



HAL
open science

D5.1 Report on Integrated Service!Needs: DARIAH (in kind) contributions - Concept and Procedures

Lisa de Leeuw, Femmy Admiraal, Matej Ďurčo, Nicolas Larrousse, Michael Mertens, Francesca Morselli, Mike Priddy, Paulin Ribbe, Carsten Thiel, Lars Wieneke

► **To cite this version:**

Lisa de Leeuw, Femmy Admiraal, Matej Ďurčo, Nicolas Larrousse, Michael Mertens, et al.. D5.1 Report on Integrated Service!Needs: DARIAH (in kind) contributions - Concept and Procedures. [0] DARIAH. 2017. hal-01628733v1

HAL Id: hal-01628733

<https://hal.science/hal-01628733v1>

Submitted on 4 Nov 2017 (v1), last revised 20 Dec 2017 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



D5.1 Report on Integrated Service Needs:
DARIAH (in kind) contributions –
Concept and Procedures

HaS-DARIAH

INFRADEV-3-2015-Individual implementation and operation of ESFRI projects
Grant Agreement no.: 675570

Date: 12 October 2017

Version: 1.0



Project funded under the Horizon 2020 Programme

Grant Agreement no.:	675570
Programme:	Horizon 2020
Project acronym:	HaS-DARIAH
Project full title:	Humanities at Scale: Evolving the DARIAH ERIC
Partners:	KONINKLIJKE NEDERLANDSE AKADEMIE VAN WETENSCHAPPEN – KNAW DIGITAL RESEARCH INFRASTRUCTURE FOR THE ARTS AND HUMANITIES CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE GEORG-AUGUST-UNIVERSITAET GOETTINGEN STIFTUNG OEFFENTLICHEN RECHTS
Topic:	INFRADEV-3-2015
Project Start Date:	01-09-2015
Project Duration:	28 months
Title of the document:	D 5.1 Report on Integrated Service Needs
Work Package title:	Basic Services Cluster 1: Integrating and Scaling National DARIAH contributions
Estimated delivery date:	30-09-2017
Lead Beneficiary:	KONINKLIJKE NEDERLANDSE AKADEMIE VAN WETENSCHAPPEN – KNAW
Author(s):	Lisa de Leeuw [lisa.de.leeuw@dans.knaw.nl], Femmy Admiraal [femmy.admiraal@dans.knaw.nl], Matej Durco [Matej.Durco@oeaw.ac.at], Nicolas Larrousse [Nicolas.Larrousse@huma-num.fr], Mike Mertens [mike.mertens@dariah.eu], Francesca Morselli [francesca.morselli@dans.knaw.nl], Mike Priddy [mike.priddy@dans.knaw.nl], Paulin Ribbe [paulin.ribbe@huma-num.fr], Carsten Thiel [thiel@sub.uni-goettingen.de], Lars Wieneke [lars.wieneke@uni.lu]
Quality Assessor(s):	Matej Durco [Matej.Durco@oeaw.ac.at], Mike Mertens [mike.mertens@dariah.eu], Lars Wieneke [lars.wieneke@uni.lu]
Keywords:	assessment, contributions, in kind, reference architecture, contribution types, DARIAH Hallmark, Services, Activities

Revision History

Version	Date	Author	Beneficiaries	Description
1.0	15-10-2017	HaS WP5	DANS, UGOE-SUB, CNRS, DARIAH	Final version to be approved by JRC and discussed in NCC and GA. Current version in use now can be found on the DARIAH-EU wiki: in kind Contributions¹

Table of Content

Executive Summary	3
Goal	5
Introduction.....	5
Documentation.....	6
Stakeholders.....	7
Main Concepts.....	8
Overall Procedure for collecting, assessing and selecting contributions.....	10
Timeline for collection of the in kind contributions.....	11
Assessment process	12
About the DARIAH Hallmark (DHM).....	13
Types of contributions	13
The DARIAH-EU reference architecture for national contributions	16
Mapping of current contribution types to the new contribution types.....	17
Core metadata for contribution	18
Assessment criteria per (sub) type	19
1. Services.....	19
<i>Services - General Assessment Criteria.....</i>	19
<i>Services - Additional Type specific Assessment Criteria.....</i>	24
2. Activities.....	31
<i>Activities - Type specific Assessment Criteria.....</i>	31
In kind contributions	38
Calculation of in kind contribution's cost.....	38
The future of (in-kind) contributions within DARIAH.....	39
The DARIAH Strategic Action Plan (STRAPL).....	39
Prognoses for basic upkeep costs of the contribution tool.....	40
Recommendations for further development	40

¹ <https://wiki.de.dariah.eu/display/DARIAH/In-kind+Contributions>

Executive Summary

Within the DARIAH research infrastructure the need to collect, disseminate and monitor the contributions offered to the infrastructure has always been present. Work package 5 of the DARIAH Humanities at Scale (HaS) project has explored and described the activities and structure required to sustain these needs, resulting in the DARIAH contributions concept and procedure including a supporting online tool.

The contributions to the DARIAH infrastructure consist either of services, being repeatable actions, or activities, which have a more unique/ one-time character. They are collected for several reasons. First of all, dissemination to the Arts & Humanities (A&H) community; contributions offered to the infrastructure need to be visible to the users within the community, so they can be found and used. Secondly the compatibility of the contributions to the infrastructure; how compatible are the contributions to the infrastructure, how easily can they be (re)used and/or combined with other services and last but not least monitoring of the variety and maturity of contributions as it is essential for the strategic planning and future development of the infrastructure.

The HaS project has developed an online tool that supports the collection of the contribution so they can be disseminated to the community. The contributions will be visible in the online tool, which in time will feed into the 'marketplace' that DARIAH is developing. Next to the contributions, the DARIAH marketplace will give access to a wide range of information relevant to the A&H community.

Furthermore, the tool enables contributors to self-assess their contribution according to criteria provided by DARIAH in order to see how compatible their contribution is to the infrastructure. These self-assessments are then peer-reviewed, on the basis of trust, by reviewers appointed by the infrastructure, which – when approved - are given a score that results in one of three tiers of compatibility to the infrastructure.

Finally, the tool can produce real time reports to monitor the diversity and capability of offered contributions, which subsequently can be utilised in outlining next steps for the infrastructure and its evolution.

Nature of the deliverable

✓	R	Document, report
	DEM	Demonstrator, pilot, prototype
	DEC	Websites, patent fillings, videos, etc.
	OTHER	

Dissemination level

	P	Public
✓	CO	Confidential only for members of the consortium (including the Commission Services)
	EU-RES	Classified Information: RESTREINT UE (Commission Decision 2005/444/EC)
	EU-CON	Classified Information: CONFIDENTIEL UE (Commission Decision 2005/444/EC)
	EU-SEC	Classified Information: SECRET UE (Commission Decision 2005/444/EC)

Disclaimer

The Humanities at Scale is project funded by the European Commission under the Horizon 2020 programme. This publication reflects the views only of the author, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

Goal

To establish a DARIAH ecosystem, in order to make available, assess, and evaluate national (in kind) contributions to the DARIAH infrastructure by taking into account existing best practises and quality requirements.

DARIAH members can contribute to the DARIAH community with a diverse range of resources, services and activities. Through an online tool the contributions will be collected and disseminated. Furthermore, the contributions will be evaluated using a self-assessment and peer review process which verifies the compatibility of the service or activity to the infrastructure, therefore allowing researchers to find matching services and activities more easily. Additionally, the tool will also be used to collect, value and monitor national in-kind contributions.

Introduction

DARIAH is a network, a research infrastructure. It connects several hundreds of scholars and dozens of research facilities in currently 17 European countries, the DARIAH member countries. In addition, DARIAH has several cooperating partner institutions in non-member countries, and strong ties to many research projects across Europe. People in DARIAH provide digital tools and share data as well as know-how. They organize learning opportunities for digital research methods, like workshops and summer schools, and offer training materials for the Digital Humanities.

The DARIAH members have agreed to contribute to DARIAH-EU in cash and in kind (see [DARIAH ERIC statutes²](#) Article 18).

DARIAH members can contribute to the DARIAH community with a diverse range of resources, services and activities. The Board of Directors has prepared [Towards a Strategy for DARIAH³](#) as a guideline for DARIAH activities in the coming years. It articulates a core mission for DARIAH and sets out high-level principles and a series of priority areas for the next two years. It is important that contributions offered to DARIAH are compliant with the core mission and the high-level principles of DARIAH.

This document elaborates on the process and procedure for offering contributions to the DARIAH community, how contributions are assessed, and how they become in-kind contributions. It is based on work done by the former Chief Integration Officer (CIO) team (Henk Harmsen and Lisa de Leeuw) and Sophie David (former Senior Management Team member) as well as on the feedback received from the participants at the DARIAH

² <http://www.dariah.eu/library/general-documents.html>

³ <https://hal.inria.fr/hal-01150112/document>

Paris meetings in November 2014, in particular from the National Coordinators Committee (NCC).

Work Package 5 in the Humanities at Scale (HaS) project has developed the concept further since, using input from the subsequent NCC and General Assembly (GA) meetings, as well as an analysis of the in-kind contributions gathered during 2015 and 2016.

Furthermore, the HaS project will develop an online tool to support and enforce the policy and process, as well as making the contributions more visible/ accessible to the DARIAH community and beyond.

At the end of the project, the maintenance of this policy and the online tool will be handed over to the DARIAH Coordination Office (DCO) to ensure sustainability. Therefore, this document needs to be considered a 'living document', that represents the concepts and procedures at current stage. The DARIAH DCO will be responsible to keep it up-to-date by including any further developments in the future. Some suggestions for enhancements were already discussed in the various meetings, and are presented in the final section of this document [The future of \(in kind\) contributions within DARIAH](#).

Besides serving as a deliverable for the HaS project, this document intends to provide a basis for:

- A. DARIAH Partner Institutions providing a structured description of their individual contribution and thereby preparing the self-assessment of (in kind) contributions.
- B. The Joint Research Committee (JRC) and other experts in the field on how to review the contributions in a concise manner.
- C. National Coordinators in collecting and submitting the in kind contributions for their country.
- D. Reporting to the DARIAH Board of directors, National Coordinators Committee and the General assembly.

Documentation

- Types of in kind contribution v1.3 (v1.0). This document was approved by the SMT in May 2014 and is now an integral part of this document.
- In kind contributions collected during 2015 and 2016.
- [CoreTrustSeal⁴](#)
- CESSDA SaW: [Reuse Readiness Criteria⁵](#)

⁴ CoreTrustSeal (10-10-2017): <https://www.coretrustseal.org/>

⁵ CESSDA SaW reuse readiness criteria (09-10-2017): (https://docs.google.com/document/d/1otafwrmW05oHE9Ju06tf3l_zzqS9FifKMNrVhD7frcU/eDit#heading=h.axq57kkk08mh)

- Capability models: [ITIL⁶](#), [FitSM⁷](#), [CMMI⁸](#)
- DASISH: [Reference model for the social sciences and humanities⁹](#)
- E-ARK: [A Maturity Model for Information Governance¹⁰](#)
- CoPoRI: [In-kind contribution¹¹](#)

Stakeholders

Listing of involved parties and their respective kind of interaction with the online tool:

DARIAH all & broad audience	Catalogue & Visualisations via the DARIAH-EU website and the online contributions tool.
Partner Institution	Offer their contributions to DARIAH via the contributions tool in order to get exposure of their contributions in the DARIAH community.
Peer reviewers; include Virtual Competency Center (VCC) heads and Working group (WG) leaders.	View, assess, approve and/or reject self-assessed contributions via the contributions tool.
National Coordinator (NC) (gathered in / represented by the National Coordinators Committee - NCC)	View, select and submit annual in kind contributions per country via the contributions tool.
Senior Management Team (SMT)	View, approve and/or reject in kind contributions per country.
General Assembly (GA)	View in kind contributions per country.
DARIAH Coordination Office (DCO)	Content management, maintenance and further development of the contributions tool through full back-office access in order to be able to (technically) assist in all stages of the process.

⁶ ITIL (09-10-2017): <https://www.axelos.com/best-practice-solutions/itil>

⁷ FiTSM (09-10-2017): <http://fitsm.itemo.org/>

⁸ CMMI (09-10-2017):

<https://www.google.nl/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&cad=rja&uact=8&ved=0ahUKewim5bez8sDNAhWHPxQKHRuxCUoQFggeMAA&url=http%3A%2F%2Fwww.sei.cmu.edu%2Freports%2F10tro34.pdf&usg=AFQjCNHiHbh7E77h9OX2oQJtkyliNkM85Q>

⁹ Reference model for the social sciences and humanities (09-10-2017):

<https://sites.google.com/a/dans.knaw.nl/reference-model-for-ssh-data-infrastructure/home>

¹⁰ A maturity Model for Information Governance (10-10-2017): <http://www.eark-project.com/resources/project-deliverables/19-d71-e-ark-a-maturity-model-for-information-governance-initial-version>

¹¹ CoPoRI: In-kind contribution (10-10-2017): <http://www.copori.eu/1369.php>

Main Concepts

Contribution	<p>Any relevant research output or activity a DARIAH partner submits as DARIAH-related work.</p> <p>See The types of contributions and assessment criteria heading for typology and criteria.</p>
In kind contribution (IKC)	<p>A contribution selected by National Coordinator as in kind for a given member country and year.</p> <p>In kind contributions are one type of support for non-profit organizations. Conversely to cash contributions, in kind contributions represent the provision of goods or services to an organisation by one of its members, valued in monetary terms according to rules agreed upon beforehand by the members of the organisation, and accounted for as part of the member's contribution to the budget.</p> <p>An in kind contribution can consist either of the direct provision of a tangible asset to the infrastructure or of expenditure incurred directly by the contributor, which benefits the infrastructure and satisfies its objections. They include goods, use of services and facilities, professional services or expertise in the form of staff time, provision of or access to equipment, special materials. They are regarded as necessary to carry out the tasks and achieve the goals commonly agreed by the members. They would have to be paid for if they were not provided by a member. From the point of view of DARIAH, in kind contributions represent a stream of revenue, though they are not monetary. In kind contributions are commonly used in RI consortia and may represent a large portion of the RI's revenue.</p> <p>In kind donations build relationships within an organization as the provider can support the mission and programs without investing cash.</p>
Self-assessment	<p>Contributions offered by DARIAH partners can be self-assessed against pre-set criteria depending on the type of contribution, which is then peer reviewed.</p>

Assessment criteria	Set of predefined criteria which are used to evaluate a contribution on their compatibility to the infrastructure. The assessment criteria will be evaluated every year and must be agreed upon by the DARIAH Senior Management Team (SMT). A type of contribution can be assessed and reviewed through a set of assessment criteria, of which some are based on the CoreTrustSeal (CTS) guidelines ¹² .
DARIAH Hallmark (DHM)	A recognition of the fulfilment of given criteria for the contributions received after the peer review of the contributions self-assessment. The DARIAH Hallmark has various tiers showing the compatibility of a contribution to the infrastructure.
In Kind Contribution (IKC) Monetary Value	IKC are assigned a monetary value based on the annual costs incurred by the DARIAH partner offering the contribution. See below for the calculation of in kind contribution's cost heading.

¹² CTS guidelines (10-10-2017): <https://www.coretrustseal.org/why-certification/requirements/>

Overall Procedure for collecting, assessing and selecting contributions

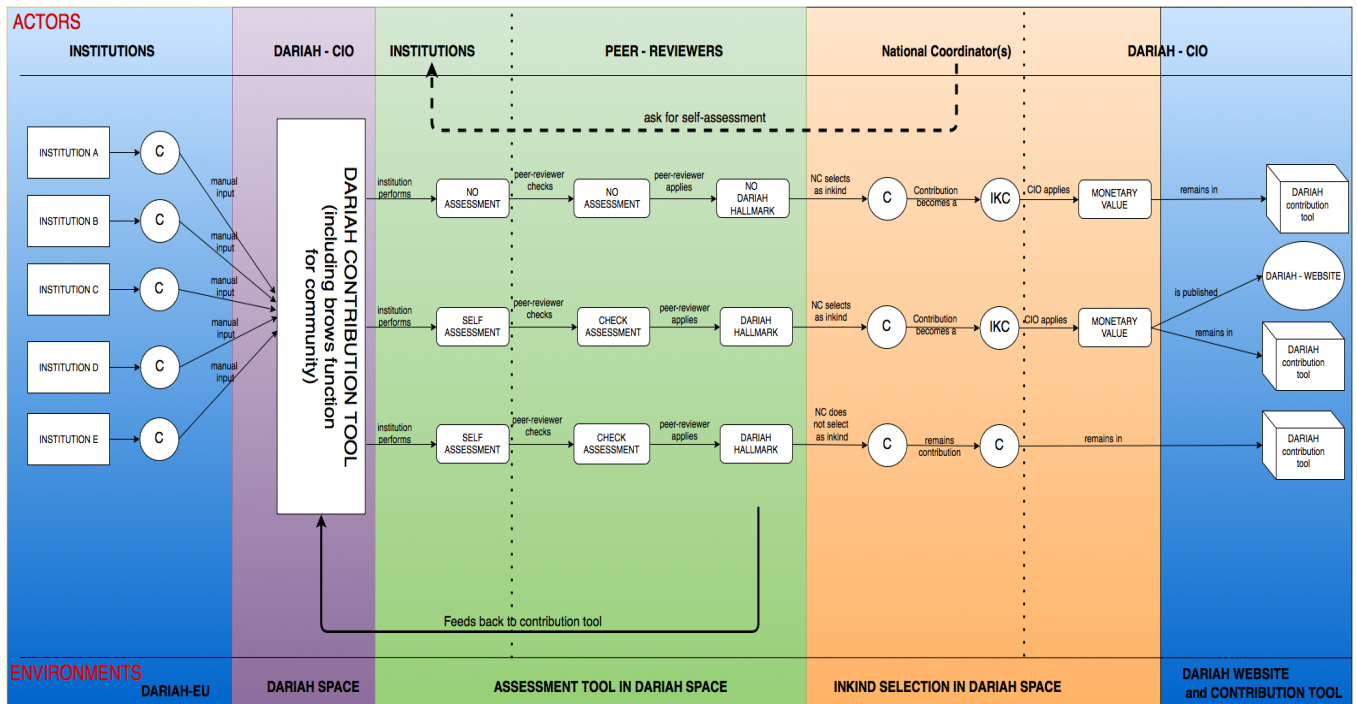


Figure 1: In kind Contributions collection, assessment and selection

As a first step in the process, the DARIAH partner institution (from now on ‘institution’) offers a contribution to the DARIAH community through the DARIAH online contribution tool (from now on ‘tool’) by providing a core set of basic information (core metadata). This can be done at any moment in time.

Once the contribution is submitted, the institution may choose to fill in a self-assessment and submit it via the tool. If needed the DARIAH National Coordinator (NC) could prompt the institution but that will not be supported by the tool. The assessment criteria (see the heading [Types of contributions](#)) are provided by the tool. Assessment criteria are offered for each type of contribution. The overall score per contribution is an average of the scores on the individual criteria. The average score will determine the tier of the DARIAH Hallmark (DHM), as specified in the [Assessment process](#) heading. If the submitting institutions chooses not to fill in a self-assessment, the contribution will not be awarded a DARIAH Hallmark, but it may still be selected as an in-kind contribution (see below). Any submitted self-assessment is peer reviewed by the appropriate peer review group via the tool (each contribution type has a different reviewers pool), as specified under the [Assessment process](#) heading. A DARIAH hallmark (DHM) is awarded according to the positive review of the contribution. The results of the self-assessment, peer review and related tier will be made public via the tool.

The National Coordinators select and submit the contributions they want to offer as in-kind contributions to DARIAH via the tool (see also the [Timeline for collection of the in kind contributions](#) heading). NC's will have access to the contributions and related monetary value for their specific country.

The tool will generate real time reports on the in-kind status including monetary value per country for the Senior Management Team (SMT) and the members of the National Coordinators Committee (NCC). The Board of Directors (BoD) will present the results to the General Assembly.

Notes:

- Submitting a self-assessment and going through the peer review is optional for the time-being and will become mandatory after a transition period.
- Self-assessments and peer reviews can be done all year round.
- Any changes to the contribution types and assessment criteria enter into force in the following calendar year.

[Timeline for collection of the in kind contributions](#)

In kind contributions for 2017 and further

Period: 1st January 2017 – 31st December 2017

- 2016 – 1st November: General Assembly; budget approval: cash and in kind (based on calculation as stated in the DARIAH statutes Article 18 – The contribution of each Member and each Observer shall consist of two parts. One part shall be the cash contribution and the other the in kind contribution. These two parts contribute to a percentage of the annual cash budget and the annual in kind budget of DARIAH ERIC, and shall be based on the GDP figures of each country)
- 2018 – 1st of January: start of in kind submissions for 2017 by NC
- 2018 – 31st March: Deadline for in kind submissions by NC
- 2018 – April: The Chief Integration Officer will send the in-kind contributions report to the Board of Directors to be shared with National Coordinators Committee and General Assembly.

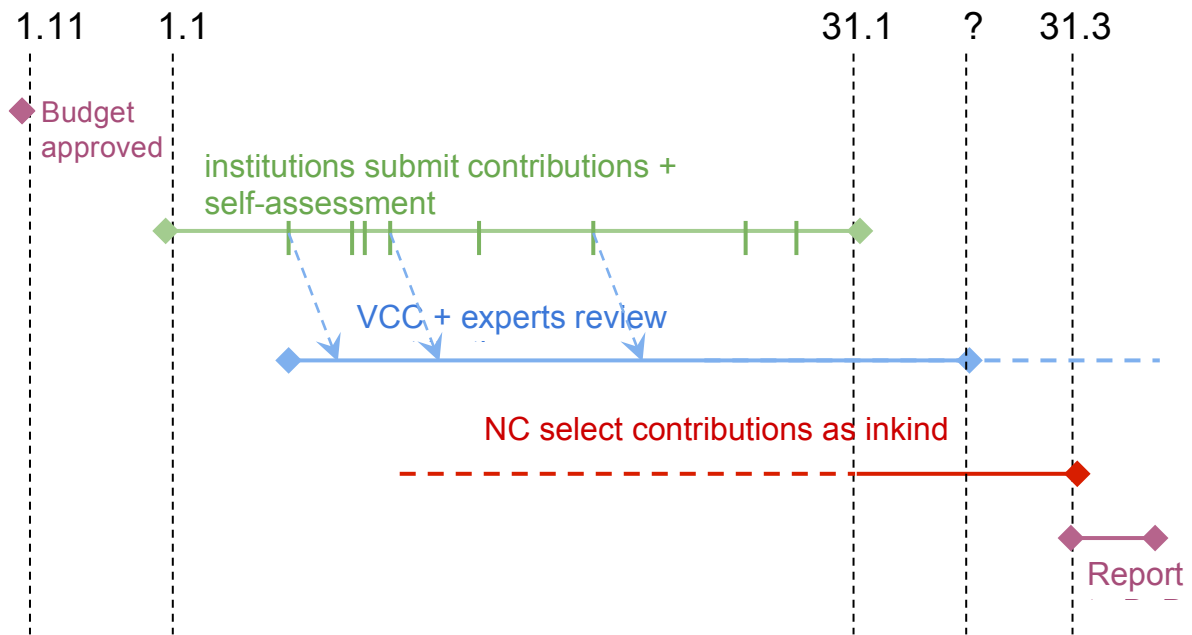


Figure 2: collection of in kind contributions timeline

Assessment process

Institutes provide contributions to DARIAH-EU through the DARIAH contribution tool. Contributions will be visible through the contribution tool and, if assessed and selected as in-kind contribution, also via the DARIAH website.

If an institute would like to get a DARIAH Hallmark for their contribution, they turn to the online system. For their type of contribution, the institute provides a score and explanation/evidence per assessment criterion via a self-assessment. A single contribution can be entered into the tool several times under a different contribution type and will then have to be assessed accordingly. The overall score per contribution is an average of the scores on the individual criteria. The average score will determine the tier of the DARIAH Hallmark (DHM). The DARIAH Hallmark is tiered to show the level of compatibility with the DARIAH infrastructure.

On completion, two peer reviewers review the explanations/evidence given by the institute. The first peer reviewer will be an expert from the field and the second reviewer is a Virtual Competency Centre (VCC) head. If the peer reviewers deem the explanations/evidence satisfactory, a hallmark is distributed that indicates the contributions compliance to the assessment criteria and in which tier. The VCC head has the final word when a conflict arises between the first and the second reviewer.

When the peer reviewers do not accept the institute's entries and score for the assessment criteria, the institute will have the opportunity to amend their explanations until the peer reviewers are fully satisfied with the entries or finally refuse acceptance.

Peer reviewers will be drawn from a pool made up of Joint Research Committee (JRC) members, DARIAH working group participants, and experts from the field. Reviewers are assigned to a self-assessment by the back-office, who makes sure no conflict of interest will arise.

If the applicant raises an objection to the verdict of the two peer reviewers, escalation is possible. The self-assessment and peer reviewers' verdict will be considered by selected members from the DARIAH Senior Management Team (SMT) and the DARIAH Scientific Advisory Board. If they agree with the peer reviewers the applicant will have to change his entry. If they disagree with the peer reviewers, the contribution is accepted and a DARIAH Hallmark is handed out to the contribution/applicant. The verdict of the DARIAH Senior Management Team (SMT) and the DARIAH Scientific Advisory Board is final.

About the DARIAH Hallmark (DHM)

The distributed DARIAH Hallmarks are valid for a limited time and are no longer valid after a maximum of 3 years depending on the funding period of the contribution. During this time, a contribution with DHM can be chosen as a national in kind contribution. In the months before the Hallmark expires the applicant will be contacted automatically and may decide to re-apply for a DHM to show that their contribution conforms to the most recent assessment criteria.

When a DHM has been distributed the assessment will be closed. When the DHM needs updating a new assessment should be initiated. A closed assessment cannot be changed or reopened.

Types of contributions

The contributions to DARIAH-EU will be assessed on their match/fit with the DARIAH infrastructure.

In order to assist and structure the assessment procedure, assessment criteria have been developed for two types of contributions: Services and Activities, which in turn consist of a number of sub-types. Services are repeatable actions, which will be disseminated to the DARIAH community and beyond through the online tool and DARIAH website. Activities, on the other hand, have a more unique/one-time character and will be available in the online tool more for administrative/reporting purposes. Figure 3 shows how Activities

and Services relate to each other: It distinguishes between efforts and products - efforts being either Activities (with a limited duration) or Services (ongoing effort, no a priori duration set), products being either Resources (in the broad sense of research data) or Software. The stroked lines indicate the default (required) flow. Thus “2.4 Resource creation” activity leads to the existence of a “Resource” which needs to be made available through some kind of “1.4 Resource access”. The dashed lines indicate optional relations, e.g. an “2.1 Event” can be accompanied by creation of training material (“2.4 Resource creation”), the created training material being a Resource in its own right. Also, a software is already a product in its own right (very relevant for the Marketplace), but it could be optionally deployed as a “1.2 Service” and/or deposited at a “1.1 Hosting services”. The 1.1 and 1.4 services are similar in nature, both making resources available online, however 1.1 Hosting service offers this capability to third-party content providers, whereas 1.4 applies when the content provider publishes the resources with their own means. Hence also the optional path from resource to 1.1. Hosting service - the resource creator can decide to deposit the resource at an external hosting service. The diagram also indicates which aspects are potentially relevant for the “DARIAH Marketplace”, a platform on which tools and service relevant to the DARIAH community will be showcased (see [The future of \(in kind\) contributions within DARIAH](#) heading).

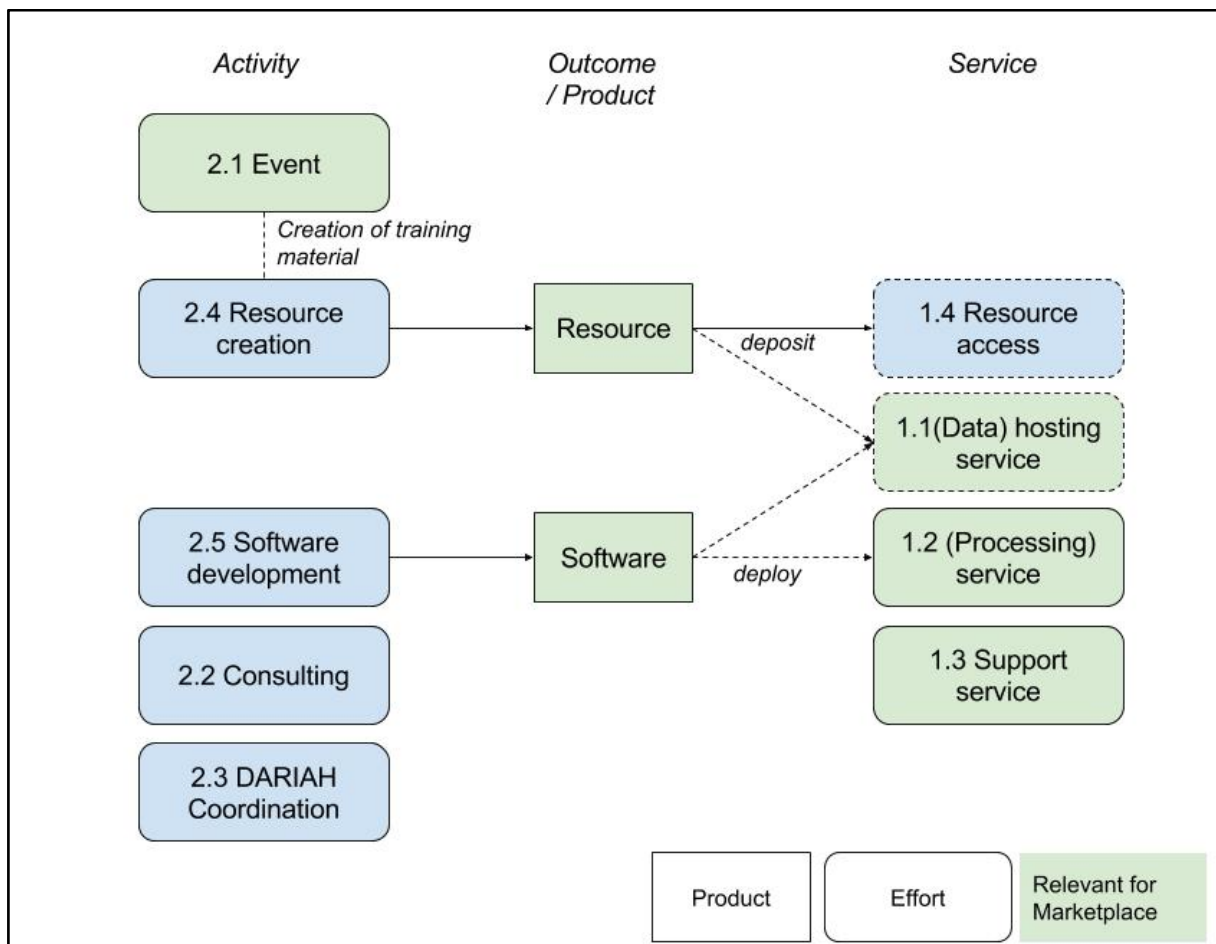


Figure 3: relationships between activities and services

The basis for developing the assessment criteria are the current assessment criteria in use by DARIAH, however, the following existing systems/models have also been consulted and/or partly re-used: Data Seal of Approval, Reuse readiness levels, Capability models (ITIL, CMMI, FitSM), alongside work done in the CESSDA-SaW and DASISH projects (see the [Documentation](#) heading).

Contributions made to DARIAH will be added to the online tool according to the subtypes described below as to facilitate submission and searching the contributions more easily.

Types of contribution	Subtype	Remarks
1. SERVICE (repeatable upon demand or continuous)		
	1.1. DATA HOSTING SERVICE (IT based)	e.g. hosting software or data.
	1.2. PROCESSING SERVICE (IT based)	
	1.3. SUPPORT SERVICE	e.g. helpdesk, software maintenance
	1.4 ACCESS TO RESOURCES	e.g. educational resources, data resources... metadata creation & enrichment
2. ACTIVITY (one-time discrete action)		
	2.1 EVENT	e.g. summer school, webinar, training
	2.2 CONSULTING	
	2.3 DARIAH COORDINATION	
	2.4 RESOURCE CREATION	e.g. educational resources, data ...
	2.5 SOFTWARE DEVELOPMENT	

The DARIAH-EU reference architecture for national contributions

In line with the division of the (in kind) contributions over services and activities, the HaS project has worked on a [reference architecture](#)¹³ for the submission of national contributions to DARIAH EU. This reference architecture further develops the definition and components by using a formal language, borrowed from the theoretical models used in ICT systems.

The main goals of the reference architecture are to:

1. Communicate and to describe contributions in a common language and to have a common understanding of what components should be described. The in kind contributions have never been formally described before, including their creation processes. The result is that up to now the contributions have been interpreted and labelled in different ways by each contributing country, by making the in kind contributions registry highly heterogeneous and unreliable from an end user perspective.
2. Facilitate the description of the interactions between components of the contributions, according to the services or functions they provide. How do the components of the contributions interact? How are they connected? What this reference architecture brings to life is not only a reliable description of the DARIAH contributions, but their description as processes and evolving systems, rather than immovable objects. It is therefore important to describe these changes and interactions with the environments in which the contributions evolve.
3. Offer an architectural reference for other projects and infrastructures that wish to offer similar services to their (research) communities.
4. Mitigate and reduce risks: we consider this reference architecture as a way to reduce the risk mainly derived from point 1: that of interpreting the (in kind) contributions in different ways, to describe them without a common reference model. This risk exists not only among partners inside the same research infrastructure but also for who accesses the DARIAH contributions from other infrastructures.

The DARIAH reference architecture for the national contributions is available as Annex 1 to this concept and procedure document and will also be made available via the DARIAH website.

¹³ Reference architecture for a SSH infrastructure : <https://sites.google.com/a/dans.knaw.nl/reference-model-for-ssh-data-infrastructure/part-1/introduction>

Mapping of current contribution types to the new contribution types

Current in kind contribution types	New contribution types/subtypes	Remarks / conversion rules
Access	1.1. HOSTING SERVICES 1.4 ACCESS TO RESOURCES	
Expertise	2.2 CONSULTING	Contributions under this type are only applicable when there was actual activity. General statements of available/potential expertise are not evaluable and won't be accepted as contribution
Interoperability	1.2. PROCESSING SERVICES 1.3. SUPPORT SERVICES	
Content Hosting	1.1. HOSTING SERVICES	
Tools and Software	2.5 DEVELOPMENT OF SOFTWARE	
Event	2.1 EVENT	
Training	2.1 EVENT	
Summer School	2.1 EVENT	
Cooperation	2.3 DARIAH COORDINATION	
Educational Resources	1.4 ACCESS TO RESOURCES 2.4 CREATING RESOURCES	Depending on contribution: If giving access to -> 1.4 If creating resource -> 2.4 If both, then it is two separate contributions
Data	1.4 ACCESS TO RESOURCES 2.4 CREATING RESOURCES	Depending on contribution: If giving access to -> 1.4 If creating resource -> 2.4 If both, then it is two separate contributions
DARIAH coordination	2.3 DARIAH COORDINATION	
Others	-	Removed as a type

Core metadata for contribution

The following core metadata are required to be able to disseminate/discover the contributions made to DARIAH.

- Identifier (created by tool)
- Country
- Submission date (submission of core metadata by contributor)
- Assessment date (submission of review by reviewer)
- Year (year of submission as in kind by NCC)
- Contribution subtype (type is added automatically, derived from the subtype)
- Title
- Description in English (including project it originated from, if applicable)
- URL to contribution
- Contact institution name
- Contact person name
- Contact person email
- Related contributions - arbitrary (typed?) links between contributions
- “tags/keywords” (a simple folksonomy way to classify the contributions)
- TADIRAH activities
- TADIRAH objectives
- TADIRAH techniques
- Disciplines
- Costs (see heading [Calculation of in kind contribution's cost](#))

Assessment criteria per (sub) type

1. Services

Services - General Assessment Criteria

- Extra metadata:
 - Support Contact -> provide contact details (email address, or other means of soliciting support, like issue tracker, helpdesk, feedback form)
 - Underlying software. (If the software itself is being developed and maintained then it is a separate contribution: see heading [2.5 Software development](#))

- Maturity Level

Score	Level	Description
1	Ad hoc	Proof of concept
2	Repeatable	Best effort basis, could change or go offline
3	Managed	running stable + commitment to support for extended period

- Support Level

Score	Level	Description
0	None	No defined way of support for users. Use as is.
2	Partial	There is a defined Support Contact, but served on best-effort basis with no dedicated capacity
3	Comprehensive	There is a defined Support Contact, with dedicated capacity to process support requests

- Documentation available (in English, relates to the software underlying the service)

- a. For the user (ideally as contextual help in the application or online)

Score	Level	Description
0	None	No usable user documentation available
2	Partial	Some user documentation available, potentially not

		complete or not fully up-to-date
3	Comprehensive	Comprehensive up-to-date user documentation available (online or as contextual help in the application)

- b. For the developer (inline in the code and/or comprehensive documentation of the system architecture, functions and APIs as part of the software package or online)

Score	Level	Description
0	None	No usable developer documentation available
2	Partial	Some documentation of the system, its APIs and architecture available, potentially not complete or not fully up-to-date
3	Comprehensive	Comprehensive up-to-date developer documentation available

- c. For the administrator - installation, deployment, monitoring

Score	Level	Description
0	None	No usable documentation regarding installation, deployment or monitoring
2	Partial	Some documentation for administrators available
3	Comprehensive	Comprehensive up-to-date administrator documentation available

- Multilingual User Interface in English (working language of DARIAH) and/or other languages, support for localisation, multi-script (allows to switching scripts in one input field) and right-to-left text (where applicable)

Score	Level	Description
0	None	(Almost) no support for multilingual UI, localisation etc. UI text is hard-coded in the application and there is no simple way to change the language.
2	Partial	Underlying software supports localisation / multilingual UI,

		but the service is available only in one language; OR the localisation/multilingual support is not implemented systematically and may be difficult to extend to other languages.
3	Comprehensive	Underlying software supports localisation / multilingual UI in a systematic manner (dictionary files), it is clearly defined and documented how to add further languages. (only applies if there is an UI) Next to that the application also supports right-to-left text and multiple writing systems for the UI and for the input if applicable.
NA	Not applicable	Given the nature of the service there is no UI to translate (e.g. web service)

- Is there an Impressum (+ Info on use of cookies) & Terms of Use

Score	Level	Description
0	None	No, Impressum OR info on use of cookies missing
2	Complete	Yes, Impressum, Terms of Use AND info on use of cookies present
NA	Not applicable	Given the nature of the service there is no Impressum or info on cookies to show (e.g. web service)

- Services exposes an API - programmatic access to (read and write) data; typically, RESTful web service, OAI-PMH endpoint, Feeds, but also an option to call the service/application via URL passing data as parameter

Score	Level	Description
0	None	no structured API exposed
2	Partial	An API is available, but it is experimental/incoherent/unstable, or the documentation is missing or incomplete
3	Comprehensive	a well-documented standard-conforming (ideally RESTful) API offering read and possibly write access to data (or invocation of the service) available
NA	Not applicable	Given the nature of the service there is no programmatic access possible/sensible

- The service is monitored, (DARIAH services as in-kinds available)
Note: When up-to-date status about the service up-time is exposed publicly, it is helpful information for the user, for the service provider (can act quickly upon downtime) and the infrastructure administration (gets overall picture of the status of the services constituting the infrastructure)

Score	Level	Description
0	None	No automatic/systematic monitoring in place
2	Partial	The service is being monitored automatically locally (either directly by service provider, or the corresponding computing centre or similar)
3	Comprehensive	Monitored by DARIAH monitoring service: (added value of being monitored by DARIAH is the sense of the service being part of one coherent infrastructure and an independent control instance)

- User statistics (DARIAH services as in kinds available). Note: User statistics via DARIAH allow to view the service in the context of the bigger infrastructure.

Score	Level	Description
0	None	No user statistics are being collected
2	Partial	Own solution for user statistics
3	Comprehensive	User statistics via DARIAH user statistic service

- Service supports federated DARIAH login (Single-Sign-On, AAI via Shibboleth <https://de.dariah.eu/en/aa1>), and/or via OpenID

Score	Level	Description
0	None	No AAI (just local login)
2	Partial	OpenID or Shibboleth login with DARIAH account OR some institutional accounts possible, however comprehensive coverage missing

3	Comprehensive	Shibboleth (and/or OpenID) login possible with DARIAH account and any institutional account (in one of the eduGain countries)
NA	Not applicable	No login required

Services - Additional Type specific Assessment Criteria

1.1 Data hosting service

Hosting or deposition service, a repository, i.e. the ability and willingness of a provider (like a data centre) to store and (optionally) publish resources of other parties for a defined period of time.

Scope Note: This covers mainly traditional repositories for publications and research data, but also any other service that allows to store and publish resources of any kind.

This service is similar to 1.4 Resource access, with the distinction that in 1.1 access is offered to data from third parties, thus the provider offers only the means to store and publish resources, not the resources themselves.

Examples: [HAL¹⁴](#), [Zenodo¹⁵](#), [NARCIS¹⁶](#), [OpenEdition¹⁷](#)

Apart from taking responsibility for the contents it hosts, a platform may :

- Inform users or other platforms about the contents it hosts (via OAI-PMH for instance)
- Provide views of the contents. That is somehow, publish the contents.
- Provide interfaces for other platforms or tools to connect and access the contents
- Manage mechanisms of access control

Assessment Criteria:

- [Core metadata for contribution](#)
- [Services - General Assessment Criteria](#)
- There is help provided by the repository/platform describing what information is needed for others to assess the quality of the data, compliance with disciplinary and ethical norms and (alt)metrics about the use of the information. [[CoreTrustSeal-Requirement¹⁸](#) 4, 11 & 14]

Score	Level	Description
1	Theoretical (initial)	We have a theoretical concept.
2	In progress (partial)	We are in the implementation phase.

¹⁴ HAL (10-10-2017): <https://hal.archives-ouvertes.fr/>

¹⁵ ZENODO (10-10-2017): <https://zenodo.org/>

¹⁶ NARCIS (10-10-2017): <https://www.narcis.nl/?Language=en>

¹⁷ Open Edition (10-10-2017): <https://www.openedition.org/?lang=en>

¹⁸ CoreTrustSeal (CTS) requirements (10-10-2017): <https://www.coretrustseal.org/why-certification/requirements/>

3	Implemented (comprehensive)	This guideline has been fully implemented for the needs of our service.
---	--------------------------------	---

- There is a list of the formats accepted or promoted by the repository/platform. So that data producers are able to provide the data in formats recommended by the data repository. [[CoreTrustSeal-Requirement¹⁹ 8](#)]

Score	Level	Description
1	Theoretical (initial)	We have a theoretical concept.
2	In progress (partial)	We are in the implementation phase.
3	Implemented (comprehensive)	This guideline has been fully implemented for the needs of our service.

- There is a description of the means the repository/platform uses to ensure compliance with legal regulations and contracts including, when applicable, regulations governing the protection of human subjects. [[CoreTrustSeal-Requirement²⁰ 2&4](#)]

Score	Level	Description
1	Theoretical (initial)	We have a theoretical concept.
2	In progress (partial)	We are in the implementation phase.
3	Implemented (comprehensive)	This guideline has been fully implemented for the needs of our service.

¹⁹ CoreTrustSeal (CTS) requirements (10-10-2017): <https://www.coretrustseal.org/why-certification/requirements/>

²⁰ CoreTrustSeal (CTS) requirements (10-10-2017): <https://www.coretrustseal.org/why-certification/requirements/>

- The processes and procedures for managing data storage are documented [[CoreTrustSeal-Requirement²¹ 9](#)]

Score	Level	Description
1	Theoretical (initial)	We have a theoretical concept.
2	In progress (partial)	We are in the implementation phase.
3	Implemented (comprehensive)	This guideline has been fully implemented for the needs of our service.

- The ways provided for data re-users to discover and use the data and refer to them in a persistent way are provided. [[CoreTrustSeal-Requirement²² 11&13&14](#)]

Score	Level	Description
0	None	No search options nor PIDs provided.
2	Partial	The UI either enables data search or it offers a persistent way (PID) of referring to the data.
3	Comprehensive	The UI enables data search and offers a persistent way (PID) of referring to the data as well.

- Is the service quality formalised?

Score	Level	Description
0	None	Not formalised
2	Partial	Best effort
3	Comprehensive	SLA available

²¹ CoreTrustSeal (CTS) requirements (10-10-2017): <https://www.coretrustseal.org/why-certification/requirements/>

²² CoreTrustSeal (CTS) requirements (10-10-2017): <https://www.coretrustseal.org/why-certification/requirements/>

Examples:

[OpenEdition](#)²³, [HAL](#)²⁴, [EHRI](#)²⁵, [EASY \(DANS\)](#)²⁶, [NARCIS](#)²⁷, [GAMS Graz](#)²⁸.

Cléo's platforms (Calenda, Hypothèses, Revues.org, OpenEdition), hosting services for digitized content/assets.

1.2 Processing Service

A (digital) service that applies some algorithmic processing (statistical analysis, annotation) on given data, or provides means to edit/curate data (e.g. web interface for collaborative editing).

scopeNote: Processing service is distinct from content hosting service in that it does not store data, but generates new data based on given input.

Ideally, processing services are connected seamlessly with hosting service to appear to a user as one (virtual research) environment.

Assessment criteria:

- [Core metadata for contribution](#)
- [Services - General Assessment Criteria](#)
- Provide sufficient information to assess the quality of the software, its compliance with disciplinary and ethical norms and (alt)metrics about the use of the information. [[CoreTrustSeal-Requirement](#)²⁹ 4,11 &14]

Score	Level	Description
1	Theoretical (initial)	We have a theoretical concept.
2	In progress (partial)	We are in the implementation phase.
3	Implemented (comprehensive)	This guideline has been fully implemented for the needs of our service.

²³ Open Edition (10-10-2017): <https://www.openedition.org/?lang=en>

²⁴ HAL (10-10-2017): <https://hal.archives-ouvertes.fr/>

²⁵ EHRI (10-10-2017): <https://www.ehri-project.eu/>

²⁶ EASY: <https://easy.dans.knaw.nl/ui/home>

²⁷ NARCIS (10-10-2017): <https://www.narcis.nl/?Language=en>

²⁸ GAMS (10-10-2017): <http://gams.uni-graz.at/context:gams>

²⁹ CoreTrustSeal (CTS) requirements (10-10-2017): <https://www.coretrustseal.org/why-certification/requirements/>

- Use due diligence to ensure compliance with legal regulations and contracts including, when applicable, regulations governing the protection of human subjects. [[CoreTrustSeal-Requirement³⁰](#) 2&4]

Score	Level	Description
1	Theoretical (initial)	We have a theoretical concept.
2	In progress (partial)	We are in the implementation phase.
3	Implemented (comprehensive)	This guideline has been fully implemented for the needs of our service.

Examples:

Conversion services ([oxgarage³¹](#)), Digitisation/OCR, Stylometric analysis, Annotation & Enrichment, Collaborative text editors (etherpad, wiki), [tokenEditor@ACDH-OEAW³²](#), Enrichment

1.3 Support service

Support services support the use of a given service, e.g. through a helpdesk.

Assessment criteria:

- [Core metadata for contribution](#)
- [Services - General Assessment Criteria](#)

1.4 Access to resources

(Online) access to resources (datasets) offered by the content provider. This can be raw data, (web) applications offering rich access to data, or web services that allow a programmatic access to the data (API). The granularity of the described resources is up to the content provider, but we encourage a high-level description of whole coherent collections of resources.

scopeNote: This type of contribution does not include (national or domain-specific) aggregators, or metadata catalogues, i.e. services that collect information about resources from multiple providers and allow to browse and search in it. These will/can ideally become a source of information about contributions.

³⁰ CoreTrustSeal (CTS) requirements (10-10-2017): <https://www.coretrustseal.org/why-certification/requirements/>

³¹ Oxgarage (10-10-2017): <http://www.tei-c.org/oxgarage/>

³² TokenEditor: <https://www.oeaw.ac.at/acdh/tools/tokeneditor/>

Neither it includes repositories, i.e. pure hosting services that offer storing and publication of data from third parties. (see 1.1. Data hosting service)

It also does not include web services processing data. (see Processing services contribution)

Assessment criteria:

- [Core metadata for contribution](#)
- [Services - General Assessment Criteria](#)
- Provide sufficient information for others to assess the quality of the data, compliance with disciplinary and ethical norms and (alt)metrics about the use of the information. [[CoreTrustSeal-Requirement³³](#) 4, 11 & 14]

Score	Level	Description
1	Theoretical (initial)	We have a theoretical concept.
2	In progress (partial)	We are in the implementation phase.
3	Implemented (comprehensive)	This guideline has been fully implemented for the needs of our service.

- Use due diligence to ensure compliance with legal regulations and contracts including, when applicable, regulations governing the protection of human subjects. [[CoreTrustSeal-Requirement³⁴](#) 2&4]

Score	Level	Description
1	Theoretical (initial)	We have a theoretical concept.
2	In progress (partial)	We are in the implementation phase.
3	Implemented (comprehensive)	This guideline has been fully implemented for the needs of our service.

³³ CoreTrustSeal (CTS) requirements (10-10-2017): <https://www.coretrustseal.org/why-certification/requirements/>

³⁴ CoreTrustSeal (CTS) requirements (10-10-2017): <https://www.coretrustseal.org/why-certification/requirements/>

- Enable the users to discover and use the data and refer to them in a persistent way. [[CoreTrustSeal-Requirement³⁵](#) 11, 13 & 14]

Score	Level	Description
1	Theoretical (initial)	We have a theoretical concept.
2	In progress (partial)	We are in the implementation phase.
3	Implemented (comprehensive)	This guideline has been fully implemented for the needs of our service.

- Terms of use, Availability, License Conditions of access to and reuse of the resource(s) are clearly indicated, and ideally support/encourage open access and reuse. If the resources (datasets) contain elements under different licenses the more restrictive license applies.

Score	Level	Description
0	None	No license information available
2	Restricted	License information available, but restricted license applies, that does not allow reuse.
3	Open	License information available, open and/or standard licences³⁶ that allows reuse.

³⁵ CoreTrustSeal (CTS) requirements (10-10-2017): <https://www.coretrustseal.org/why-certification/requirements/>

³⁶ Open Source Initiative (10-10-2017): <https://opensource.org/licenses/category>

2. Activities

Unlike the Services, Activities have no General assessment criteria, just subtype specific assessment criteria and /or extra meta data fields.

Activities - Type specific Assessment Criteria

2.1 Event

Events organised in the field of digital arts and humanities and cultural heritage, whose aim is to:

- offer a theoretical insight of the methods, practices, history and future developments in the field of digital arts and humanities
- showcase case studies in the field of digital arts and humanities
- offer training and guidance in the application of digital methodologies in traditional humanities curricula

The duration of the event doesn't represent a selection criteria: events of the duration of few hours (e.g. workshop) or few weeks (e.g. summer schools) can be submitted as DARIAH contributions, as long as they meet the requirements stated above.

Please note that English is the preferred working language for the events. While DARIAH recognises the importance of a multilingual development of the digital humanities, these DARIAH contributions represent nonetheless an invaluable resource for international researchers and institutions. The adoption of English during your event and its documentation, could help to disseminate your work among other communities.

Assessment criteria:

- [Core metadata for contribution](#)
- Extra metadata:
 - Date of the event
 - Location of event
 - Duration of event
 - Number of participants at event
- Documentation or report on targeted public is provided. Documentation includes:
 - An English version or abstract of the documentation
 - main discipline/ area of participants
 - level of expertise
 - Nationally/ Internationally oriented

Score	Level	Description
1	Initial	No information or report about the target audience has been created and provided

2	Partial	Partial information about the target audience has been created and provided
3	Complete	The target audience has been fully documented and reported as part of the in-kind contribution

- Documentation of the event is provided:
 - in English
 - Though a stable URL where the documentation is available
 - and can include: abstract, programme, slides, blogs etc.

Score	Level	Description
1	Initial	No documentation of the event available
2	Partial	Partial documentation of the event available online
3	Complete	The event has been fully documented; it is fully available for reuse at a stable link

Examples:

Lecture, symposium, workshop, conferences, training day(s), summer/ spring/ winter school, academic course, academic program, MOOCs and other forms of online education

2.2 Consulting

Expertise that has been provided as a service to another DARIAH member or the DARIAH community. This may cover advice including consulting, audit, design or other activities where expertise forms the basis of the service. Transfer of knowledge activities should be only considered if they do not fit in the training categories (see headings [2.1 Event Training / Summer school](#) or [2.4 Creating Resources](#)).

Assessment criteria:

- [Core metadata for contribution](#)
- Extra metadata: If the project has been completed also indicate a targeted public (user, producer, repository manager, etc.) that could make use of the results in the future.
- Indication of the consumer of the service:
 - Name of the corresponding institutions and their national affiliation,
 - Name of the contact persons OR DARIAH-EU including the name of a contact person

Score	Level	Description
-------	-------	-------------

1	Initial	No or only very generic information is available
2	Partial	Only partial information (e.g. only the name of an institution/national affiliation) is available
3	Complete	All the required information is provided

- If the contribution is still in process please provide a description of the existing results in English (summary, problem, table of content, content (type of formats, metadata, etc.) and draft the next steps as well as the final delivery date.

Score	Level	Description
1	Initial	No or only very generic information is available
2	Partial	Only partial information (e.g. only a summary) is available
3	Complete	All the required information is provided

- If the project has been completed provide the existing documentation/report about software, methods, etc., which should be in English or in English and any other relevant language(s), that specifies the results.

Score	Level	Description
1	Initial	No or only very generic information is available
2	Partial	Only partial information is available
3	Complete	All the required information is provided

Examples:

- DARIAH.eu wants to assess the overall impact of IKC within the DARIAH community and asks DANS to perform this evaluation as well as to provide a final report.
- The Austrian Academy of Sciences wants to implement a new single-sign-on solution and asks DAASI for advice on choosing the appropriate solution. The result of this consultation is documented in a final report

2.3 DARIAH Coordination / Cooperation

This type of contribution involves the coordination of DARIAH activities.

For instance, coordination of or cooperation in:

- The different governance bodies of DARIAH
- Central offices (DCO, CIO)
- National Coordinator includes: national coordination at the financial, administrative or communication level done by the National Coordinating Institution
- VCC Head
- Working Group
- Ambassadors

Assessment criteria:

- [Core metadata for contribution](#)

2.4 Resource creation

The contribution is the creation of any kind of resource. This could be educational resources accompanying a training event, or digitisation of historic material. The outcome is a new resource (made available to the public as a service contribution 1.4.)

Examples for this type of contribution are:

- Tutorials
- Guidelines
- Teaching material
- Software documentation (users)
- bibliography or registries of digital humanities resources
- New representation of existing historic material (cultural heritage objects)
- Structured metadata for resources

The new resources can be practically in any format, such as:

- text (as a Word or PDF document)
- audio/video recording
- interactive resources (e.g. quizzes, interactive maps and diagrams etc.)
- code
- images
- structured data (TEI, 3D-objects)

Please note that English is the preferred language for the contents of (educational) resources. While DARIAH recognises the importance of a multilingual development of the digital humanities, the DARIAH contribution represent nonetheless an invaluable resource for international researchers and institutions. The adoption of english for your training material could help to disseminate your work among communities.

Assessment criteria:

- [Core metadata for contribution](#)
- Extra metadata field: Type of the Resource (e.g. video recordings; digitized manuscripts)

- English (verbose) description of the resource created (as opposed to structured metadata that could also include a verbose description):
 - context (organisational, research)
 - what research need it answers (motivation)
 - how the resource is organized and presented

Score	Level	Description
0	None	No description of the resource is provided
2	Partial	Partial description of the resources is provided
3	Complete	The description of the resource is fully provided and contextualised

- Terms of use, Availability, License Conditions of access to and reuse of the resource(s) are clearly indicated, and ideally support/encourage open access and reuse.

Score	Level	Description
0	None	No licence information available
2	Partial	Restricted license applies
3	Comprehensive	License information available, open and/or standard licences ³⁷ are used that allow reuse.

- A stable URL where the resource is available

Score	Level	Description
0	None	URL is not provided
2	Partial	A link to the resource is provided, but the URL is not stable
3	Complete	A stable URL is provided (PID, DOI)

- Metadata about the resource is provided (in a structured format)

Score	Level	Description
0	None	No metadata available

³⁷ Open Source Initiative (10-10-2017): <https://opensource.org/licenses/category>

2	Partial	Metadata available just via web interface or in an unstructured format (word document, excel sheet)
3	Complete	Metadata conforming to a metadata schema is publicly available and machine readable

2.5 Software development

Availability of software, i.e. executable code that can be installed and run by other partners. The contribution should include the source code (not just the binaries).

Scope Note: We distinguish between software and service as separate contributions. Every (digital) service is an activation of some software. If both a service and the underlying software are provided, then it counts as two distinct (though related) contributions.

The code can be in any programming language, it can also be only a simple script dedicated to one specific task, as long as it is working and documented.

Assessment criteria:

- [Core metadata for contribution](#)
- Terms of use, Availability, License Conditions of reuse of the software are clearly indicated, ideally open source licenses apply.

Score	Level	Description
0	None	No licence information available
2	Partial	Restricted licence applies
3	Comprehensive	License information available, open and/or standard licences ³⁸ that allows reuse.

- Code is maintained under version control (and available via a public repository)

Score	Level	Description
0	None	There is no defined way of code versioning and sharing with a broader audience
2	Partial	a) There is a code repository with versioning, but it is not available to the public. b) The source code (or just the executables) is shared online, but not via a code repository.

³⁸ Open Source Initiative (10-10-2017): <https://opensource.org/licenses/category>

3	Complete	The code is available via a public code repository (sourceforge, GitHub, etc.). It could be also institutional repository, but it needs to make the code available without restrictions.
---	----------	--

- Specification/Documentation available

Score	Level	Description
0	None	no (shareable) documentation of the application is available
2	Partial	There is some documentation (either online or packaged with the software), which can be incomplete, or partly outdated, but should contain at least somewhat useful information about the service
3	Complete	There is comprehensive up to date documentation for users, developers and maintainers available (either online or packaged with the software)

- Commitment level

Score	Level	Description
0	None	This is an abandoned/orphaned project, which still may be useful for someone to take up, but there is not really anybody to talk to about it.
2	Partial	A prototype or a project, we may work on in the spare time, but (currently) don't have dedicated resources to commit to. Though there is someone knowledgeable of the code, who can be contacted by interested parties and might implement bug fixes. And we are interested in cooperation on this.
3	Complete	This is a critical piece of software for us, we (and others) use it actively, we have resources available to maintain and further develop. We have developers knowledgeable of the software available as contact persons.

Examples:

XSLT-scripts, Java applications, source code of web applications, software newly developed for or during the digitization enrichment

In kind contributions

The DARIAH-ERIC states that the yearly country contribution to DARIAH is divided in a cash contribution and an in kind contribution. In kind is a specific effort made by a country in relation to and in accordance with DARIAH activities. In kind for a specific year are expenses that have occurred in that specific year, e.g. in kind offered for 2017, must be executed in 2017.

Next to the assessment procedure, the online system will also allow the National Coordinator (NC) to select and submit national in kind contributions. According to a calculation a value will be put on a contribution turning it into an in-kind contribution. If a contribution is selected as an in-kind and it has a DARIAH Hallmark it will be visualised on the DARIAH-EU website as a DARIAH service.

The budget for a certain year for in-kind contributions is set in November of the previous year as stated in the DARIAH-EU statutes (Article 18 and ANNEX II). The collection of the in-kind contributions will take place after the fiscal year has passed and is the report of the contributions realized in that fiscal year.

Calculation of in kind contribution's cost

The cost of an in kind contribution will be calculated according to the annual costs incurred by the provider for allowing access to keep the service and or activity up and running. The amount of the contribution should be reported as actual costs according to the accounting rules agreed upon by the members in the statutes or bylaws of the organisation or they can be calculated according to H2020 project accounting rules. The costs are entered into the system on a basis of trust and will not be submitted to an audit. The value, or in other words the annual costs, is entered during the filling of the core metadata of the contribution.

The future of (in-kind) contributions within DARIAH

The DARIAH Strategic Action Plan (STRAPL)

- One of the 25 actions that DARIAH will undertake according to its STRategic Action Plan (STRAPL) is to create and populate a marketplace platform for exposing tools and services in the DARIAH network. This would involve building on the contribution tool's API. Also included would be other extant services such as the Course Registry and a new Project Registry. Models for 'hallmarking' the tools and services searchable via the marketplace to indicate applicability for certain research methods or disciplines would also be part of this development, as well as visualising the range, type, as well as state of development and current financing of the tools and services included.
- Sustainability; how will DARIAH ERIC keep this up and running? One option for the contribution tool as such is that it is financed directly by DARIAH-EU (see projected costs below); other options could involve the tools being accepted as an in-kind itself from DARIAH-NL, going out to tender within the DARIAH network, or going out to tender to involve other publicly-funded providers as well as eventually commercial ones. The last alternative may have to be the case, given EU procurement rules in the context of the projected costs. In terms of the marketplace, it is recommended that DARIAH-EU conducts some purely economic and impact research to see where meaningful exchanges of monetary and sustainability value could be made within the DARIAH network imagined as a functional 'common market'.
- A major element of such a business model would imply moving from auditing the nominal monetary value of contributions made by DARIAH members (the current 'in-kind' model), to a framework of measuring user satisfaction and feedback (an 'out-kind' model) with a corresponding Hallmark to readily indicate the major capabilities of any tool or service presented within the proposed marketplace. This would lead to a much greater ability to assess the validity of tools and services, their feasibility, whether they should be deprecated, and where and when the demand for new tools and services arises. Being able to measure these values or assessments would represent a major plank in ensuring DARIAH-EU's sustainability, since there would be a further direct way of being grounded in those user communities who represent the network's public or audience.

Prognoses for basic upkeep costs of the contribution tool

Based on the effort it takes for DANS to upkeep the Data Seal of Approval tool we estimate the yearly costs for the running of the in kind contribution tool will amount to around 25,000 euro on an annual basis.

Please note that this does not include any major changes to the system like new functionality or major upgrades; extra budget will have to be set aside for this.

Staff/ Material	Annual Costs	Total
Back office personnel/ helpdesk (DCO)	300 hrs x 50 euro	15,000 euro
Technical staff for bug fixing and small changes	180 hrs x 50 euro	9,000 euro
Material costs (servers)		1,000 euro
	Total	25,000 euro

Recommendations for further development

After the completion of the HaS project, the DARIAH DCO will be responsible for the maintenance and future development of the DARIAH national contributions concept, procedure and online tool.

During the course of this project recommendations for further development have already been discussed and suggested, but could not be accomplished within the project's lifespan. We have therefore listed them below:

- Possibility for the user to provide feedback (for general public and/or between the producer of the resource)
- Possibility when searching for contributions have an option to show results in your region.
- To determine ranking of search results according to 'weighting' certain aspects, such as language of service interface or re-use status of underlying software, etc.
- Link new contributions to related existing resource(s) and/or contributions.

- Interoperability of the registry with other Research Information Systems (for example, combining results from the Contribution Tool API with those drawn from the Europeana API, to present appropriate software and services alongside certain types and genres of digitised material held by European GLAM institutions).
- Visualization of the tiers of the DARIAH Hallmark in the DARIAH marketplace, showing the compatibility of a contribution to the infrastructure according to the self-assessment and peer review done in the contributions collection and assessment online tool.



D5.1 Report on Integrated Service Needs:
Annex 1
DARIAH Integrated Service References Architecture

HaS-DARIAH

INFRADEV-3-2015-Individual implementation and operation of ESFRI projects
Grant Agreement no.: 675570

Date: 12 October 2017

Version: 1.0



Project funded under the Horizon 2020 Programme

1. Introduction	5
1.1 Why a reference architecture for DARIAH-EU?	5
1.2 Methodology	6
1.3 Further Development	9
2. Glossary	9
3. In-kind contributions	12
3.1 What is an in-kind contribution?	12
3.2 In-kind contributions for DARIAH-EU	13
3.3 Classes of DARIAH-EU in-kind contributions	13
4. Services	14
4.1 DARIAH Service Provision	14
4.1.1 Enterprise Viewpoint	14
DARIAH Service Provision Roles	14
DARIAH Service Provision Behaviours	15
4.1.2 Information Viewpoint	16
Information Object types	16
4.1.3 Computational Viewpoint	18
CV Presentation Objects	20
Service Manager UI	20
Process Manager UI	21
Consumer UI	21
CV Service Objects	22
Administration Functions	22
User Authentication Service	22
Data Identification Service	22
Catalogue Service	23
Broker Service	23
Service Functions	24
Service Subsystem	24
Process Subsystem	25
Coordination Service	26
CV Component Objects	26

Data Curation Service	26
Provenance Data Subsystem	27
User Metadata Subsystem	27
Data Provision Subsystem	27
Data Creation Subsystem	28
Data Processing Subsystem	29
Data Transfer Subsystem	30
CV Back End Objects	31
Storage System.....	31
Service Registry	32
4.2 Data Hosting Service	32
4.2.1 Enterprise Viewpoint.....	33
4.2.2 Information Viewpoint.....	34
4.2.3 Case Study - EASY Data Archive	34
4.3 Processing Services.....	35
4.3.1 Enterprise Viewpoint.....	35
4.3.2 Information Viewpoint.....	36
4.3.3 Case Study: TextGrid.....	37
4.4 Support Services	37
4.4.1 Enterprise Viewpoint	38
4.4.2 Information Viewpoint.....	38
4.5 Access to Resources	39
4.5.1 Enterprise Viewpoint.....	39
4.5.2 Information Viewpoint.....	40
4.5.3 Case Study: the CENDARI AtOM	41
5. Activities	41
5.1 DARIAH Activity Provision.....	41
5.1.1 Enterprise Viewpoint	42
5.1.2 Information Viewpoint	43
5.1.3 Computational Viewpoint	45
CV Presentation Objects	45
Activity Management UI.....	45

Support Service UI	46
Consumer UI	46
CV Service Objects	46
Administration Functions	46
User Authentication Service	47
Activity Subsystem	47
Coordination Service	48
Support Services Service	49
CV Component Objects	49
Storage Controller	49
CV Back End Objects	50
Storage System	50
Activity Registry	50
5.2 Events	51
5.2.1 Enterprise Viewpoint	51
5.2.2 Information Viewpoint	53
5.3 Development of Software	54
5.3.1 Enterprise Viewpoint	54
5.3.2 Information Viewpoint	54
5.4 Consulting	55
5.4.1 Enterprise Viewpoint	56
5.4.2 Information Viewpoint	56
5.5 DARIAH Coordination	57
5.5.1 Enterprise Viewpoint	58
5.5.2 Information Viewpoint	58
5.5.3 Case Study	59
5.6 Creating Resources	59
5.6.1 Enterprise Viewpoint	60
5.6.2 Information Viewpoint	61
5.6.3. Case Study: Creation of linguistic data	62

1. Introduction

1.1 Why a reference architecture for DARIAH-EU?

The project Humanities at Scale (<http://has.dariah.eu/>) is currently developing a service for the submission of the in-kind contributions for the DARIAH-ERIC member states. Currently, the in-kind contributions delivered to DARIAH-EU are submitted via a FileMaker database, which has been customized by the DARIAH-EU team in the last years. This system fulfilled its goals in the starting years of the DARIAH ERIC, but it doesn't meet the requirements of stability and scalability that are requested in this growing phase of the infrastructure.

This document aims to develop a reference architecture¹ for the submission of in-kind contributions of DARIAH EU. This reference architecture develops in parallel with the current documentation on the in-kind contributions developed by WP5: in fact, it further develops their definition and components by using a formal language, borrowed from the theoretical models used in ICT systems.

The main goals of such reference architecture are:

1. To communicate and to describe contributions in a common language and to have a common understanding of what components should be described. The in-kind contributions have never been formally described before, including their creation processes. The result is that up to now the contributions have been interpreted and labelled in different ways by each contributing country, by making the in-kind registry highly heterogeneous and unreliable from an end user perspective.
2. To facilitate the description of the interactions between components of the in-kinds, according to the services or functions they provide. How do the components of the contributions interact? How are they connected? What this reference architecture brings to life is not only a reliable description of the DARIAH contributions, but their description as processes and evolving systems, rather than immovable objects. We think therefore that it is important to describe these changes and interactions with the environments in which the contributions evolve.
3. To offer an architectural reference for other projects and infrastructures that wish to offer similar services to their (research) communities.
4. To Mitigate and Reduce the risk. We consider this reference architecture as a way to reduce the risk mainly derived from point 1: that of interpreting the in-kind contributions

¹ Reference Architecture for a SSH Infrastructure: <https://sites.google.com/a/dans.knaw.nl/reference-model-for-ssh-data-infrastructure/part-1/introduction>

in different ways, to describe them without a common reference model. This risk exists not only among partners inside the same research infrastructure but also for anyone who accesses the DARIAH contributions from other infrastructures.

1.2 Methodology

In the process of creating a Reference Architecture for the DARIAH in-kind contributions, we were able to count on the support of already existing reference models and reference architectures, both in the fields of ICT, humanities and social sciences

The models that we used as a reference for our architecture are well known models implemented in the design of ICT processes like the RM-ODP (*Reference Model - Open Distributed Processing*). In addition, a recently developed model as the RM-SSH (*Reference Model for Social Science and Humanities*) for social sciences and humanities developed by the DASISH project provides a more specific reference model for our work on the DARIAH ERIC, whose field of action is in the digital arts and humanities.

A. RM-ODP Reference Model

Reference Model of Open Distributed Processing (RM-ODP) is a reference model in computer science which provides a framework to describe the architecture of open distributed processing (ODP); whenever possible RM-ODP uses a [formal description techniques](#) for specification of the [architecture](#), in order to guarantee consistency and reliability of such description.

This framework is composed of two main approaches:

- **Object modelling:** *Object-oriented modelling* (OOM) is an approach used to model applications and systems by applying the object-oriented paradigm to the entire system. Object modelling usually applies to the modelling of dynamic systems (e.g. business processes) as well as to static structures (components of a system)
- **The description of the system in different viewpoints** (*Viewpoint Modelling*): it is an effective approach when dealing with complex systems as it represents the subdivision of the specification of a complete system. Despite being relatively independent, the viewpoints are interconnected as some of the key items are identified in other viewpoints.

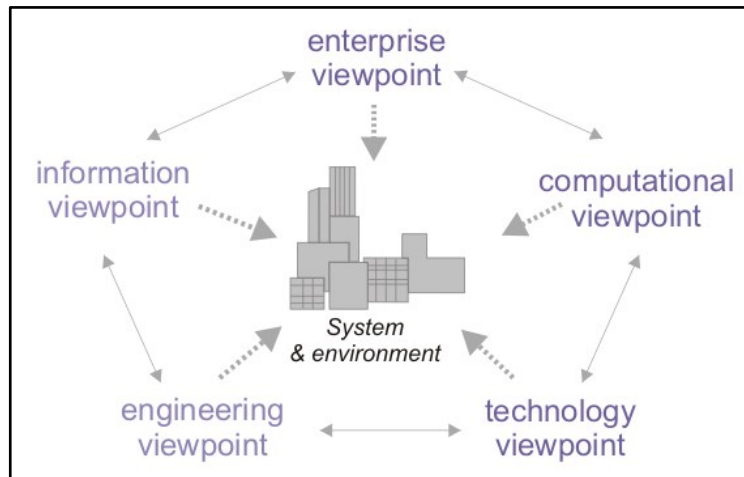


Figure 1: overview of the RM-ODP Reference Model viewpoints

Enterprise viewpoint: allows to focus on the purpose, scope and policies of a system by describing interrelated communities. These communities are described by community contracts which contain the objectives, roles, policies and behaviour. The SSH Data Infrastructure is composed of the following communities:

- **Data Creation Community**, who gather data from deployed instruments,
- **Data Management Community**, who ingest, administer and curate data,
- **Data Provision Community**, who provide discovery of, and access to data,
- **Data Processing Community**, who generate processed data from deployed processes,
- **Data Identification Community**, who provide resolution of identifiers for research data,
- **User Authentication Community**, who provide federated authentication services for users.

Information viewpoint: the focus is on the information object without considering representation, implementation or distribution details. The aim of the information viewpoint is to provide a common understanding of the model of the shared information for all stakeholders of the reference model. The information viewpoint is independent from functions that transform and manipulate the data and the computational interfaces. This viewpoint defines a configuration of information objects, their behaviours, actions that can be performed upon the information object and constraints. Those information objects are:

- **Data:** Data information object types can be content objects (observation, concept, and metadata), or container objects (dataset, inventory) grouping similar content object types.
- **Agents:** Agents are typically human users or machines acting on behalf of human users.
- **Service:** Services can be human- or automated services and are described by their behaviour or interaction.

- **Contracts:** Contracts usually exist between two agents, whereas standards exist within a community. They are used to describe agreements between two agents or within a community.

Computational viewpoint: describes the functionality of a system. It decomposes the system into objects performing specific functions with specific interfaces. The viewpoint is expressed using an object model which defines the interfaces that objects can have, how these can be bound with other interfaces, the interaction that can take place at these interfaces and the actions that an object can perform (such as creation of new objects, interfaces and bindings).

[In our deliverable, we won't work on the engineering viewpoint. Probably we won't take it into consideration for DARIAH/HaS.

The engineering viewpoint: focuses on the mechanisms and functions required to support distributed interactions between objects in the system]

B. DASISH SSH Social Sciences and Humanities Data Infrastructure Reference Model

The DASISH project² developed a Reference Model based on the previous observation that there was a lack of understanding of research infrastructures and the relationships between their entities: as Research Infrastructures are distributed infrastructures, they gather numerous stakeholders and have complex organisations. Their structure and functioning is therefore not always clear to an external user.

In this context, RM-ODP appears to be the most pertinent model to describe them, by providing definition of the core entities and their archetypal distribution.

² <http://dasish.eu/>. DASISH brought together all 5 ESFRI research infrastructure initiatives in the social sciences and humanities (SSH) represented each by some centers: [CLARIN](#), [DARIAH](#), [CESSDA](#), [ESS](#), [SHARE](#). Its goal was to determine areas of possible synergies in the infrastructure development and to work on a few concrete joint activities. The project ran from 1st January 2012 and ending on 31st December 2014.

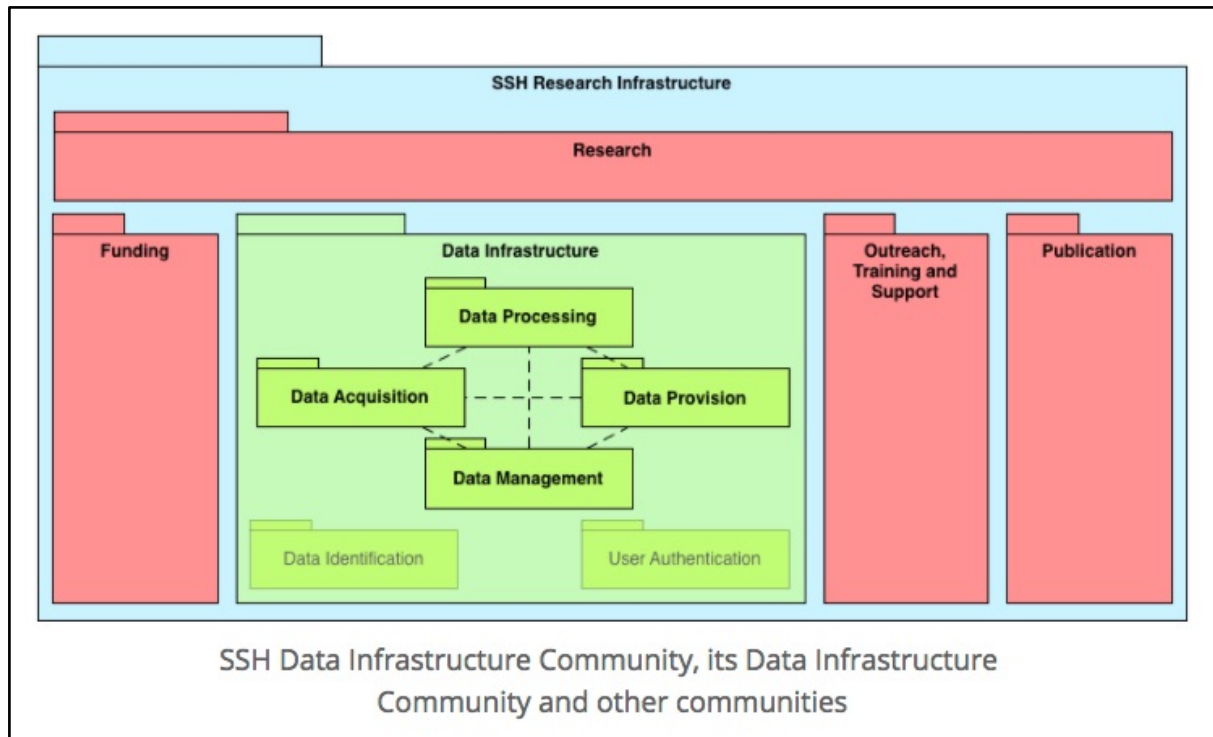


Figure 2: SSH Data Infrastructure Community, its Data Infrastructure Community and other communities

1.3 Further Development

This reference architecture is the first attempt, and a work-in-progress, to create a reference architecture in the Social Sciences and Humanities and the first based upon the SSH reference model. As such, it required validation against existing and future architectures for services and activities in DARIAH, and elsewhere.

Moreover, it requires further development of the viewpoints, and in particular the computational viewpoint, for DARIAH contribution subtypes (see chapter 3). This will aid the understanding of the computational viewpoint with concrete examples of how the described computational objects interact. Naturally the glossary needs refinement, because, as with all first drafts, it is never complete.

Furthermore, feedback from users of the reference architecture is essential to the further development as well as a governance structure for future releases.

2. Glossary

Activity: In general terms, an activity can be described as one or a set of actions, performed by an agent (single or institution) in order to achieve a result or product that can be shared with another agent. Differently from a service, an activity is discrete, as it is marked by a beginning and an end.

Cultural Heritage Object: analogue and original physical object or its digital representation of an artefacts from a cultural heritage institution

Cultural Heritage Institution: Institution that preserves the cultural heritage objects and that makes available digitally cultural object and/ or metadata of specific collections to researchers.

Archival Information Package: The definition of the term Archival Storage in OAIS includes the services and functions necessary for the storage of the Archival Information Package (AIP). Archival storage encompasses data management and includes processes such as storage media selection, transfer of AIP to storage system, data security and validity, backup and data restoration, and reproduction of AIP to new media. (IASA)

DARIAH-EU: Digital Research Infrastructure for Arts and Humanities. DARIAH is a European research infrastructure (ERIC). It connects several hundreds of scholars and dozens of research facilities in currently 17 european countries, the DARIAH member countries. People in DARIAH provide digital tools and share data as well as know-how. They organize learning opportunities for digital research methods, like workshops and summer schools, and offer training materials for the Digital Humanities.

DARIAH (affiliated) Institutions and Projects: initiatives whose activity and even existence are closely related on the technical and strategic background established by DARIAH (and vice-versa)

Data: digitized version of an existing object (when referred to a cultural heritage object); it can be research data (input or outcome of research, conducted through digital means); it can be information (e.g. data about a certain topic, which can be collected and stored)

Information object: Information objects model the data about entities in the real world, their state and behaviour and interactions with other information objects. An information object has a type that share a common set of features and behaviours (**RM- SSH, DASISH**)

Information object type: The information types are identified from an inventory of the communities. These were then abstracted to common information types that are independent of the communities into four categories: Data, Agent, Service and Contract.

In-kind contribution (IKC): “the provision of goods or services to an organisation by one of its members, valued in monetary terms according to rules agreed upon beforehand by the members of the organisation, and accounted for as part of the member’s contribution to the budget.”

Knowledge transfer: act of sharing and disseminating knowledge from one organisation to another, or from one field of expertise to another.

Metadata: Metadata is additional data about an informational object and can be "descriptive" or "structural" in nature. A standard is normally employed to define the form of the metadata takes.

OAI-PMH: The Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) is a low-barrier mechanism for repository interoperability. It is a protocol developed for harvesting (or collecting) metadata descriptions of records in an archive so that services can be built using metadata from many archives. An implementation of OAI-PMH must support representing metadata in Dublin Core, but may also support additional representations.

OAIS: [Reference Model for an Open Archival Information Systems](https://public.ccsds.org/pubs/650x0m2.pdf)
<https://public.ccsds.org/pubs/650x0m2.pdf>

Resource Creation Community: equivalent to the [Data Creation Community](#) in the [RM-SSH Enterprise Viewpoint](#)

Data Hosting Community: equivalent to the [Data Management Community](#) in the [RM-SSH Enterprise Viewpoint](#)

Research Infrastructure (RI): distributed organisation which provides services (e.g. tools, hosting, processing, archiving, displaying data), access to research data and community support for researchers.

Research Object: Research data part of a scientific investigation. It can be identified (DOI), described (metadata) and aggregated.

Service: a service can be described as one or multiple objects, systems or activities that an agent makes available to another agent, in order to enable the service using agent to perform a specific activity.

Service Provider: An organisation or group, whom together, provide a specific service for a designated community.

Service Consumer: an agent that requests, receives and utilises a service.

Software: programming interfaces used in order to produce code.

3. In-kind contributions

3.1 What is an in-kind contribution?

In-kind contributions are a form of support for not-for-profit organisations, communities, or infrastructures. Such contributions represent the provision of services or activities to an organisation by one of its members, valued in monetary terms according to rules agreed upon beforehand by the members of the organisation, and accounted for as part of the member's contribution to the budget.

The contribution can be in the form of a direct provision of a tangible asset to the infrastructure or an expenditure incurred directly by the member, which benefits the infrastructure and satisfies its high-level principles or objectives³.

In-kind contributions include goods, access to and use of services and facilities, expertise in the form of staff time, provision of or access to equipment. They must be viewed as necessary to carry out tasks of and in the infrastructure and meet the objectives commonly agreed by the members. The value of the contribution should be assessed according to the accounting rules agreed upon by the members in the infrastructure through a governance structure. These rules may rely on the costs incurred by the contributor or on standard cost equivalents defined to ensure fairness among partners (this applies in particular to the accounting of personnel costs). For research infrastructures, in-kind contributions represent a revenue stream. Even though they are not monetary contributions, they may represent a large portion of the research infrastructure's revenue⁴.

In-kind contributions help to build relationships within an infrastructure as the providers support the mission and activities without the need for large investments of cash and can collaborate to provide contributions jointly. During the construction of the infrastructure contributors become partners in that they may take over the responsibility of construction tasks and equipment, and effective interfacing with other parts of the infrastructure⁵. Fixed-term projects with specific goals and set of partners is an effective method in the solving of start-up difficulties. Ensuring that a provider is informed of what is done with their in-kind contribution encourages them to take on a greater role in development of the infrastructure⁶.

³ <http://www.copori.eu/1369.php>

⁴ Ibid

⁵ Ibid

⁶ Ibid

3.2 In-kind contributions for DARIAH-EU

The new submission and assessment system of contributions, which is currently in development by the project Humanities at Scale, differentiates between *contributions* and *in-kind contributions*.

Contribution	Any relevant research output or activity a DARIAH partner submits as DARIAH-related work. See The types of contributions and assessment criteria for typology and criteria.
In-kind contribution (IKC)	A contribution selected by National Coordinator as in-kind for given member country and year.

All the research outputs created by DARIAH affiliated institutions are considered as contributions: this means that the whole DARIAH research community can benefit from such services (e.g. digitized resources, tools, registries etc.) without them having been recognised as “in-kind contributions”, which represents a further level of selection and assessment.

This selection is performed by the country’s national coordinators, who decides which contributions will be officially submitted to DARIAH-EU as part of their annual participation to the DARIAH-ERIC. In a second phase, the selected contributions are assessed both for their formal aspects (e.g. responding URL) as well as for the quality of the content, performed by peer reviewers among the working groups and the VCC heads. Both submission and assessment of the contributions are performed in the DARIAH Contributions Tool, whose scope and functioning is explained in detail in the document “DARIAH (in-kind) contributions – Concept and Procedures”.

3.3 Classes of DARIAH-EU in-kind contributions

Type of contribution	Subtype	Remarks
SERVICE		
	DATA HOSTING SERVICE	e.g. data repository service, data deposit service, software repository.
	PROCESSING SERVICE	E.g. NERD service
	SUPPORT SERVICE	e.g. helpdesk, software maintenance
	ACCESS TO RESOURCES	e.g. educational resources, data resource, enriching/creating metadata not normally available ...
ACTIVITY		

	EVENT	e.g. summer school, webinar, training
	CONSULTING	
	DARIAH COORDINATION	
	CREATING RESOURCES	e.g. educational resources, data resource ... controlled vocabularies/thesauri
	DEVELOPMENT OF SOFTWARE	

4. Services

4.1 DARIAH Service Provision

Generally speaking, a service can be described as one or multiple objects, systems or activities that an agent makes available to another agent, in order to enable a user to perform a specific activity.

In the context of DARIAH a service represents an action that one or more institutions affiliated to DARIAH (the service provider/s) offers to another (DARIAH) institution or single researcher in order to enable the user of the service to reach a certain objective (related to the enrichment, dissemination and sustainability of research outputs).

4.1.1 Enterprise Viewpoint

The DARIAH Service Provision comprises the following community:

- **Generic Service Provision community:** the objective of the Generic Service Provision community is to provide value to an agent through bringing about results that the agent want to achieve, through the provision of a service environment, service delivery, and service product.

DARIAH Service Provision Roles

The following roles are identified in the DARIAH Service Provision community:

- **Service Provision Subsystem** (passive role): the community component representing the generic service provision community.
- **Service Owner** (active role): has the overall responsibility for the Service Management Subsystem. The Service Owner sets the key objectives and provides overall direction for the Service Management Subsystem through governance protocol, procedure and service agreements made.
- **Service Provider** (active role): an agent, usually an institution, which makes a service available for use by the Service Consumer in accordance with a service agreement.

- **Service Consumer** (active/passive role): an agent that requests, receives and utilises the service. Access to the service may be direct, or mediated via a third-party, but in accordance with a service agreement.
- **Service Designer** (active role): a person who designs, defines, and documents a service based upon the objectives set by the **Service Owner**.
- **Process Designer** (active role): a person who defines, designs and documents a process based upon the objectives of the **Process Owner**.
- **Service Manager** (active role): The Service Manager coordinates all efforts in the day-to-day running, planning, implementing, monitoring, reviewing and improving the service provision. The Service Manager is the point of contact for issues with the service
- **Service Management Subsystem** (passive role): the system and/or policies used to run a service provision.
- **Process Owner** (active role): Has overall responsibility for the process, sets the key objectives, and overall direction, goals, & governance for the process.
- **Process Manager** (active role): Coordinates all efforts in the day-to-day running, planning, implementing, monitoring, reviewing and improving the process. The Process Manager is the point of contact for issues with the process and manages incidents or changes to the process.
- **Process Implementer** (active/passive role): An agent which is involved in the execution one or more activities within a process. The agent carries out defined activities according to the defined process and, as applicable, its exception procedures and reporting incidents to the Process Manager.

DARIAH Service Provision Behaviours

The following behaviours are identified as part of the generic service provision:

- **Service Deployment:** the behaviour of the service subsystem performed by the Service Manager, the designer and the service management subsystem, whereby an service is requested, build, instantiated and deployed.
- **Service Provision Management:** the behaviour of the service subsystem performed by the Service Manager in taking care of the day-to-day running, planning, implementing, monitoring and reviewing the service provision.
- **Process Management:** behaviour of the process manager, who coordinates all efforts in the day-to-day running, planning, implementing, monitoring, reviewing and improving the process.
- **Activity Execution:** behaviour of the Process Implementer, who is involved in the execution of activities within a process.

4.1.2 Information Viewpoint

In the Information viewpoint, the information transmitted between the different stakeholders in the context of service providing within the infrastructure is represented. It is specified by Information Objects and their interrelationships.

The objective of the Information Viewpoint is to provide a common model for generic DARIAH Activities that occur in the infrastructure. It defines a set of Information Viewpoint objects and the set of actions acting upon those objects. The Information Viewpoint specifies the types of Information Objects and their interrelationships.

Platform-specific, or implementation, details are not considered here. The information viewpoint is independent of the Computational Viewpoint interfaces and functions that manipulate the objects, and the technology used to provide the technical infrastructure upon which the objects are stored or transmitted.

Information Object types

Data:

- **Data Object Type** (e.g. Datasets or digitized cultural object): the minimum available information object type that can be shared from an institution with an end user
- **Metadata Object Type**: the information object type that describes and contextualizes the data object type
- **Web applications** (Service Object Type) (e.g. web interfaces, portal, faceted research): the information object type that translates the data object type (in raw format) into a comprehensible information object for the end user.
- **Transfer Service Object Type** (e.g. API, OAI-PMH, FTP): the information object subsystem that makes it possible for the researcher to access the data
- **AAI response** (Object derived from the authentication service)
- **Service Documentation**: the documentation component of a service (e.g. procedure, list of documents used for the planning and execution of a service).

Agents: Agents are typically human users or machines acting on behalf of human users.

An [Agent refers to the information object type defined in the RM-SSH](#). There are a number of agents that may be considered part of the generic DARIAH Service Provision.

- **Service Owner**: the agent who owns the service and its running
- **Service Provider**: the agent who gives instruction to the implementer of a service
- **Service Implementer**: the agent that executes a certain action on behalf of the service provider.
- **Service Consumer**: an agent that requests, receives and utilises the service.
- **Service Designer** (active role): a person who designs, defines, and documents a service based upon the objectives set by the **Service Owner**.

- **Process Owner** (active role): Has overall responsibility for the process, sets the key objectives, and overall direction, goals, & governance for the process.
- **Process Designer** (active role): a person who defines, designs and documents a process based upon the objectives of the **Process Owner**.
- **Process Manager** (active role): Coordinates all efforts in the day-to-day running, planning, implementing, monitoring, reviewing and improving the process. The Process Manager is the point of contact for issues with the process and manages incidents or changes to the process.
- **Process Implementer**: the agent that executes a certain action on behalf of the process owner.
- **Activity Implementer**: the agent that executes a certain activity.
- **Service Management Subsystem** (passive role): the system and/or policies used to run a service provision.

Service:

Services can be human or automated services and are described by their behaviour or interaction.

- **AAI** (authentication subsystem): system to authenticate the user in order to access the resources

Contracts:

- **Service Level Agreement**: agreement defining the levels of service and their management between the Service Provider and the Service Consumer.

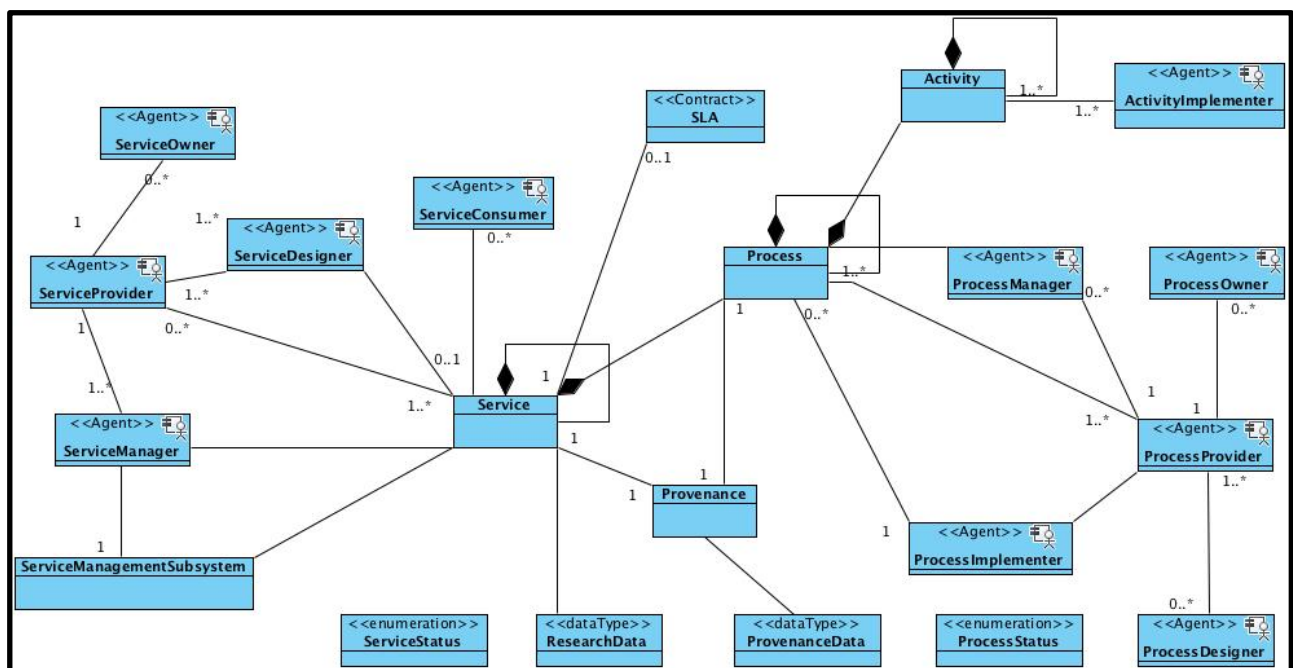


Figure 3: Service Information Object Types and Roles

4.1.3 Computational Viewpoint

A research Infrastructure (RI) such as DARIAH provides services which facilitate scholars to create, interact, process and store, or deposit, data. Other services maybe provided by the DARIAH community. Here we consider a general form of a service.

The computational viewpoint (CV) describes the functionality of a system. It decomposes the system into objects encapsulating specific functionality that are implemented by a service or activity. Specific interfaces bind the computational objects together, thus allowing the functionality to be distributed.

This viewpoint is expressed using an object model in which computational object (CV Object) are abstractions, in this instance, of a research infrastructure's systems & functions, and the interfaces (CV Interfaces) the computational objects can have, are defined.

Computational objects can relate to other computational objects in the following ways:

- **Interaction:** Two objects exchange information as either an operation (request/response), as a (continuous) flow, or as a signal.
- **Instantiation:** An object creates a new object via instantiation.
- **Inheritance:** An inherited object has the same interface(s) as its parent object and extends it with additional interface(s).

Specific **types of computational objects** are used here:

- **Binding object:** An object to support more complex interactions between two or more interfaces. These binding objects coordinates this interaction by providing bindings to all required interfaces.
- **Proxy object:** An object that provides all external interactions of a subsystem and takes care of managing the internal objects of that subsystem.
- **Controller object:** An abstract object to represent the external interaction with custom objects. It is used to represent specific instruments and processing.⁷

Research infrastructures, including DARIAH, tend to have a **service-orientated architecture** (SOA) with core functionality encapsulated and accessed via externally facing services and user interfaces which act as brokers.

The CV for a generalised DARIAH service contribution describes set of computational objects that could be expected to constitute a service within the infrastructure. For each of the specific contributing service types a subset of computational objects would normally be expected to be present in the SOA for the functionality of the contribution.

⁷ Reference Architecture for a SSH Infrastructure: <https://sites.google.com/a/dans.knaw.nl/reference-model-for-ssh-data-infrastructure/part-2/computational-viewpoint> [Accessed 2017.08.30]

Four groups of computational objects are described below which correspond to architectural layers in a SOA-like structure⁸:

- **Presentation Objects:** computational objects that provide access to the service by human users (user interfaces)
- **Service Objects:** computational objects that are service intermediaries that broker and orchestrate the use of the component objects (which may be distributed).
- **Component Objects:** computational objects that manage and provide access to data
- **Back End Objects:** computational objects that manage and maintain the data of the service

Figure 4 below shows the grouping of CV objects for a generalised DARIAH Service Contribution.

⁸ Broker and orchestration services have been included in the Service Object layer rather than a separate intermediate layer.

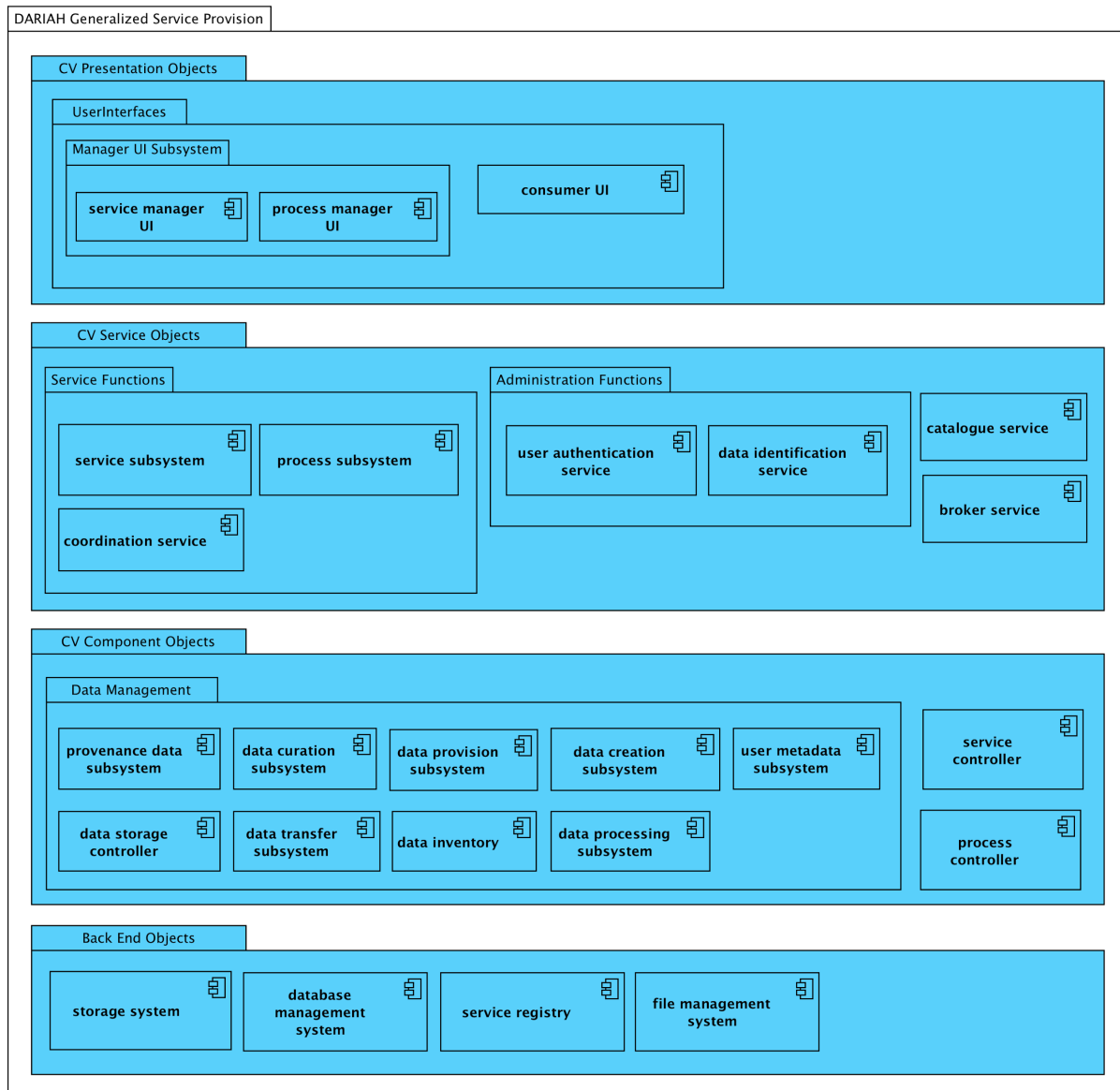


Figure 4: DARIAH Generalized Service Provision CV objects

CV Presentation Objects

CV Presentation objects are the entry points for human users to the services provided. The user interfaces consist of both service management and researcher (consumer) access points, or gateways.

Service Manager UI

The Service Manager UI provides an user interface to access the Service Subsystem for authenticated human users with management credentials.



Figure 5: Computational Viewpoint Service Manager User Interface object

Process Manager UI

The Process Manager UI provides an user interface to access the Process Subsystem for authenticated human users with management credentials.



Figure 6: Computational Viewpoint Process Manager User Interface object

Consumer UI

A community portal for interacting with a service, or services, of the distributed infrastructure.

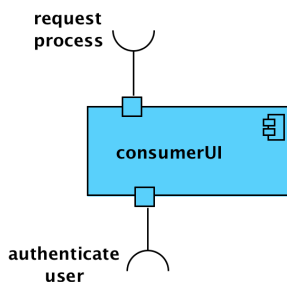


Figure 7: Computational Viewpoint Consumer User Interface object

The Consumer UI encapsulates the functions required to interact with the services provided by the research infrastructure externally to the infrastructure. It supports the following interactions:

- **request process (client):** takes input from the researcher to request and start a service.
- **authenticate user (client):** a client interface which requests confirmation of the identity of the user agent.

This object may take the form of a Virtual Research Environment (VRE), providing a persistent context for a researcher, or group of researchers, service component objects, research data, and interactions between these. Alternatively, it may be an object that facilitates access to a single discreet service.

CV Service Objects

Administration Functions

These objects may be internal or external to the RI and services, provided by an external agent such as a national authority.

User Authentication Service⁹

The **User Authentication Service** is a proxy object which confirms the identity of an agent when an agent makes a request to the service (or infrastructure). As part of the (generalised) service provision it is only composed of an authentication service object.

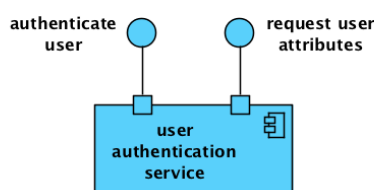


Figure 8: Computational Viewpoint User Authentication Service object

The **User Authentication Service** is a proxy object for all actions needed to verify the identity of an agent. It supports the following interactions:

- **Authenticate user (server):** is a public interface for determining whether the agent is who it claims to be.
- **Request user attributes (server):** is a public interface for requesting provided attributes from the agent.

Data Identification Service¹⁰

The **Data Identification Service** provides global identification and location functions for research data resources. It is composed of a persistent identification (PID) service object.

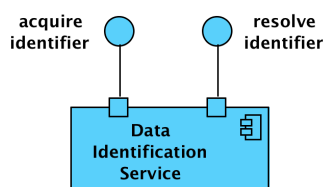


Figure 9: Computational Viewpoint Data Identification Service object

⁹ From the Reference Architecture for a SSH Infrastructure: <https://sites.google.com/a/dans.knaw.nl/reference-model-for-ssh-data-infrastructure/part-2/computational-viewpoint/user-athentication-subsystem>

¹⁰ From the Reference Architecture for a SSH Infrastructure: <https://sites.google.com/a/dans.knaw.nl/reference-model-for-ssh-data-infrastructure/part-2/computational-viewpoint/data-identification-subsystem>

The **Data Identification Service** is a proxy object for all actions needed to register, maintain, and retrieve the location of an identified object. It supports the following interactions:

- **acquire identifier (server)**. Public interface for registering new identifiers for data resources and updating the current location.
- **resolve identifier (server)**. Public interface for retrieving the current location of an identified data resource.

Catalogue Service

The Catalogue Service object provides resource discovery functions within the research infrastructure.

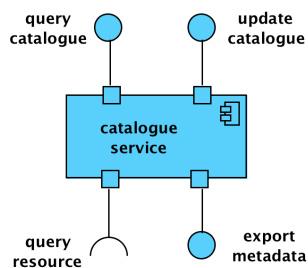


Figure 10: Computational Viewpoint Catalogue Service object

The **Catalogue Service** is a proxy object, which encapsulates all all the actions needed to allow research resources (data or otherwise) to be discovered. It supports the following interactions:

- **Update Catalogue (server)**: is an interface for registering, updating and indexing resources in the catalogue.
- **Query Catalogue (server)**: is an interface for querying the catalogue for registered/indexed research resources.
- **Query Resource (client)**: retrieves metadata from data stores or other catalogues and sources
- **Export Metadata (server)**: provides functionality to gather metadata.

Broker Service

The Broker Service facilitates data access and storage of research data. It facilitates data transfers, query requests, annotation and curatorial services. The broker service validates all requests and verifies the identity and privileges of agents making data requests.

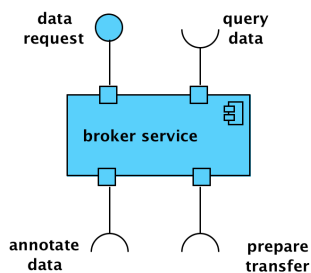


Figure 11: Computational Viewpoint Broker Service object

A **broker service** is a proxy object, which encapsulates all the actions required to access data. It supports the following interactions:

- **data request (server):** provides functions for requesting the import or export of datasets, the validation of the request and agent making the request.
- **annotate data (client):** is used to request annotation of data held within the data curation subsystem.
- **prepare data transfer (client):** is used to initiate and negotiate data transfers with the **data curation subsystem** or the **data provision subsystem**.
- **query data (client):** is used to forward queries onto the data provision object receive the results.

Service Functions

CV service functions offer access to distributed systems and resources (internal and external). This allows building of customer (researcher) services using both internal and external sourced components used in the research and data lifecycles.

Service Subsystem

The **Service Subsystem** provides service functions and is composed of a service subsystem object, a service workbench object and a service controller object. Each request to deploy a service on behalf of an agent, an instance of the service controller object is instantiated by the service workbench.

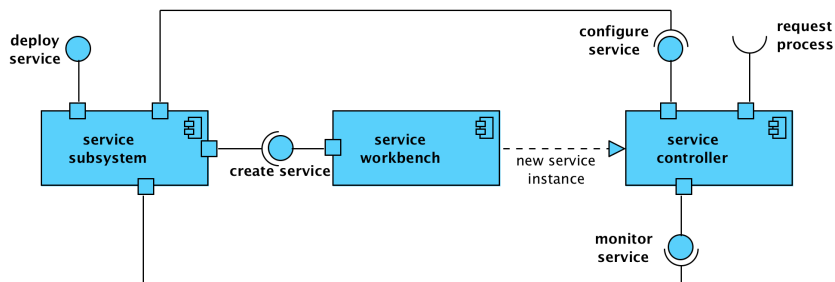


Figure 12: Computational Viewpoint Service Subsystem

The **Service Subsystem** is a proxy object for managing the service, input and resulting data in the service. It supports the following interactions:

- **deploy service (server):** the public interface for the deployment of a new service controller.

The **Service Workbench** is a computational object that instantiates service controller objects. It supports the following interactions:

- **create service controller (server):** is the interface for requesting a new service controller.
- **new service controller (instantiation):** Instantiation of a new service controller object by the service workbench.

The **Service Controller** is a controller object that runs the service. It supports the following interactions:

- **configure service (server):** the public interface for configuring and managing the service.
- **monitor service (server):** is the public interface for monitoring the service.
- **request process (client):** is used to deploy a process within the service workflow and controller from the **process subsystem**.

Process Subsystem

The **Process Subsystem** provides access to process functions and is composed of a process subsystem object, a process workbench object and a process controller object. For each request to deploy a process from a service controller, an instance of the process controller object is instantiated by the process workbench. A process can request to deploy a process by which a workflow (chain) or processes can be deployed within a service.

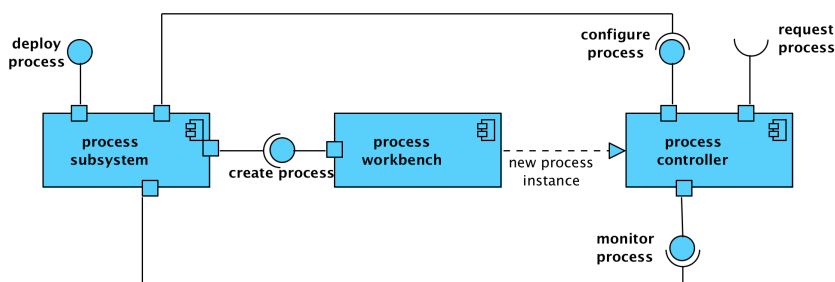


Figure 13: Computational Viewpoint Process Subsystem

The **Process Subsystem** is a proxy object for managing the process, input and resulting data in the service. It supports the following interactions:

- **deploy process (server):** the public interface for the deployment of a new process controller.

The **Process Workbench** is a computational object that instantiates a new process controller objects. It supports the following interactions:

- **create process controller (server):** is the interface for requesting a new process controller.
- **new process controller (instantiation):** Instantiation of a new process controller object by the process workbench.

The **Process Controller** is a controller object that runs the process. It supports the following interactions:

- **configure process (server)**: the public interface for configuring and managing the process.
- **monitor process (server)**: is the public interface for monitoring the process.
- **request process (client)**: is used to deploy a process within the service workflow and controller from the **process subsystem**.

Coordination Service

The **coordination service** delegates all servicing and processing tasks, coordinates multi-stage workflows and initiates execution and manages data flow within a service.

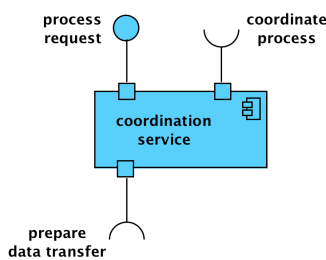


Figure 14: Computational Viewpoint Coordination Service object

The **coordination service** is a proxy object for processing tasks deployed on infrastructure execution resources. It supports the following interactions:

- **process request (server)**: which provides functions for scheduling the execution of process tasks.
- **coordinate process (client)**: is used to coordinate the execution of data processing tasks on execution resources presented by process controllers. Complex workflows involving many sub-tasks and processes may be orchestrated in this way.
- **prepare data transfer (client)**: This is used if data is required for or generated in a service and is moved into and out of the data store via the **data transfer service**.

CV Component Objects

Data Curation Service¹¹

The **Data Curation Service** allows data curation by monitoring data and performing transfer, processing and provisioning of the data. The subsystem is composed of a Data Curation Service object.

¹¹ From the Reference Architecture for a SSH Infrastructure: <https://sites.google.com/a/dans.knaw.nl/reference-model-for-ssh-data-infrastructure/part-2/computational-viewpoint/data-management-subsystem>

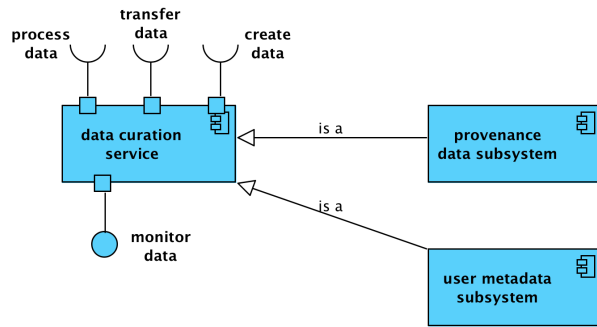


Figure 15: Computational Viewpoint Data Curation Service, Provenance Data Service and User Metadata Service

The **Data Curation Service** provides only one computational function, which is the monitoring of data. When needed, it consumes other subsystems to perform the actual curation. It provides the following interfaces:

- **monitor data (server)**: this is a public interface that allows data to be monitored.
- **transfer data (client)**: this is an interface for transferring data from e.g. data provisioning to data processing.
- **process data (client)**: the interface for processing the monitored data. This may include the calculation of checksums, or the transformation to new data formats.
- **create data (client)**: the interface for the creation of new data. This may include the creation of new metadata.

Provenance Data Subsystem

The **provenance data subsystem** is a specialised form of the **data curation service** which deals with the capture and maintenance of provenance data about the creation, processing, and dissemination of data within a service. See Figure 15: Computational Viewpoint Data Curation Service, Provenance Data Service and User Metadata Service

User Metadata Subsystem

The **user metadata subsystem** is a specialised form of the **data curation service** which deals with the capture and maintenance of user metadata within a service, for example user and groups privileges within a VRE instance of a VRE service. See Figure 15: Computational Viewpoint Data Curation Service, Provenance Data Service and User Metadata Service

Data Provision Subsystem¹²

The **Data Provision Subsystem** provides data sharing, data discovery and data access functions and is composed of a data provision service object, a **data inventory** service and a **data storage controller**.

¹² From the Reference Architecture for a SSH Infrastructure: <https://sites.google.com/a/dans.knaw.nl/reference-model-for-ssh-data-infrastructure/part-2/computational-viewpoint/data-provision-subsystem>

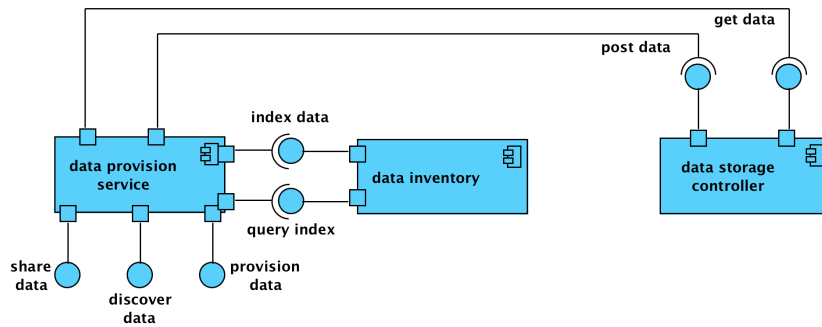


Figure 16: Computational Viewpoint Data Provision Subsystem

The **Data Provision Service** is a proxy service for the Data Provision Subsystem. It supports the following interactions:

- **share data (server)**: is a public interface for allow data to be shared.
- **discover data (server)**: is a public interface for searching provided data.
- **access data (server)**: is a public interface for accessing provided data.

The **Data Storage Controller** is a controller object that persists the published data. It supports the following interactions:

- **post data (server)**: is an interface for persisting data to the data storage controller.
- **retrieve data (server)**: is an interface for getting data from the storage controller.

The **Data Inventory** is a computational object that allows data to be discovered. It supports the following interactions:

- **index data (server)**: is an interface for registering/indexing data in the inventory .
- **query inventory (server)** is an interface for querying the inventory for registered/indexed data.

Data Creation Subsystem¹³

The **Data Creation Subsystem** creates data by recording observations. The subsystem is composed of a data creation service object, a creation instrument workbench object and an creation instrument controller object.

¹³ From the Reference Architecture for a SSH Infrastructure: <https://sites.google.com/a/dans.knaw.nl/reference-model-for-ssh-data-infrastructure/part-2/computational-viewpoint/data-creation-subsystem>

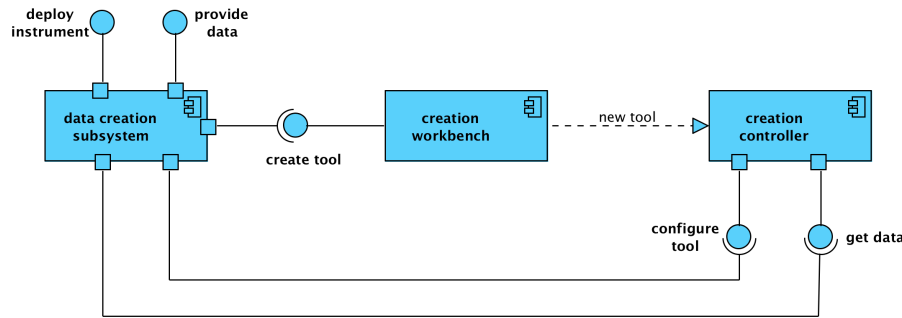


Figure 17: Computational Viewpoint Data Creation Subsystem

The **Data Creation Service** is a proxy object for managing the instruments and the resulting data. It supports the following interactions:

- **deploy instrument (server)**: is a public interface for the deployment of a new tool.
- **configure tool (server)**: is a public interface for controlling the tool.
- **provide data (server)**: is a public interface for retrieving the created data from the tool.

The **Creation Instrument Workbench** is a computational object that instantiates creation tool controller objects. It supports the following interactions:

- **create tool controller (server)**: is an interface for requesting a new instrument controller.
- **new creation tool controller (instantiation)**: an instantiation of a new tool controller object by the creation instrument workbench.

The **Creation Tool Controller** is a controller object that records observations. It supports the following interactions:

- **configure tool (server)**: is an interface for configuration of the tool controller.
- **get data (server)**: is an interface for requesting the data created by the controller.

Data Processing Subsystem¹⁴

The **Data Processing Subsystem** provides data processing functions and is composed of a data processing service object, a data process workbench object and an data process controller object.

¹⁴ From the Reference Architecture for a SSH Infrastructure: <https://sites.google.com/a/dans.knaw.nl/reference-model-for-ssh-data-infrastructure/part-2/computational-viewpoint/data-processing-subsystem>

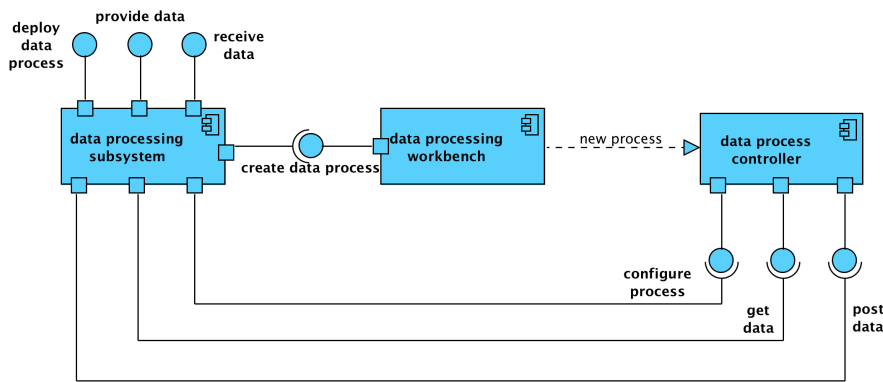


Figure 18: Computational Viewpoint Data Processing Subsystem

The **Data Processing Service** is a proxy object for managing the processing controllers and the data that is (to be) processed. It supports the following interactions:

- **deploy process (server)**: is a public interface for requesting a new process.
- **get data (client)**: is a public interface for retrieving the processed data.
- **post data (client)**: is a public interface for providing data to the process.

The **Data Process Workbench** is a computational object that instantiates data process controller objects. It supports the following interactions:

- **create data process controller (server)**: is an interface for requesting a process controller object.
- **new data process controller (instantiation)**: is the instantiation of the new data process controller object.

The **Data Process Controller** is a controller object that processes data. It supports the following interactions:

- **configure process (server)**: the interface for configuring the process.
- **get data (server)**: is the interface for retrieving processed data from the data process controller object.
- **post data (server)**: is the interface for providing data to the data process controller object.

Data Transfer Subsystem¹⁵

The **Data Transfer Subsystem** provides transfer services within the service and RI. The subsystem is composed of a Data Transfer Service and an Abstract Data Transfer Object, which may be instantiated as e.g. a Data Importer Object or a Data Exporter Object.

¹⁵ From the Reference Architecture for a SSH Infrastructure: <https://sites.google.com/a/dans.knaw.nl/reference-model-for-ssh-data-infrastructure/part-2/computational-viewpoint/data-transfer-subsystem>

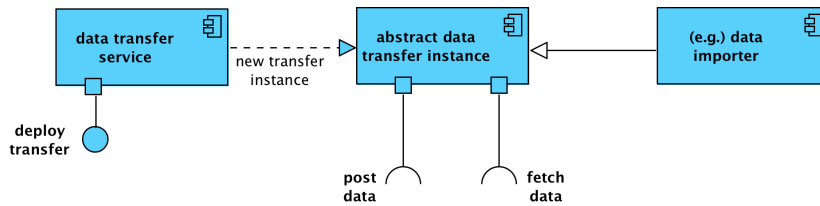


Figure 19: Computational Viewpoint Data Transfer Subsystem

The **Data Transfer Service** is a proxy object for providing data transfers. It supports the following interface:

- **deploy transfer (server):** is a public interface for requesting a data transfer.

-

The **Abstract Data Transfer Object** is an abstract object that provides the following interfaces.

- **fetch data (client):** is a client interface for retrieving data from a source.
- **post data (client):** is a client interface for providing the data to a target.

The example **Data Importer** implements the Data Transfer Object for specific environments. In this case, the data importer may specify how the data is fetched from a remote source, whether specific processing needs to take place (via a Data Processing Subsystem) and how the outcome should be provided to the target system.

CV Back End Objects

Back End Objects are objects which encompass the systems and resources provided for preserving, publishing, and processing research data through user accessible services.

Storage System

The **Storage System** is a system that manages and stores data and metadata of the research infrastructure.

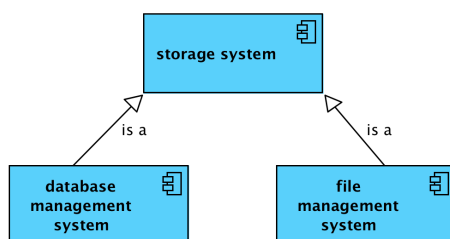


Figure 20: Computational Viewpoint Storage Systems

The **File Management System** manages the storage and retrieval of data as files in a computer system.

The **Database Management System** manages the storage and retrieval of data and metadata in logically structured repository systems.

Service Registry

The **service registry** is an information system for registering services within the research infrastructure.

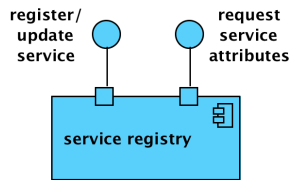


Figure 21: Computational Viewpoint Service Registry object

The **Service Registry** is a proxy object that encapsulates all actions needed to register, update and request service information. It supports the following interactions:

register/update service (server): is a public interface for registering a service and any maintenance events.

request service attributes (server): is a public interface for requesting provided attributes for a registered service.

4.2 Data Hosting Service

DEFINITION:

A service provider manages, oversees, and has responsibility for infrastructure, software and administrative tasks and makes this system available as a **hosted service** to users, usually, but not necessarily, over the Internet. A hosted service provider may consist of one organisation, or a collaboration between more organisations and/or sub-contracted specialist service providers. The hosted service may be provided on a dedicated or shared-service model¹⁶.

Hosting services are most often used for hosting Websites, but can also be used for hosting research data objects such as: audio-video qualitative data, quantitative data, software (and software configuration), digital documents, or information about research object and similar content. In the context of DARIAH IKC all research data object types are considered as **research data**, and hence services that concentrate on a specific object type as a specialisation of a hosting service.

A hosting service generally comprises of a repository, a depositing (acquisition) service, a discovery service, and (optionally) a provision service that gives access to the deposited data resources - even if the provision service per se is more connected with the contribution type “Access to resources”.

¹⁶ In a dedicated hosting environment, the service provider reserves technological infrastructure elements for each client. In shared hosting, services are provided to multiple clients from pooled resources.

4.2.1 Enterprise Viewpoint

COMMUNITIES

The **Hosting Service** IKC comprises the following communities:

Data Hosting Community: The Data Hosting Community is that who acquires, administers and curate those data created or collected by the **Resource Creation Community**. Within DARIAH, this community is represented by an institution which has the technical capacity of hosting research data on its servers, for a defined period of time.

Data Hosting Roles

- **Data Hosting Subsystem** (passive role): the community component representing the data hosting community as a whole.
- **Data Owner** (active role): a person or organisation that owns the data and determines the conditions for deposit/access with the Data Manager.
- **Data Manager** (active role): a person that is responsible for ensuring the quality of the research data during its ingest and management.
- **Preservation Manager** (active role): a person that is responsible for generating data products, and entering/registering them in the data administration service.
- **Data Administration System** (passive role): A system that is responsible for registering the data products for management purposes.

Data Hosting Behaviour

The following behaviours are identified in the Data Hosting Community:

- **Acquire Data:** the behaviour performed by the data manager, the data owner and the data administration system whereby they negotiate a submission agreement, transfer a submission information package to the data administration system and appraise it.
- **Curate Data:** the behaviour performed by the data manager, the preservation manager, and the data administration system, whereby the data manager check the quality of the data and removes faults. The data manager then adds meaningful documentation¹⁷ (metadata) to the data. The preservation manager transforms the data to a preservation format or migrates it to another platform, creates an Archival Information Package and stores it in a data administration system.
- **Administer Data:** the behaviour performed by the data manager and the data administration System whereby the required information for data management is administered.

¹⁷ Minimally technical and administrative metadata. The addition of descriptive metadata is considered here as a behaviour of the **data provisioning community**.

Communities that typically collaborate with the data hosting community related are: **resource creation community**, **data provisioning** community for retrieving (well-documented) data and for making the data accessible, and the **data processing community** for e.g. transforming data.

4.2.2 Information Viewpoint

Differently from the Enterprise viewpoint (where the community was central), for this type of in-kind contribution the **information type** represents the core of the investigation. The information type needs to be considered in itself, with no reference from the community that created or is responsible for it.

Information Object Types:

Data

For this type of contribution, the “data” are represented by the humanities research outputs that need to be hosted. These data can be organised in structured/ unstructured data; digitized cultural objects and its related (descriptive) metadata; digital and enhanced publications

Agents

The administrative roles that describe who can perform an action in the hosting system and what kind of role he/she can perform (e.g. administrator, editor, user). Three are the main agents identified at the information viewpoint level for the Hosting Services: The **Data Owner**, who owns and is responsible for the quality of the hosted data; the **Service owner** is responsible for the hosting service provided; the **Service Interface** represents the interface between the service owner/ provider and the data owner.

Service

The process of hosting data for an agent by a service provider.

Contract

It usually describes the relation between two or more agents. In the case of hosting of humanities data, the contract can establish the following responsibilities: for how long will the data be maintained? how will they be preserved? will they be enriched? how will they be accessed? what are the responsibilities of the agents? In terms of contracts, a **Hosting Agreement** can be identified between the Data Owner and the Service Owner.

4.2.3 Case Study - EASY Data Archive

Easy is the data archive for the social science and humanities in the Netherlands. Such data archive has been developed and is currently maintained and enriched by DANS, an institute of the Dutch Royal Academy of Science, that has a leading role (nationally and internationally) in the development of data quality policies and research data management plans.

As such, EASY offers to the research community the possibility to host their data after they

have finished their research and therefore can be considered as a candidate for a DARIAH contribution Data Hosting Service.

In terms of Enterprise Viewpoint, the **Data Hosting community** is represented by the DANS team, which is the community who acquires, administers and curates the dataset created by the **Resources Creation Community** (the researcher depositing the datasets). The team at DANS **acquires the data** (with the support of a **submission agreement**) through the submission of information packages.

To all the archived data, DANS applies the principles of the **Preservation Policies Data Archiving and Networked Services**, which draws the lines of the agreement between the Data Hosting Community and the Resource Creation Community.

In terms of information Viewpoint, the research outputs and the data delivered to EASY in order to be archived are the main **information object type** for this type of contribution. According to the recommendation of DANS, the data need to be structured according a series of preferred data and metadata formats and they need to include reference to their accessibility and reuse possibility. A **contract** (another data object type) will then establish duties and rights for the Data Hosting Community and the Data Resource Community by answering questions like: for how long will the data be maintained? what are the responsibilities of the stakeholders involved?

4.3 Processing Services

DEFINITION: A partner offers a (digital) service that applies some algorithmic processing (statistical analysis, annotation) on given data, or provides means to edit/curate data (e.g. web interface for collaborative editing).

Processing service is distinct from content hosting service in that it does not store data, but generates new data based on given input. Ideally, processing services are connected seamlessly with hosting service to appear to a user as one (virtual research) environment.

4.3.1 Enterprise Viewpoint

Communities

- **Data Creation Community:** The Data Creation Community is represented by whom creates or collects data. These data can be the results of digitization of cultural heritage objects or research outputs from scientific investigation. Within DARIAH-EU, this community can be represented by a single or a group of researchers, as well as by a single or a group of institutions.
- **Data processing Community:** this community allows the transformation (editing, curation, visualization) or the enrichment of the data created by the Data Creation Community by creating and applying an algorithmic process.

Roles

Roles within the Data Creation Community:

- **Data Creator:** the agent or agents responsible for the creation of the data
- **Instrument Designer:** the agent who is responsible to design, build and maintain the instrument that is used for data collection
- **Instrument:** a system that is used by the data collector to collect the data
- **Data Collector:** The agent that is responsible to collect the data created by the Data Creator

Roles within the Data Processing Community:

- **Processor Designer:** an agent who builds, maintains and configures an algorithmic processor.
- **Processor:** piece of software used by the Process consumer that transforms the data.
- **Process Consumer:** an agent that is responsible for applying the algorithmic process to the data created by the Data Creation Community.

Behaviours: the following behaviour is identified in the **Data Processing Community:**

Process Data

the behaviour performed by the processor (software) which processes the original data created by the data Creation Community; the behaviour performed by the Process Consumer (agent), when applies the algorithmic process to the data created by the data Creation Community.

4.3.2 Information Viewpoint

Data

For the in-kind contribution ‘Processing Services ‘ data are represented by the **algorithm** that is necessary to process the input data (Code) and the **data** themselves, that are the object of the the transformation. Finally, the **processed data** are also a data type in the information viewpoint and represent the data that are object of the transformation, at the end of the processing.

Agents

Three are the main agents identified at the information viewpoint level for the Processing Services: the **Data Owner**, who owns and is responsible for the quality and outcome of the data after their processing. The **service owner** is responsible for the processing service provided as well as the application of the processing algorithm. The **service interface** represents the interface between the service owner/ provider and the data owner.

Service

The **algorithmic processing** lies at the heart of this in-kind contribution, as it represents the core activity and service that service owner can provide to a data owner.

Contract

In terms of contracts, a **Processing Agreement** can be identified between the Data Owner and the Service Owner.

4.3.3 Case Study: TextGrid

TextGrid is a virtual research environment for researchers in the text-based humanities and cultural studies. Since 2006, TextGrid has been funded by the German Federal Ministry of Education and Research and, today, forms an integral part of DARIAH-DE. TextGrid has a large and growing community of users that roughly fall into the categories of individual researchers and research projects of varying sizes.

Among other features, it supports the creation of digital editions using free, project-specific and expandable tools and services. Tools, data and methods can be used independently, regardless of operating system, software equipment and location.

TextGrid consists of the TextGrid Laboratory (TGLab), a desktop application for all major operating systems, and the TextGrid Repository (TGRep). TGLab is the central **Service Interface** which is used as the **Instrument** by which a **Data Owner** ingests her research data provided by the **Data Creator** into the TextGrid storage, which is managed by DARIAH-DE infrastructure partners as the **Data Collectors**. The data can be used further in TGLab's tools, the **Processors**, operated, maintained and supported by DARIAH-DE development partners as **Service Owners**. Within TGLab the user can use tools such as the Text-Image-Link Editor, a **Processor**, to enrich images with text annotation and link the image with (parts of) its plain text transcription. This **Processed Data** can, ultimately, be published to the **Hosting Service** TGRep.

4.4 Support Services

DEFINITION:

Support services are designed to help for the use of a tool or a service. They also contribute to the update and maintenance for a service an institution provides.

This type of contribution includes helpdesks (services charged to answer requests from the users). It can be an automated helpdesk (email address).

On a second level, support services can be the components helping to maintain the configuration for a software. It can be human support, on the base of a service level agreement, or a helpdesk providing the information for the maintenance of a service.

4.4.1 Enterprise Viewpoint

Community

- **Support Provision Community:** Community providing and maintaining the support services.
- **Support Receiving Community:** Community requesting the support services.

Roles

- **Support provider:** the agent who legally provides the support activity (e.g. institution)
- **Support implementer:** the agent/or system who performs the support service
- **Support receiver:** subject of the consultancy: the agent who instantiates the request for support and receives the outcome of it.
- **Support Provision System (passive):** the system and/or policies used to run a support service.

Behaviours

- **Support management:** the behaviour of the Support provider, who overviews the running of the support services, and ensures the maintenance of the service.
- **Support execution:** the behaviour of the support implementer, who maintains the service.

4.4.2 Information Viewpoint

Data Object Types:

Data

- **Support Documentation:** the documentation component of the support service (Helpdesk)

Agents

- **Support provider:** the agent who legally provides the support activity (e.g. institution)
- **Support implementer:** the agent/or system who performs the support service
- **Support receiver:** the agent who instantiates the request for support and receives the outcome of it.

Service

- **AAI (authentication subsystem):** system to authenticate the user in order to access the support service.

Contracts

- **Support Service Contract:** Service Level Agreement between the Support provider and the Support implementer, that provides a support activity

4.5 Access to Resources

DEFINITION:

(Online) access to resources (datasets). This can be raw data, (web) applications offering rich access to data, or web services that allow a programmatic access to the data (API).

This type of contribution does not include (national or domain-specific) aggregators, or metadata catalogues, i.e. services that collect information about resources from multiple providers and allow to browse and search in it. These will/can ideally become a source of information about contributions.

It also does not include web services processing data (see Processing services contribution).

4.5.1 Enterprise Viewpoint

Community

Data Provision Community: who provides discovery of, and access to data. Within DARIAH this community can both be represented by a community of researchers/research institutions (that through means of communications, disseminates information about the deposited data, or provides interfaces for the discoverability of the deposited data) or by the same community that provides the storage for the research data.

Roles

- **Data Provision Subsystem** (passive role): the community component representing the data provision community.
- **Data Owner:** who owns the data
- **Data Provider** (active role): a person or organization that is responsible for making the research data from the data owner available for the data consumer.
- **Data Discovery System** (passive role): a system facilitating discovery of research data and its context.
- **Documentalist** (active role): a person who documents data with metadata according to widely accepted standards of **Data Management community**.
- **Data Retrieval System** (passive role): a system that enables retrieval of research data. Data retrieval includes access control (and thus, authorization).
- **Data Consumer** (passive/active role): an agent that requests and receives the research data (for further use). Access to the data may be direct, or mediated via a third-party.

Behaviours

- **Data Publication:** A behaviour of the data provider, the Data Discovery System and the Data Retrieval System for making research data discoverable and accessible.
- **Data Discovery:** A behaviour of the data consumer and the Data Retrieval System whereby the data consumer uses the Data Retrieval System to discover data of interest.
- **Data Retrieval:** A behaviour by the Data Consumer and the Data Retrieval System whereby the Data Consumer request data or metadata and the Data Retrieval System

responds, after successful authorization, with the requested metadata and/or data. The metadata may be used for referral.

Communities that typically collaborate with the data provision community related are: **data management community** for improving metadata for discovery, **data processing community** for harmonisation or just-in-time conversion during access, **data identification community** for global identification or referral, **user authentication community** for global authentication of data consumers and **data creation community** for capturing enrichments to the provisioned data.

4.5.2 Information Viewpoint

Data

- **Data Object Type** (e.g. Datasets or digitized cultural object): the minimum available information object type that can be shared from an institution with an end user
- **Metadata Object Type**: the information object type that describes and contextualizes the data object type
- **Web applications** (Service Object Type) (e.g. web interfaces, portal, faceted research): the information object type that translates the data object type (in raw format) into a comprehensible information object for the end user.
- **Transfer Service Object Type** (e.g. API, OAI-PMH, FTP): the information object subsystem that makes it possible for the researcher to access the data
- **AAI response** (Object derived from the authentication service)

Agents

- **Data Owner**: the agent who owns the data that are made accessible
- **Data Consumer**: the agent who access the data
- **Data Manager**: the agent who makes sure that the quality of the data is met and that the user experience is satisfactory

Services

- **User guide/ manual**: information on how to access the resources
- **AAI** (authentication subsystem): system to authenticate the user in order to access the resources

Contracts

- **Contract**: e.g. reuse licenses and intellectual property rights clearance
- **Agreement**: An agreement (e.g. a standard) defines the choice by a community to use a certain standard in order to harmonize the access to the data and metadata.

4.5.3 Case Study: the CENDARI AtOM

The CENDARI project ran from January 2012 to January 2016 funded under the FP7 grant scheme. CENDARI created a data space whose data have been created in two different ways. On the one hand memory institutions and aggregators delivered their data to the CENDARI federated data space. On the other hand, CENDARI made available hundreds of the so-called “hidden collections” from European archives: these collections were not aggregated by third institutions, nor were they digitised - in fact they often belong to small archives whose digitisation resources are often lacking. Rather, the CENDARI researchers manually encoded in a structured format the metadata describing those collections, by making these collections discoverable and reusable for the first time.

From an enterprise viewpoint, the **data provision community** in this case the CENDARI researchers - played a crucial role in this contribution, by providing a first and unique access to resources that otherwise would have never been accessible. The data provision community related directly with the **data owner**, which in this case is represented by the Archive, who owns and archives the original documents and artefacts.

The **data discovery system** - the tool AtOM in our case, facilitates the discovery of the research data and their context; AtOM also acts as the **Data Retrieval system**, which is a passive role as its actions are not tangible directly, but nonetheless allow actions like user-control and thus authorisation.

From an information viewpoint, the user interface represents the information object “**web applications**”, as it translated the data object type into a comprehensible information object for the end user.

In the case of AtOM, then, the Data object type, can be described in tight connection with the metadata object type, as in our case the object is purely made of metadata, without any object attached to it.

In terms of service...

5. Activities

5.1 DARIAH Activity Provision

In general terms, an activity can be described as one or a set of actions, performed by an agent (single or institution) in order to achieve a result or product that can be shared and valued by another agent.

In the context of DARIAH, an activity is interpreted as a collection of people, work items, communications and processes to achieve a specific goal¹⁸, whose outputs can be made

¹⁸ IBM Knowledge Centre

https://www.ibm.com/support/knowledgecenter/SSKTWP_8.0.1/com.ibm.openactivities.client.doc/r_oa_c_what_is_an_activity.html

accessible and valued by the DARIAH community (e.g. single researchers as well as research institutions or cultural heritage institutions).

Differently to the services, an activity is discrete, as it is marked by an agreed beginning and an end. An activity can be part of a service or of a process.

The examples listed below provide an overview of different activities performed in the DARIAH-ERIC, by displaying different types of data, different communities involved in the creation as well as in the usage/ consumption of the outputs of such activities.

5.1.1 Enterprise Viewpoint

Who are the communities involved in the performing of DARIAH activities?

- **Data Creation Community:** The Data Creation Community is involved in the DARIAH activities at different levels and different in-kinds contributions, especially in the “Development of Software” and “Creating resources in-kinds”. This community creates the research data, which in the DARIAH case are represented by digitized cultural objects, as well as training resources.
- **Data Management Community:** also in this case, this community is involved in different DARIAH activities, especially in the DARIAH “Consulting” in-kinds. The Data management community has two main aims: on the one hand, it makes sure that the data quality is as high as possible. It also administers and curates the research data.
- **Data Provision Community:** It’s involved in the DARIAH activities, especially at the level Consulting and Coordination. The ultimate goal of the DARIAH Coordination is that of making the data and services in DARIAH recognized among the research community. Similarly, as part of the Consulting Activity, the main aim might be that of making the DARIAH resources and activities well known and accessible.

Roles

The following roles are identified in the Activity community

- **Activity Creation Subsystem** (passive) the community component representing the activity creation community
- **Activities Owner:** (single or institution) the agent who owns and is ultimately responsibility of the owned activity
- **Activity Provider (different from the activity implementer):** the agent (individual or institution) that performs a certain action in the context of DARIAH
- **Activity Implementer:** the agent that executes a certain action on behalf of the activity provider.
- **Activity Consumer:** the agent (individual/ institution) that receives, consumes or experiences the outputs created during an activity.
- **Activity Manager:** has a day to day overview of the running of the activity. This role is also the point of contact for issues with the expected running of the activity

- **Activity Management System** (passive): the community component representing the activity management community (e.g. management tools and object buro)
- **Activity Support:** agent who provides support for the planning and running of an activity. This can be represented by a service (link to service here).

Behaviours

The following behaviours are encountered:

- **Activity management:** the behaviour by the activity manager, who overviews the running of the activities on a daily basis.
- **Activity support:** the behaviour of the activity support agent, offering support for the planning and running of an activity.
- **Activity execution:** the behaviour of a number of agents (activity provider, activity implementer) performing certain actions in the context of an activity.

5.1.2 Information Viewpoint

The objective of the Information Viewpoint is to provide a common model for generic DARIAH Activities that occur in the infrastructure. It defines a set of IV objects and the set of actions acting upon those objects. The IV specifies the types of Information Objects and their interrelationships.

Platform-specific, or implementation, details are not considered here. The information viewpoint is independent of the Computational Viewpoint interfaces and functions that manipulate the objects, and the technology used to provide the technical infrastructure upon which the objects are stored or transmitted.

Information Object types

Data

- **Activity Documentation:** the documentation component of an activity (e.g. procedure, list of documents used for the planning and execution of an activity).

Agent

An [Agent refers to the information object type defined in the RM-SSH](#)¹⁹. There are a number of agents that may be considered part of the generic DARIAH Activity Provision.

- **Activity Owner:** the agent who owns the activity and its running
- **Activity Provider:** the agent who gives instruction to the implementer of an activity

¹⁹ <https://sites.google.com/a/dans.knaw.nl/reference-model-for-ssh-data-infrastructure/part-2/information-viewpoint>

- **Activity Implementer:** the agent that executes a certain action on behalf of the activity provider.
- **Activity Consumer:** the agent (individual/ institution) that receives, consumes or experiences the outputs created during an activity.
- **Activity Manager:** has a day to day overview of the running of the activity. This role is also the point of contact for issues with the expected running of the activity
- **Activity Management System:** the community component representing the activity management community (e.g. management tools and project officer)

Service

- **Support Service:** series of services that make an event possible (e.g. catering, AV recording)

Contract

- **Activity Agreement:** between the activity provider and the activity owner, and the activity consumer.
- **Support Service Contract:** between the activity provider and service provider, that provides a support activity (e.g. catering, venue)

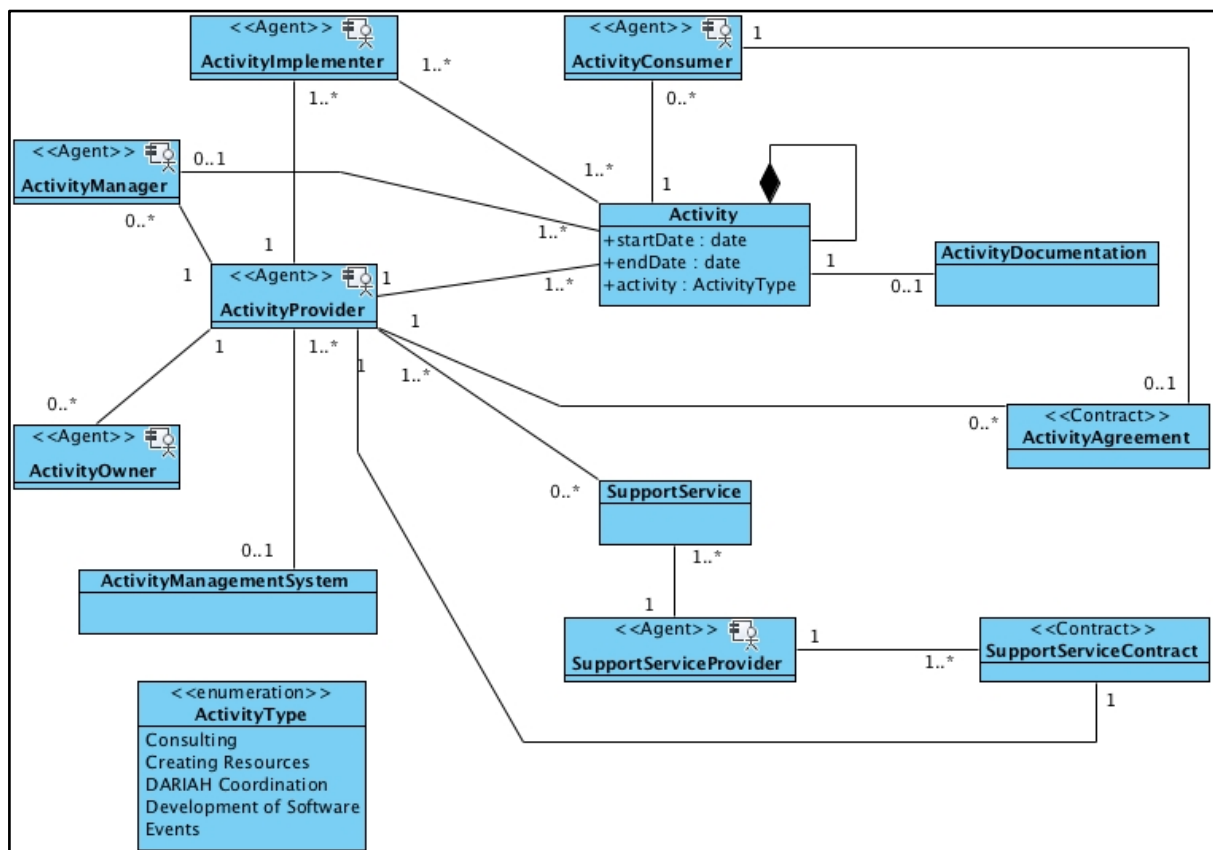


Figure 22: Information Viewpoint - Activity Objects and Roles

5.1.3 Computational Viewpoint

The CV for a generalised DARIAH activity contribution describes set of computational objects that could be expected to constitute an activity within the infrastructure. For each of the specific contributing activity types a subset of computational objects would normally be expected to be present in the SOA for the functionality of the contribution.

As per the CV for a generalised service, there are four groups of computational objects as can be seen in figure 7 below.

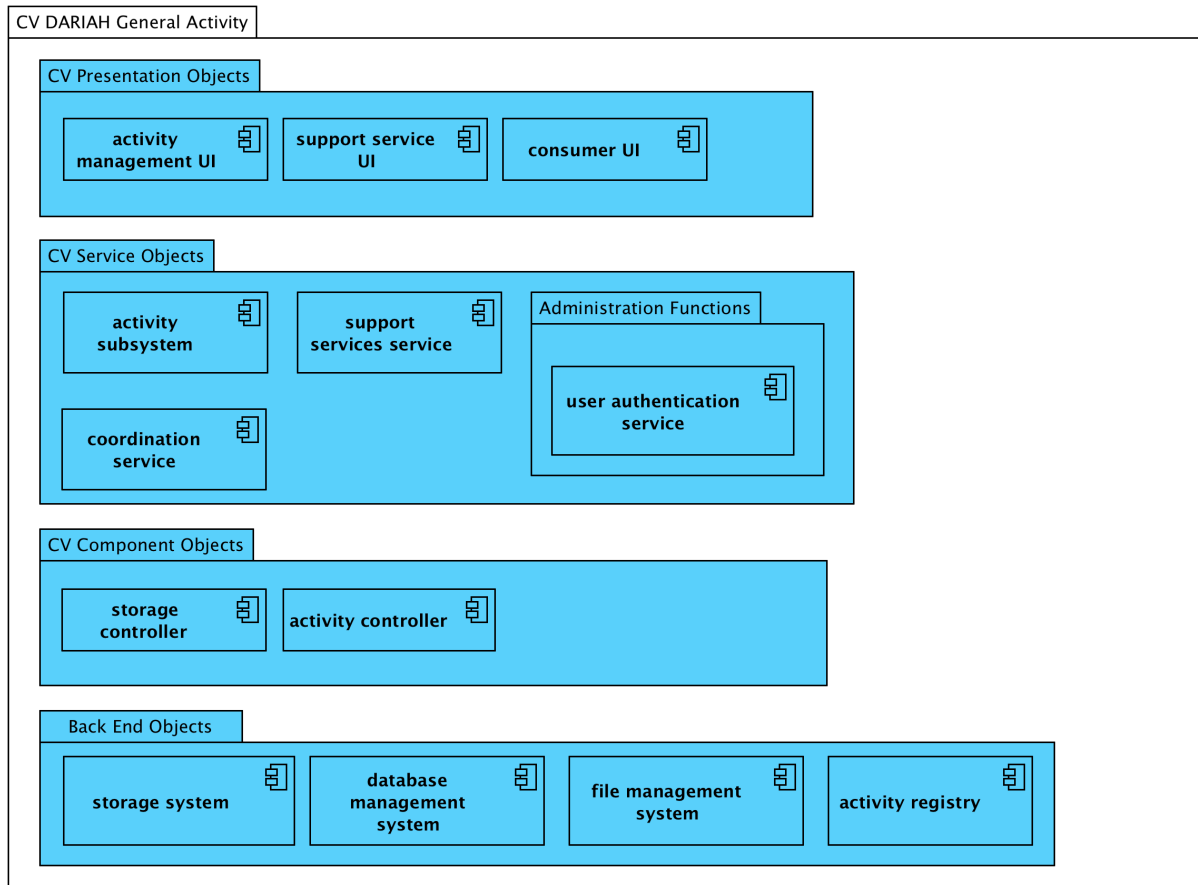


Figure 23: Computational Viewpoint DARIAH General Activity

CV Presentation Objects

Presentation objects are the entry points for human users to the services provided. The user interfaces consist of both service management and researcher (consumer) access points, or gateways.

Activity Management UI

The Activity Management User Interface provides an user interface to access the **Activity Subsystem** for authenticated human users with management credentials.



Figure 24: Computational Viewpoint Activity Management User Interface object

Support Service UI

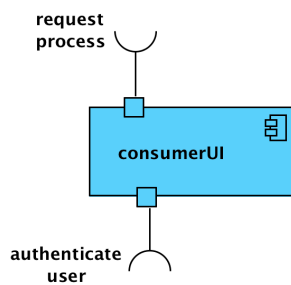
The **Support Service User Interface** provides a user interface to access the **Support Services Service** for authenticated human users with management credentials.



Figure 25: Computational Viewpoint Support Service User Interface

Consumer UI

A community portal for interacting with an activity, or activities, of the distributed infrastructure.



The Consumer UI encapsulates the functions required to interact with the activities provided by the research infrastructure externally to the infrastructure. It supports the following interactions:

- **request process (client):** takes input from the researcher to request and start a service.
- **authenticate user (client):** a client interface which requests confirmation of the identity of the user agent.

CV Service Objects

Administration Functions

These objects may be internal or external to the RI and activities, provided by an external agent such as a national authority.

User Authentication Service²⁰

The **User Authentication Service** is a proxy object which confirms the identity of an agent when an agent makes a request to the service (or infrastructure). As part of the (generalised) service provision it is only composed of an authentication service object.

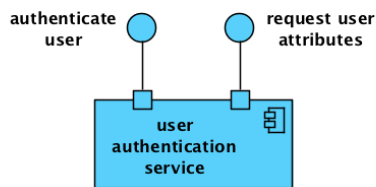


Figure 5: Computational Viewpoint User Authentication service object

The **User Authentication Service** is a proxy object for all actions needed to verify the identity of an agent. It supports the following interactions:

- **Authenticate user (server)**: is a public interface for determining whether the agent is who it claims to be.
- **Request user attributes (server)**: is a public interface for requesting provided attributes from the agent.

Activity Subsystem

The **Activity Subsystem** provides access to activity functions and is composed of an activity service object, an activity workbench object and a activity controller object. For each request to deploy an activity from a service controller, an instance of the activity controller object is instantiated by the activity workbench.

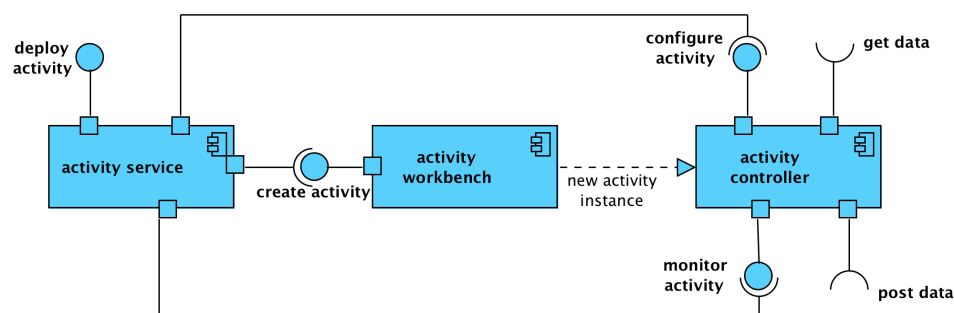


Figure 26: Computational Viewpoint Activity Subsystem

The **Activity Service** is a proxy object for activity subsystem, managing the activity. It supports the following interactions:

²⁰ <https://sites.google.com/a/dans.know.nl/reference-model-for-ssh-data-infrastructure/part-2/computational-viewpoint/user-authentication-subsystem>

- **deploy activity (server):** the public interface for the deployment of a new activity controller.

The **Activity Workbench** is a computational object that instantiates a new activity controller objects. It supports the following interactions:

- **create activity controller (server):** is the interface for requesting a new activity controller.
- **new activity controller (instantiation):** Instantiation of a new activity controller object by the activity workbench.

The **Activity Controller** is a controller object that runs the activity. It supports the following interactions:

- **configure activity (server):** the public interface for configuring and managing the activity.
- **monitor activity (server):** is the public interface for monitoring the activity.
- **get data (client):** is the interface to get data for use in the activity.
- **post data (client):** is the interface to receive data created in the activity.

Coordination Service

The **coordination service** delegates all activity tasks. It coordinates multi-stage workflows and initiates execution and manages any data flow within an activity.

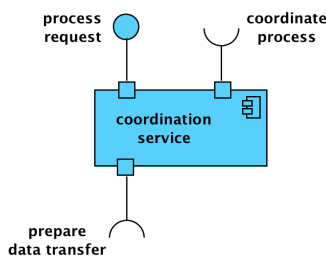


Figure 27: Computational Viewpoint Coordination Service object

The **coordination service** is a proxy object for activity tasks deployed on infrastructure execution resources. It supports the following interactions:

- **process request (server):** which provides functions for scheduling the execution of activity tasks.
- **coordinate process (client):** is used to coordinate the execution of activity tasks on execution resources presented by activity controllers. Activity workflows may be orchestrated in this way.
- **prepare data transfer (client):** This is used if data is required for or generated in a service and is moved into and out of the data store via the **data transfer service**.

Support Services Service

The **support services service** provides access to services provided externally to the research infrastructure and activity. For example, catering services for an event, or software publishing services (e.g. GitHub).

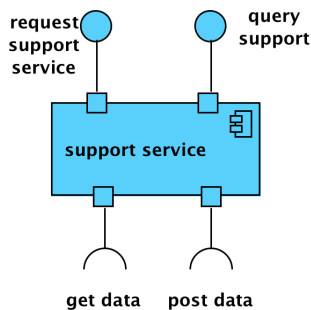


Figure 28: Computational Viewpoint Support Services Service object

The **support service** is a proxy object for support services service for external services. It supports the following interactions:

- **request support service (server):** is the interface for initiating a support service for an activity.
- **query support (server):** is the interface which provides status information about the support service.
- **get data (client):** is the interface to get data for use in the support service.
- **post data (client):** is the interface to receive data created in the support service.

CV Component Objects

Storage Controller

Data storage needs for an activity are significantly less complex than potentially for a service, for example, a data hosting and preservation service, thus an activity requires storage controller objects to interface with storage system used by the activity.

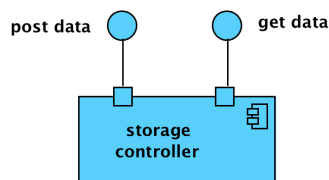


Figure 29: Computational Viewpoint Storage Controller object

The **Storage Controller** is a controller object that manages the storage used in an activity. It supports the following interactions:

- **get data (client):** is the interface to get data from the storage system.
- **post data (client):** is the interface to post data to the storage system.

CV Back End Objects

Back End Objects are objects which encompass the systems and resources provided for preserving, publishing, and processing research data through user accessible services.

Storage System

The **Storage System** is a system that manages and stores data and metadata of the activity in a research infrastructure.

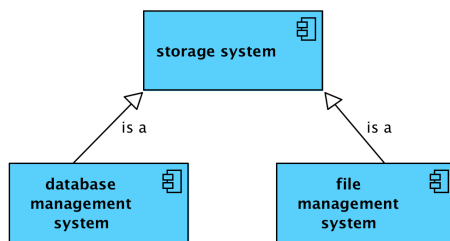


Figure 30: Computational Viewpoint Storage Systems

The **File Management System** manages the storage and retrieval of data as files in a computer system.

The **Database Management System** manages the storage and retrieval of data and metadata in logically structured repository systems.

Activity Registry

The activity registry is an information system for registering activities within the research infrastructure.

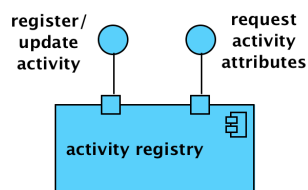


Figure 31: Computational Viewpoint Activity Registry

The **Activity Registry** is a proxy object that encapsulates all actions needed to register, update and request activity information. It supports the following interactions:

register/update activity (server): is a public interface for registering an activity and any maintenance events.

request activity attributes (server): is a public interface for requesting provided attributes for a registered activity.

5.2 Events

DEFINITION:

This type of in-kind contribution describes **events** organised in the field of digital arts and humanities and cultural heritage, whose aim is to:

- offer a theoretical insight of the methods, practices, history and future developments in the field of digital arts and humanities
- showcase case studies in the field of digital arts and humanities
- offer training and guidance in the application of digital methodologies in traditional humanities curricula

Examples for this type of contribution are:

- lecture
- Symposium
- Conferences
- workshop
- training day/s
- summer/ spring/ winter school
- academic course
- academic program
- MOOCs and other forms of online education

5.2.1 Enterprise Viewpoint

Communities

The “Event” in-kind contribution comprises the following communities:

Outreach, Training & Support Community: who provides community services to the SSH research community. They interact with the data infrastructure community for developing standards or for providing training for or about the data infrastructure community. This community is not directly involved in the data creation process, but it rather disseminates and gives training on the data developed and processed.

Roles

- **Event owner:** the agent who initiates and is responsible for the event; it can be an individual or the organization
- **Event host:** institution hosting the event
- **Event manager:** the agent who supervises the organization of the event on a daily basis in view of its successful outcome
- **Event deliverer:** the agent who delivers the event, (e.g. a conference or an academic course)
- **Event participant:** the agent who attends the event or training

- **Event Support:** the agent who provides support with organizational aspects of the event or training
- **Event documentation System (Passive):** written material that accompanies the event or the training material provided in support of a training
- **Event Registration System (Passive):** the system that allows the registration of the participants to the event
- **Event Assessor:** the agent who i.e. assesses the papers, abstract for the conference // or assesses the test at the end of the training; in training, i.e. the assessor assesses the quality of the outcomes and maybe gives a certificate
- **Event Assessment System (Passive):** the community component representing the event assessment community

EXAMPLE BEHAVIOURS IN THE PREPARATION OF A DH EVENT

1. **Paper submission:** the behaviour of the event deliverer, who submits a proposal to be presented at the DH event.
2. **Registration attendees:** the behaviour of the event participants, who register for the event via the event registration system in order to attend the event itself
3. **Running the event:** the behaviour of the Event Manager, who makes sure that the event runs smoothly and according to the programme
4. **(Post Event activities):** the behaviour of a number of stakeholders (event manager, event host and event owner) that reflect on the organised event, focussing in particular on the successes and failures.

Example of Event from the Enterprise viewpoint perspective

An event is a system instantiated by an **Event Owner**. The last DH2016 conference was held in Krakow and organized jointly by a **Programme Committee** (international) and a **local organization team**. The local organization team also acts as the **Event Manager**, supervising the successful outcome of the event.

The **event documentation system** includes written material that accompanies the event (brochure, maps, abstracts). One important aspect of such big event as the DH2016 is represented by the **event registration system**. The **event assessment system** is also very important and comprises one or more **Event Assessors, who are responsible for assessing and accepting the proposals submitted**.

The **Event Participants** are the conference attendees, while the **Event Deliverer** are the people presenting at the conference.

5.2.2 Information Viewpoint

This paragraph provides an overview of the category type: “event” from an Information Viewpoint.

Data

Information that is used as support or training material during an event (including training event). The data can consist of research outputs, both in the form of raw data or publications, standardized in order to describe consistently the observation and analysis. Data information objects can be content objects (observation, concepts and metadata) or container objects (dataset, inventory).

Agents

- **Event owner:** the agent who initiates and is responsible for the event; it can be an individual or the organization
- **Event host:** institution hosting the event
- **Event manager:** the agent who supervises the organization of the event on a daily basis in view of its successful outcome
- **Event deliverer:** the agent who delivers the event, (e.g. a conference or an academic course)
- **Event participant:** the agent who attends the event or training
- **Event Support:** the agent who provides support with organizational aspects of the event or training
- **Event documentation System (Passive):** written material that accompanies the event or the training material provided in support of a training
- **Event Registration System (Passive):** the system that allows the registration of the participants to the event
- **Event Assessor:** the agent who i.e. assesses the papers, abstract for the conference // or assesses the test at the end of the training; in training, i.e. the assessor assesses the quality of the outcomes and maybe gives a certificate
- **Event Assessment System (Passive):** the community component representing the event assessment community

Service

- **Information material about the event:** advertising, but also training material distributed to support the behaviour of the event deliverer
- **Event support:** support during the organization of the event

Contract

- **Standards:** standards that exist between a community (e.g. a certain standard use in an exercise during the training)
- **Access conditions:** access conditions that regulate the participation to the event
- **Attendance Certificate:** the certificate that proves the participation of an attendee to the event

5.3 Development of Software

DEFINITION: Availability of software, i.e. executable code that can be installed and run by other partners. The contribution should include the source code (not just the binaries).

ScopeNote: We distinguish between software and service as separate contributions. Every (digital) service is an activation of some software. If both a service and the underlying software are provided, then it counts as two distinct (though related) contributions.

The code can be in any programming language, it can also be only a simple script dedicated to one specific task, as long as it is working and documented.

5.3.1 Enterprise Viewpoint

Communities

- **Data Creation Community:** the community that creates the code necessary to install and run a software
- **Data Provision Community:** The Community who provides discovery and access to the software developed by the Data Creation Community.

Roles

- **Software owner:** the agent who owns the software and it is ultimately “responsible” for its functioning
- **Software developer:** the agent who develops the source-code of the software
- **Software manager:** the agent who is responsible of the release and good functioning of the software
- **Software architect:** the agent who is responsible of the overall architecture of the software, from its design to the completion of the software
- **Software Tester:** the agent who is responsible for the testing of the software during the production phase and before its release
- **Documentation/ training subsystem (passive):** the documentation that will facilitate the comprehension of the software by the end users once it is released
- **Software publisher:** the agent who publishes and makes available the software
- **Software development subsystem (passive):** the underlying system supporting the software developer
- **Software publishing subsystem (passive):** the system supporting the publication and release of the software, initiated by the software publisher

5.3.2 Information Viewpoint

Data

- **Code**
- **Software:** programming interfaces used in order to produce code

- **Documentation/ training subsystem** (passive): the documentation that will facilitate the comprehension of the software by the end users once it is released

Agents

- **Software owner:** the agent who owns the software and it is ultimately “responsible” for its functioning
- **Software developer:** the agent who develops the source-code of the software
- **Software manager:** the agent who is responsible of the release and good functioning of the software
- **Software architect:** the agent who is responsible of the overall architecture of the software, from its design to the completion of the software
- **Software Tester:** the agent who is responsible for the testing of the software during the production phase and before its release
- **Data Owner:** the agent who owns the code and it is ultimately “responsible” for its correct functioning
- **Data Consumer:** the agent who will make use of the software produced
- **Data Manager:** the agent who makes sure that the quality of the code is met and that the user experience is satisfactory
- **Software publisher:** the agent who publishes and makes available the software

Services

- **User guide/ manual:** information on how to use the software
- **Software development subsystem** (passive): the underlying system supporting the software developer
- **Software publishing subsystem** (passive): the system supporting the publication and release of the software, initiated by the software publisher

Contracts

- **Contract:** e.g. reuse licenses and intellectual property rights clearance

5.4 Consulting

DEFINITION

Consulting is a time-based support activity about a situation or a problem offered to another DARIAH member or the DARIAH community. This may cover advice including audit, design or other activities where expertise support and knowledge transfer form the basis of the service.

5.4.1 Enterprise Viewpoint

Community

- **Consultancy provider community:** the community which legally provides the consultancy activity. It holds the expertise to provide and perform it.
- **Consultancy receiver community:** the community requesting consultancy and receiving it.

Roles

- **Consultancy provider:** the agent who legally provides the consultancy activity (e.g. institution)
- **Consultancy implementer:** the agent who performs the consultancy activity
- **Consultancy consumer:** subject of the consultancy: the agent who instantiates the request for an activity and receives the outcome of the activity.
- **Consultancy Provision System (passive):** ((how do we define this object when put somewhere else))

Behaviours

- **Consultancy Management:** the behaviour of the consultancy provider, who overviews the running of the consultancy activity, and ensures the transfer of knowledge.
- **Consultancy Support:** the behaviour of the consultancy implementer, giving his expertise
- **Consultancy Execution:** the behaviour of a number of agents (Consultancy provider, Consultancy implementer) performing certain actions in the context of consulting.

5.4.2 Information Viewpoint

Data

- **Raw Data:** notes
- **Documentation:** procedure, list of documents used for the planning and execution of the consultancy
- **Report:** official output resulting of the consultancy

Agents

- **Consultancy provider:** the agent who legally provides the consultancy activity (e.g. institution)
- **Consultancy implementer:** the agent who performs the consultancy activity
- **Consultancy receiver:** subject of the consultancy: the agent who instantiates the request for an activity and receives the outcome of the activity.

Services

- **Supports of communication:** external (DARIAH website), internal (team collaboration software such as Confluence Wiki, Basecamp)

Contracts

Very important in this type of in-kind is in fact the **agreement** and **contract** side of the consultancy, which is essential to stipulate the detail of the performance, e.g. whether the consultancy is for free or not, how long will it be etc.

5.4.4. Case Study

The Data Seal of Approval enounces Guidelines that are of interest to data producers and institutions that create digital data, to organizations that archive data, and to consumers of data.

The objectives of the Data Seal of Approval are to safeguard data, to ensure high quality and to guide reliable management of data for the future without requiring the implementation of new standards, regulations or high costs.

An institution requests consultancy to DANS on how to get the Data Seal Of Approval Certification. This institution requiring consultancy is the **Consultancy Receiver Community**. DANS represents the **Consultancy Provider Community**. As the institution providing consultancy, DANS is the **Consultancy Provider**. The institution receiving consultancy is the **Consultancy Receiver**.

5.5 DARIAH Coordination

DEFINITION: This type of contribution involves the coordination of DARIAH activities. The coordination roles that are paid by DARIAH-ERIC can not be considered as in-kind contributions.

For instance coordination of or cooperation in:

- National Coordinator coordinates DARIAH activities in his or her country. In the NCC they meet regularly to integrate their national DARIAH activities at the European level. Each DARIAH Member Country has its own National Coordinator.
- VCC Heads: their work is based on pro-bono activities. A Memorandum of Understanding defines the amount of hours that a VCC head should dedicate to DARIAH per year.
- Working Groups (coordinators): their work is based on pro-bono activities. Their work is regulated by a tacit contract

5.5.1 Enterprise Viewpoint

Communities

- **DARIAH Coordination Community**

Roles

- **Coordination Activity Subsystem** (passive):
- **Coordination Activity owner:** (single or institution) the agent that has the responsibility of the coordination activity.
- **Coordination Activity provider:** an agent (institution) that performs a certain action in the context of the DARIAH coordination
- **Coordination Activity Implementer:** the agent who performs the activity on the behalf of the Coordination Activity provider
- **Coordination Activity consumer/ attendee** (active/ passive role): an agent (individual/ institution) that receives, consumes or experiences the outputs created during an activity
- **Coordination Activity Manager:** the agent that has a day to day overview of the running of the activity. This role is also the point of contact for issues with the smooth running of the activity

Behaviours

- **Coordination Activity management:** the behaviour of the Coordination Activity Manager, who has a day to day overview of the running of the activities.
- **Coordination Activity support:** the behaviour of the support agent, whose task is to support the coordination activities
- **Coordination Activity execution:** the behaviour of the Coordination Activity Implementer, who is responsible to perform a certain action on behalf of the Coordination Activity Provider.

5.5.2 Information Viewpoint

Data

Data object type: information useful for the participants, **internal project documentation** (shared documents such as action plans for example), **official outputs** (official documents)

Agents

- **Coordination Activity owner:** (single or institution) the agent that has the responsibility of the coordination activity.
- **Coordination Activity provider:** an agent (institution) that performs a certain action in the context of the DARIAH coordination
- **Coordination Activity Implementer:** the agent who performs the activity on the behalf of the Coordination Activity provider

- **Coordination Activity consumer/ attendee** (active/ passive role): an agent (individual/ institution) that receives, consumes or experiences the outputs created during an activity
- **Coordination Activity Manager**: the agent that has a day to day overview of the running of the activity. This role is also the point of contact for issues with the smooth running of the activity

Services

- **Communication Support**: external (DARIAH website), internal (team collaboration softwares such as Confluence Wiki, Basecamp)
- **Processing tool**: contribution tool

Contracts

- **Memorandum of understanding**: defines the amount of hours a VCC Head should dedicate to DARIAH a year.

5.5.3 Case Study

A DARIAH institution offers a contribution. The institute fills a self-assessment and submits it via the Contribution Tool. The assessment criteria are provided by the tool. Assessment criteria are offered for each type of contribution.

This self-assessment is peer reviewed by the appropriate peer review group via the tool, composed of VCC Heads and members of the Working Groups. When the VCC Heads peer-review in-kind contributions, they act as **Coordination Activity implementer**. The **Coordination Activity provider** is the peer-reviewing group (composed of VCC heads and Working groups members). It acts on the behalf of DARIAH, which is the **Coordination Activity owner**. The **Coordination Activity Consumer** is also DARIAH, since it will beneficiate of the expertise of this group to review in-kind contributions. The **Coordination Activity subsystem** is represented by the Contribution Tool. DARIAH and the VCC belong the **DARIAH Coordination Community**.

5.6 Creating Resources

DEFINITION:

This type of in-kind contribution describes the resources that an agent (institution, individual) creates in order to make them digitally available with the DARIAH community. In this context a “resource” is considered as an **asset** that supports a certain research activity (e.g. can be the digitized document that was not previously available or the training material to support an educational trajectory).

The resources can also have different forms for example:

- Text

- Images
- Data
- Video recordings
- Audio recording
- Interactive resources (e.g. quizzes, interactive maps and diagrams etc.)

If the resource proposed is code, its specificity will make it affiliated with the in-kind contribution “Development of software”.

5.6.1 Enterprise Viewpoint

Data Creation Community: the objective of the data creation community is to create and/ or collect research data using specific instruments and methodologies.

Within DARIAH this community can be represented by the community of cultural heritage institutions that make available digitally cultural object and/ or metadata of specific collections to researchers. It can also be represented by the organization of a conference that creates ancillary material to support the conference (e.g. abstracts) or by the training material created for an online course (e.g. test, readings etc.)

Data Management Community: The main objective of the data management community is to maintain or improve the quality of the research data for verification and reuse. In the context of this in-kind contribution, the data management community is involved only for what concerns the communication between the data creation community and the data management community about the appropriateness of the data created for its storage and for making sure that the data management community will take good care of the data.

Data Provision Community: the objective of the Data Provision community is to disseminate and to provide access to the data. Their main objective is to make data accessible and discoverable.

Roles

- **Data Creation Subsystem (passive)**
- **Data Creator:** person responsible for the data creation
- **Data Collector:** person or organization responsible for the collection of existing data
- **Metadata Creator:** person or organization responsible for creating metadata related to the data examined, according to agreed metadata standards.
- **Data Manager:** person responsible for communicating with the Data Creator, and maintaining the quality of the data on a technical perspective (keeping/improving it to defined quality/standards).
- **Data Owner:** a person or organisation that owns the data and determines the conditions for deposit/access with the Data Manager.
- **Data Provider:** person responsible for making the data accessible and discoverable.

- **Data Discovery System (passive role):** a system facilitating discovery of research data and its context.
- **Data Consumer:** an agent that requests and receives the research data (for further use).

Behaviours

- **Collect Data:** the behaviour of the data creation subsystem performed by the collector and the instrument whereby the data collector records and aggregates observations from the instrument into a dataset.
- **Curate Data:** the behaviour performed by the data manager whereby he checks the quality of the data and removes faults. He then adds meaningful documentation (metadata) to the data.
- **Data Publication:** A behaviour by the data provider, the Data Discovery System and the Data Retrieval System for making data for making research data discoverable and accessible.
- **Data Discovery:** a behaviour by the data provider, to making the created data discoverable and accessible.

5.6.2 Information Viewpoint

Data

- **Data Object Type:** (e.g. digitized cultural object or training material for a course): the minimum available information object type that can be shared from the data creator to the data consumer
- **Metadata Object Type:** the information object type that describes and contextualizes the data object type

Agents

- **Data Owner:** the agent who ultimately owns the data
- **Data Creator:** the agent who creates the data
- **Data Collector:** the agent who collects already existing data, creating new type of resources
- **Data Manager:** agent who makes sure that the data/ resources are of good quality
- **Data consumer:** the agents who accesses the data created

Services

- **Digitization service:** information object type that translates the analogue object type into a data object type

Contracts

In this case (creating resources) the use of a standard might be crucial in the harmonization of the data and metadata of a research or heritage institution. The agreement and eventually the contract for this in-kind is represented by the agreement between the resource owner and the data creator (in the sense of digitisation) if they are two different agents. Another agreement can

be thought in the case of teaching material between students and teacher, where the agreement is about the students studying/ reading the material.

5.6.3. Case Study: Creation of linguistic data

EASY²¹ is the data archive for the social science and humanities in the Netherlands. Such data archive has been developed and is currently maintained and enriched by DANS, an institute of the Dutch Royal Academy of Science, that has a leading role (nationally and internationally) in the development of data quality policies and research data management plans.

As such, EASY offers to the research community the possibility to host their data after they have finished their research and therefore can be considered as a candidate for a DARIAH contribution Data Hosting Service.

In terms of Enterprise Viewpoint, the Data Hosting community is represented by the DANS team, which is the community who acquires, administers and curates the dataset created by the Resources Creation Community (the researcher depositing the datasets). The team at DANS acquires the data (with the support of a submission agreement) through the submission of information packages.

To all the archived data, DANS applies the principles of the Preservation Policies Data Archiving and Networked Services, which draws the lines of the agreement between the Data Hosting Community and the Resource Creation Community.

In terms of information Viewpoint, the research outputs and the data delivered to EASY in order to be archived are the main information object type for this type of contribution. According to the recommendation of DANS, the data need to be structured according a series of preferred data and metadata formats and they need to include reference to their accessibility and reuse possibility. A contract (another data object type) will then establish duties and rights for the Data Hosting Community and the Data Resource Community by answering questions like: for how long will the data be maintained? what are the responsibilities of the stakeholders involved?

²¹ <https://dans.knaw.nl/en/about/services/archiving-and-reusing-data/easy>