



# In-kind Contributions for two ERICs in one National Initiative: Practices and Experiences in CLARIAH-DE

September 2021

Version 1.0

Date 17.09.2021

Authors Stefan Buddenbohm<sup>a</sup>, Melina Jander<sup>a</sup>, Thorsten Trippel<sup>b</sup>, Nathalie Walker<sup>b</sup>,

Lukas Weimer<sup>a</sup>

Peer review Edward Gray<sup>c</sup>, Nanette Rißler-Pipka<sup>a</sup>, Andreas Witt<sup>d</sup>

Project CLARIAH-DE

Funder Bundesministerium für Bildung und Forschung

Grant Agreement No 01UG1910A bis I

Duration 01.03.2019 bis 30.09.2021

<sup>a</sup>Göttingen State and University Library, <sup>b</sup>University of Tübingen, <sup>c</sup>DARIAH-EU/Huma-Num, <sup>d</sup>Leibniz-Institut für Deutsche Sprache

#### <u>Abstract</u>

An ERIC (European Research Infrastructure Consortium) operates and provides research infrastructure offerings, usually within a disciplinary scope, for researchers regardless of their national or institutional background. Apart from direct funding, an ERIC is built upon the national contributions from partner countries, notably cash and in-kind contributions. This concept, the procedures, advantages and options for future enhancement are described in this paper from the German perspective focussing on the in-kinds. The paper is intended to support an informed discussion on the advancement of the in-kind contribution concept. The overall aim is to contribute to the efficiency and uptake of the ERIC's research infrastructure offerings.



# Table of Contents

Introduction	4
Definition of in-kinds, their role, ownership and formats	5
CLARIN-D	6
DARIAH-DE	8
Comparison of in-kind practices in CLARIN-D and DARIAH-DE	10
Outlook for Germany	11
Annex	14
TaDiRAH Taxonomy	14
Bibliography	15



#### Introduction

Why are in-kind contributions to ERICs important, how are they determined on the national level, and how can users from across Europe access them? What, precisely, are in-kind contributions, and how can they be described in a standardised way and compared with one another? These are some of the questions addressed in this paper, based on experiences of CLARIAH-DE, the German merger of CLARIN-D and DARIAH-DE. The paper is intended to serve for discussing best practices and synergies for the description and evaluation of in-kind contributions in the European research infrastructure consortia (ERIC).

ERICs are by now (2021) a seasoned legal framework for the coordination and provision of research infrastructures in the European Research Area (ERA) and – in the case of CLARIN<sup>1</sup> and DARIAH<sup>2</sup> - much more than just "full legal entities under European Union law"<sup>3</sup>. CLARIN and DARIAH are full-fledged research infrastructures on a non-economic basis, particularly addressing the humanities and relating to a constantly growing user base across the ERA. Especially the affiliated communities of researchers - be they providers or users of resources - form the character of an ERIC, and eventually, are mirrored in the in-kind contributions, which will be described later on.

What constitutes a research infrastructure? From a formal point of view, an ERIC is endowed with annual cash contributions by its member countries and with in-kind contributions provided by participating partner institutions (organised on a national level, usually proportionally to the GDP of the member countries). The cash contributions are subject to a negotiation process involving administrative decision-makers rather than researchers, whereas the in-kind contributions from the national branches - e.g., DARIAH-DE or CLARIN-D - are shaping the ERICs substantially through their practical and user-involving level. Though the cash contributions are equally important, the in-kind contributions materialise research infrastructure components literally and are, therefore, the focus of this paper.

The following paper, written from the German national perspective, discusses the differing in-kind practices CLARIAH-DE is facing as legacy from CLARIN-D and DARIAH-DE. CLARIAH-DE is the merger project<sup>4</sup> of the formerly separately developed research infrastructures CLARIN-D and DARIAH-DE and takes over the responsibility for reporting the in-kinds to the ERIC level in the future. On this occasion the following paper could serve as discussion impulse on how to guarantee for high quality in-kinds with a minimum of management overhead.

<sup>&</sup>lt;sup>4</sup> Buddenbohm, Stefan, Friedrichs, Sonja, & Walker, Nathalie. (2020). CLARIAH-DE - Added value of the infrastructure merger. Scholarly Primitives - DARIAH Annual Event 2020, Zagreb, Croatia. Zenodo. https://doi.org/10.5281/zenodo.4247243



<sup>&</sup>lt;sup>1</sup> https://www.clarin.eu/

<sup>&</sup>lt;sup>2</sup> https://www.dariah.eu/

<sup>&</sup>lt;sup>3</sup> https://en.wikipedia.org/wiki/European Research Infrastructure Consortium

# Definition of in-kinds, their role, ownership and formats

An individual ERIC is in large part composed by the contributions of its member countries<sup>5</sup>. Each member of the ERIC commits itself to make annual contributions to the ERIC, consisting of cash contributions and in-kind contributions. Cash contributions are monetary contributions directly paid to the ERIC, usually by the national funding institution of the member country; for Germany, this is the BMBF, the Federal Ministry of Education and Research<sup>6</sup>. The annual amount of the cash contributions is – in the case of CLARIN and DARIAH – determined through an agreed mechanism based on the member countries' GDP<sup>7</sup>.

ANNEX III - PROVISIONAL BUDGET AND CONTRIBUTIONS

Countries	Length of Commitment	Estimated Cash Year 1 (euro)	Estimated In-kind Year 1 (euro)	Estimated Total Year 1 (euro)
Austria **	3 years	28,250.00	231,250.00	259,500.00
Belgium	5 years	27,700.00	227,000.00	254,700.00
Croatia	5 years	3,400.00	28,000.00	31,400.00
Cyprus	5 years	1,300.00	11,000.00	12,300.00
Denmark	5 years	18,000.00	147,000.00	165,000.00
France	5 years	149,200.00	1,221,000.00	1,370,200.00
Germany **	until 26.2.2016	242,125.00	1,981,250.00	2,223,375.00
Greece	5 years	15,600.00	128,000.00	143,600.00
Ireland	5 years	11,800.00	96,000.00	107,800.00
Italy	5 years	118,100.00	966,000.00	1,084,100.00
Luxembourg	5 years	3,200.00	26,000.00	29,200.00
Malta	5 years	500.00	4,000.00	4,500.00
The Netherlands	5 years	45,100.00	369,000.00	414,100.00
Serbia	5 years	2,500.00	21,000.00	23,500.00
Slovenia	5 years	2,700.00	23,000.00	25,700.00
	Total	669,475.00	5,479,500.00	6,148,975.00

<sup>1)</sup> In the following years, the annual contribution is the contribution of the previous year with an annual increase of 2% cent to compensate for inflation and increase in costs (cf. principle 7, Annex II).

Figure 1: Cash and in-kind contributions by the member countries to the DARIAH ERIC in 2015 according to the DARIAH ERIC statutes, p.19 (note: figures have substantially changed since 2015)

In-kinds may take diverse shapes, such as research tools, resources like data or collections, development capabilities for the improvement of research tools, basic services, e.g., AAI (Authentication and Authorisation Infrastructure) or PID (Persistent Identifier) services, scholarly resources such as online courses, teaching courses, or training efforts aimed at

<sup>&</sup>lt;sup>7</sup> Compare Article 18 of the DARIAH ERIC Statutes: https://www.dariah.eu/wp-content/uploads/2017/02/170405 DARIAH ERIC Statutes.pdf



<sup>2) \*\*</sup> As the length of commitment is less than 5 years the increase of 25% has been included (cf. principle 10, Annex II).

<sup>&</sup>lt;sup>5</sup> Other sources of inflows may be successful tenders for funded projects, e.g., H2020 or EOSC.

<sup>&</sup>lt;sup>6</sup> See: <u>https://www.bmbf.de/</u>

research tools, methods, or resources. Such a resource may qualify as in-kind once it is created in the respective context (e.g., CLARIN or DARIAH) and corresponds to the topics of a digital European research infrastructure in the humanities and cultural sciences. In-kind contributions have a monetary value fixed to them, based upon the costs incurred in their creation. These values are then used to calculate the overall national in-kind contribution, which is set by the ERICs. The overall framework governing the relations among the CLARIN and DARIAH member countries, partners and towards the ERIC are manifested in the statutes<sup>8</sup>.

The main differences of the varying in-kind collection practices in CLARIN-D and DARIAH-DE will be described in depth in the following two sections.

#### **CLARIN-D**

The German structure of CLARIN was established before the creation of the ERIC in 2012. Funded until 2020 as a national project, CLARIN-D was the German consortium for the CLARIN ERIC. Since 2020, CLARIAH-DE has fulfilled this role on the national level, taking over many of the responsibilities and services previously provided by CLARIN-D, its partners and the community. From the ERIC point of view, the in-kind process requires a national consortium consisting of at least one CLARIN centre and a national coordinator. Apart from this, no further formal structures are required. The General Assembly, which consists of delegates of every member country, is the highest decision making body. The General Assembly is responsible for the CLARIN Agreements which codify the relations, budget and work plan for a fixed period of time (e.g., three years). The executive body is the board of directors, which also coordinates with the National Coordinators' Forum. The National Coordinators' Forum is the committee of all national coordinators correlating the national activities, including in-kind contributions. On the national level, the consortia consist of centres (see below) to maintain the resources and services. As part of the community and contributed by members of the community, the centres provide data and services that are accessible for the whole community and contribute this as part of the national in-kind contributions. For communication and cooperative strategic discussion, formats such as working groups exist as well. Projects funded nationally or on a European level may also contribute in this regard, for instance the aforementioned CLARIAH-DE project.

With CLARIN-D, the process and presentation of the in-kinds are institutionally low-threshold as they are bound to the individual centres. The CLARIN in-kinds are determined in the CLARIN Agreements, which define the in-kinds to be provided by the national members. The restriction to a specific time period is of small importance for data, for example, and hence does not require an annual process. Ongoing activities that are part of CLARIN's knowledge infrastructure and dissemination activities such as workshops, teaching courses, conferences, and contributions to working groups or boards are usually

 $<sup>\</sup>frac{https://www.dariah.eu/wp-content/uploads/2021/06/DARIAH-ERIC\_Statutes\_May-2021.pdf}{https://www.dariah.eu/wp-content/uploads/2021/06/DARIAH-ERIC\_Statutes\_May-2021.pdf}{https://www.dariah.eu/wp-content/uploads/2020/11/IRP-Version-November-2020.pdf}{https://www.dariah.eu/wp-content/uploads/2020/11/IRP-Version-November-2020.pdf}{https://www.dariah.eu/wp-content/uploads/2021/06/DARIAH-ERIC\_Statutes\_May-2021.pdf}{https://www.dariah.eu/wp-content/uploads/2020/11/IRP-Version-November-2020.pdf}{https://www.dariah.eu/wp-content/uploads/2020/11/IRP-Version-November-2020.pdf}{https://www.dariah.eu/wp-content/uploads/2020/11/IRP-Version-November-2020.pdf}{https://www.dariah.eu/wp-content/uploads/2020/11/IRP-Version-November-2020.pdf}{https://www.dariah.eu/wp-content/uploads/2020/11/IRP-Version-November-2020.pdf}{https://www.dariah.eu/wp-content/uploads/2020/11/IRP-Version-November-2020.pdf}{https://www.dariah.eu/wp-content/uploads/2020/11/IRP-Version-November-2020.pdf}{https://www.dariah.eu/wp-content/uploads/2020/11/IRP-Version-November-2020.pdf}{https://www.dariah.eu/wp-content/uploads/2020/11/IRP-Version-November-2020.pdf}{https://www.dariah.eu/wp-content/uploads/2020/11/IRP-Version-November-2020.pdf}{https://www.dariah.eu/wp-content/uploads/2020/11/IRP-Version-November-2020.pdf}{https://www.dariah.eu/wp-content/uploads/2020/11/IRP-Version-November-2020.pdf}{https://www.dariah.eu/wp-content/uploads/2020/11/IRP-Version-November-2020.pdf}{https://www.dariah.eu/wp-content/uploads/2020/11/IRP-Version-November-2020.pdf}{https://www.dariah.eu/wp-content/uploads/2020/11/IRP-Version-November-2020.pdf}{https://www.dariah.eu/wp-content/uploads/2020/11/IRP-Version-November-2020.pdf}{https://www.dariah.eu/wp-content/uploads/2020/11/IRP-Version-November-2020.pdf}{https://www.dariah.eu/wp-content/uploads/2020/11/IRP-Version-November-2020.pdf}{https://www.dariah.eu/wp-content/uploads/2020/11/IRP-Version-November-2020/11/IRP-Version-November-2020/11/IRP-Version-November-2020/11/IRP-Version-November-2020/11/IRP-Version-November-2020$ 



-

 $<sup>^{\</sup>rm 8}$  As illustration the DARIAH ERIC Statutes as of May 2021:

not listed separately as part of the in-kind contributions, but expressed in the activities of the national consortia or individual centres. These contributions are nevertheless regarded as important and are expected from the national partners. The evaluation of the contributed services and processes is regularly repeated in the CTS certification process and the CLARIN Centre assessment.

Therefore, the CLARIN Centres are of core importance. They stand for a reliable system to support data management, including reference data sets provided by the hosting institutions and the community (e.g., general and specialized corpus collections, lexical resources such as the German Reference Corpus DeReKo and the wordnet resource GermaNet), and services (for example query tools, and analysis services like endpoints for federated content search or language processing services provided in WebLicht). The resources and services are attributed to the individual CLARIN Centres but often branded as part of CLARIN. For each individual resource, a specific CLARIN Centre maintains ownership, responsibility, and has an inherent interest to enhance and maintain the resources and services.

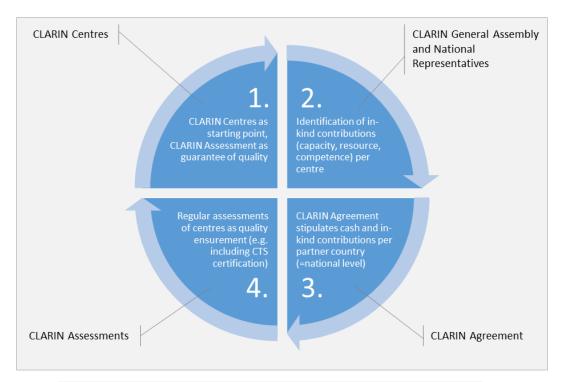


Figure 2: Formation process of the CLARIN in-kind contribution and the important role of the CLARIN centres (Source: own presentation)

How is the quality and reliability of these resources guaranteed, when there is no annual process and no further assessment of the individual in-kinds like in DARIAH? This question can be answered when taking a look at the requirements for the CLARIN Centres. CLARIN Centres have to apply for certifications<sup>9</sup>, such as:

CLARIN B-Centre: The service-providing centres and thus the backbone of the CLARIN research infrastructure. B-Centres are certified centres providing data and

<sup>&</sup>lt;sup>9</sup> The current requirements to become a CLARIN Centre are laid down as of February 2020: https://office.clarin.eu/v/CE-2012-0037-centre-types-v08.pdf



services on a high level of reliability to the community. The application procedure of the B-Centres is the most sophisticated compared to the other centre types<sup>10</sup>, which is important as the services provided by these centres range from research-specific services to contributing to basic infrastructure components such as an AAI.

- CLARIN C-Centres do not require this certification process, but are recognised for providing valuable data and services. These centres provide the metadata for their data to support findability. Such centres usually offer data collections, like corpora. The focus is on the interoperability of these resources to B-Centre provided data and services.
- CLARIN K-Centre or Knowledge Centres are a third type of centre, which bundle crucial expertise required by the community and form part of the CLARIN Knowledge Infrastructure. They are formally recognised by the CLARIN Knowledge Infrastructure Committee<sup>11</sup>.

For data management centres and service providers, the trust level achieved by transparent certification criteria is essential and part of the in-kind contribution. For data centres, this includes the Core Trust Seal (CTS) to ensure a high level of quality, sustainability, and interoperability. By passing through the CTS process, a centre displays, in a transparent way, compliance to a set of quality criteria, which may vice versa be applied to the provided resources and services. From this perspective, a CTS certified centre may be counted as in-kind without the need for an annual assessment.

#### **DARIAH-DE**

DARIAH-DE was established as a Digital Research Infrastructure for the Arts and Humanities in 2011 – also a few years before the ERIC – and has since then been expanded with resources, services, partners, and the community at large. When looking at the other national DARIAH partners, DARIAH-DE is a pioneer in the provision of sustainable resources and services and has been involved in the in-kind contribution-process since 2015, when the DARIAH ERIC was founded.

DARIAH-DE is characterised by its community-driven approach, which boils down to the fact that the in-kinds take very granular shapes regarding format, ownership, and quantity. Usually each DARIAH-DE partner contributes with at least one in-kind, which may either stay stable over some years (such as an AAI or a research data collection) or may be dynamic or transient (such as a workshop or a development activity for a service). This is also a reason for the decision to use the TaDiRAH taxonomy to describe the in-kinds. TaDiRAH comes with two main categories: activities and services, which underline well the examples mentioned above.

Taking this into account, the in-kind contributions in DARIAH-DE differ significantly from those in CLARIN-D. In an annually iterative process, DARIAH-DE's in-kind contributions are evaluated through a predetermined procedure, forming a cycle of four core steps:

Overview of the different centre types: <a href="https://www.clarin.eu/content/clarin-centres">https://www.clarin.eu/content/clarin-centres</a>
 Application procedure for a CLARIN K-Centre: <a href="https://www.clarin.eu/content/knowledge-centres">https://www.clarin.eu/content/knowledge-centres</a>



- 1. A decision of the DARIAH General Assembly with the National Representatives determining the figure of the annual monetary equivalents, that have to match with the national contributions (cash and in-kind).
- 2. The collection and description of the in-kinds according to the TaDiRAH contribution subtypes 'Services' and 'Activities'. A service may be a data hosting service (IT based), a processing service (IT based), a support service, or providing access to resources. Activities may be events, consulting activities, the DARIAH coordination, resource creation and curation, or software development.
- 3. The calculation of a monetary equivalent for each in-kind and transferring of the in-kind descriptions in a European DARIAH database, the 'In-kind Contribution Tool' (<a href="https://contrib.dariah.eu/">https://contrib.dariah.eu/</a>). This is a separate step in the process as the in-kinds are reported by the partners, whereas the calculation is done centrally by the DARIAH-DE coordination office.
- 4. Representation of the various national contributions at the level of the DARIAH ERIC, use of the resources by the community, and often further refinement of in-kinds, and flowing into the discussion and decision process for the next annual in-kind contributions.

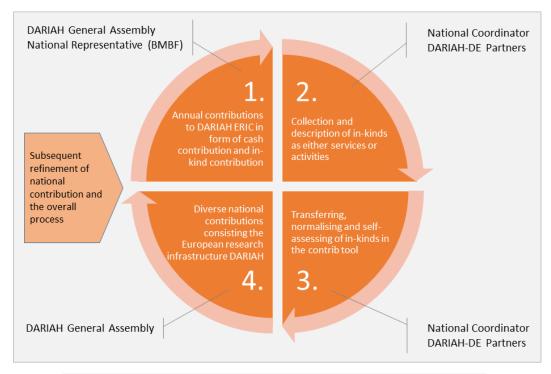


Figure 3: Process for the description of the DARIAH-DE in-kinds highlighting the collection process (Source: own presentation)

Many of the specific descriptions and the respective monetary values change only slightly throughout the years – with the exception of many activities – but there are also cases of service terminations, which lead to punctually more noticable changes. With this iterative process, however, an up-to-date overview of the current national landscape is generated that allows for a mutual consultation with the other national in-kinds. Thus, the different types of in-kinds are transparent and, theoretically, the assessment process allows for



evaluation or ranking measures. The current assessment of the in-kinds is being conducted by the Joint Research Committee (JRC) and the Senior Management Team (SMT) on an annual basis.

Although this circular procedure is beneficial on the levels mentioned above, the annual process of editing and refining the in-kinds is time-consuming. The workflow can be accelerated by reliable documentation of each service; here, the corresponding project partners are responsible to keep their service documentation up-to-date and accessible to their colleagues.

In sum, the collection process of the DARIAH-DE in-kinds can be described as follows:

- It takes place on the level of the individual in-kind, which is organised in one of the TaDiRAH categories, either as a service or as an activity.
- Ideally, it leads to a European tableau of comparable and transparent in-kinds among the DARIAH member countries.
- It is work intensive on an annual basis, both for the national and the European level.
- The in-kind data once collected and described in a standardised way may be made interoperable and accessible for other measures. Registries or discovery systems like the SSH Open Marketplace or the EOSC Marketplace could harvest the DARIAH in-kinds via an API.

# Comparison of in-kind practices in CLARIN-D and DARIAH-DE

With the merger of CLARIN-D and DARIAH-DE until 2021, CLARIAH-DE is faced with varying practices regarding cash and in-kind contributions. With reference to the cash contributions, merging the procedures of CLARIN-D and DARIAH-DE is a straight-forward practice. To this end, CLARIAH-DE can offer the formal body for processing the annual payments from the national funding level (the BMBF) to the European level (the ERICs). This formal body will be taken over by the "Geistes- und kulturwissenschaftliche Forschungsinfrastrukturen e.V.", an association under German law, which will be operational for this purpose by the end of 2021. This association functions also as a discussion arena for the in-kind process as such and will be used for the refinement and enhancement of the German contributions.

The differing practices regarding in-kinds represent both CLARIN-D's and DARIAH-DE's varying agreements among the partners and member countries in their respective ERICs. On this level, each practice has its advantages: while CLARIN ensures its in-kinds through its centre structure – to become a B centre, a CTS certification is obligatory – DARIAH presents a granular approach on the individual in-kinds which allows for a transparent traceability and easier re-use of the resource. Furthermore, these varying procedures also represent different organisational cultures and structures. For CLARIN, the centre-oriented structure makes use of specific research institutions that are firm parts of the community, e.g., the Leibniz Institute for the German Language (IDS Mannheim), university institutions,



academies, and individual researchers. They provide resources reflecting their respective research areas, ranging from experimental data, collected and annotated corpora, to lexical resources, software and models for machine learning applications operating on these types of data.

For DARIAH-DE, as a largely community-driven research infrastructure, many resources are created and provided within research-near, usually smaller and distributed contexts, e.g., in research groups. Larger infrastructure or service providers, like data centres of universities, play an important role in DARIAH-DE but it is not uncommon that they serve as enablers for smaller structures like a research group, which makes use of the data centre's infrastructure but IS NOT the data centre or part of it. The same applies for the DARIAH Data Federation Architecture<sup>12</sup> which is a framework of data-related services and makes all in all well over 2.2 million research data records visible and available to researchers but DOES NOT own most of the research data. Another aspect of DARIAH's approach is to be very flexible towards a range of partner countries and partners. Even smaller or very recent partner countries are able to contribute according to their capabilities due to the specific in-kind concept of activities and services and the granular description. However, as a result, the effort to collect and describe the resources is considerably higher in DARIAH as compared to CLARIN.

These findings may pose a challenge<sup>13</sup> on the national level since with CLARIAH-DE, the formerly separate structures of CLARIN-D and DARIAH-DE cease to exist and for the sake of efficiency, a unitary procedure would be advantageous. This does not imply that the varying practices of collecting and describing in-kind contributions have to be brought in line -CLARIAH-DE is with regards to the ERICs not in a position to do so - but it is a valuable opportunity to compare, learn from each other, and possibly to suggest improvements.

# Outlook for Germany

Although a coherent procedure may pose a challenge due to the diversity of the in-kinds and their transmission to the ERICs, several possibilities to facilitate work through collaboration can be identified. Such a strategy would ideally minimise the workload on the national level and would contribute to improving the data offerings on the European level, thus finally benefiting the users of the infrastructure.

To this end, it is possible to jointly coordinate the in-kinds either under the label of CLARIAH-DE or, perspectively, within the aforementioned association for research infrastructures in the humanities and cultural sciences (Geistes- und Kulturwissenschaftliche Forschungsinfrastrukturen e.V.). According to its statutes, the association shall serve as a national node for European research data infrastructure networks including, of course, the ERICs. The association thus offers a forum for all national participants in an ERIC and facilitates joint coordination. Furthermore it may serve as a low threshold entry point for interested new partners and function as an important advocate for the European research

<sup>&</sup>lt;sup>13</sup> An introduction into the merger challenges of CLARIAH-DE seen from an infrastructural and technological point of view: Buddenbohm, Stefan, & Eckart, Thomas. (2021, March 23). Merging Subject-Specific Searches of CLARIN and DARIAH in CLARIAH-DE: Challenges of Technical Integration. Zenodo. https://doi.org/10.5281/zenodo.4628889



<sup>&</sup>lt;sup>12</sup> See the German documentation of the DFA: <u>https://dfa.de.dariah.eu/doc/dfa/</u>

infrastructure level. With the end of the national funding for CLARIAH-DE, this is important as the association ensures continuity of outreach and recruitment.

From the point of view of the ERICs, the advantage is that the association is a fixed contact point in Germany, even if there is a change of participating institutions or national coordinators. Since it is intended to pay the cash contributions also via the association, the duality of in-kind and cash contributions would no longer apply; here, too, the procedures will become easier. As a national node, the association thus takes an important step towards standardisation, coordination and simplification of work.

It is obvious that such a joint catalogue of CLARIAH-DE has to meet requirements on at least three dimensions:

- 1. The joint catalogue has to be interoperable in the growing environment of machine-readable catalogues and registries, such as the SSH Open Marketplace or the EOSC Portal. In this regard the joint catalogue could showcase an important share of the German humanities and cultural science related research resources. This machine-to-machine interoperability would also serve valuable for added value services, for instance the CLARIN Switchboard or knowledge graphs, in consequence enhancing the visibility, uptake, and standardisation of the resources.
- 2. The joint catalogue should serve as a convenient and useful information resource for researchers, that is the <u>functionality for human users</u>. This is an important requirement even if the the association (for research infrastructures in the humanities and cultural sciences) does not perceive itself as a research infrastructure addressing end users.
- 3. The joint catalogue has to serve as means to <u>leverage administrative synergies</u> for CLARIN-D and DARIAH-DE. As mentioned, these two infrastructures have merged in CLARIAH-DE but retain a formal legacy as they still report to the ERICs on the European level. As a third important group of participating institutions the formerly not in CLARIAH-DE involved institutions have to be considered.

Within the association, a joint catalogue of offers will be devised which will then represent the entirety of the services, tools and resources and will thus offer a comprehensive overview of the available resources to the users. Within such a catalogue, researchers from national or international backgrounds, members of ERICs or other infrastructures, and even service providers can browse for the offering that appeals to their specific research questions or daily work. This is not only of relevance within CLARIAH-DE or the association, but relates to the work on cross-cutting topics conducted within the German National Research Data Infrastructure (NFDI<sup>14</sup>) and the MoU group<sup>15</sup>, in which consortia within the humanities and cultural sciences have agreed to work together on common topics. Such a common topic relates, for instance, to basic services and their development; a theme which re-occurs in various contexts and could be tackled more easily if a common ground on existing offers were to be created.

<sup>&</sup>lt;sup>15</sup> Sabine Brünger-Weilandt, Kai-Christian Bruhn, Alexandra Busch, Erhard Hinrichs, Gerald Maier, Johannes Paulmann, Andrea Rapp, Philipp von Rummel, Eva Schlotheuber, Dörte Schmidt, Torsten Schrade, Holger Simon, Regine Stein, & Elke Teich. (2020). Memorandum of Understanding by NFDI Initiatives from the Humanities and Cultural Studies. Zenodo. https://doi.org/10.5281/zenodo.4045000



-

<sup>&</sup>lt;sup>14</sup> Website of the German National Research Data Infrastructure, NFDI: <a href="https://www.nfdi.de/">https://www.nfdi.de/</a>

Integration is a crucial topic in CLARIAH-DE – it is, after all, the cornerstone of the project's very being. As such, CLARIAH-DE has extensive experience in integrating services, tools, resources, knowledge structures, infrastructures and other relevant aspects. To name only one example, CLARIAH-DE addressed the problem of integrating various search applications, while at the same time dealing with the risks and possibilities that arise in the context of harmonization, especially (but not exclusively) on the technical side.<sup>16</sup>

Consequently, devising a joint catalogue of offers is a task that is rooted in the heartland of the CLARIAH-DE project and its spirit. Nonetheless, devising such a catalogue is no easy task and various challenges arise:

- 1. Finding a common structure for heterogeneous offerings may pose a problem.
- 2. Some offerings may be difficult to quantify and qualify.
- 3. A catalogue would have to be constantly maintained and updated to remain relevant.
- 4. The catalogue needs to be interoperable in the growing environment of machine-readable registries.

In the association, these challenges will be taken on by a designated working group through building on the resources that have already been established in CLARIAH-DE. The aim is to coordinate services and tools which are frequently updated according to the community's needs and trends in the field. The structures of the association and its statutes serve as the legal frame to realise this undertaking.

<sup>&</sup>lt;sup>16</sup> Thomas Eckart, Tobias Gradl, Robin Jegan, Eliza Margaretha, Antonina Werthmann, Felix Helfer, Stefan Buddenbohm (2021): CLARIAH-DE Cross-Service Search: Prospects and Benefits of Merging Subject-specific Services". DARIAH-DE Working Papers Nr. 41. Göttingen: DARIAH-DE, 2021. URN: <a href="https://www.urn:nbn:de:gbv:7-dariah-2021-1-9">urn:nbn:de:gbv:7-dariah-2021-1-9</a>. address the challenges related to the cross-domain search and use of research data.



## Annex

## TaDiRAH Taxonomy

The TaDiRAH<sup>17</sup> taxonomy (Taxonomy for Digital Research Activities in the Humanities) is used as a metadata schema for the description of the DARIAH in-kinds. Each in-kind has to be described according to the following properties.

meta date	description		
identifier	identifier of the contribution, automatically created by the contribution tool		
country	country of the contribution		
submission date	date of submission of the contribution		
assessment date	date of assessment of the contribution		
year	year in which the contribution was provided		
contribution subtype	subtype of the contribution  There are two types of contributions: services and activities. Services are sustainable and repeatable and contain the following subtypes:		
title	name of the contribution		
description	precise short description of the contribution		
contact institution name	institution to which the contribution is assigned		
contact person name	primary responsible for the contribution, in doubt the national coordination office or the national coordinator		
contact person mail	mail of the primary responsible		
url to contribution	URL where the contribution can be found		
related contributions	links and relations to other contributions		
TaDiRAH Object	research objects categorisation of the contribution according to the TaDiRAH taxonomy		
TaDiRAH Activity	research activities categorisation of the contribution according to the TaDiRAH taxonomy		
TaDiRAH Technique	research techniques categorisation of the contribution according to the TaDiRAH taxonomy		
discipline	subject area to which this contribution is to be assigned		
keywords/tags	(further) keywords and tags to further classify the contribution		
costs	costs incurred by the contribution in the relevant year. These include working hours, personnel costs, hardware costs, license fees, etc.		

<sup>&</sup>lt;sup>17</sup> TaDiRAH: <a href="https://github.com/dhtaxonomy/TaDiRAH">https://github.com/dhtaxonomy/TaDiRAH</a> and Christof Schöch, Jody Perkins, Luise Borek (2015): TaDiRAH: Release version 0.5.3 (Version v.0.5.3). Zenodo. <a href="https://doi.org/10.5281/zenodo.32492">https://doi.org/10.5281/zenodo.32492</a>.



# Bibliography

Andrea Scharnhorst, Femmy Admiraal, Dirk Roorda (2019): Visual Analytics of the DARIAH in-kind contributions. <a href="https://dariah-ae-2019.sciencesconf.org/261649/document">https://dariah-ae-2019.sciencesconf.org/261649/document</a>

Christof Schöch, Jody Perkins, Luise Borek (2015): TaDiRAH: Release version 0.5.3 (Version v.0.5.3). Zenodo. http://doi.org/10.5281/zenodo.32492

DARIAH-EU (2020): IRP - Internal Rules of Procedures and Policies (November 2020). <a href="https://www.dariah.eu/wp-content/uploads/2020/11/IRP-Version-November-2020.pdf">https://www.dariah.eu/wp-content/uploads/2020/11/IRP-Version-November-2020.pdf</a>

DARIAH-EU (2017): Statutes of DARIAH ERIC. https://www.dariah.eu/wp-content/uploads/2017/02/170405 DARIAH ERIC Statutes.pdf

Lisa de Leeuw, Femmy Admiraal, Matej Ďurčo, Nicolas Larousse, Michael Mertens, et al.. (2017): D5.1 Report on Integrated Service!Needs: DARIAH (in kind) contributions - Concept and Procedures. [Other] DARIAH. 2017. hal-01628733v2

Mike Priddy (2017): <a href="https://priddy.github.io/DARIAH-RA/in-kind/">https://priddy.github.io/DARIAH-RA/in-kind/</a>

Peter Wittenburg, Dieter Van Uytvanck, Thomas Zastrow, Lene Offersgaard (2020-02-18): CLARIN Centre Types:

https://office.clarin.eu/v/CE-2012-0037-centre-types-v08.pdf

Sabine Brünger-Weilandt, Kai-Christian Bruhn, Alexandra Busch, Erhard Hinrichs, Gerald Maier, Johannes Paulmann, Andrea Rapp, Philipp von Rummel, Eva Schlotheuber, Dörte Schmidt, Torsten Schrade, Holger Simon, Regine Stein, Elke Teich (2020): Memorandum of Understanding by NFDI Initiatives from the Humanities and Cultural Studies. Zenodo. https://doi.org/10.5281/zenodo.4045000

Stefan Buddenbohm, Sonja Friedrichs, Nathalie Walker (2020): CLARIAH-DE - Added value of the infrastructure merger. Scholarly Primitives - DARIAH Annual Event 2020, Zagreb, Croatia. Zenodo. <a href="https://doi.org/10.5281/zenodo.4247243">https://doi.org/10.5281/zenodo.4247243</a>

Stefan Buddenbohm, Thomas Eckart (2021, March 23): Merging Subject-Specific Searches of CLARIN and DARIAH in CLARIAH-DE: Challenges of Technical Integration. Zenodo. <a href="https://doi.org/10.5281/zenodo.4628889">https://doi.org/10.5281/zenodo.4628889</a>

Thomas Eckart, Tobias Gradl, Robin Jegan, Eliza Margaretha, Antonina Werthmann, Felix Helfer, Stefan Buddenbohm (2021): CLARIAH-DE Cross-Service Search. DARIAH-DE Working Papers No. 41. Göttingen: DARIAH-DE, 2021. URN: <a href="https://www.urn.nbn:de:gbv:7-dariah-2021-1-9">urn.nbn:de:gbv:7-dariah-2021-1-9</a>

