Title	CLARIN Virtual Collection Registry (VCR) Specification
Version	1.0
Author(s)	Willem Elbers
Date	2015-12-31
Status	Final
Distribution	Public
ID	CE-2015-0700

1. Introduction

The virtual collection registry has been originally developed at the IDS in Mannheim and was further developed by the CLARIN ERIC into the current version which is running in production at the the IDS. The goal within the CLARIN+ project is to evaluate the current version and investigate how we can best improve the usability of the system for end user's, by making the workflow of creating and managing a virtual collection as easy as possible. A second goal is the look into the integration of the VCR with the Language Resource Switchboard.

The CLARIN+ project proposal already includes some requirements in order to improve the usability, but a review of the data model to see where complexity can be reduced is needed as well. A more simple data model will result in an easier to use and less complex user interface and an optimized workflow. This enables us to better promote the virtual collection registry.

The high level requirements, which follow from the CLARIN+ workplan and feedback received from users, are the following:

R1: Automatic resource type recognition when adding a digital object to a virtual collection.
R2: Analyse possibilities to integrate the use of ORCID user identifiers
R3: Analyse possibilities to publish virtual collection metadata to the DataCite metadata store.
R4: Improve the robustness and usability, especially providing a lightweight procedure for authenticated users to create a new virtual collection.
R5: Add an administration interface to the backend, to enable the correction of small errors in existing virtual collections.
R6: Allow users to process cited data objects with the Language Resource Switchboard (outcome of subtask 2.3.1).
R7: Clarin.eu vhost (e.g. vcr.clarin.eu).
R8: CLARIN branding.

2. Planning

Year	2016												2017						
Month	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6	7
Usability	▓	▓	▓	▓															
M1 alpa	▓	▓	▓	▓															
M1 beta / rc					▓														
M1 stable						M													
M2 alpha													▓	▓	▓	▓			
M2 beta																▓	▓		
M2 rc																		▓	
M2 stable																			D

Usability
This task consists of two parts. (1) Assess the data model complexity and investigate ways to reduce complexity. The expectation is that this will result and a simplification of the user interface. (2) Assess the user interface and investigate improvements to improve the user experience. Possible improvements can be (but are not limited to) choosing sensible default settings with the possibility of overriding, hiding complex options where possible and providing a simple default, automatic detection of mime types (controlled vocabulary, must be editable by the user), investigate using markdown in the description field, ...

Output: The output of this task is a short document outlining the results of both assessments and a list of tickets with concrete task descriptions which will be taken up in M2.

This task runs in parallel with the development phase of the M1 alpha version.

Milestones
Both milestones consist of three phases: (1) alpha, this where development takes place, (2) beta, this is where development of new features is stopped and only testing and bug fixes takes place and finally (3) stable where no development should take place, documentation is finished and the production deployment is done.

Milestone 1 (M1): VirtualCollectionRegistry-1.1
This task aims to deliver the first release of the VCR focussed on R4, R5, R7 and R8. In this task a selection of high priority tickets which currently exist for the VCR will be implemented, see appendix A. This task focuses on an intermediate release and therefore the beta/rc phase is

short to allow for maximum development time. This task will also take care of the migration of the codebase to GitHub.

Output: The output of this task is the tagged codebase on GitHub (with associated documentation) and the release deployed in production.

Milestone 2 (M2): VirtualCollectionRegistry-1.2
This task aims to deliver the second and final release of the VCR, focussing on R1, R2, R3 and R6. In this task the output of the usability assessment will be included together with a set of tickets defined in appendix B. This task has a more extensive beta and rc phase to allow for better testing and more time to fix bugs in order to deliver the final production release and main deliverable for this task.

Output: The output of this task is the tagged codebase on GitHub (with associated documentation) and the release deployed in production.

Appendix A: Milestone 1 tickets

Ticket	Summary	Priority	
#602	Layout tweaking	minor	R5
#603	Auto-recognize links in description field	minor	-
#605	Make it possible to re-use authors that have been entered before	minor	R5
#609	check validity reference links	major	-
#615	Add a 'frozen' state for public collections	major	-
#623	Upgrade Wicket	major	-
#626	Bootstrap styled pagination controls	minor	-
#627	Help tooltips on activation of text input	minor	R4
#653	Admins should be able to delete published collections	critical	R5
#654	Improve dealing with failed PID minting	critical	-
#677	Add an eye-catching Cite this button major	major	R4

#685	Make display of 'beta' badge in UI configurable through context parameter	trivial	-
#733	details does not work in Firefox 35	minor	-
#747	Include xml.xsd schema in source tree	major	-
#748	Inject OAIProvider instance	major	-
#766	Add canonical link to base page	minor	-
#838	Default settings in UI	major	R4
#839	Administration UI	major	R5
#840	CLARIN branding	major	R8
#841	vcr.clarin.eu subdomain	minor	R7

Appendix B: Milestone 2 tickets

Ticket	Summary	Priority	
#611	prefill the mimetype for ResourceProxy	major	R1
#612	add creator lookup based on ORCID minor	major	R2
#837	Automatically detect mimetype of references	major	R1
#842	DataCite metadata store	major	R3
#843	Language Resource Switchboard	major	R6