.org/cover/iso639a.html>
- de = Deutsch
- en = Englisch
- Beispiel zur Schreibweise: ISO 639: 1988 de

2.4.14 DC: Relation

Derzeit nicht vergeben.

2.4.15 DC: Coverage

Bei allen Büchern, die eindeutig einer geographischen Region zuordenbar sind oder einer zeitlichen Epoche (nicht Erscheinungsjahr!) vergeben. Schreibweise:

- Place: Innsbruck
- Time: Befreiungskrieg 1809

2.4.16 DC: Rights

Derzeit nicht vergeben.

2.5 Technische Metadaten

- Name des Mitarbeiters für das Scannen
- Typ: BookEye
- Software: BSC2
- Achtung: 300x300 dpi eingeben
- Bei reUSE Büchern: PDFs, etc. ev. Name der Software, mit dem das PDF erzeugt wurde

2.6 Kontrollblick in ALO

Trotz der Kontrolle gleich nach der Eingabe der Metadaten können Fehler passieren. Um das frisch hochgeladene Buch auf solche Fehler zu untersuchen, einfach in ALO einsteigen (www.literature.at) und unter „neu in alo“ das Buch anschauen (stichprobenartig durchklicken). Erkannte Fehler mit der Funktion „change“ korrigieren. Dazu wird das Buch automatisch aus ALO genommen, muss nach Korrektur der Fehler aber wieder mit „send to server“ hochgeladen werden. Die objectID bleibt so dieselbe.

Beispiele

Schloß Kostenitz.



Novelle

von

Ferdinand von Saar.



Heidelberg.
Georg Weiß, Verlag.
1893.

Originalversion

dc:title

- ⇒ Schloß Kostenitz
- ⇒ Novelle

dc:author

- ⇒ Saar, Ferdinand von

dc:publisher

- ⇒ Georg Weiß
- ⇒ Heidelberg

dc:date

- ⇒ 1893

dc:identifier

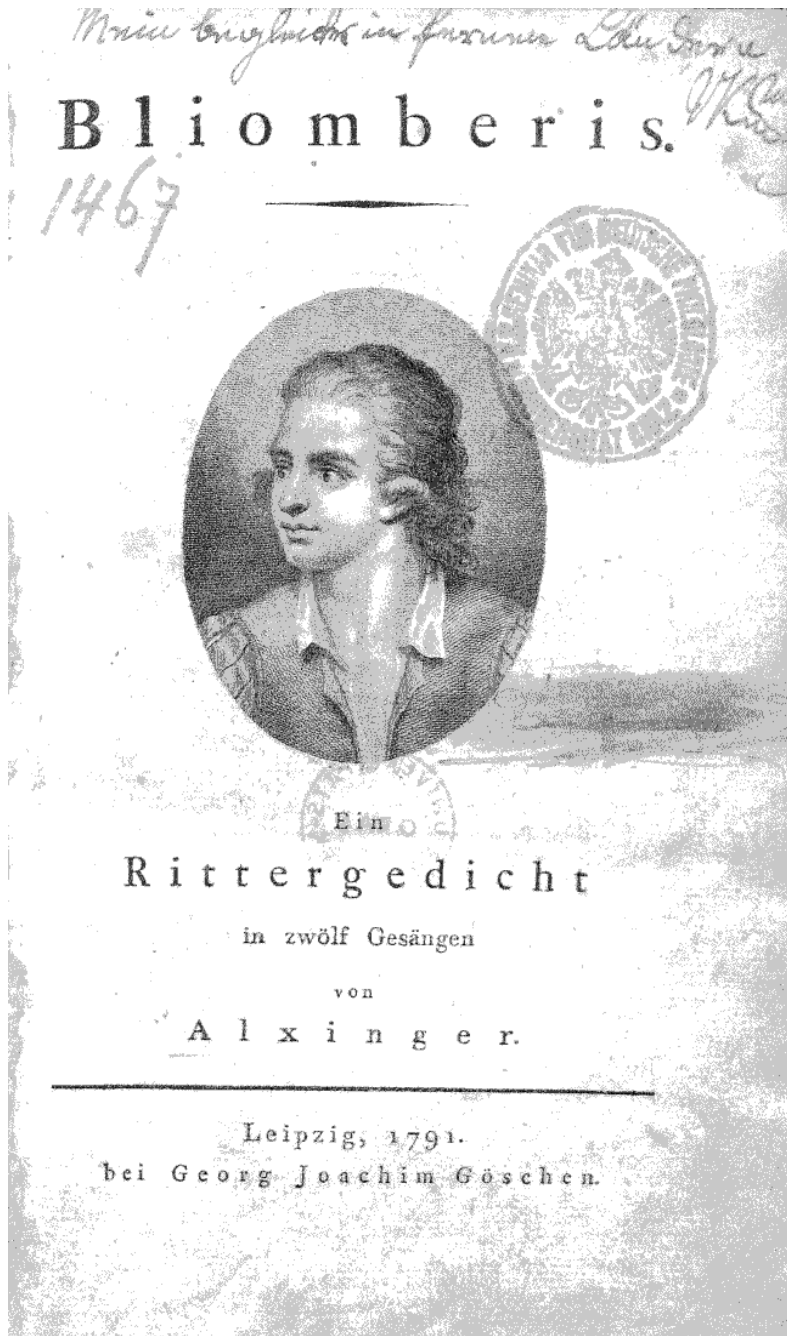
- ⇒ UBG ...

dc:type:

- ⇒ Novelle oder Roman

dc:coverage

- place: Steiermark



Originalversion

dc:title

- ⇒ Bliomberis
- ⇒ Ein Rittergedicht in zwölf Gesängen

dc:author

- ⇒ Alxinger, Johann Baptist von [kvk nachschauen!]

dc:publisher

- ⇒ Georg Joachim Göschen
- ⇒ Leipzig

dc:date

- ⇒ 1791

dc:identifier

- ⇒ UBG ...

dc: type

- Versdichtung

dc: coverage

- Österreich

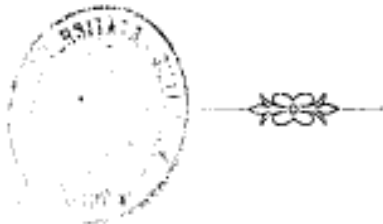
Der Tharerwirt,

ein Tiroler Held vom Jahre 1809.

Geschichtliche Erzählung

von

Johann Steck.



Innsbruck.

Verlag der Wagner'schen Universitäts-Buchhandlung.

1893.



Originalversion

dc:title

- ⇒ Der Tharerwirt, ein Tiroler Held vom Jahre 1809
- ⇒ Geschichtliche Erzählung

dc:author

- ⇒ Steck, Johann

dc:publisher

- ⇒ Wagner'sche Universitäts-Buchhandlung
- ⇒ Innsbruck

dc:date

- ⇒ 1893

dc:identifier

- ⇒ UBI ...

dc: type

- place: Tirol
- time: Befreiungskriege

Andreas Hofer.

Historischer Roman

von

L. Mühlbach.

Erster Teil.

Oesterreichs Erhebung.

1. Heft.



Dresden.

Druck und Verlag von Hermann Schoenfeld.

Originalversion

dc:title

- ⇒ Andreas Hofer
- ⇒ Historischer Roman
- ⇒ Erster Teil
- ⇒ Oesterreichs Erhebung
- ⇒ 1. Heft

dc:author

- ⇒ Mühlbach, Louise [nachschaun und ergänzen!]

dc:publisher

- ⇒ Hermann Schoenfeld
- ⇒ Dresden

dc:date

- ⇒ 1893

dc:identifier

UBI ...

dc: type

- roman

dc: coverage

- place: Tirol
- time: Befreiungskriege

3. Standardaufnahme

3.1 Mehrbändige Werke

Auf der linken Seite steht meistens das Titelblatt der Reihe (Sammlung). Auf der rechten Seite findet sich das Haupttitelblatt mit den Infos zum vorliegenden Buch

Adolf Bichler

Gesammelte Werke

Vom Verfasser für den Druck vorbereitet

Band I

Zu meiner Zeit

München und Leipzig
bei Georg Müller
1905

Adolf Bichler

Zu meiner Zeit

Schattenbilder aus der Vergangenheit

Zweite ergänzte und berichtigte Auflage

Mit einer biographischen Einleitung

von E. M. Prem



München und Leipzig
bei Georg Müller
1905

AUFNAHME:

Sammlung anlegen – Daten der linken Titelseite

dc:title

Pichler, Adolf: Gesammelte Werke in 17 Bänden (1905-1909)

Einzelne Bücher anlegen – Daten der rechten Titelseite (= Haupttitelseite)

dc:title

- ⇒ Zu meiner Zeit
- ⇒ Schattenbilder aus der Vergangenheit
- ⇒ Zweite ergänzte und berichtigte Auflage
- ⇒ Mit einer biographischen Einleitung von S. M. Prem

dc:author

- ⇒ Pichler, Adolf

dc:publisher

- ⇒ Georg Müller
- ⇒ München
- ⇒ Leipzig

dc:date

- ⇒ 1905

dc:identifier

UBI ...

dc: type

Roman

dc: coverage

place: Tirol

time: Befreiungskriege

F r a n z G r i l l p a r z e r
S ä m t l i c h e W e r k e

Österreichische Hofbibliothek
1831 Ueberführung
des Handschriftenbestandes der Universität
im Auftrag der Stadt Wien
herausgegeben von
Rudolf Sauer (†)
fortgesetzt von Heinrich Badmann

Dritte Abteilung
Dichter Denk



Wien
Verlag von Anton Schroll & Co.
Deutsch-österreichischer Verlag für Jugend und Volk

F r a n z G r i l l p a r z e r

B r i e f e u n d D o k u m e n t e

Dritter Teil



Wien
Verlag von Anton Schroll & Co.
Deutsch-österreichischer Verlag für Jugend und Volk



AUFNAHME:

⇒ Zuerst wieder die **Sammlung** anlegen, relevant ist idR die linke Titelseite, hier:

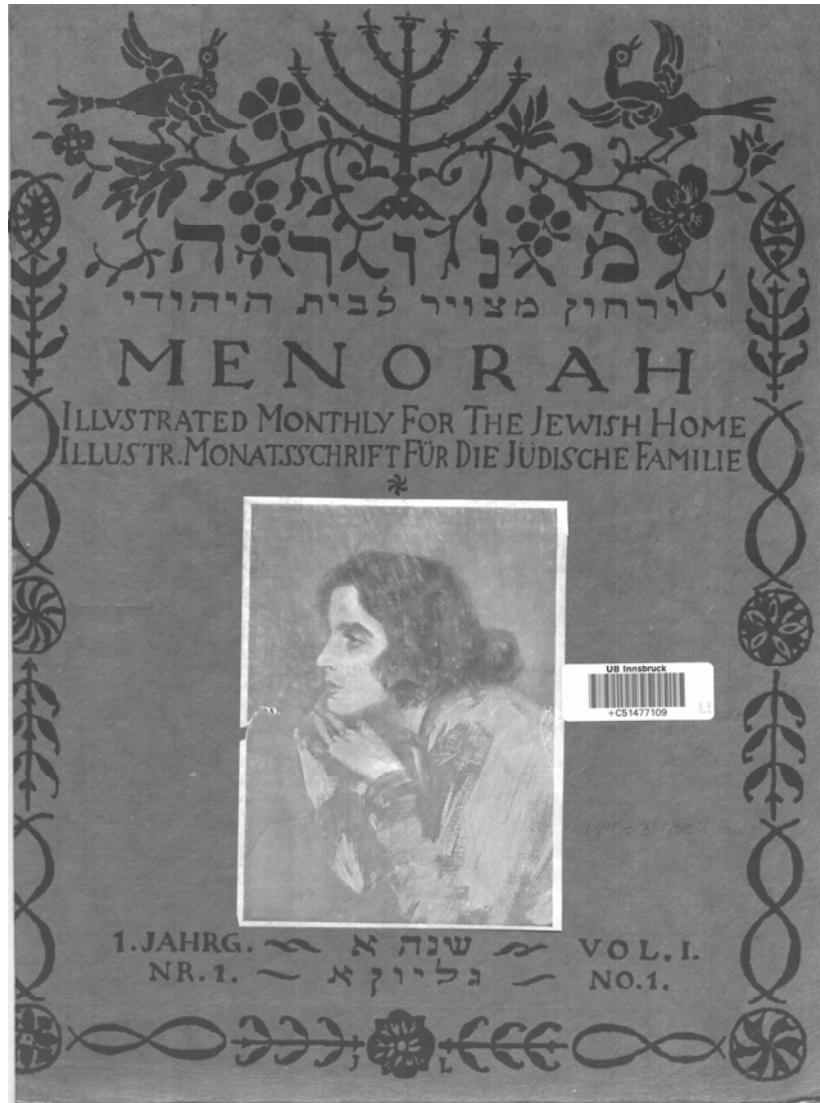
Dc:title Grillparzer, Franz: Sämtliche Werke in 42 Bänden (1909 – 1941)
Falls Umfang der Sammlung oder die Jahreszahlen bekannt sind, kurze online-Recherche durchführen
(zB www.zvab.com, www.ubka.uni-karlsruhe.de/kvk.html oder in ALEPH)

Die Werke sind – wie man aus der linken Titelseite erkennen kann – in Abteilungen und Kapitel unterteilt. Derzeit wird dieser Sachverhalt so gelöst, dass innerhalb der Sammlung eine Untersammlung angelegt wird (NEW → COLLECTION) und der Dublin Core-Eintrag für den Titel lautet: „Erste Abteilung“. Dies wird für alle weiteren Abteilungen wiederholt.

⇒ Dann werden wieder die **einzelnen Bücher** innerhalb der Untersammlung angelegt, lt. der rechten Titelseite, hier:

Dc:title Briefe und Dokumente
Vierter Teil
Dc:author Grillparzer, Franz
Dc:publisher Anton Schroll & Co. (Deutscher Verlag f. Jugend & Volk kann vernachlässigt werden)
Wien
Dc:date 1935 (auf der nächsten Seite – nicht abgebildet)
Dc:identifizier UBI-GER III-554 (auf dem Buchumschlag – nicht abgebildet)
Dc:language ISO 639: 1988 de (Standard)

3.2 Zeitschriften



Collection

dc:title

- ⇒ Menorah (1923-1932)
- Illustrierte Monatsschrift für die jüdische Familie

Originalversion

dc:title

- ⇒ Menorah
- 1. Jahrgang
- 2. Für die bessere Zählung: 01, 02, ..

dc:author

- ⇒ Feld bleibt leer

dc:publisher

- ⇒ Verlag der Monatsschrift „MENORAH“ (letzte Seite, nicht abgebildet)

dc:date

- ⇒ 1928


dc:identifier

- ⇒ UBI-HB ...

dc: type

Familienzeitschrift

I r i s.



Taschenbuch

f ü r d a s J a h r 1840.



Herausgegeben

von

Johann Graf Mailáth und Dr. S. Saphir.

Erster Jahrgang.

(Mit sechs Stahlstichen.)

P e s t h.

Verlag von Gustav Heckenast

Collection

dc:title

⇒ Iris. Taschenbuch. Jahrgang 1 bis 9. (1840-1848)

Originalversion

dc:title

⇒ Iris
⇒ Taschenbuch für das Jahr 1840
⇒ Erster Jahrgang
⇒ Mit sechs Stahlstichen

dc:author

⇒ Mailáth, Johann
⇒ Saphir, S.

dc:publisher

⇒ Gustav Heckenast
⇒ Pesth

dc:date

⇒ 1840

dc:identifier

⇒ UBG ...

dc: type

Jahrbuch