



D4.2

Joint infrastructure services

Document information

Title	Joint infrastructure services
ID	CLARINPLUS-D4.2 (CE-2017-0985)
Author(s)	Claus Zinn, Twan Goosen, Marie Hinrichs, Emanuel Dima, Willem Elbers, Dieter Van Uytvanck, Dirk Goldhahn, Thorsten Trippel, Jozef Mišutka
Responsible WP leader	Erhard Hinrichs
Contractual Delivery Date	2017-02-28
Actual Delivery Date	2017-02-28
Distribution	Public
Document status in workplan	Deliverable

Project information

Project name	CLARIN-PLUS
Project number	676529
Call	H2020-INFRADEV-1-2015-1
Duration	2015-09-01 – 2017-08-31
Website	www.clarin.eu
Contact address	contact-clarinplus@clarin.eu

Table of contents

1	Executive Summary	2
2	Introduction	3
3	Cooperation with EUDAT	4
3.1	B2ACCESS and B2SAFE	4
3.1.1	B2ACCESS	4
3.1.2	B2SAFE	5
3.2	B2DROP and the CLARIN Language Resource Switchboard	6
3.2.1	B2DROP	6
3.2.2	The CLARIN Language Resource Switchboard	7
3.2.3	Goals for Connecting the LRS with B2DROP	7
3.2.4	Using B2DROP as Alternative to the MPG server	7
3.2.5	Creating a bridge between B2DROP and the Switchboard	8
3.3	EUDAT's General Execution Framework.....	9
4	Cooperation with other e-Research infrastructures	9
4.1	GÉANT	9
4.2	RDA	10
4.2.1	Data Foundation and Terminology (DFT).....	10
4.2.2	Dynamic Data Citation	10
4.2.3	RDA/WDS Certification of Digital Repositories.....	10
4.2.4	Federated Identity Management.....	10
4.2.5	Legal Interoperability	10
4.2.6	Data Fabric.....	10
4.2.7	Domain Repositories	11
4.2.8	Group of European Data Experts (GEDE).....	11
4.3	DARIAH	11
4.4	EUROPEANA	12
4.5	LAPPS Grid.....	13
5	Conclusion	14
	References	15

1 Executive Summary

The CLARIN-PLUS partners actively follow the progress of a number of other neighbouring infrastructure projects. To support the cross-fertilization between projects, it has been our aim to implement joint services with these infrastructures, whenever possible. In this deliverable, we report on joint work carried out in cooperation with EUDAT, GÉANT, RDA, DARIAH, EUROPEANA, and the LAPPS Grid. The joint work touches issues such as secure authentication and authorisation, data management policies, trusted exchange of data, sharing of metadata for language-related resources, and the availability of tools across communities.

2 Introduction

This deliverable reports on progress on work package 4 of the CLARIN-PLUS proposal. In brief, the objectives for WP4 are to:

- strengthen the ties with other research infrastructure initiatives, inside and outside the EU;
- obtain a higher degree of synergy, by re-using other infrastructural services within CLARIN and by promoting the usage of CLARIN services in different contexts;
- enhance the visibility of CLARIN's infrastructure, paving the way for future collaborations and eventually the growth of the CLARIN ERIC member base (see WP5); and
- provide more and better services to the CLARIN user community.

In this deliverable, we report the work toward achieving the objectives with regard to the European infrastructure projects EUDAT, GÉANT, DARIAH, EUROPEANA as well as RDA and the LAPPS Grid.

3 Cooperation with EUDAT

In this section, we discuss the use of EUDAT's B2 services for the CLARIN infrastructure. In Section 3.1, we focus on our contributions to B2ACCESS and B2SAFE. Section 3.2 describes the interface between B2DROP and the CLARIN Language Resource Switchboard. In Section 3.3, we discuss the ongoing cooperation between EUDAT and CLARIN staff to develop the General Execution Framework (GEF), and the inclusion of or access to CLARIN tools.

3.1 B2ACCESS and B2SAFE

3.1.1 B2ACCESS

B2ACCESS service is “an easy-to-use and secure Authentication and Authorization platform developed by EUDAT. B2ACCESS is versatile and can be integrated with any service. When B2ACCESS is integrated with a given service, the user may log in by using different methods of authentication”, see <https://eudat.eu/services/b2access>. An architectural overview of the B2ACCESS service is shown in Figure 1.

To give the CLARIN user access to the EUDAT infrastructure, the CLARIN identity provider (IdP) has been integrated with B2ACCESS. Now, CLARIN users can use their account to access EUDAT services, *e.g.*, to deposit data with B2SHARE.

B2ACCESS has been integrated with the eduGAIN¹ IdPs as well. However, this integration is suffering from eduGAIN's opt-in policy that prevents users to log in to the EUDAT services with their home organization account if their NREN² has an opt-in policy and their IdP did not opt in to B2ACCESS yet. This issue is one of the things solved by the CLARIN service provider federation (SPF).

Effort has been put into getting B2ACCESS in the SPF for approximately a year now, without much progress. A long process of legal advice and feedback on the possibility of computing centres in EUDAT joining the SPF has proven to be the bottleneck. Since CLARIN promotes usage of the home organization accounts, this is a major issue for adoption of the EUDAT services within CLARIN.

In parallel, CLARIN ERIC is also investigating the use of “unityIDM”, the core component of B2ACCESS, as the central identity management solution, replacing the current setup [3]. In this context, Charles University has developed an LDAP endpoint for unityIDM³, which is needed to integrate the CLARIN website as well as developer-centred services such as Trac, SVN and the Nexus repositories. This setup has already been deployed in a test environment and we are currently planning the migration.

¹ See http://www.geant.org/Services/Trust_identity_and_security/eduGAIN.

² A **National Research and Education Network (NREN)** is a specialised internet service provider dedicated to supporting the needs of the research and education communities within a country, see https://en.wikipedia.org/wiki/National_research_and_education_network.

³ “unityIDM”, a software package for identity, federation and inter-federation management, see <http://unity-idm.eu>.

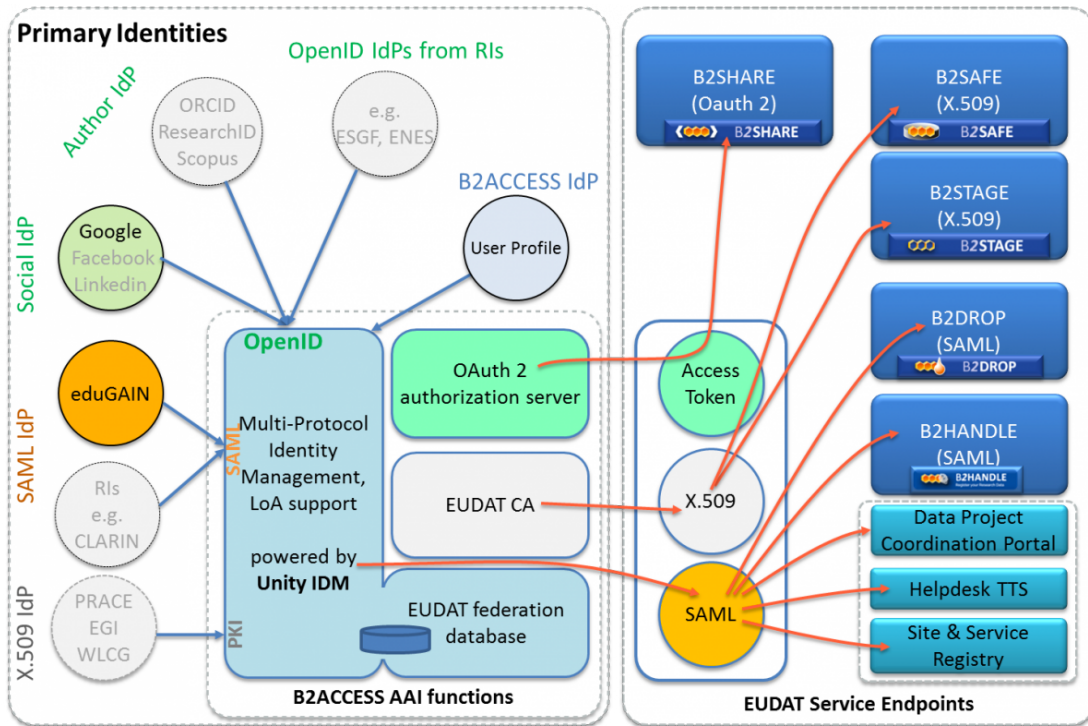


Figure 1. B2ACCESS Architectural Diagram

The LDAP endpoint can play an important role for the EUDAT integration as well. Based on discussions with B2ACCESS and B2DROP developers, it has become clear, however, that a seamless integration of B2DROP for CLARIN users, using the same credentials within both CLARIN and EUDAT, might not be possible policy-wise. Currently, the best alternative solution seems to be a dedicated B2DROP instance for CLARIN that is directly connected to the CLARIN LDAP server.

3.1.2 B2SAFE

B2SAFE service is “a robust, safe and highly available service which allows community and departmental repositories to implement data management policies on their research data across multiple administrative domains in a trustworthy manner”, see <https://eudat.eu/services/b2safe>, and the illustration given in Figure 2.

In the last quarter of 2015, a survey has been distributed across all CLARIN centres to explore their interest in the integration with B2SAFE. Responses were received from 10 centres. To get all participants to a basic level of knowledge about the service, a one day workshop⁴ at the end of 2015 was organised. Based on the participation in this workshop and because of the urgency to have an off-site copy of the repository, an initial plan to integrate about 8 centres was devised.

All CLARIN centres expressed interest in using B2SAFE (instead of joining B2SAFE), either using the *icommands* or *GridFTP* if secure transfers are needed.

Currently, four centre integrations have been either completed or suspended. The MPI-PL and SOAS integrations have started earliest. Especially the SOAS integration with B2SAFE was urgent because of the lack of proper offsite backups for their repository. Both integrations, however, have been problematic mainly because of a lack of human

⁴ See <https://www.clarin.eu/event/2015/clarin-b2safe-workshop>

resources. Both are connected to an EUDAT data centre on the technical level, but both are lacking a proper implementation of the backup scripts on the repository side. At the time of writing this report, SOAS has resources available again and work on implementing the backup scripts has started again. The other two centres, CLARIN-AT and Språkbanken, have been integrated smoothly, each with a lead-time of about two months from initial discussions to clarify the requirements to running the backup scripts in production.

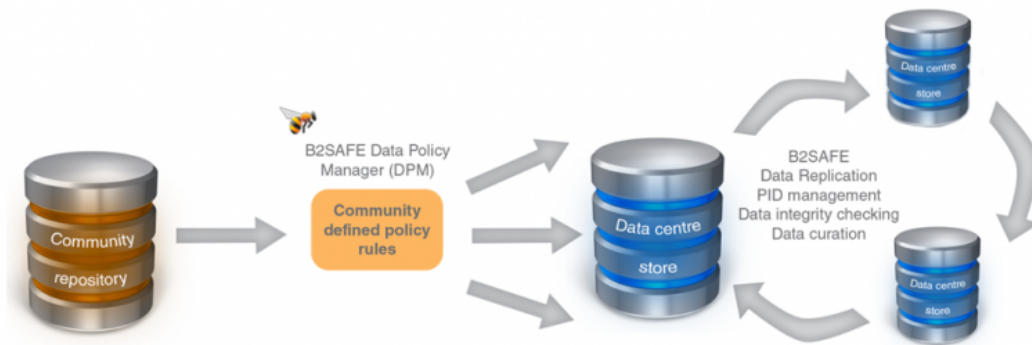


Figure 2. B2SAFE Overview.

Currently, three centres are in the process of being integrated, namely, Meertens, CELR and FIN-CLARIN. This leaves the LINDAT, CLARIN-PL and CMU centres on the list with lowest priority. Integration for these centres will be planned when Meertens, CELR and FIN-CLARIN have been fully integrated.

3.2 B2DROP and the CLARIN Language Resource Switchboard

3.2.1 B2DROP

Following the description of B2DROP on <https://eudat.eu/services/b2drop>, B2DROP is “a secure and trusted data exchange service for researchers and scientists to keep their research data synchronized and up-to-date and to exchange with other researchers.”

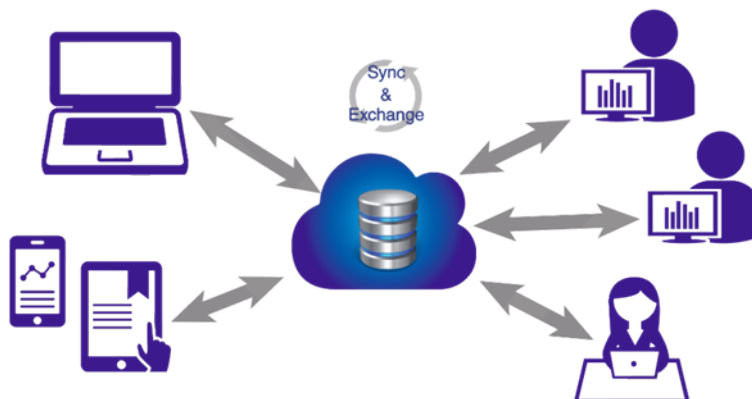


Figure 3. The B2DROP Usage Scenario.

B2DROP allows individual users to store 20G of research data in the cloud, and to exchange such data with selected colleagues, over a given amount of time. B2DROP

offers services for synchronizing multiple versions of data across devices and users, given configurable file permissions.

B2DROP is built upon ownCloud, see <http://owncloud.org> , which in turn is written in the PHP programming language, see <https://owncloud.org/blog/owncloud-and-php/> . For more details, on B2DROP, see the deliverable on EUDAT services [2], and consult its user documentation, see <https://eudat.eu/services/userdoc/b2drop#UserDocumentation-B2DROPUsage-Documentdata> .

3.2.2 The CLARIN Language Resource Switchboard

The LR Switchboard (LRS) is being developed in WP2 of the CLARIN-PLUS project. It aims at easily connecting users and their resources with the tools that can process them, see [1]. The LRS has been connected with the Virtual Language Observatory (VLO), see <https://vlo.clarin.eu>. Here, users can easily invoke the switchboard from VLO's resource viewer, which in turn, suggests applicable tools for the resource in question. In the future, the LRS will also be connected to CLARIN's Virtual Collection Registry (VCR).

The switchboard is also offered as a standalone version, see <https://www.clarin.eu/switchboard>. Here, users can simply upload a resource from their local filesystem to the LRS, for which then applicable tools are being identified.

In the standalone version, the file uploaded by the user is currently stored on a file storage server at the MPCDF computing centre in Garching.

3.2.3 Goals for Connecting the LRS with B2DROP

The main goals for connecting the LRS with B2DROP are as follows:

- To re-use other infrastructural services within CLARIN.
- To promote the usage of CLARIN services in different contexts.
- To increase the visibility of the CLARIN infrastructure.

There are two possible avenues to work toward these goals:

1. In the standalone version of the LRS, use B2DROP rather than the file storage server in Garching for the storage of resources so that tools connected to the switchboard can access them, see Section 3.2.4.
2. Offer a bridge between B2DROP and the switchboard, therefore giving B2DROP users access to the CLARIN tool space, see Section 3.2.5.

3.2.4 Using B2DROP as Alternative to the MPG server

When users of the standalone version of the LRS upload a resource, it is temporarily stored at an external file storage server (MPCDF Garching). This server has a number of drawbacks. In particular, a very limited amount of disk space is available, and more gravely, there is little access control in place so that users that know the server address can easily view and access all uploads. The server, hence, acts like a "public dropbox".

To address privacy concerns, it is necessary to better restrict access to file uploads. As a first improvement to the situation, the MPG server could be replaced by using a B2DROP instance as alternative storage device. In this scenario, the B2DROP instance will have a single user account, say, *switchboardAdmin*, which is operated by the LR switchboard: whenever a user of the LR switchboard (standalone version) uploads a file to the switchboard, it will be transferred to the *switchboardAdmin* account of the B2DROP instance. Using B2DROP's API, the admin user will associate a *shared link* to the resource in question, possibly with a set expiration date. Tools connected the switchboard will be given this link to access the file.

We have developed a prototypical implementation of this scenario, which is being tested. A B2DROP instance associated with the LR switchboard is currently hosted on the same server than the LR switchboard, which also helps tackling CORS-related issues.⁵

3.2.5 Creating a bridge between B2DROP and the Switchboard

To promote the usage of CLARIN-related tools across communities, we are creating a bridge from B2DROP and the CLARIN Language Resource Switchboard. Figure 4 depicts a screenshot of an initial prototype of B2DROP with a LRS plugin. When users click on the '...' button, a menu opens giving access to a range of actions connected to the resource. This menu has been extended with an action "Switchboard". When users select this option for a resource, the LR switchboard application opens in a new browser tab; the switchboard is invoked with a B2DROP reference to the resource in question, mirroring the connection between the VLO and the LRS.

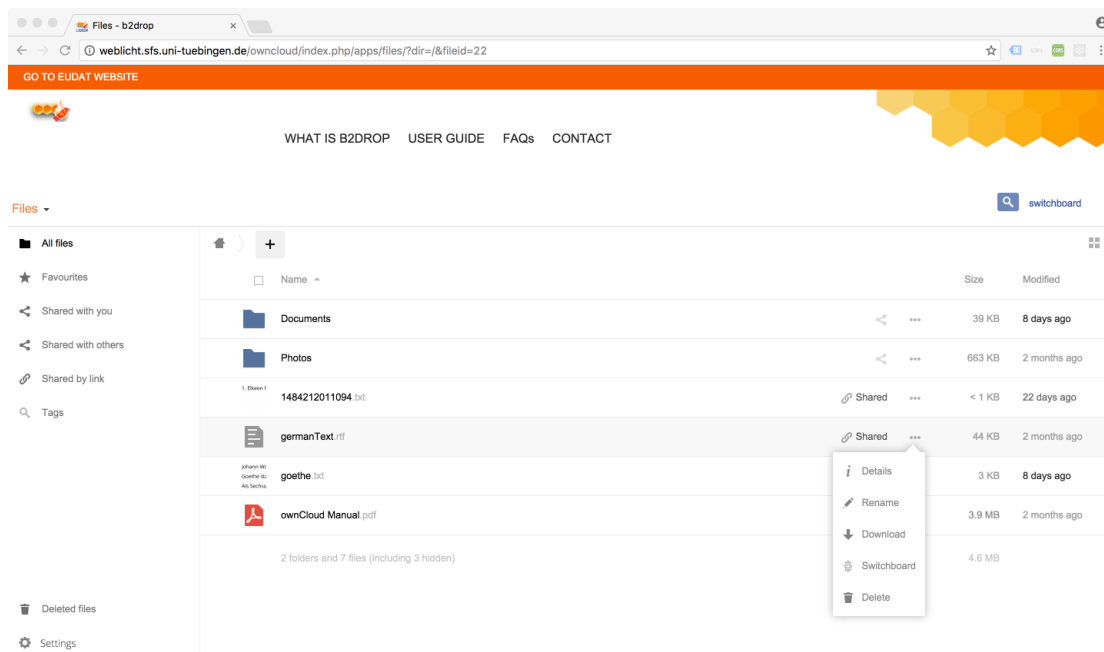


Figure 4. Bridge between B2DROP and the LRS.

⁵ **Cross-origin resource sharing (CORS)** is a mechanism that allows restricted resources on a web page to be requested from another domain outside the domain from which the first resource was served, see https://en.wikipedia.org/wiki/Cross-origin_resource_sharing.

3.3 EUDAT's General Execution Framework

The General Execution Framework (GEF) is a Docker-based platform⁶ that aims at enabling the execution of scientific workflows close to the data. The central idea of GEF is to encapsulate a scientific processing tool in a Docker “image”, making it portable and allowing its enactment in various suitable locations, where the access to the primary scientific data is fast and efficient.

A GEF service is defined as the encapsulation of a scientific (non-interactive) tool, together with some required metadata. The metadata specifies human-oriented information (service name and description) but also operational parameters such as the expected file system locations for the input and output.

The GEF platform provides a web based user interface that allows authorised users to build new services, and regular users to run existing services and to inspect and download the results. The web interface is completely based on an HTTP API, which is also available to the users for programmatic access to the service.

GEF is developed in the frame of the EUDAT2020 project and is currently under active development. It is expected to reach a public testing phase at around the middle of March 2017 and be used in community use cases by Summer 2017. The source code of the project is available on GitHub: <https://github.com/EUDAT-GEF/GEF>.

In the CLARIN context, a GEF use case for the computer-supported annotation of data with WebLicht is being devised.

4 Cooperation with other e-Research infrastructures

4.1 GÉANT

The CLARIN centres, and especially the members of the Authentication and Authorisation Infrastructure (AAI) task force have regular contacts with GÉANT representatives about the functioning of eduGAIN and in general the experiences with SAML-based authentication. Examples of practical outcomes of this interaction are:

- The Attribute Aggregator⁷ service, described in CLARIN-PLUS D2.2, providing insight in the attribute release policies from the individual Identity Providers.
- The eduGAIN opt-in dashboard⁸ – showing the percentage of identity providers per country that are connected to eduGAIN. This and the previous service are very instrumental in deciding on how to connect CLARIN service providers to national identity providers.
- The eduGAIN Connectivity Check Service⁹ that has been inspired by CLARIN's Shibboleth IdP QA tool¹⁰.

⁶ Docker is an open platform for developers and sysadmins to build, ship, and run distributed applications, see <https://www.docker.com>.

⁷ See <https://lindat.mff.cuni.cz/services/aaggreg/>

⁸ See <https://technical.edugain.org/isFederatedCheck/Federations/>

⁹ See <https://technical.edugain.org/eccs/>

¹⁰ See <https://github.com/ufal/lindat-aii-shibbie>

Next to such interactions (*e.g.*, at eduGAIN town hall meetings, FIM4R and TNC conferences) CLARIN ERIC is also participating in GÉANT's international user advisory committee¹¹.

4.2 RDA

The role of the Research Data Alliance (RDA) in bringing together players in the research data and research infrastructure field on a world-wide scale is indisputable. Therefore, CLARIN has been participating in RDA activities from the beginning. In this section, we provide an overview of RDA outputs and ongoing activities to which CLARIN contributed and is contributing.

4.2.1 Data Foundation and Terminology (DFT)

CLARIN contributed to the DFT recommendations¹² about defining a consistent terminology for data management. It was also among one of the first adopters.

4.2.2 Dynamic Data Citation

The working group on Dynamic Data Citation has come up with recommendations¹³ on how to reliably cite data sets that are changing over time, using persistent identifiers and timestamps.

4.2.3 RDA/WDS Certification of Digital Repositories

This interest group¹⁴ prepared the harmonization between the Data Seal of Approval and World Data Systems certification procedures for data repositories. CLARIN provided input from its experience with the DSA procedure and its own centre certification.

4.2.4 Federated Identity Management

The interest group on Federated Identity Management¹⁵ was a forum (partially overlapping with the FIM4R initiative¹⁶) for exchanging experiences and best practices. CLARIN contributed with several presentations on the extension of its Service Provider Federation.

4.2.5 Legal Interoperability

Getting a better understanding of the legal frameworks to enable (research) data exchange and interoperability is the aim of this interest group¹⁷.

4.2.6 Data Fabric

The Data Fabric interest group¹⁸ – with several subgroups – works on the topic of registered data objects and making these machine-actionable. It pertains to metadata, persistent identifiers, repositories, registries of repositories and data typing. CLARIN is actively involved, providing insights it has gained from experience with *e.g.* the centre registry and the LR switchboard.

¹¹ See http://www.geant.net/Users/Pages/User_Advisory_Committee.aspx

¹² See <http://dx.doi.org/10.15497/06825049-8CA4-40BD-BCAF-DE9F0EA2FADE>

¹³ See <http://dx.doi.org/10.15497/RDA00016>

¹⁴ See <https://www.rd-alliance.org/groups/rdawds-certification-digital-repositories-ig.html>

¹⁵ See <https://www.rd-alliance.org/groups/federated-identity-management.html>

¹⁶ See <https://indico.cern.ch/event/605369/>

¹⁷ See <https://www.rd-alliance.org/groups/rdacodata-legal-interoperability-ig.html>

¹⁸ See <https://www.rd-alliance.org/group/data-fabric-ig.html>

4.2.7 Domain Repositories

This interest group¹⁹ about domain-specific repositories and data management plans receives regular input from CLARIN. One of the goals is work on a protocol for language data management plans.

4.2.8 Group of European Data Experts (GEDE)

This expert umbrella group²⁰ tries to identify common recommendations based on the outputs of other RDA groups and international bodies (like the ITU). Its current activities focus mostly on persistent identifiers.

4.3 DARIAH

In CLARIN-PLUS, CLARIN is in constant consultation with DARIAH, which is also an European Research Infrastructure Consortium (ERIC) for the Humanities. Although the focus of both infrastructures is in different areas of the Humanities, there are obvious connections as CLARIN focuses on language related analysis, both as a method for answering questions in the Humanities and Social Sciences and as an objective for research such as in Linguistics. The cooperation extends to technical backend activities to avoid reduplication of work, such as in the area of persistent identification of digital objects in repositories and core components for user identification and management. Parts of these activities are conducted by national partners, others are developments on the European level.

As both research infrastructures have a partially overlapping user base, it was decided early on to enable users of the one infrastructure to use components of the other as well. The use is especially important with regard to the reuse of research data, licensed data and editions, which may be used for language based analysis in CLARIN as well as in other contexts of the Humanities such as spatial recreations, object descriptions *etc.*

As a policy, CLARIN uses the Shibboleth system, which is also used by libraries; hence scholars from institutions from countries participating in CLARIN can immediately use the services whenever their institution provides a Shibboleth-based connection. To allow DARIAH users to access CLARIN services, the internal DARIAH Identity Provider was connected to the CLARIN Service Provider Federation via eduGAIN. Likewise, there are also plans to connect the CLARIN Identity Provider to the DARIAH services.

Finding the resources already available in one of the infrastructures is also an important requirement by the user community, who wish to have a single point of entry to find adequate material for reuse. This requires that the metadata is shared and interoperable. First important steps toward the integration of metadata from both CLARIN-D and DARIAH-DE have been taken. Based on the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH), a prototype system for metadata exchange is in place. On the CLARIN-side, DARIAH-DE Dublin Core based metadata has been converted to and integrated into CLARIN's Component MetaData Infrastructure (CMDI). The data has been integrated into a test instance of the Virtual Language Observatory where it is already searchable.

For the integration of CLARIN-metadata into the search engine operated by DARIAH-DE, it is required to create schema mappings first. Examples of these have been developed and are currently being tested; also, a more general mapping procedure is being devised.

¹⁹ See <https://www.rd-alliance.org/groups/domain-repositories-interest-group.html>

²⁰ See <https://www.rd-alliance.org/groups/gede-group-european-data-experts-rda>

DARIAH and CLARIN also collaborate on building a sustainable registry of courses and other educational material related to the use of digital language resources and tools in HSS research. For more details, see the upcoming CLARIN-PLUS deliverable D5.2 (Operational course and education material registry).

4.4 Europeana

The Europeana Digital Service Infrastructure (DSI) is funded under the Connecting Europe Facility (CEF) with the goal to develop [Europeana](#) into a widely-recognised platform of services and resources, not only for metadata references, but also for access to cultural content, tools and technologies, projects and other services. During the time span of CLARIN-PLUS, CLARIN participates in WP2 of the second phase of DSI (DSI-2)²¹. This work package is concerned with the “design of end-user products & services”. Task 2.6.3, “Data sharing with third parties”, assigned to CLARIN, closely matches the CLARIN-PLUS WP4 objectives of strengthening the ties with other research infrastructures and re-using other infrastructural services within CLARIN.

The work plan²² for task 2.6.3 of Europeana DSI-2 describes the following actions towards the goal of integrating Europeana resources into the CLARIN infrastructure and increasing the visibility and ease of access of Europeana and its data within the CLARIN community and vice versa:

1. Data sets relevant to CLARIN’s community are identified out of the full set available from Europeana’s OAI harvester.
2. Implementation of a conversion from Europeana’s EDM format to CLARIN’s CMDI.
3. Inclusion of obtained metadata in the Virtual Language Observatory (VLO).
4. Selection of tools from CLARIN’s infrastructure to be included in a processing workflow based on Europeana resources.
5. Adaptation of the CLARIN infrastructure for increased load (to accommodate for a significant increase in the number of harvested and indexed metadata records) where necessary.
6. Inclusion of Europeana APIs of potential interest to CLARIN’s target audience CLARIN’s “language resource and tool inventory”.

As of the completion of the present deliverable, the following concrete results have been achieved:

1. A CMDI profile for EDM has been created, see https://catalog.clarin.eu/ds/ComponentRegistry#/?itemId=clarin.eu%3Acr1%3Aap_1475136016208®istrySpace=public
2. An EDM-CMDI conversion stylesheet has been implemented, see <https://github.com/clarin-eric/metadata-conversion>
3. A selection of data sets to harvest has been made, building on work carried out by CLARIN in the first phase of DSI in 2015 (DSI-1).
4. Test OAI harvests and VLO imports have been carried out with a smaller selection of sets containing a total of about 3 million records.

Work on preparing the infrastructure components (OAI harvester, VLO) for the increased load is ongoing. In the months following the completion of this deliverable, possibilities for processing Europeana resources by means of the currently available pipeline for discovering resources (using the VLO), finding matching tools (using the LRS) and carrying out linguistic analyses (using applicable tools provided by CLARIN

²¹ See <https://www.clarin.eu/group-page/europeana-dsi-2>

²² See <https://www.clarin.eu/file/3932>

centres) will be investigated. A number of scenarios involving publicly available resources and tools will be prepared for demonstration purposes.

Additionally, a selection of Europeana APIs and services are planned to be described and registered in the Language Resource Inventory²³. Finally, one or more online publications (weblog posts) on Europeana's and CLARIN's web portals are planned towards the end of DSI-2 (third quarter of 2017) to provide mutual exposure and present the practice and potential of this integration.

Parallel to DSI-2, Europeana currently participates in an EUDAT pilot as a part of which it investigates, with support from CLARIN, the possibility of integrating selected CLARIN resources into the Europeana ecosystem²⁴.

4.5 LAPPS Grid

There is on-going coordination with members of the LAPPS²⁵ Consortium (Brandeis University and Vassar College) about technical coordination of workflows between CLARIN (WebLicht) and LAPPS (since December 2015). The joint funding proposal involving LAPPS partners Brandeis University and Vassar College, and CLARIN Centres at Charles University Prague and University Tübingen was submitted to the Mellon Foundation in April 2016 and approved for funding in September 2016.

A preliminary meeting of the project PIs took place at the COLING conference in December 2016. Work on all work packages of the project started in January 2017, including implementation of software to convert between the internal data exchange formats used, mapping of linguistic terms to ensure semantic compatibility between the projects, and investigation of authentication and authorization issues. The kick-off meeting will take place in Prague in March 2017.

²³ See <https://www.clarin.eu/content/language-resource-inventory>

²⁴ The "EUROPEANA Data Pilot Meeting" was on the agenda of the EUDAT user forum in Helsinki in January 2017 (<https://www.eudat.eu/events/user-meetings/eudat-helsinki-meeting-23-27-january-2017-helsinki-finland>).

²⁵ See <http://www.lappsgrid.org/>

5 Conclusion

The CLARIN-PLUS project partners closely follow the progress of a number of neighbouring infrastructure projects, namely EUDAT, GÉANT, RDA, DARIAH, EUROPEANA, and the LAPPS Grid. In this deliverable, we have described all contributions of the CLARIN-PLUS partners to maximise cross-fertilization between projects. We have been working together with regard to identity management (authentication and authorization), see B2ACCESS, GÉANT and DARIAH, the implementation of data management policies, see B2SAFE, secure data exchange, see B2DROP, and the execution of scientific workflows, see EUDAT's GEF.

The CLARIN community makes available its resources and tools to other communities such as EUROPEANA and DARIAH, solving technical issues such as metadata conversion from Europeana's EDM format to CMDI, or tool inclusion in Europeana-based processing workflows. Also, B2DROP users will soon be able to connect the resources of their cloud drive with the CLARIN Language Resource Switchboard, thus giving these users access to tools that can process their resources. As the deliverable shows, the CLARIN-PLUS partners are in a productive contact with the main e-Research infrastructures, including RDA and the LAPPS Grid.

As a result, CLARIN-PLUS is well prepared to meet the challenge of realising an e-Research infrastructure where researchers can easily and smoothly search for, manage, and process research data across institutional, national, and technological boundaries.

References

- [1] C. Zinn. LR Switchboard (Software). Deliverable of the CLARIN-PLUS project, CLARINPLUS-D2.5, retrieved from: https://office.clarin.eu/v/CE-2016-0881-CLARINPLUS-D2_5.pdf
- [2] M. van de Sanden, C. Staiger, C. Cacciari, R. Mucci, C. J. Hakansson, A. Hasan, S. Coutin, H. Thiemann, B. von St. Vieth, J. Jensen (2015). D5.3: Final Report on EUDAT Services. Retrieved from: <http://hdl.handle.net/11304/2433d23a-6079-49a6-9010-ca534f6e348d>.
- [3] J. Mišutka. Robust SPF 1: workflow and monitoring. Deliverable of the CLARIN-PLUS project, CLARINPLUS-D2.2, retrieved from: https://office.clarin.eu/v/CE-2016-0809-CLARINPLUS-D2_2.pdf