**A Network of Excellence forging the**

**Multilingual Europe Technology Alliance**

**META-SHARE V2 Provider Manual**

**Author(s):** Victoria Arranz, Christian Federmann, Ioanna Giannopoulou, Olivier Hamon, Marc Schröder

**Date:** March 07th, 2012

**Table of Contents**

1.Executive Summary 3

2.Provider-oriented documentation 3

2.1Upgrading a user account to a provider account 3

2.2Add a Language Resource 3

2.3Manage Language Resources 7

2.4Add and change metadata information 8

2.5Import Language Resources 9

2.5.1Prepare the import file 9

2.5.2Upload/Import in META-SHARE 9

2.5.2.1Import script 9

2.5.3Conversion of existing LRs 10

2.6Add Downloads to Resources 10

2.6.1Adding a Download 10

2.6.2Known Limitations 11

# Executive Summary

This document is a guide for the users of META-SHARE interested in providing Language Resources (LRs) through META-SHARE.

# Provider-oriented documentation

## Upgrading a user account to a provider account

As per V2 of META-SHARE, only user accounts marked as “staff” users have the permission to create or edit LRs. For any given META-SHARE node, each “staff” user has the right to edit all language resources on that node. This means that only trusted individuals should be granted staff status.

Future versions of META-SHARE will provide a more fine-grained system of access rights.

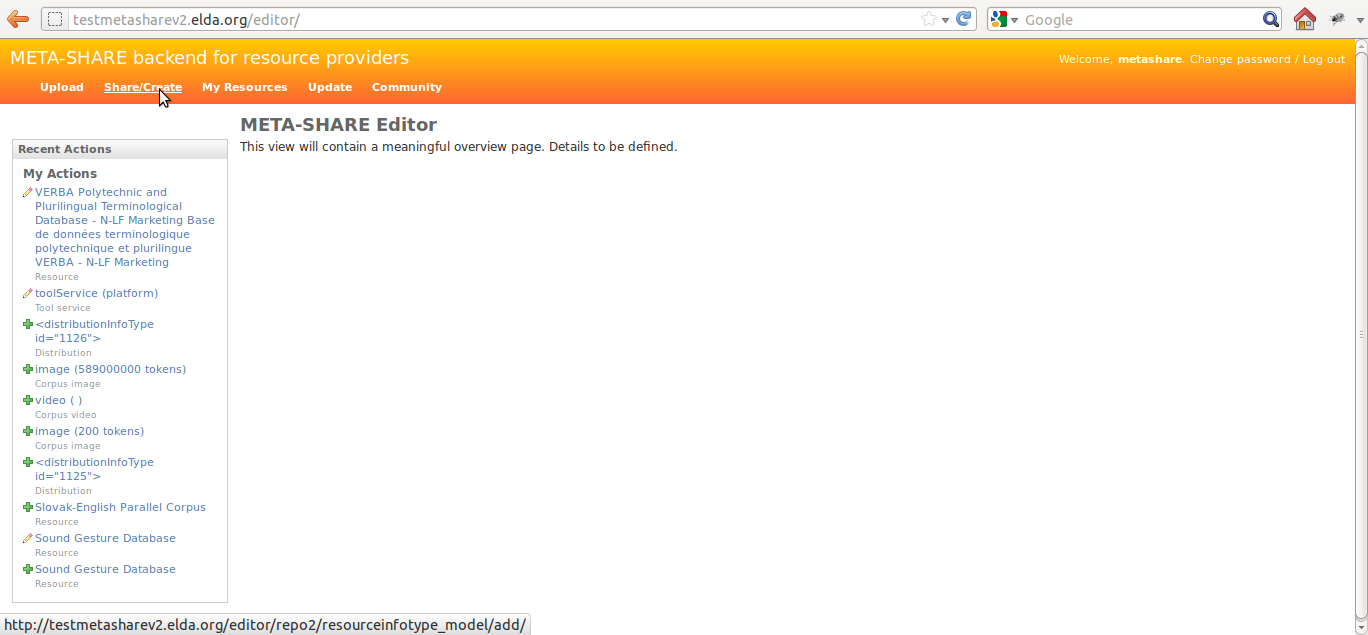
In order to endow a user account with provider access, the following steps are necessary.

1. The user account should be created as documented in the general User Manual.
2. An administrative user must upgrade this account to a “staff” account. As described in the Installation Manual, at least one Administrator account is created as part of the installation process. This administrator can log into the admin interface at <metashare-url>/admin. Under “Users”, go to the details page for the user in question, tick the Permissions box “Staff status”, and confirm with “Save”.
3. The user should now be able to log in to the META-SHARE portal as described in the User Manual and have the permissions required to add and edit LRs as described below.

## Add a Language Resource

The provider can add a new LR by inserting the metadata that describes this LR. The provider must be registered and logged in so as to add new LRs, and then do as follows:

1. Click on the “Editor” button on the top right of the main page of META-SHARE.
2. In the result page (“META-SHARE backend for resource providers”), click on the “Share/Create” link at the top left (see Figure 1) . The result page presents a dropdown menu, listing all the available resource types (corpus, lexical conceptual resource, language description, tool/service).



**Figure 1: META-SHARE backend for resource providers**

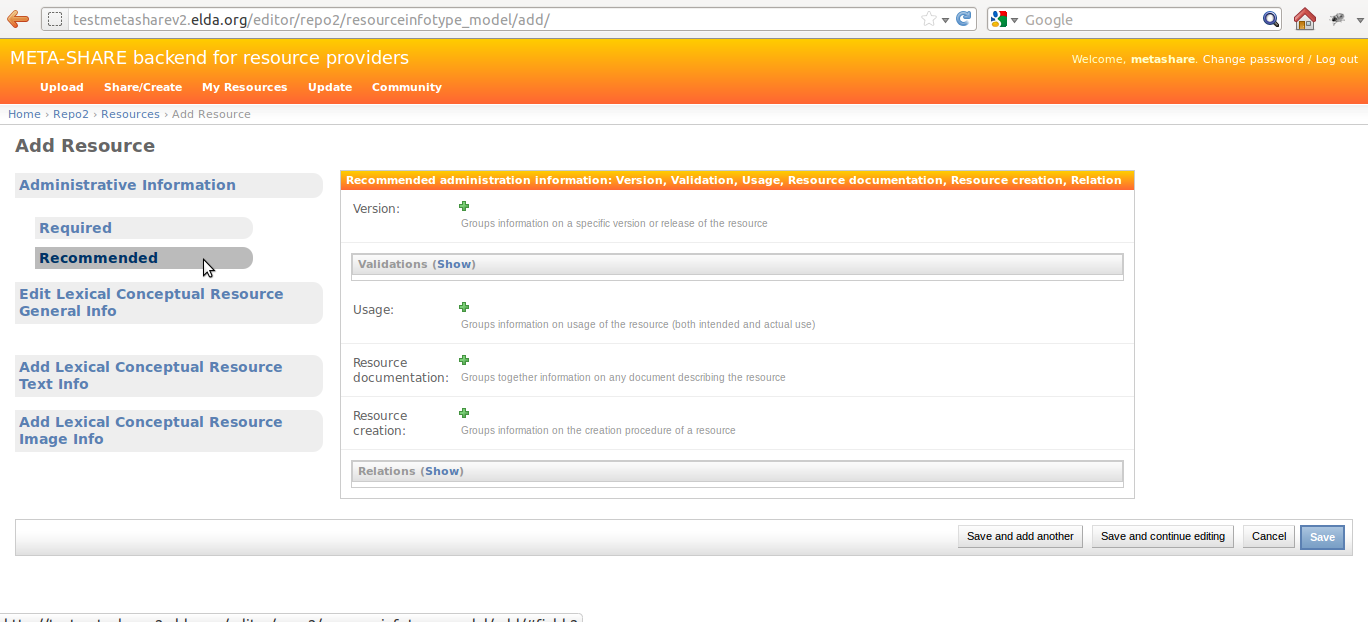
1. Choose a resource type for the LR. (A LR can only have one resource type). If your choice is “tool/service”, then click the “proceed” button.

If your choice is one of the other three resource types, select from the list that appears, the media type(s) for the LR.

Note: In the case of “lexical conceptual resource” and “language description” resource types, “text” is a required media type and, as such, it is selected by default.

1. Click on the “proceed” button.

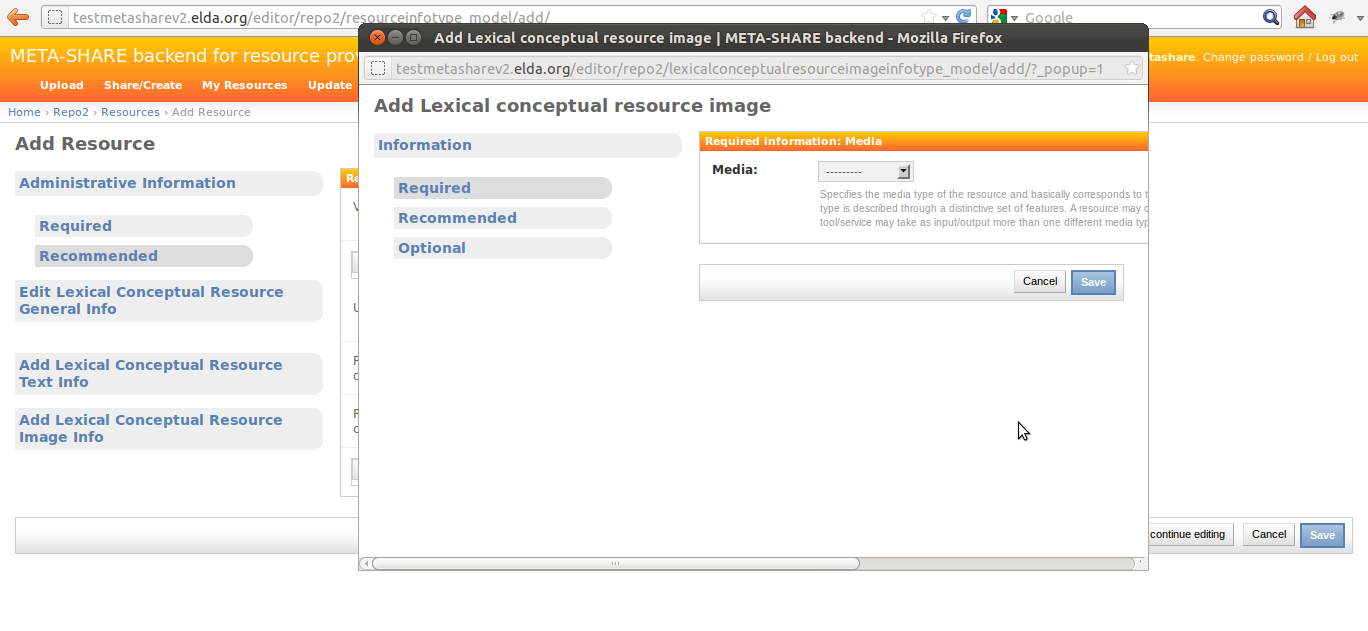
In the result page, all the metadata information to be filled in, is summarized under a left hand side menu (see section 2.4 for the description of the metadata fields). On the top of this menu, there is the administrative information categorized as required, recommended (and optional). A click on any of these options, will lead in expanding the appropriate metadata block on the right hand side (see Figure 2).



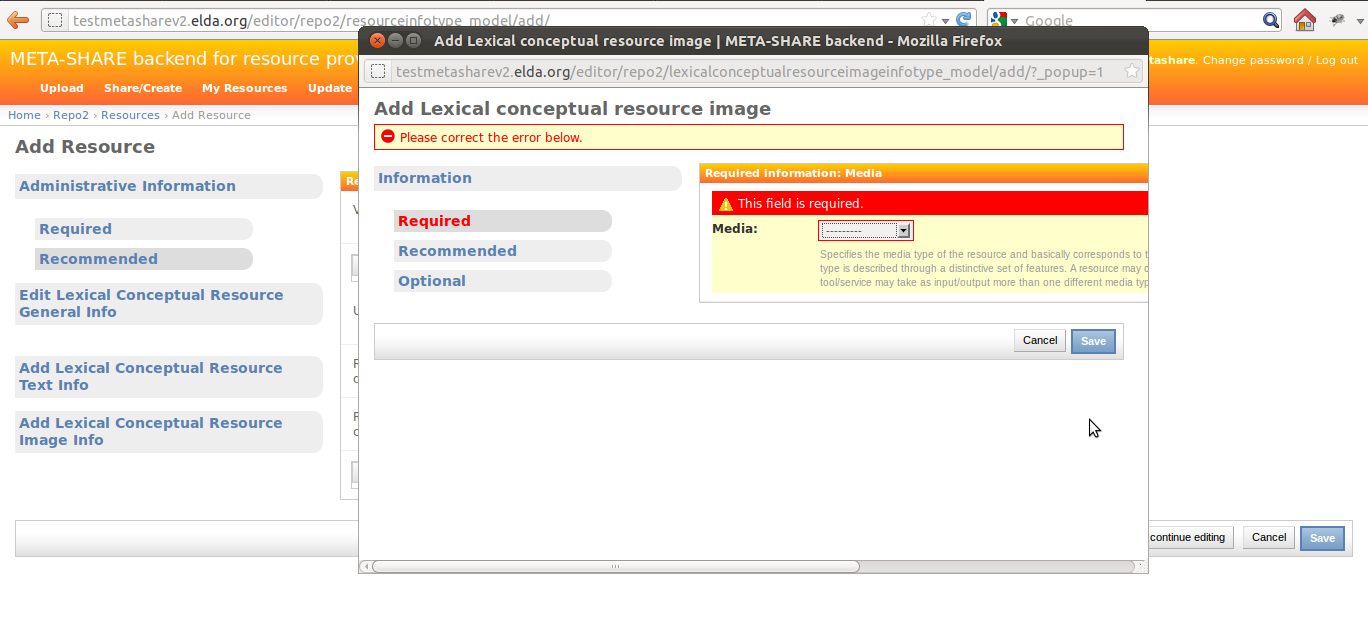
**Figure 2: Administrative Information metadata blocks**

1. Fill in the administrative metadata fields. (Only the required information is necessary for a LR to be successfully saved.)

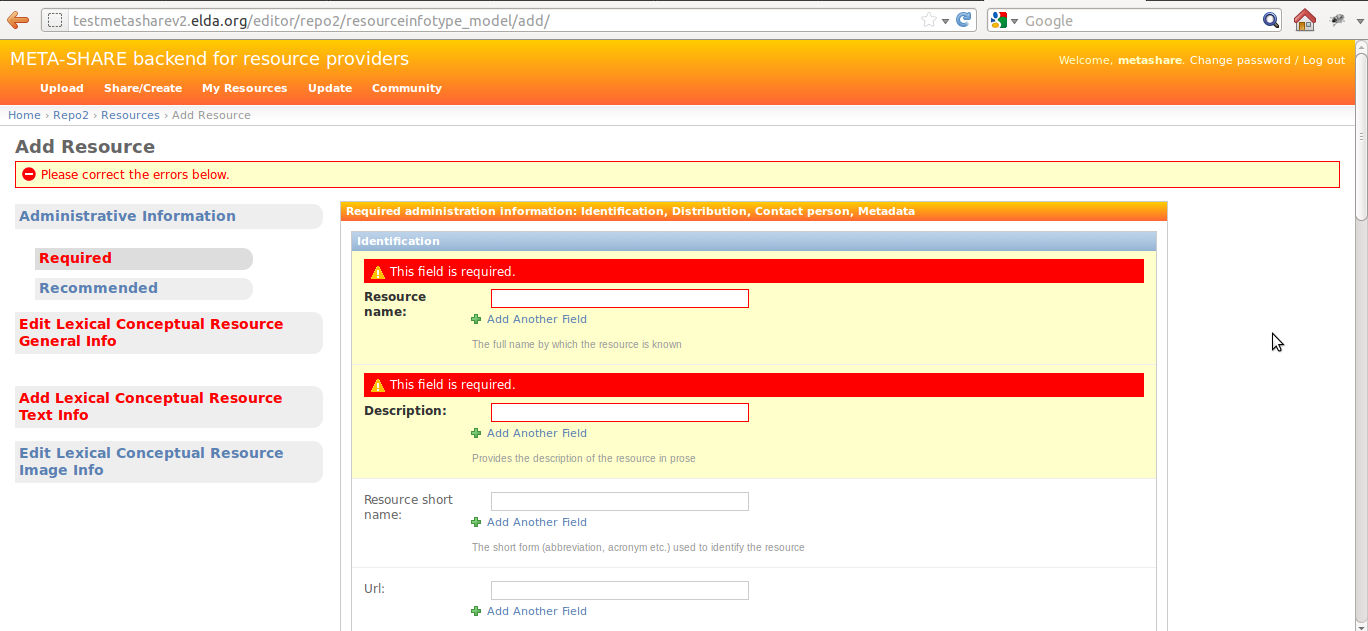
On the same level of the left hand side menu as the administrative information, there are the resource-type-specific information. Clicking on one of these, will leed in opening a popup window, containing the blocks of metadata to be filled in (see Figure 3 for an example). These popup windows follow a similar menu structure (required/recommended/optional) as the main page.

**Figure 3: Resource-type-specific popup window**

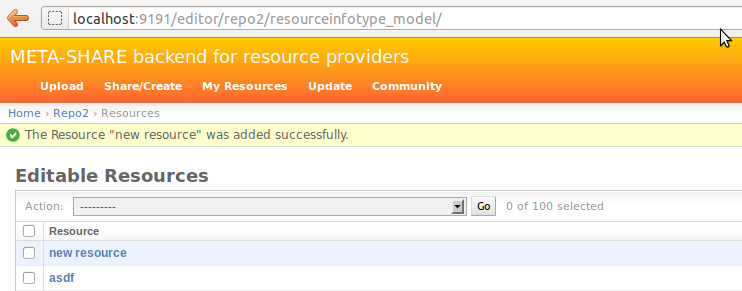
1. For each one of the popups, fill in the metadata fields. (Again, only the required information is necessary for a LR to be successfully saved.) Once done with filling in information within a popup window, click on the “Save” button at the bottom of the page. In case of errors, a message will be displayed on the top of the page, the error will be reflected on the menu option and the problematic fields (i.e. fields which either contain an error or are empty) will be highlighted in red (see Figure 4 for an example). If there are no errors, the “Save” button will close the popup window.

**Figure 4: Errors in popup window**

1. Once done with filling in information both in the main page and in all the popup windows, click on the “Save” button at the bottom of the main page. If there are errors, they will be indicated in a similar way as in the popup windows (see **Figure** 5 for an example). If there are no errors, clicking on “Save” will result in a new page, displaying the message “The Resource 'resource name' was added successfully”, and listing all the LRs added in META-SHARE, included the newly created one (see **Figure** 6).



**Figure 5: Errors in main window**



**Figure 6: Successfully saved LR**

## Manage Language Resources

The provider can rview, edit, delete and publish/unpublish/ingest all the resources he has provided in META-SHARE. The user can access the management page for his LRs by clicking on the “My Resources” link at the top left of the “META-SHARE backend for resource providers” page. The result page will list all the provider's LRs.

The provider an action from the dropdown menu of actions (delete/publish/unpublish/ingest) and then select the LR(s) he wants to apply this action to. Clicking on “Go button” will execute the selected action.

If the provider wishes to edit a LR, he can click on its name. This will result in a page where he can edit all the metadata fields for this resource, in exactly the same way as described in section 2.2.

## Add and change metadata information

Metadata information are organised in two levels, the first being Administrative / Type-specific, and the second being required / recommended / optional.

The tabbed menu described in section 2.2 is making the two levels clear to the provider. A description of the metadata fields is following:

* + Content: Information on the content of the resource;
  + Identification: Information needed to identify the resource;
  + Metadata: Information on the metadata record itself;
  + Distribution: Information on the distribution of the resource;
  + Person: Information on the contact person of a resource;
  + Usage: Information on usage (both intended and actual use);
  + Version: Information on version/release of the resource;
  + ResourceCreation: Information on the creation procedure, tools, etc. of a resource;
  + ResourceDocumentation: Information on relevant documentation (papers, etc.) of the resource;
  + LexicalConceptualResource: Information specific to lexical/conceptual resources, such as the type of lexical/conceptual resource (word list, thesaurus, ontology, etc.) as well as further details about the creation of the resource;
  + Text: Information on the text component of a corpus/resource, such as the language, the text format, the encoding or the annotation;
  + Validation: Information on the validation of a resource.

In order to insert or modify metadata, the provider has an access to different kinds of fields:

* Text field: the provider may add/modify free text that corresponds to the description of the metadata elements;
* Drop-down list box: the provider may choose one item from the list provided;
* Multiple-choice Drop-down list box: the provider may chose one or several items from the list provided;
* Repeatable text field: the provider may add or delete all the text field he/she wants to by clicking on the “+” or “–“ buttons, respectively.

## Import Language Resources

### Prepare the import file

To be able to import new LRs, the provider must create XML files for each LR that follow the META-SHARE metadata. This compliance can be validated using the META-NET XML schema (XSD), the current version being under the misc/schema/v2.0 folder.

### Upload/Import in META-SHARE

A provider can upload an XML metadata file by clicking on the “Upload” link at the top left of the “META-SHARE backend for resource providers” page. Uploading works as follows:

* Choose a file by clicking on the “Browse button”
* Click on “Upload Terms” checkbox to accept them.
* Click on the “Upload” button.

Importing XML metadata is also performed from the command line. It describes the import of existing LRs, but the process is the same for LRs converted from different representations as described below.

#### Import script

Importing of resources is possible using import\_xml.py which can be found within the metashare installation folder.  
Basic usage information is given when called without any arguments:

$ python import\_xml.py

usage: import\_xml.py <file.xml|archive.zip> \

[<file.xml|archive.zip> ...]

The script reports errors and gives import statistics when finished.

A single xml file can be imported as follows:

$ python import\_xml.py my\_LR.xml

Done. Successfully imported 1 files into the database, \

errors occurred in 0 cases.

Multiple LRs can be imported using the “\*” wildcard or by packaging the xml file into a zip file:

$ python import\_xml.py metashare-v1-resources.zip

Done. Successfully imported 1,277 files into the database, \

errors occurred in 0 cases.

Importing a large file takes some time; this is normal.

After successful import of these files, you should have the new resources available from your META-SHARE node.  
The last line of the import\_xml.py output should report no errors.

### Conversion of existing LRs

A tool has been specifically developed within META-NET to help providers convert their own LRs from a specific metadata model to the META-SHARE one. This tool can be found in the tools below the META-SHARE installation folder.

The conversion of the existing LRs from several different schemata to META-SHARE schema has been made with xsl transformations.

First step was to map one by one all the elements of old schemata to the elements of META-SHARE schema.

In some cases the tool Altova-MapForce was used for the mapping between schemata and the production of the xsl files.

## Add Downloads to Resources

In previous versions of META-SHARE, we used the downloadLocation attribute to encode the URL from which a resource could be downloaded.

Due to the addition of a dedicated *storage layer*, the handling of downloads has changed. Each StorageObject has a local folder attached to it which is used to store associated downloadable data. For version v1, only *one* downloadable file can be attached to a resource; its name is fixed to archive.{'zip', 'tar.gz', 'gz', 'tgz', 'tar', 'bzip2'}.

### Adding a Download

You can add a download to an existing resource by following these steps (in order, of course):

1. Login as staff user
2. Access the LR view, you should see the StorageObject's identifier on top (64-character hex String).  
   We will refer to this id as YOUR\_ID in subsequent steps.
3. Inside your META-SHARE folder, call

$ python storage\_admin.py folder YOUR\_ID

This will print the storage folder path for the respective resource.

1. Copy your archive to the storage folder and update the corresponding checksum using

$ python storage\_admin.py checksum YOUR\_ID

1. You should now see the Download button, even for non-staff users.

### Known Limitations

It is mandatory to follow the naming scheme. Make sure that your downloadable data is contained within *one* archive file named archive.{'zip', 'tar.gz', 'gz', 'tgz', 'tar', 'bzip2'} .