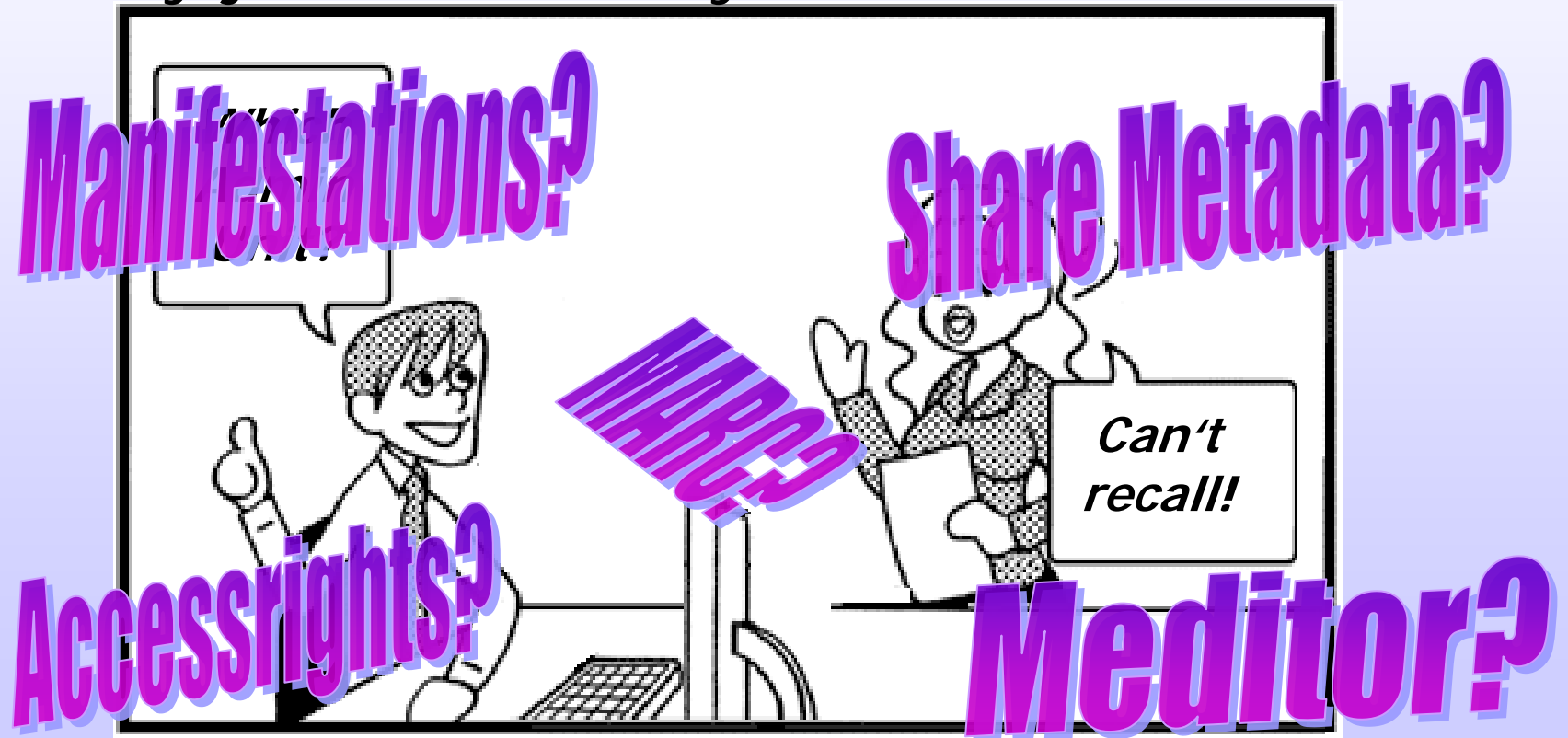


Get objects ↔ support librarians!



Matthias Gross (Bavarian Library Network)

Importing objects into Digitool is something they just do time by time.



The screenshot shows the BVB Ingest web application interface. At the top, there is a navigation bar with 'Folders | **New Ingest Activity** | Ingest Flows | Task Chain | Upload'. The user is identified as 'SCHROEDERP' and the admin unit as '[DTL02] Test-Unit'. A progress indicator shows steps 1, 2, 3, and 4, with step 3 highlighted. The main content area is titled 'Parameters' and contains a 'Transformers' section with file location and upload options. A red circle highlights the text 'Upload now for converting German style Excel csv-file'. Below this is the 'Tasks' section, specifically 'Add Metadata', which is enclosed in a red box. A 'Windows Internet Explorer' error dialog is displayed over the 'Add Metadata' form, stating: 'Please select a metadata type in task number 1 ('Add Metadata').' The dialog has an 'OK' button. At the bottom of the interface, there are navigation buttons: '< Back', 'Next >', and 'Cancel'. An arrow points to the 'Next >' button.

The screenshot shows the 'Web Ingest - Schritt 3/4: Parameters' interface. The main content area is titled 'Parameters' and includes sections for 'Transformers' and 'Tasks'. The 'Transformers' section has a dropdown for 'Please select file location:' set to 'Local Files', and input fields for 'The csv file' and 'The mapping xml'. The 'Tasks' section is titled 'Add Metadata' and includes a 'Metadata Type:' dropdown set to 'MARC - by Verbund'. A red box highlights the 'Share this Metadata?' section, which contains a radio button for 'no, I want to...' and a 'Warning' icon. Below this are fields for 'Metadata ID to share:', 'Add metadata to COMPLEX/METS Objects only:' (set to 'false'), and three filter fields for file extensions and sizes. At the bottom, there are 'Loading' and 'Core Digital Entity Loader' sections. A red arrow points from the 'Next >' button at the bottom right of the interface to the 'Next >' button in the browser window's navigation bar.

The screenshot shows the BVB Ingest web interface. At the top left is the BVB BibliotheksVerbund Bayern logo. The top right features a 'DTL02 Test-Unit' badge and a user profile icon. Below the header is a navigation bar with 'Folders', 'New Ingest Activity', 'Ingest Flows', 'Task Chain', and 'Upload'. The main content area displays a table of activities with columns for '#', 'Id', 'Activity Name', 'Task Chain Name', 'Assign To', and 'Actions'. A sidebar on the left shows status counts: 'Not Scheduled (4)', 'Scheduled (0)', 'Running (0)', 'Success (34)', and 'Failed (0)'. An orange arrow points from the 'BVB+' icon in the 'Actions' column of the first row to a red-bordered box containing text about pre-ingest services.

#	Id	Activity Name	Task Chain Name	Assign To	Actions
1.	inq6915	Deutsches Museum -Demo	chain_04.08.09.15.39.53.51	SCHROEDERP ...	
2.	inq6616	ubrmeddisstest	chain_19.06.09.15.45.15.638	SCHROEDERP ...	
3.	inq6332	Kartegross_NDPN	chain_26.04.09.14.57.0.469	SCHROEDERP ...	
4.	inq5720	sdsds	chain_07.03.09.18.28.1.819	SCHROEDERP ...	

Pre-Ingest Services

- creating Mets-Files

Ingest Services after Upload

- semi-automatic ingest of newspapers

German law: Libraries are allowed to offer digital versions of their books within their reading rooms

- digitization of own copy
- one concurrent user per physical copy

Great chance to get objects into our repository

⇒ One licensed object per library

⇒ One catalogue record on the network level but many URLs

Catalogue expert

*Only
1 URL
per
catalogue
record !*



The Solution: URL to a Gateway object

Gate Object

no stream/any stream

usage type: view_main

Partition A: bvb_gate

**only the URL
of the Gate
(manifestation
leader) will be
passed to
ALEPH**

**the MARC data are shared
DTL-ALEPH Sync**

Manifestation 1

PDF

IP range limited

holder: Library A

Manifestation 2

METS

IP range limited

holder: Library B

Gate Object

no stream/any stream

usage type: view_main

Partition A: bvb_gate

**The Gate
manifestation
can be
handled by our
gateway script
ALEPH**

**the MARC data are shared
DTL-ALEPH Sync**

Manifestation 1

PDF

IP range limited

holder: Library A

Manifestation 2

METS

IP range limited

holder: Library B

Gateway viewer = „Find-Appropriate-Manifestation Viewer“

JSP using DigiTool Java Libs

1. get DE by pid (pid from viewer url)
2. get all manifestation usage_type=VIEW
3. traverse DEs
any access right matches user credentials?
yes : get the url and redirect
no : continue
4. finally: display fallback text

Fallback page, dynamically created



The screenshot shows a Mozilla Firefox browser window with the address bar containing the URL: `http://bvbm1.bib-bvb.de:1802/webclient/bvb_gateway.jsp?/2009/02/09/file_1/31575&usePid1=true&usePid2=true`. The main content of the page is titled "BVB-Multimediaserver Gateway" and features the BVB BibliotheksVerbund Bayern logo. The text explains that the user has been redirected to a page that could not find any licensed objects for their IP address [193.174.96.1]. It lists several libraries: Deutsches Museum, BSB Bayerische Staatsbibliothek, STAATSBIBLIOTHEK BAMBERG, Universitätsbibliothek Regensburg, and Universitätsbibliothek Augsburg (UBA). The page concludes with a message advising the user to contact their library for access restrictions.

BVB-Multimediaserver Gateway

BVB BibliotheksVerbund
Bayern

By the address you have sent to our server
http://bvbm1.bib-bvb.de:1802/webclient/DeliveryManager?pid=31575&custom_att_2=simple_viewer
you could have been redirected to objects that were provided by one of the following libraries

Deutsches Museum

BSB Bayerische
Staatsbibliothek

StB STAATSBIBLIOTHEK
BAMBERG

UR Universitätsbibliothek
Regensburg

Universitätsbibliothek
Augsburg **UBA**

Unfortunately our server could not find any licensed object for your IP address [193.174.96.1].

The libraries listed above are responsible for the definition of their access restrictions. Contact your library if you think you should have been given access from your computer.

Fertig

Fallback page, dynamically created

- DirectStreamViewerPreProcessor
- URL created by VPP

.../webclient/bvb_gateway.jsp?/2009/02/09/file_1/31575&usePid1...



PID of Gate object

Extending the extension

Some objects presented outside of DigiTool ?

Create DigiTool object that contains the corresponding URL

Add access rights metadata on it

Add something in partition a to control to the appropriate viewer

Write a simple redirect-to-URL viewer (perl/cgi)

- reads the URL from the file
- redirects to the URL

Serials in DigiTool

Bavarian Library Network Center

Munich, at the day of Mary of Bethany 2009

One PDF per issue

=> We love METS because the navigation by a table of contents on the left hand side seems to be the natural way to present a serial

Right now we have ...

University Library of Bamberg

one weekly journal with access rights

State Library of Neuburg/Donau

3 daily newspapers with access rights

Bavarian State Library

34 serials of statistical reports of the Bavarian State Office for Statistics, free

Mets Viewer - Windows Internet Explorer

STÄATLICHE BIBLIOTHEK
NEUBURG AN DER DONAU

Ansichtsart: Jahrgangsübersicht(Logical) Donaukurier / Ausgabe Neuburg

>> Jahrgangsübersicht > 2009 > März > Mo., 30.03.2009

Ansichten: Donaukurier / Ausgabe Neuburg 2009, Nr. 74

Suchen

1 / 36 48,2%

DONAUKURIER
UNABHÄNGIGE ZEITUNG FÜR INGOLSTADT, NEUBURG UND REGION

Nr. 74, Montag, 30. März 2009 www.donaukurier.de Einzelpreis 1,20 €

Kyoto-Ziel erreicht: Deutschland reduziert Treibhausgase Seite 4

Andrang im Stadtmuseum: Schau über Raimund Kerner Seite 19

Pflichtsieg: DFB-Team schlägt Liechtenstein mit 4:0 Seite 29

LOKALES

Handwerk setzt auf Qualität
Neuburg (DK) Gemischte Gefühle herrschen bei der Freisprechung von 17 jungen Leuten. Während die frisch gebackenen Gelehrten zumeist für den neuen Lebenschnitt freilassen, gab es auch nachdenkliche Worte, denn die Krise kommt langsam auch beim Handwerk an. Dennoch sei man zuversichtlich, setze auf Qualität. „Wir sind keine Spulanten, sondern schickten Werte“, heißt es. Seite 19

Klärungsbedarf beim TSV
Neuburg (DK) Auf ein erfolgreiches Jahr blickte der TSV bei der Delegiertenversammlung zurück. Umwelts ist hingegen ein Zukunftsspektrum, die Beteiligung an der Sanierung der Maffachturnhalle im Grund abgelehnt von offenen Fragen bei der Stadt, herrscht auch intern Klärungsbedarf. „Wir wissen noch nicht, ob wir darau wollen.“ Seite 21

BYERN

Wechsel bei Katholiken
Albert Scheidel ist neuer Vorsitzender des Landeskirchenrats der bayerischen Katholiken. Das Komitee wählte den 63-jährigen Rügenburger auf der Prälatenversammlung in Bamberg am Samstag zum Nachfolger von Hans-Joachim Mergel (70). Seite 11

Protesttag gegen die Krise
Berlin/London (AP) Am Samstag haben europaweit Zehntausende gegen die Vorzeichen der Wirtschaftskrise und das fehlende Krisenmanagement der Regierungen demonstriert. Die größte Demonstration gab es mit 10 000 Teilnehmern in London - Schauplatz des G-20-Gipfels in dieser Woche - aber auch in Berlin und anderen Metropolen gingen Zehntausende auf die Straße. Verschiedene Organisationen forderten von dem Gipfel verbindliche und globale Regeln für die internationalen Märkte. In London findet ab Nachmittags das erste G-20-Gipfel im November 2009 in Washington statt, bei dem sich die Teilnehmer auf eine stärkere Regulierung der Finanzmärkte verständigen sollen. Die EU hat für den Londoner Gipfel einen Fiskalpaktskizze für die nächste Woche angekündigt. Seite 4

Bund steigt bei HRE ein
Bankenrettungsfonds kauft in erstem Schritt Anteil von 8,7 Prozent

München (AP) Der Bund macht sich für die wir sehr dankbar sind, nicht der Bundesbank hätte eine In-schuld die Bundesrepublik schmerzschlimmere Konsequenzen

ANZEIGE

.. rather manually (without any magic)?

First import of already available issues

by METS ingest

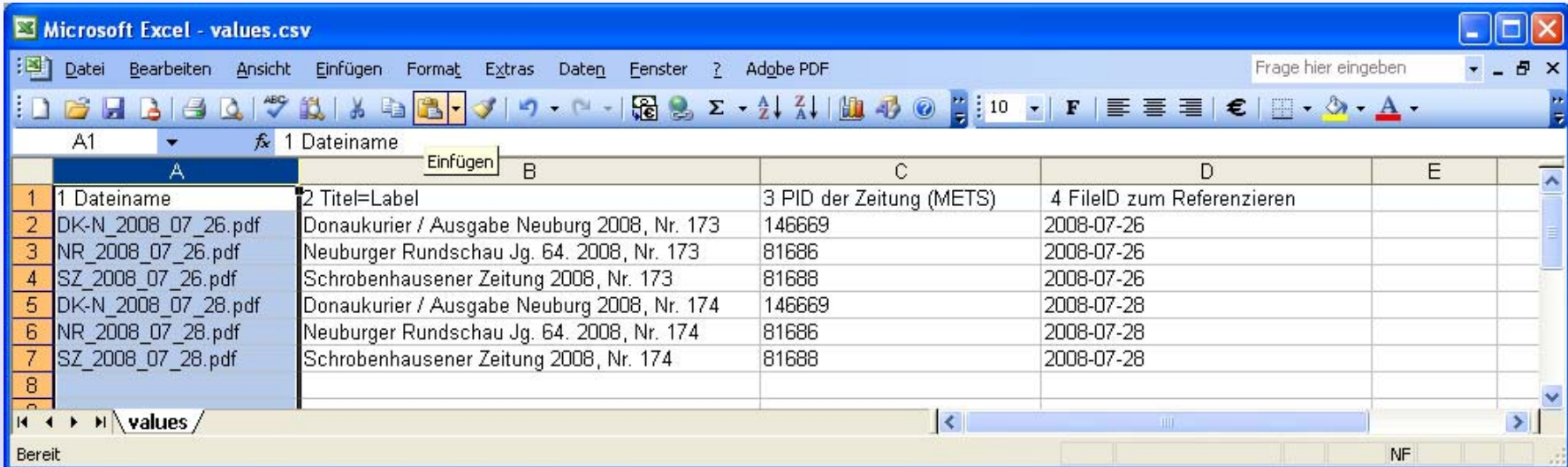
=> create a mets file

=> webingest with access rights if necessary

Import of new issues

by CSV ingest and editing the METS fileSec and structMap

Step 1: creating a CSV file



The screenshot shows a Microsoft Excel window titled "Microsoft Excel - values.csv". The spreadsheet contains a table with 4 columns and 7 rows of data. The columns are labeled A, B, C, and D. The data in the table is as follows:

A	B	C	D
1 Dateiname	2 Titel=Label	3 PID der Zeitung (METS)	4 FileID zum Referenzieren
DK-N_2008_07_26.pdf	Donaukurier / Ausgabe Neuburg 2008, Nr. 173	146669	2008-07-26
NR_2008_07_26.pdf	Neuburger Rundschau Jg. 64. 2008, Nr. 173	81686	2008-07-26
SZ_2008_07_26.pdf	Schrobenhausener Zeitung 2008, Nr. 173	81688	2008-07-26
DK-N_2008_07_28.pdf	Donaukurier / Ausgabe Neuburg 2008, Nr. 174	146669	2008-07-28
NR_2008_07_28.pdf	Neuburger Rundschau Jg. 64. 2008, Nr. 174	81686	2008-07-28
SZ_2008_07_28.pdf	Schrobenhausener Zeitung 2008, Nr. 174	81688	2008-07-28

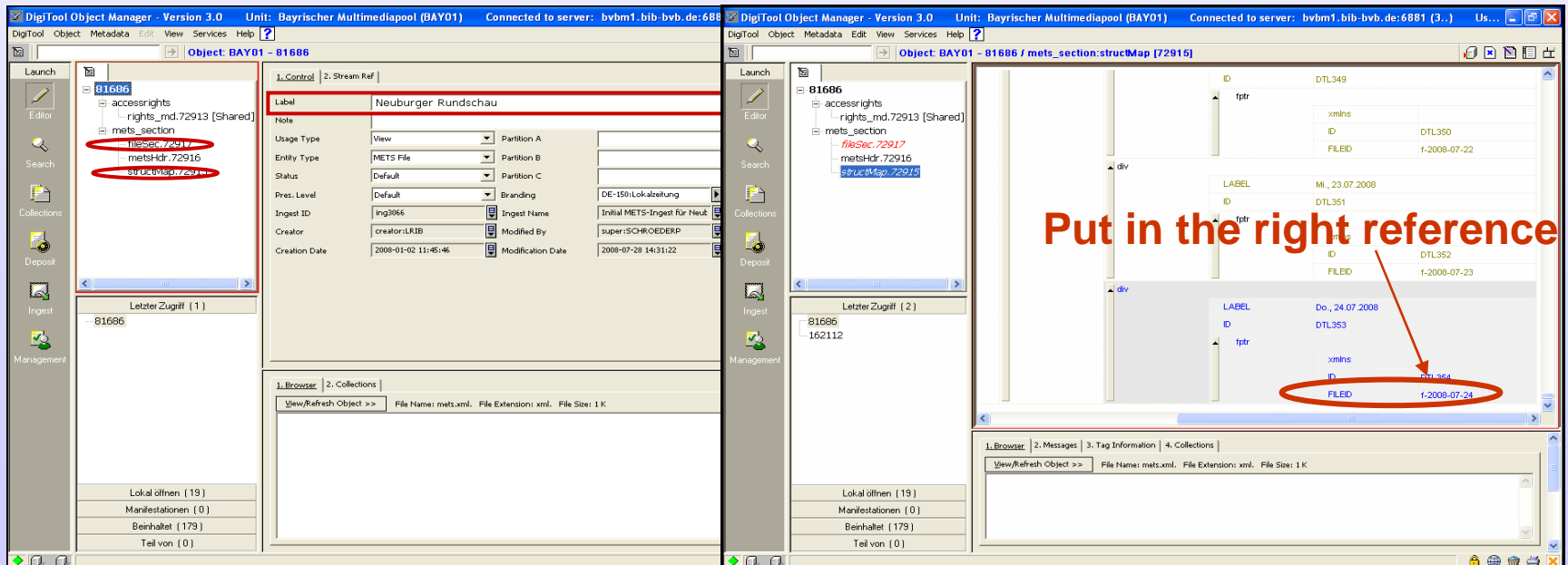
filename, label for the single object, PID of parent object, DTL fileid to be referred to in the METS fileSec

Step 2:

CSV ingest with a saved task chain

Step 3:

Editing FileSec and StructMap in Mediator



fileSec : refer to the fileID specified in csv file and create an ID for this file

```
- <file ID="f-2008-07-25" MIMETYPE="application/pdf" GROUPID="g-2008-07-25">  
  <FLocat xlink:href="2008-07-25" LOCTYPE="URL"/>  
</file>  
- <file ID="f-2008-07-26" MIMETYPE="application/pdf" GROUPID="g-2008-07-26">  
  <FLocat xlink:href="2008-07-26" LOCTYPE="URL"/>  
</file>  
- <file ID="f-2008-07-28" MIMETYPE="application/pdf" GROUPID="g-2008-07-28">  
  <FLocat xlink:href="2008-07-28" LOCTYPE="URL"/>  
</file>
```

structMap: create a div node and refer to the ID given in the fileSec

```
<fptr ID="DTL356" FILEID="f-2008-07-25"/>  
</div>  
- <div LABEL="Sa. &#47;So., 26.&#47;27.07.2008" ID="DTL357">  
  <fptr ID="DTL358" FILEID="f-2008-07-26"/>  
</div>  
- <div LABEL="Mo., 28.07.2008" ID="DTL359">  
  <fptr ID="DTL360" FILEID="f-2008-07-28"/>  
</div>  
- <div LABEL="Di., 29.07.2008" ID="DTL361">  
  <fptr ID="DTL362" FILEID="f-2008-07-29"/>
```

This workflow won't work for metadata larger than 50,000 characters ☹ due to the limited ability of Meditor ...

So we were forced to improve the import in November 2008.

... more automatically and with some magic?

The concept

Providing a script that

- a) creates a csv file automatically,
- b) starts the ingest,
- c) supplements the structMaps and the fileSecs.

First we are going to show the result ...

1. The user uploads files by Webingest
2. calls a script with parameter workingdirectory and username.

1. Upload

2. Start a script, the workingdirectory where the files are uploaded to will be passed

#	Id	Activity Name	Task Chain Name	Assign To	Actions
1.	ing6540	files uploaded at: Mon Jun ...	Tageszeitungen	SCHROEDERP	BVB+

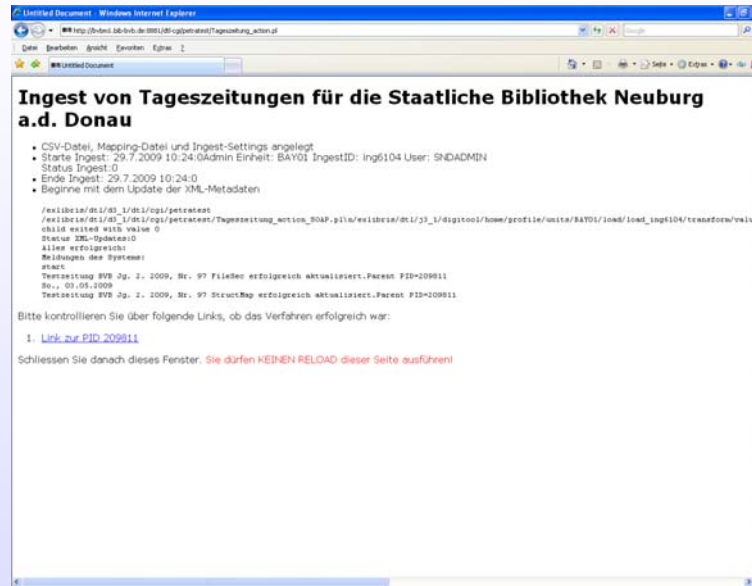
3. fills in a form
filenames, parent pid, access rights mid are filled in by the script
other information is filled in automatically (guessed from the corresponding filename) but can be edited
4. clicks the start button for starting the ingest

Herzlich Willkommen Staatliche Bibliothek Neuburg a.d. Donau!

	Date for StructMap leaf	DTL- Label for child object	Month and year for StructMap nodes			
DK-N_2009_02_28.pdf	Sa./So., 28.02./01.03.2009	Donaukurier / Ausgabe Neuburg 2009, Nr. 49	Februar	2009	146669	219826
DK-N_2009_03_02.pdf	Mo., 02.03.2009	Donaukurier / Ausgabe Neuburg 2009, Nr. 50	März	2009	146669	219826
DK-N_2009_03_03.pdf	Di., 03.03.2009	Donaukurier / Ausgabe Neuburg 2009, Nr. 51	März	2009	146669	219826
NR_2009_02_28.pdf	Sa./So., 28.02./01.03.2009	Neuburger Rundschau Jg. 65. 2009, Nr. 49	Februar	2009	81686	219826
NR_2009_03_02.pdf	Mo., 02.03.2009	Neuburger Rundschau Jg. 65. 2009, Nr. 50	März	2009	81686	219826
NR_2009_03_03.pdf	Di., 03.03.2009	Neuburger Rundschau Jg. 65. 2009, Nr. 51	März	2009	81686	219826
SZ_2008_03_04.pdf	Di., 04.03.2008	Schrobenhausener Zeitung 2008, Nr. 54	März	2008	81688	72913
SZ_2009_02_28.pdf	Sa./So., 28.02./01.03.2009	Schrobenhausener Zeitung 2009, Nr. 49	Februar	2009	81688	72913
SZ_2009_03_02.pdf	Mo., 02.03.2009	Schrobenhausener Zeitung 2009, Nr. 50	März	2009	81688	72913

Ingest mit diesen Werten starten

Start ingesting



5. The user gets a report about success or failure, which he has to check

Finished !

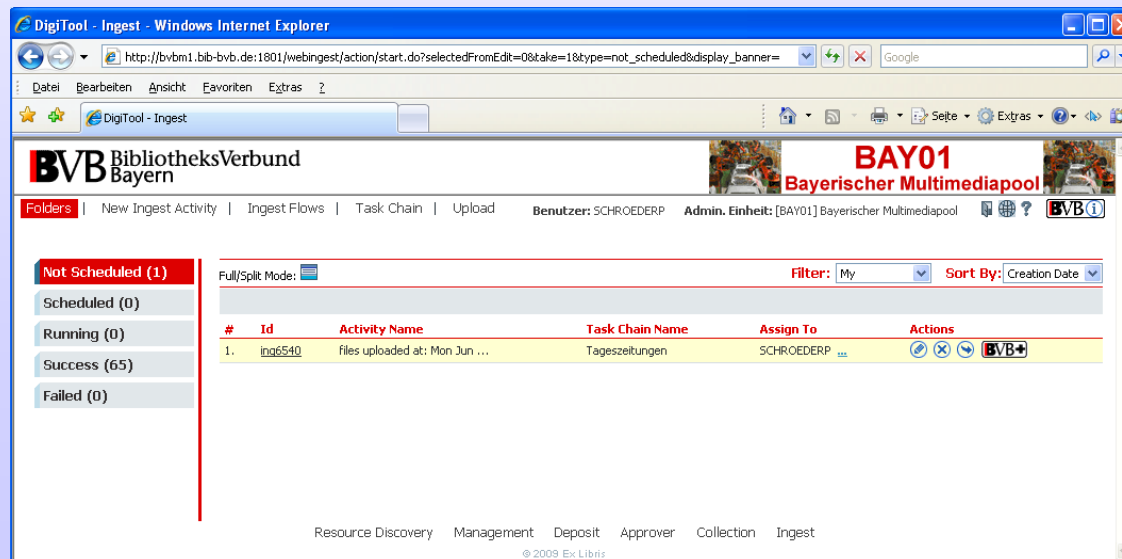
For creating the fill-in form and for performing the ingest you have to know

- a) the working load directory
- b) the username

You may either ask the user or just take it from the jsp script

digitool-mng.ear/digitool-webingest.war/views/admin/main.jsp

If you are interested in details, see next page. If you aren't, just skip it.



How to get the directory where the files have been loaded into

In *digitool-mng.ear/digitool-webingest.war/views/admin/main.jsp*
there are usefull variables:

- `${row.appId}` = ingestID
- session variable `adminUnitOfLoggedInUser` = Admin Unit
- `${row.assignTo}` = Username
- from javascript you can get the port by `window.location.port` for the DTL slot;

You can mix this together in order to create a call like:

`http://servername_and_port/path_to_script?workingLoadId=/exlibris/dtl/j3_1/
digitool/home/profile/units/DTL02/load/load_ing123&user=ITSME`

For the form:

1. order the files by date (in this case simply alphabetical) so that the sequence in the StructMap(s) will be correct
2. provide fields for the branches you have in your structMap
here two: year and month
3. provide fields for all information that isn't shared by all objects
here parent-PID, accessrights-mid
4. any information that will be shared by all objects is defined in the mapping.xml

Herzlich Willkommen Staatliche Bibliothek Neuburg a.d. Donau!

Hoffentlich klappt alles. Korrigieren Sie die Felder und schicken Sie das Formular ab.

Der Schrägstrich '/' muss nicht mehr als / codiert werde. Das macht das Programm alleine.

DK-N_2009_02_28.pdf	Se./So., 28.02./01.03.2009	Donaukurier / Ausgabe Neuburg 2009, Nr. 49	Februar	2009	146669	219826
DK-N_2009_03_02.pdf	Mo., 02.03.2009	Donaukurier / Ausgabe Neuburg 2009, Nr. 50	März	2009	146669	219826
DK-N_2009_03_03.pdf	Di., 03.03.2009	Donaukurier / Ausgabe Neuburg 2009, Nr. 51	März	2009	146669	219826
NR_2009_02_28.pdf	Se./So., 28.02./01.03.2009	Neuburger Rundschau Jg. 65, 2009, Nr. 49	Februar	2009	81686	219826
NR_2009_03_02.pdf	Mo., 02.03.2009	Neuburger Rundschau Jg. 65, 2009, Nr. 50	März	2009	81686	219826
NR_2009_03_03.pdf	Di., 03.03.2009	Neuburger Rundschau Jg. 65, 2009, Nr. 51	März	2009	81686	219826
SZ_2008_03_04.pdf	Di., 04.03.2008	Schrobenhausener Zeitung 2008, Nr. 54	März	2008	81688	72913
SZ_2009_02_28.pdf	Se./So., 28.02./01.03.2009	Schrobenhausener Zeitung 2009, Nr. 49	Februar	2009	81688	72913
SZ_2009_03_02.pdf	Mo., 02.03.2009	Schrobenhausener Zeitung 2009, Nr. 50	März	2009	81688	72913

Processing the fill-in-form ...

1. create the values.csv file in the working directory
2. copy some files into the working directory:
mapping.xml
ingest_settings.xml
3. start tasker.sh by calling the shell
4. start a script that is adding the entries to fileSec and structMap
the necessary information about the branches and the DTL-fileIDs
can be taken from values.csv

We are processing these structural metadata by help of soap calls.
Have a look at *Repository Web Services* in the *General Configuration Guide*

This is what we are currently working on to present digitized serials.

Our idea:

1. Do one METS-Ingest per issue
2. Provide a HTML page where you provide the links to the single issues. They should open in a new window. Maybe you use handles instead of urls, so that the links will be persistent.
3. Alternative: Provide a METS file for the TOC of the serial. On the right side the link or handle to the issue will be shown. Extending the viewers you can do a lot of magic things.

What we would love

The capability of DigiTool to handle METS that is referencing METS-objects.
Making Ex Libris to provide an ingest that will add files to METS-Objects.
Will we ever be able to convince Ex Libris??? Perhaps we can manage that together.