



ELRC-SHARE Repository Guidelines for CEF generic services projects



Author(s): Miltos Deligiannis (ILSP)
Stelios Piperidis (ILSP)
Maria Giagkou (ILSP)

Dissemination Level: Public

Version No.: V1.0

Date: 2020-02-08



Contents

1	<u>Introduction</u>	3
2	<u>Guidelines</u>	3
2.1	Registration	3
2.2	Contribution of resources	4
2.3	Description of resources: metadata completion	4
2.4	Publication of resource descriptions	5

1 Introduction

ELRC-SHARE (<https://elrc-share.eu>) is the European Commission's repository of language data and tools for the purposes of the CEF Automated Translation platform.

This document is addressed to the consortia members of CEF-funded projects (generic services projects). It is a concise step-by-step guide to contributing language resources through the ELRC-SHARE repository. Detailed guidelines and the documentation of the metadata schema are also available on the [ELRC-SHARE Help webpage](#) and appropriately linked to herein, where relevant.



Important notes:

- 1. The responsibility to upload, appropriately describe the datasets with metadata, and publish metadata records that can be browsed through the ELRC-SHARE public inventory lies with the CEF-funded project itself.**
- 2. ELRC will by no means edit or curate an external project's datasets and/or metadata hosted in ELRC-SHARE.**
- 3. ELRC and ELRC-SHARE do not impose any criteria on the size, content and quality of the datasets delivered by the CEF generic services projects. All these aspects are exclusively addressed by the projects' grant agreements.**
- 4. ELRC and ELRC-SHARE will not perform any kind of content validation of the uploaded datasets, except for a check for malicious content.**
- 5. ELRC will not perform any legal validation of the uploaded datasets. The licence, terms and conditions under which the datasets are made available through the ELRC-SHARE are under the full control and responsibility of the depositor.**
- 6. ELRC-SHARE cannot officially issue certificates of delivery for the projects' contracting authority. It does not substitute or replace the formal delivery mechanisms specified in the projects' grant agreements.**

2 Guidelines

2.1 Registration

1. Register to ELRC-SHARE through the respective menu item on <https://elrc-share.eu>. Note that all project team members who wish to use ELRC-SHARE for contributing resources should be registered users.
2. Contact the ELRC-SHARE repository administrators at elrc-share@ilsp.gr to request **Editor** access rights. State the username you created during registration, your affiliation and the CEF-funded project you represent. If more than one users have

registered for a project, indicate the user who will act as the Editors group manager on ELRC-SHARE¹.

2.2 Contribution of resources

Datasets / Language Resources can be contributed to the ELRC-SHARE repository through the following three alternative channels:

- a. the online [Data Contribution Form](#) (see detailed guidelines in the [Walkthrough for Contributors](#)) or
- b. the [ELRC-SHARE documentation environment](#) (see detailed guidelines in the [Walkthrough for Editors](#)) or
- c. the **ELRC-SHARE-client API** (the API and its documentation are available on [Gitlab](#))

Options (a) and (b) above are recommended for creating a single language resource at a time. Option (c) is recommended for batch creating multiple resources or for batch updating the metadata of your previously created resources.

2.3 Description of resources: metadata completion

The ELRC-SHARE metadata schema is an application profile of the META-SHARE schema, appropriately modified for the requirements of ELRC. The full metadata schema and its documentation are available [here](#).

Once created through any of the channels described in section 2.2 above, the contributed resources can be further described through the ELRC-SHARE metadata editor:

1. Access the full list of your resources from the top menu Manage resources -> Manage your own resources.
2. Select any of the resources
3. Edit the metadata² and save.
4. Publish the metadata record (see section 2.4 below)

Important note: The following metadata elements are either mandatory or highly recommended and should be filled in:

- Resource name
- Resource description
- Distribution elements group (where the Licence(s), the IPR owner and any information on the terms and conditions under which the resource is being shared are defined)

¹ The Editors group manager has full metadata editing rights to all the resources created by the group members.

² If the contribution was performed through the online contribution form, the info you have provided in the form automatically populates the metadata record (apart from the legal tags), but typically you will need to provide more information for an as complete as possible metadata record.

- **Funding project** (from left menu Administrative information -> Recommended): This is very important for the traceability of your contribution
- Size and data format (for corpora and lexical resources)

Detailed guidelines for the ELRC-SHARE documentation environment are available in the [Walkthrough for Editors](#).

2.4 Publication of resource descriptions

Upon their creation, the metadata records are automatically set to an “Internal” publication status, i.e. they are only accessible by and visible to their creator(s). In order to make the resource description(s) visible through the ELRC-SHARE public inventory, you need to change their publication status:

1. Access the full list of your resources from the top menu Manage resources -> Manage your own resources.
2. Select one or more internal resources from the list
3. From the “Action” drop-down menu select “Ingest selected internal resources”
4. Select one or more ingested resources from the list
5. From the “Action” drop-down menu select “Publish selected ingested resources”

Note: It is important that you publish your resource(s) description(s), if you would like your contracting authority to verify the delivery of your resources to ELRC-SHARE.