Project acronym: LISE

Grant Agreement Number: 270917

Project title: Legal Languages Interoperability Services

DELIVERABLE

D4.2 Evaluation Plan

Dissemination Level: PU

Version No. 1.1

25/07/2011

Project co-funded by the European Commission within the ICT Policy Support Programme.

L!SE

Document Information

Deliverable Number	D4.2
Deliverable title	Evaluation Plan
Due date of deliverable according to DOW	M6
Actual submission date of deliverable	29^{th} July 2011
Main Author(s)	Joeri Van de Walle
Participants	
Reviewer	Gerhard Budin
Work package	WP4
Work package leader	CL
Dissemination level	PU
Version	1.1

Revision History

Revision	**Date**	**Author**	**Organisation**	**Description**
1.0	**2011-07-07**	**Joeri Van de Walle**	**CL**	**Initial draft**
1.1	**2011-07-25**	**Joeri Van de Walle**	**CL**	**Incorporated feedback from partners**

Statement of originality:
This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.

Contents

1. Purpose of This Document

The purpose of this document is to describe how the outcome of the tool development of the LISE project will be evaluated.

As the details of the Evaluation Plan depend heavily on the outcome of a number of deliverables from other work packages (see below), the consortium partners decided during the EMB meeting of July 4th that the Evaluation Plan will be delivered in two phases. The first version will be delivered at M6 (July 2011, this version), as stipulated in the Description of Work, and the final version will be delivered at M18 (July 2012).

This version of the Evaluation Plan will contain a high-level description of the evaluation and validation strategy the consortium intends to adopt for the LISE project.

The final version of the deliverable will contain the detailed Evaluation Plan that will draw on the outcome of the following deliverables from work packages 2 and 3:

- D2.2.2 Second LISE Service Version (M12)
- D3.1 Report Analysis of existing Terminology Workflows (M12)
- D3.2 Report Workflow adaption for LISE (M18)

The Evaluation Plan that will be delivered at M18 will describe the adopted evaluation and validation strategy in relation to these deliverables.

2. High-level Description of the Approach

The two key elements in our evaluation approach will be user profiles and workflow steps. In this document we will start by listing the different user profiles that the consortium envisions may benefit from the proposed tools and terminologies. We will then analyse the workflows described in D3.1 and break them down into workflow steps. The next step will be to create a matrix of user profiles and workflow steps that shows which workflow steps are potentially relevant for which user profiles.

In this matrix, workflow steps supported by the LISE tools will be marked and use cases will be developed in which the tools have a role to play. A detailed description of each use case will illustrate how LISE tools can be applied to assist in the successful implementation of the workflow steps in question. For each workflow step a number of success criteria will be defined.

In addition to the profiles/workflow steps matrix, a questionnaire will be developed to help collect feedback about the requirements and tools.

The evaluation of the LISE tools will then consist of having a well-targeted group of users fill out the questionnaire and enact the workflows illustrated by the use cases.

The outcome of the evaluation will be a report that will contain both quantitative and qualitative information about the tools and services offered by LISE.

3. User Profiles

In this section we will list the user profiles that we think may be involved in terminology-related activities. User profiles listed here are provisional. The outcome of the interviews that will be conducted in the context of WP3 will determine whether this list is accurate or will need to be updated.

To describe each of the different profiles, we will give list a number of activities that are typical for the profile. At the same time, these workflow steps will already hint at the workflow steps that the profile requires.

At this time we distinguish between the following profiles:

- Terminologists
- Subject matter experts (non-terminologists)
- Translators/reviewers
- Terminology consumers

Note that in practice, the activities performed by these types of users may overlap. For example, translators/reviewers may do more than just use terminology in the context of their translation work. They may also be involved in the validation process if no subject matter experts are available. In the sections that follow, however, we try to list those activities that can be seen as the defining attributes for each profile, the characteristics that set them apart from other profiles.

1. Terminologists

By terminologists we understand users for whom working with terminology in all its facets is a core activity. This type of users usually has an in-depth knowledge of terminology as a concept and all matters relating to it. Terminologists have extensive experience with a wide variety of tasks that have to do with terminology.

Typical terminology-related tasks terminologists perform include:

- Defining terminology database models (term data, meta data to be included)
- Collecting (mono-, bi- or multi-lingual) term data for a specific subject field or customer)
- Cleaning up existing terminology databases (remove duplicate entries, entries that are too general,…)
- Migrating terminology data from one repository to another
- Devising validation procedures for new/existing terminology
- Devising access models to terminology data

- …

2. Subject matter experts

By subject matter experts we understand users that are experts in a specific subject field for their language and that are therefore often solicited as validators for verifying correctness and/or accuracy of existing or newly extracted terminology. This type of users is not necessarily familiar with terminology- or terminology management-related concepts, but they are typically very familiar with the terminology itself.

Typical terminology-related tasks subject matter experts perform include:

- Validating term data for a specific subject field or domain in a specific language
- Harmonising data coming from different sources for a specific subject field or domain in a specific language
- Standardise terms, definitions, term statuses, or notes of existing terminology

3. Translators/reviewers

By translators/reviewers we understand users that are at least bi-lingual and that work with language and terminology on a daily basis. Translators/reviewers are not only proficient in at least two languages; they may or may not have in-depth knowledge about one or more subject fields as well. More technical or terminology-oriented translators/reviewers may be familiar with the concepts relevant to terminology and terminology management. Other translators/reviewers may not be familiar with them.

Typical terminology-related tasks translators/reviewers perform include:

- Using source terms and translated equivalents for translation/validation in a translation project
- Suggesting new term candidates with their corresponding translation

4. Terminology consumers

By terminology consumers we understand users that only occasionally consult terminology databases. They are the so-called ‘guest users’. They will look up terminology in repositories available to them when the need arises, but they are typically not involved in the organisation or maintenance of terminology. This type of user typically has little or no understanding of terminology-related concepts and terminology management. A typical example of this type of user may be a marketing staff member wanting to verify a translation for a certain term.

Typical terminology-related tasks translators/reviewers perform include:

- Looking up a term to find its translation or to find meta information about it

4. Workflow steps

In this section we will list the workflow steps that we think may be part of a terminology-related workflow. Workflow steps listed here are provisional. The outcome of the interviews that will be conducted in the context of WP3 will determine whether this list is accurate or will need to be updated.

Terminology workflows will usually involve activities that aim to support one or a combination of the following activities:

- Term acquisition: Collecting new terminology. New terminology can be extracted from existing textual resources or it may be suggested by domain experts. New terminology may also come into play in the case organisations that have existing terminology get merged.
- Terminology maintenance: Once terminology has been identified and recorded, maintenance starts. Proposed terminology will have to be validated. Existing terminology may have to be cleaned, harmonised, or standardised.
- Disseminating terminology: The goal of collecting terminology and keeping it clean is of course that other users may benefit from it when creating or translating texts. The focus of dissemination activities is to make terminology visible and accessible to end users, the terminology consumers.

Each of these activities may involve different workflows, each of which in turn can be subdivided into workflow steps. Below is a non-exhaustive list of possible workflow steps:

- Defining ontologies
- Defining termbase structure
- Extracting monolingual terminology
- Extracting bilingual terminology
- Adding translations for existing terms in additional languages
- Validating (approving/rejecting) newly extracted (proposed) terminology
- Validating (approving/rejecting) existing terminology
- Cleaning terminology (removing duplicates or irrelevant entries, correcting misspellings,…)
- Harmonising existing terminologies
- Standardising existing terminology entries
- Adding meta data to existing terminology
- Publishing approved terminology
- Searching for terms in a repository
- Documenting terminology strategy

This list will be adapted based on the outcome of deliverables D3.1 (Report Analysis of existing Terminology Workflows) and D3.2 (Report Workflow adaption for LISE).

5. Evaluation

Two types of evaluations will be performed in the LISE project:

- A t*echnical evaluation* will be carried out by Cross Language and will be performed in close cooperation with the development team at ESTeam.
- The *usability evaluation* will involve potential real users of the platform and will assess in how far the provided terminology and tools meet the requirements of their potential users.

1. Technical evaluation

The objective of the technical evaluation is to ensure that the software components that get built for the LISE project are stable and bug free.

In addition to the internal testing procedures performed by the development team, Cross Language will also perform user tests on the software using the use cases described in the functional specifications document as a guideline.

Any issues discovered during the user tests will be logged in the Case Tracker component of the LISE communication platform. This will allow the testers and developers to exchange information and keep track of the status of any issues that may be reported.

2. Usability evaluation

The objective of the usability evaluation is to ensure the tools and guidelines that get produced by the LISE project meet the requirements and expectations of their potential users.

The usability evaluation will mainly consist of a usability experiment. The basis for the experiment will be the matrix that will result from combining the identified user profiles and workflow steps described above.

For each user profile we will assess which workflow steps may potentially be relevant to it. Each of the steps will also be marked to indicate whether or not the software offered by the LISE project will support it or not.

From the matrix, a number of use cases will be derived that implement workflows that involve the LISE tools. For each use case a number of success criteria will be defined to help assess whether or not it achieved its goals.

We will then bring together a number of representative users of the LISE user group and have them put the use cases into practice for one or more profiles. This will probably take place during a workshop organised by the LISE consortium. We would opt for a workshop, because that would give us the opportunity to give a short introduction about the tools and how they work and would allow us to assist the users in case they run into problems or have questions.

After enacting the use cases, participants will be asked questions to assess whether or not the success criteria that were defined at the outset of the experiment were met.

The usability experiment will be conducted twice in the course of the project: once around mid project (M15 or around that time) and once before project completion (M30).