

Title CMDI 1.2 rollout plan
Version 1
Author(s) Twan Goosen (CLARIN ERIC)
 Menzo Windhouwer (Meertens
 institute)
Date 2015-11-23
Status Draft
Distribution Public
ID CE-2015-0702



1 Planning

Year	2015			2016				
Month	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May
Toolkit								
OAI Harvester								
CLAVAS								
VLO								
Component Registry								
CMDI 1.2 release								
Specification								

2 Work

2.1 Specification

The specification for CMDI 1.2 is currently being developed by members of the CMDI taskforce in parallel to the development of the toolkit. This work can be completed in parallel without any other task within the scope of this document being dependent on it except for the final release of CMDI 1.2

This work will be done by the members of the CMDI task force and Thorsten Trippel.

2.2 Toolkit

Work on the toolkit consists out of writing the general component schema, stylesheets for conversion to profile schemas as well as stylesheets and other tools for conversion to and from the new version of CMDI. It provides the essentials creating CMDI components, profiles and instances.

A large part of this work has already been done, but some effort is still required to finish it.

In addition to the toolkit itself, a test suite is being developed simultaneously. This currently requires some extension in order to gain reasonable test coverage.

The primary 'client' of the toolkit is the Component Registry. Essentially this tool serves as a wrapper around the elements of the toolkit and exposes its functionalities (validation, schema generation, conversion) through a public interface.

Work on the toolkit has primarily been carried out by Menzo Windhouwer. Mitchell Seaton will also carry out some development work starting November 2015. Other people can contribute to the test suite by providing test cases.

2.3 VLO

When CMDI 1.2 is released, it needs to be supported by the VLO as from that point on CMDI 1.2 metadata records can come in via OAI-PMH and these should not be rejected by the VLO importer. The importer needs to be adapted to deal with the new paths and namespaces found in profile schemas and instance documents. Support for CMDI 1.1 can be kept by having the harvester convert all incoming records to CMDI 1.2 and letting the VLO importer operate on the resulting files.

On the front end side, relatively little adaptation is needed. The only part of the front end that accesses the metadata records directly is the rendering of the full record (the 'show all metadata' option)

The CMDI 1.2 related work on the VLO is ideally carried out when the toolkit has released a stable state. No other components in the scope of this document depend on these changes in the VLO. The version of the VLO with support for CMDI 1.2 can be released before the release of CMDI 1.2 itself provided that the OAI provider already ensures the conversion of CMDI 1.1 records. A work plan for the VLO up till version 4.0 has been made in the context of the CLARIN PLUS project. The changes regarding CMDI 1.2 are covered in this work plan¹.

Work on the VLO related to CMDI 1.2 will be carried out by Menzo Windhouwer, Thomas Eckart and/or Twan Goosen.

2.4 CLAVAS

CLAVAS has been developed in CLARIN-NL in 2012 - 2013, but has not seen much actual use. However, in CMDI 1.2 CLAVAS will have a clear role as a provider of open/closed external vocabularies. This will certainly spark interest in its use and we need to fresh CLAVAS up so it can meet these new demands. This includes:

- cleanup and refresh of the current vocabularies (ISO 639-3 language codes, organization names and ISocat metadata data categories), where the ISocat vocabulary should probably be deleted
- upgrade to the latest version of OpenSKOS, which is also used by the CCR

Another issue is that we can expect requests by the CMDI community to add new vocabularies to CLAVAS, but there is currently no process in place to deal with such requests. So we need to think about what would be the requirements for these vocabularies, who makes decisions about them and who does the actual work of importing them into CLAVAS.

This work on CLAVAS will be carried out by Menzo Windhouwer.

2.5 Component Registry

The Component Registry web application serves two purposes: first, it's a wrapper around the elements of the toolkit and exposes its functionalities (validation, schema generation, conversion) through a public REST interface. Second, it provides a user interface (UI) for browsing existing components and creating new ones.

¹ TODO: URL for VLO CLARIN-PLUS specification

The REST interface needs to be extended with a number of new calls that allow for the retrieval of components, profiles and schemata that comply with either CMDI 1.1 or 1.2. Since all specifications will be stored as CMDI 1.2 (existing specifications will be converted once), it must dynamically convert these to 1.1 versions using the “downgrading” stylesheet from the toolkit upon request.

In the UI, both the browser and editor parts need to be adapted. The browser needs to represent the ‘lifecycle status’ of components and profiles, and the editor needs to be made to generate valid CMDI 1.2 specifications and extended to account for the following new features:

- External vocabularies
- Mandatory attributes
- Documentation: localised and at all levels (components and attributes)
- Cues for tools [possibly in a later version]
- Auto value statements [possibly in a later version]

This work on the Component Registry can begin after the toolkit has been functionally completed and the new front end, which is currently under development, has reached a stable state as well.

The Component Registry is developed by Twan Goosen.

2.6 OAI harvester

The OAI harvester can help in the upgrade of the infrastructure to CMDI 1.2 in various ways:

- convert all the CMDI it harvests also to 1.2, which can be done by plugging in the upgrade XSLT in its configuration
- upgrade the XSLT that converts DC or OLAC metadata to CMDI to produce CMDI 1.2

The latter is covered by the first as well, but it is good to now move core parts of the infrastructure completely to CMDI 1.2 and don’t leave legacy in place, which will keep on depending on the upgrade path.

The OAI harvester is developed by Menzo Windhouwer.

2.7 Release

Since CMDI 1.2 is not a single tool or specification, its ‘release’ in fact comes down to a number of steps that need to be taken more or less simultaneously:

- Tagging of the 1.2 version of the CMDI toolkit and making this frozen version the ‘current’ version available at <https://infra.clarin.eu/cmd/1.2>²
- Upgrade of all component and profile definitions within the Component Registry database to CMDI 1.2
- Deployment of the new version of the Component Registry that uses CMDI 1.2

² <https://infra.clarin.eu/cmd/> will keep resolving to the 1.1 toolkit (as an alias to <https://infra.clarin.eu/cmd/1.1>) so as not to break existing metadata records and tools that are not adapted to the new versioning scheme

- *Deployment of the version of the OAI harvester that generates CMDI 1.2 records for incoming CMDI 1.1. metadata
- *Deployment of the version of the VLO that supports CMDI 1.2 records

The last two steps can take place before the actual release as they will not affect the handling of CMDI 1.1 records or any other parts of the infrastructure.

All of these steps need to be carried out (due to permissions and knowledge of the target environments) by CLARIN system administrators, i.e. Willem Elbers and/or Sander Maijers.

Detailed planning

