

# The Research Data Alliance and the Humanities



René van Horik  
Data Archiving and Networked Services (DANS)  
National Node: RDA in the Netherlands  
July 2019  
<https://doi.org/10.5281/zenodo.3355145>



The work described in this document has been conducted within the project RDA Europe 4.0. This project has received funding from the European Union's Horizon 2020 (H2020) research and innovation programme under the Grant Agreement no 777388. This document does not represent the opinion of the European Union, and the European Union is not responsible for any use that might be made of its content.

# Table of Contents

<b>1 INTRODUCTION</b>	<b>2</b>
<b>2 THE RESEARCH DATA ALLIANCE</b>	<b>2</b>
<b>3 THE HUMANITIES</b>	<b>5</b>
3.1 Digital Humanities	5
3.2 Digital Research Infrastructures in the Arts and Humanities	5
<b>4 THE RDA AND THE HUMANITIES</b>	<b>8</b>
4.1 RDA Working Groups and Interest Groups	8
4.2 Recommendations and outputs	11
<b>5 GET INVOLVED</b>	<b>12</b>
5.1 DARIAH and RDA	12
5.2 CLARIN and RDA	14

# 1 INTRODUCTION

The goal of this report is to provide a broad overview of the way the Research Data Alliance (RDA) and the humanities as an academic discipline can be of value to each other.<sup>1</sup> The RDA aims to improve the open and interoperable sharing of data and to build the social, technical and cross-disciplinary infrastructure for data sharing on a global scale. The humanities are academic disciplines that study aspects of human society and culture and for this digital tools, methods and services are developed and used. The intended audience for this report in the first instance are digital humanities researchers and service providers in the humanities domain. By providing basic relevant information on the RDA for this target group it is foreseen that they can benefit from the output of the RDA and increase their involvement in RDA activities. This report has the intention to be of interest for stakeholders in the RDA as well, since it describes aspects of the research data infrastructure of the humanities and this could inspire them to engage with the humanities research community.

The compilation of this report is part of the RDA Europe 4.0 project<sup>2</sup> in which DANS<sup>3</sup>, the Dutch national RDA node, acts as an ambassador of the RDA for the humanities. Since 2012, the EU has supported RDA activities within Europe, by funding Europe-specific projects such as the project RDA Europe 4.0 which was launched in March 2018 (with a duration of 27 months) and for which this report is produced as part of the ambassador track of the project.<sup>4</sup> RDA ambassadors are domain experts who can develop a two-way communication and engagement: on the one hand by disseminating, globally, RDA outputs and perspectives with data practitioners active in their specific domain of expertise, encouraging them to become contributors and adopters of RDA outputs and to join the community; on the other hand providing insight and practical contribution to RDA groups and activities from the domain community and organisations they represent.

In the next chapter an introduction on the RDA is given, followed by a chapter on research data management aspects in the humanities. The third chapter pays attention on the activities of the RDA with respect to their relevance for the humanities.

## 2 THE RESEARCH DATA ALLIANCE<sup>5</sup>

The Research Data Alliance (RDA) was launched as a community-driven initiative in 2013 by the European Commission, the United States Government's National Science Foundation and the National Institute of Standards and Technology, and the Australian Government's

---

<sup>1</sup> The author would like to thank Françoise Genova for providing comments and suggestions on an earlier version of the report.

<sup>2</sup> <https://www.rd-alliance/rda-europe>

<sup>3</sup> Data Archiving and Networked Services (DANS) is the Dutch national institute for permanent access to digital resources, promoting open sciences and FAIR data. See: <https://dans.knaw.nl/en/>

<sup>4</sup> An overview of the RDA Europe Ambassadors can be found at: <https://www.rd-alliance.org/rda-disciplines/rda-europe-ambassadors>

<sup>5</sup> This section is based on: <https://www.rd-alliance.org/about-rda>

Department of Innovation with the goal of building the social and technical infrastructure to enable open sharing and re-use of data.

The RDA has a grass-roots, inclusive approach covering all data lifecycle stages, engaging data producers, users and stewards, addressing data exchange, processing, and storage. It has succeeded in creating the neutral social platform where international research data experts meet to exchange views and to agree on topics including social hurdles on data sharing, education and training challenges, data management plans and certification of data repositories, disciplinary and interdisciplinary interoperability, as well as technological aspects. In short, the goal of the RDA is to build the social and technical infrastructure to enable open sharing and re-use of data.

In July 2019, the RDA counted over 8.700 members from about 140 different countries and from various backgrounds, including data professionals, researchers, librarians, IT specialists and policy developers. About 70% of the members are working in academic / research organisations. Individual membership to the RDA is free for anyone who supports the RDA mission and its guiding principles (see Figure 1).



- 
- ⦿ **OPENNESS**  
Membership is open to all interested individuals who subscribe to the RDAs Guiding Principles. RDA community meetings and processes are open, and the deliverables of RDA working Groups will be publicly disseminated.
  - ⦿ **CONSENSUS**  
The RDA moves forward by achieving consensus among its membership. RDA processes and procedures include appropriate mechanisms to resolve conflicts.
  - ⦿ **BALANCE**  
The RDA seeks to promote balanced representation of its membership and stakeholder communities.
  - ⦿ **HARMONIZATION**  
The RDA works to achieve harmonization across data standards, policies, technologies, infrastructure and communities.
  - ⦿ **COMMUNITY - DRIVEN**  
The RDA is a public, community-driven body constituted of volunteer members and organizations, supported by the RDA Secretariat.
  - ⦿ **NON-PROFIT**  
RDA does not promote, endorse, or sell commercial products, technologies or services.

---

Figure 1. RDA Guiding Principles<sup>6</sup>

RDA members can join Interest Groups (IGs) and Working Groups (WGs) focused on various topics related to data sharing and innovation. Both IGs and WGs undergo a review process before they are established and endorsed by the RDA. IGs consists of experts from the community and serve as a platform for exchange. IGs produce outputs such as surveys, recommendations, reports, and can initiate new WGs. While IGs can be active over longer periods of time, WGs have a fixed period of 12 to 18 months during which members work on

---

<sup>6</sup> See: <https://rd-alliance.org/get-involved/individual-membership.html>

concrete outputs. WG outputs can include, for instance, technical specifications, conceptual models or frameworks, or implemented policies. Table 1 gives an overview of where to find relevant RDA information online.

An important event at which members of the RDA come together are the RDA Plenaries.<sup>7</sup> Plenaries are held every six months in different places around the world and serve as an important point of contact where the communities are brought together and where WGs and IGs present and discuss their work.

Find information about	Web-address
RDA in general	<a href="http://www.rd-alliance.org/">www.rd-alliance.org/</a>
RDA Interest Groups	<a href="http://www.rd-alliance.org/groups/interest-groups">www.rd-alliance.org/groups/interest-groups</a>
RDA Working Groups	<a href="http://www.rd-alliance.org/groups/working-groups">www.rd-alliance.org/groups/working-groups</a>
Recommendation and Outputs	<a href="http://www.rd-alliance.org/recommendations-outputs">www.rd-alliance.org/recommendations-outputs</a>
RDA Europe 4.0 project	<a href="http://www.rd-alliance.org/rda-europe">www.rd-alliance.org/rda-europe</a>
RDA on Twitter	@resdatall   @RDA_Euorpe   @RDA_US
RDA email	enquiries@rd-alliance.org / info@europe.rd-alliance.org

*Table 1. Overview of where to find information on the RDA*

The RDA community creates 3 types of outcomes:<sup>8</sup>

1. Recommendations are the flagship outputs of RDA. They are RDA's equivalent of the "specifications" or "standards" that other organisations create and endorse. The process for creating and endorsing these is defined, including a period for community comments.
2. Supporting Outputs are the outputs of RDA WGs and IGs that are fruit of RDA work, but are not necessarily adoptable bridges. "Upon request", these sort of outputs go through a community comment period and if no major objections or gaps are identified they get the RDA brand.
3. Other Outputs include workshop reports, published articles, survey results, etc. Anything a WG or IG wants to register and report. Upon request, these are published and discoverable on the RDA website but they have no level of endorsement.

In chapter 4 of this report the Working Groups and Interest Groups of the RDA, as well as the outputs, are discussed with respect to their value for the humanities.

<sup>7</sup> See: <https://www.rd-alliance.org/plenaries>

<sup>8</sup> Based on slides of RDA-in-a-nutshell presentation:  
<https://www.rd-alliance.org/sites/default/files/attachment/RDA-in-a-nutshell-March-2019.pptx>

## 3 THE HUMANITIES

Humanities are academic disciplines that study aspects of human society and culture. The humanities use methods that are primarily critical, or speculative, and have a significant historical element as distinguished from the mainly empirical approaches of the natural sciences. Unlike the sciences, it has no central discipline.<sup>9</sup> According to the classification “fields of science and technology” published by the OECD in 2015<sup>10</sup> the broad classification “Humanities and the Arts” can be subdivided into the second-level classifications: (1) History and archaeology; (2) Languages and literature; (3) Philosophy, ethics and religion; (4) Arts (arts, history of arts, performing arts, music) and (5) Other humanities. There is a close relation between humanities and the social sciences. Some scientific fields, such as law and politics, can belong to either discipline, depending on the classification applied.

### 3.1 Digital Humanities

Researchers in the humanities have developed numerous large- and small-scale digital corpora, such as digitized collections of historical texts, along with the digital tools and methods to analyze them. Their aim is both to uncover and share new knowledge about corpora and to visualize research data in new and revealing ways. Much of this activity occurs in a field called “digital humanities”. Initially “humanities computing” was the practice of using computing for and in the humanities. After 2004 “digital humanities” became the prominent name for the field.<sup>11</sup> Examples of the application of computers in the humanities are machine translation, text analysis, text markup, creation of concordances and glossaries, image processing and geospatial analysis.

One of the main problems related to the use of computers and computational techniques in the humanities concerns the representation of data for input, processing, and output. Humanities data is often rich and complex, non-standardised in format, without common or consistent metadata and ontologies. It also can be subject to complex rights issues. Consensus and best practice regarding digitisation and metadata standards for common usage, that still retain the richness of different disciplines and data types, could enable open access to humanities data, and facilitate data exchange.

### 3.2 Digital Research Infrastructures in the Arts and Humanities

The following quotes illustrate the role of digital research infrastructures in the arts and humanities. “The culture of data sharing in the humanities follows a variety of (often ad hoc) approaches. Formalized approaches, such as institutional repositories, shared infrastructure

---

<sup>9</sup> <https://en.wikipedia.org/wiki/Humanities>

<sup>10</sup> See: [https://en.wikipedia.org/wiki/Fields\\_of\\_Science\\_and\\_Technology](https://en.wikipedia.org/wiki/Fields_of_Science_and_Technology)

<sup>11</sup> The (hi)story of the application of computing, computers and computation in the humanities is described in: “The Gates of Hell: History and Definition of Digital Humanities Computing” in: *Defining Digital Humanities* edited by: Melissa Terras, Julianne Nyhan, Edward Vanhoutte. (Ashgate, 2013).

services, etc. are used when they are available and accessible, but there is not a single common practice across the entire domain of the humanities and availability of such services is not universal.”<sup>12</sup> “The fact that often the largest obstacles to data sharing are not technical, but instead have to do with questions of governance, funding and sustainability, is one important lesson to be taken from humanities infrastructure and data sharing practices.”<sup>13</sup> The European Strategic Forum for Research Infrastructures (ESFRI) is an important factor in the improvement of the situation.<sup>14</sup> ESFRI develops a roadmap of “research infrastructures of pan-European interest”, which covers all fields of research including the humanities, and is regularly updated.<sup>15</sup> The Roadmap includes two prominent European digital research infrastructures in the field of digital humanities, which provide among other things tools and services to share data. In this respect they share a common goal with the RDA. These infrastructures are DARIAH and CLARIN.<sup>16</sup>

Representatives of both infrastructures contribute to RDA activities, e.g. by providing case-studies or by active contribution to RDA WGs and IGs and also the other way around, RDA output is used by the infrastructures. More background information on the DARIAH and CLARIN infrastructure is given below.



DARIAH, the Digital Research Infrastructure for the Arts and Humanities was established as a European Research Infrastructure Consortium (ERIC) in August 2014.<sup>17</sup> Currently, DARIAH has 17 member countries and 11 cooperating partners. Each country counts several cooperating partners. DARIAH aims to enhance and support digitally-enabled research and teaching across the arts and humanities. DARIAH is a network of people, expertise, information, knowledge, content, methods, tools and technologies from its member countries. It develops, maintains and operates an infrastructure in support of ICT-based research practices and sustains researchers in using them to build, analyse and interpret digital resources. By working with communities of practice, DARIAH brings together individual state-of-the-art digital arts and humanities activities and scales their results to a European level. It preserves, provides access to and disseminates research that stems from these collaborations and ensures that best practices, methodological and technical standards are followed.

---

<sup>12</sup> Genova, F. et al 2017 Building a Disciplinary, World-Wide Data Infrastructure. *Data Science Journal*, 16: 16, pp 1-13 DOI: <https://doi.org/10.5334/dsj-2017-016>. p 2.

<sup>13</sup> Ibid, p 3.

<sup>14</sup> <https://www.esfri.eu/>

<sup>15</sup> The current roadmap (2018) can be found at: <http://roadmap2018.esfri.eu/>

<sup>16</sup> Next to DARIAH and CLARIN other initiatives are involved in the management of research data in the humanities. Examples are: European Infrastructure for Heritage Science - E-RIHS (<http://www.e-rihs.eu/>), ARIADNE (<https://ariadne-infrastructure.eu/>) that brings together and integrates existing archeological research data infrastructures in Europe and the PARTHENOS project (<http://www.parthenos-project.eu/>) that provides joint policies and solutions for the humanities and linguistic data lifecycle. The projects are all related in some way to DARIAH and CLARIN.

<sup>17</sup> Information on DARIAH taken from <https://www.dariah.eu/>

Scholarly information activities in humanistic disciplines such as history and historical geography are often classified as “discover”, “gather”, “create” and “share”. Humanities scholars engage not only in information seeking activities, but also in research activities related to the curation of information objects. Research infrastructures must support the annotation and editing of information objects by means of scholarly writing and publication.

DARIAH operates through its network of Virtual Competence Centres (VCCs) centered on a specific area of expertise. VCC “e-Infrastructure” is responsible for technological foundations working on a digital environment that allows to share community-developed data and tools to ensure the quality, permanence and growth of eInfrastructures and technical services.<sup>18</sup>



CLARIN, which stands for “Common Language Resources and Technology Infrastructure” makes digital language resources available to scholars, researchers, students and citizen-scientists from all disciplines, especially in the humanities and social sciences, through single sign-on access.<sup>19</sup> CLARIN offers long-term solutions and technology services for deploying, connecting, analyzing and sustaining digital language data and tools. CLARIN supports scholars who want to engage in cutting edge data-driven research, contributing to a truly multilingual European Research Area.

The technical pillars of the CLARIN infrastructure are<sup>20</sup>:

- Federated identity - letting users login to protected data and services with their own institutional login and password
- Persistent identifiers - enabling sustainable citations of electronic resources
- Sustainable repositories - digital archives where language resources can be stored, accessed and shared
- Flexible metadata and concept definitions - to ensure semantic interoperability when describing language resources
- Content search - offering a search engine for a wide range of language resources
- Web service chaining - giving users the possibility to freely combine language processing services

---

<sup>18</sup> <https://www.dariah.eu/activities/competence-centers/e-infrastructure/>

<sup>19</sup> Information on CLARIN taken from <https://www.clarin.eu/>

<sup>20</sup> <https://www.clarin.eu/content/clarin-technology-introduction>



## 4 THE RDA AND THE HUMANITIES

### 4.1 RDA Working Groups and Interest Groups

The Working Groups and Interest Groups of the RDA address specific issues in the field of research data management. Figure 2, 3 and 4 provide an overview of the RDA WGs and IGs.<sup>21</sup> Based on their title and description a number of them are (potentially) relevant for humanities research and they are described in more detail.



Figure 2. RDA Active Interest Groups and Working Groups by Focus (1)

Currently the RDA does not have a “Domain Science” WG that is active in the humanities field. A couple of IGs, however, might be relevant for the humanities. These are for instance:

Digital Practices in History and Ethnography IG. This IG works to advance data standards, practices and infrastructure for historical and ethnographic research, contributing to broader efforts in the digital humanities and social sciences.<sup>22</sup>

<sup>21</sup> Slides are part of the presentation “RDA in a Nutshell” March 2019.

<https://www.rd-alliance.org/sites/default/files/attachment/RDA-in-a-nutshell-March-2019.pptx>

<sup>22</sup> <https://www.rd-alliance.org/groups/digital-practices-history-and-ethnography-ig.html>

Linguistics Data IG aims to publish a set of recommendations for citing research data in linguistics (and for which a draft version is available).<sup>23</sup>

Mapping the landscape IG has the status “completed”. The Interest Group organised meetings at 4 successive plenaries. The scope of the IG was to ‘map’ projects that are describing the landscape of activities or assets related to building Research Data Infrastructures that enable the sharing of data, tools and computational infrastructures.<sup>24</sup>

Social Science and Humanities Research Data IG is initiated early 2019 and is not yet endorsed by the RDA. The initial objectives of this interest group will be to bring together major community members seeking to coordinate international social science and humanities research data sharing. Several existing groups within the RDA are discussing topics of interest to the SSH disciplines but none focuses specifically on the social science and humanities with the key focus on data sharing.<sup>25</sup> This IG intends to fill this gap.

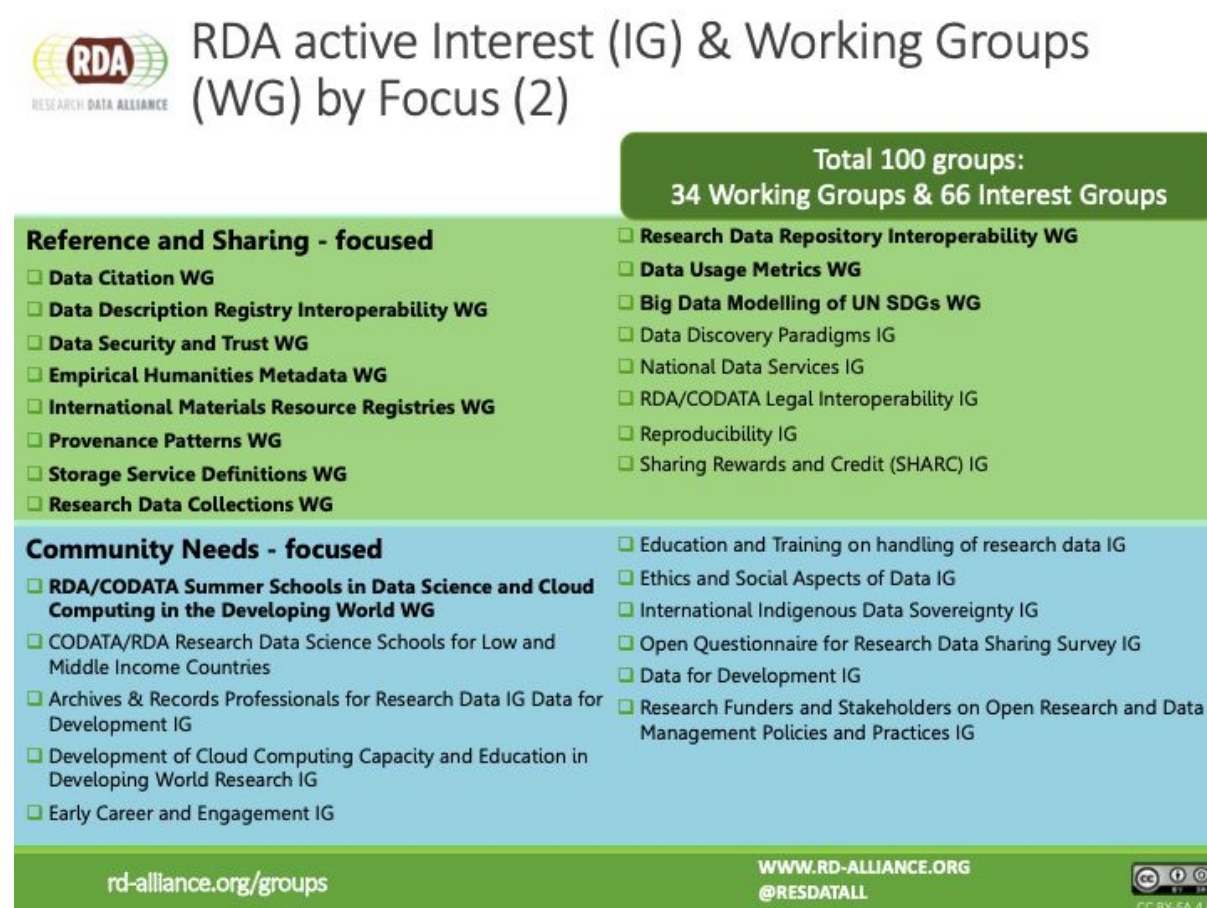


Figure 3. RDA Active Interest Groups and Working Groups by Focus (2)

<sup>23</sup> <https://www.rd-alliance.org/groups/linguistics-data-ig>

<sup>24</sup> <https://www.rd-alliance.org/groups/mapping-landscape-ig>

<sup>25</sup> <https://www.rd-alliance.org/groups/social-sciences-humanities-research-data-ig>

The WGs focused on “Reference and Sharing” are relevant for humanities research depending on the way research data is positioned in the humanities research landscape. Based on their description some relevant WGs are presented below.

Empirical Humanities Metadata WG This WG will conduct research and develop an adoptable and adaptable protocol for designing metadata management plans (MDMPs) in empirical humanities projects. A metadata management plan (MDMP), as a subcomponent of an overall data management plan (DMP), will be a formal document that outlines how metadata elements will be selected, structured, implemented, and managed within a project, and how social conventions will be developed to ensure discoverability and citability of data.

<sup>26</sup>

Research Data Repository Interoperability WG This WG will establish standards for interoperability between different research data repository platforms. These standards may include (but are not limited to) a generic API and import/export formats.<sup>27</sup>

Ethics and Social Aspects of Data IG This IG discusses ethical and social issues with respect to data archiving, sharing, and reuse. It focuses on the ethical agreements and social contracts that inform and constrain data sharing practices.<sup>28</sup>

The IGs and WGs active in the field of “Data Stewardship and Services” are relevant for humanities as they cover a wide range of activities, such as the management of data objects including metadata. They have a specialist role that incorporates processes, policies, guidelines and responsibilities for administering an organization's data in compliance with policies and/or regulatory obligations. A data steward is commonly responsible for data content, context, and associated business rules.<sup>29</sup>

Of particular interest seems to be the Domain Repositories Interest Group. This IG brings together active data repositories that serve specific scientific disciplines. Domain repositories provide archiving services tailored to the needs of scientific communities, and they mediate between domain scientists and rapid changes in information and data science. All domain repositories have shared interests in data curation, dissemination, preservation, and institutional sustainability, and this IG will provide a forum for sharing practical experience and developing joint projects. It will also develop training and mentoring relationships for new and emerging repositories.<sup>30</sup>

---

<sup>26</sup> <https://www.rd-alliance.org/group/empirical-humanities-metadata-working-group/case-statement/empirical-humanities-metadata>

<sup>27</sup> <https://www.rd-alliance.org/groups/research-data-repository-interoperability-wg.html>

<sup>28</sup> <https://rd-alliance.org/group/ethics-and-social-aspects-data/case-statement/rda-ethics-and-social-aspects-data-esad>

<sup>29</sup> [http://sedataglossary.shoutwiki.com/wiki/Data\\_steward](http://sedataglossary.shoutwiki.com/wiki/Data_steward)

<sup>30</sup> <https://rd-alliance.org/groups/domain-repositories-ig.html>



## RDA active Interest (IG) & Working Groups (WG) by Focus (3)

Total 100 groups: 34 Working Groups & 66 Interest Groups	
<b>Data Stewardship and Services – focused</b>	
<input type="checkbox"/> <b>Brokering Framework WG</b>	<input type="checkbox"/> Long tail of research data IG
<input type="checkbox"/> <b>DMP Common Standards WG</b>	<input type="checkbox"/> Physical Samples and Collections in the Research Data Ecosystem IG
<input type="checkbox"/> <b>Exposing Data Management Plans WG</b>	<input type="checkbox"/> Preservation e-Infrastructure IG
<input type="checkbox"/> <b>WDS/RDA Assessment of Data Fitness for Use WG</b>	<input type="checkbox"/> Preservation Tools, Techniques, and Policies IG
<input type="checkbox"/> <b>Data Versioning WG</b>	<input type="checkbox"/> RDA/WDS Certification of Digital Repositories IG
<input type="checkbox"/> <b>FAIR Data Maturity Model WG</b>	<input type="checkbox"/> RDA/WDS Publishing Data Cost Recovery for Data Centres IG
<input type="checkbox"/> Active Data Management Plans IG	<input type="checkbox"/> Repository Platforms for Research Data IG
<input type="checkbox"/> Data in Context IG	<input type="checkbox"/> Research Data Architectures in Research Institutions IG
<input type="checkbox"/> Data Rescue IG	<input type="checkbox"/> Research Data Provenance IG
<input type="checkbox"/> Domain Repositories IG	<input type="checkbox"/> Data policy standardisation and implementation IG
<input type="checkbox"/> Virtual Research Environments IG	<input type="checkbox"/> GO FAIR IG
<input type="checkbox"/> Libraries for Research Data IG	
<b>Base Infrastructure – focused</b>	
<input type="checkbox"/> <b>Array Database Assessment WG</b>	<input type="checkbox"/> Disciplinary Interoperability Framework IG
<input type="checkbox"/> <b>Data Type Registries WG</b>	<input type="checkbox"/> Big Data IG
<input type="checkbox"/> <b>Metadata Standards Catalog WG</b>	<input type="checkbox"/> Brokering IG
<input type="checkbox"/> <b>PID Kernel Information WG</b>	<input type="checkbox"/> Metadata IG
<input type="checkbox"/> <b>Persistent Identification of Instruments WG</b>	<input type="checkbox"/> PID IG
<input type="checkbox"/> <b>Software Source Code ID WG</b>	<input type="checkbox"/> Software Source Code IG
<input type="checkbox"/> Data Fabric IG	<input type="checkbox"/> Vocabulary Services IG
<input type="checkbox"/> Data Foundations and Terminology IG	<input type="checkbox"/> Federated Identity Management IG
	<input type="checkbox"/> Data Economics IG

Figure 4. RDA Active Interest Groups and Working Groups by Focus (3)

The IGs and WGs active in the “Base Infrastructure” discuss specific technical building blocks of research infrastructures and can be considered as discipline agnostic. Examples are the application of persistent identifiers, federated identity management and metadata and vocabulary services.

### 4.2 Recommendations and outputs

The recommendations and outputs of the RDA are published on the website.<sup>31</sup> At the time of writing the website listed twelve endorsed recommendations (of which four are classified as “ICT Technical specification”, six recommendations with RDA endorsement in process and twelve RDA supported outputs.

The recommendations and outputs cover topics such as “data citation”, “certification of repositories”, “persistent identifiers”, “metadata management”, “interoperability”, and “identity management”. They are discipline agnostic so they are as relevant to humanities as they are to any other discipline.

<sup>31</sup> <https://www.rd-alliance.org/recommendations-outputs>



The recommendation “[Metadata standards directory](#)” contains an overview of metadata standards for a wide range of disciplines, also for arts and humanities.<sup>32</sup> A second useful output for humanities is the “[23 things: Libraries for research data](#)”.<sup>33</sup> It provides an overview of free online tools and resources that libraries for research data can use to incorporate research data management into their practice of librarianship. Although addressed at libraries, information about the tools might be of relevance for humanities scholars as well. The “23 things” are an excellent starting point for humanities scholars and service providers to engage in data sharing and as a basis for further exploration of the more detailed recommendations of the RDA. It covers the topics “learning resources”, “data reference and outreach”, “data management plans”, “data literacy”, “metadata”, “citing data”, “data licencing and privacy”, “digital preservation”, “data repositories”, and “community of practice”.

“Adoption use cases”<sup>34</sup> put the theory of the recommendations into practice. CLARIN is involved in the adoption of three recommendations. In the first place in “Data foundation terminology” that ensures that researchers apply a common core data model when organising their data to make the data accessible and re-usable. Secondly in the “Metadata standards directory” that enables the discovery of metadata standards for documenting research data and that addresses issues related to coverage, ease of maintenance and sustainability. Thirdly in the “Repository audit and certification” group that creates harmonized common procedures for certification of repositories at the basic level, drawing from the procedures already put in place. DARIAH-DE is involved in adopting the recommendation of the “Research Data Repository Interoperability WG” that provides recommendations with respect to an interoperable packaging and exchange format for digital content.

## 5 GET INVOLVED

Open sharing of data is an essential aspect of the evolving research landscape. The RDA is an important platform for the development of services and standards to create the infrastructure for this. The humanities, both researchers and service providers, can benefit from the RDA as a network and a neutral forum to work on guidelines, standards and services to improve the quality of the research infrastructure and community research practices, and take advantage of the wide diversity of profiles and expertise of RDA members. Humanist scholars and service providers can help to shape this research landscape by becoming an active member of the RDA.

### 5.1 DARIAH and RDA

Every year the DARIAH ERIC organises a networking event where workgroups meet, ideas are exchanged and infrastructural solutions are discussed in an open conference setting.

---

<sup>32</sup> <https://rd-alliance.github.io/metadata-directory/subjects/arts-and-humanities.html>

<sup>33</sup> <https://rd-alliance.org/group/libraries-research-data-ig/outcomes/23-things-libraries-research-data-supporting-output>

<sup>34</sup> <https://www.rd-alliance.org/recommendations-outputs/adoption-recommendations>

During the DARIAH annual event of 2019<sup>35</sup>, as part of the RDA ambassadorship for humanities, a poster was presented aimed to lower the threshold for humanities scholars to participate in the activities of the RDA (see figure 5).



Figure 5. Poster “What can the RDA do for you?” Presented at the DARIAH Annual Event, May 2019, Warsaw.

<sup>35</sup> <https://dariah-ae-2019.sciencesconf.org/>

The first part of the poster contains information on the activities and organisational aspects of the RDA. Next, examples are given on how RDA and DARIAH can be of value for each other. The RDA provides recommendations on data sharing, funding possibilities (e.g. early career or experts grants to support participation to RDA plenaries) and open working and interest groups that can be joined. DARIAH contributes to the infrastructure development by RDA by bringing needs for the arts and humanities to the forefront. By engaging with RDA, DARIAH can share experiences with other scientific communities and gain expertise in data science. The third part of the poster provides 3 examples of RDA output that is of potential value for the DARIAH community. These are an overview of free, online resources in the field of research data management, a directory of metadata standards, and a harmonized common procedure for the certification of repositories.

## 5.2 CLARIN and RDA

A prominent example of a fruitful interaction between the RDA and the humanities field is the involvement of CLARIN representatives in the RDA and the implementation of RDA output in CLARIN. An interview with Dieter Van Uytvanck, the technical director of CLARIN illustrates this. The CLARIN case can act as an inspiration for other stakeholders in the humanities research landscape to become active in the RDA. The transcription of the interview can be found below.<sup>36</sup>

- Can you describe the relation between CLARIN and RDA?

“ It is a kind of natural relationship, a symbiosis. A lot of things we do are also relevant in the RDA. For instance, activities in relation to the certification of repositories, machine actionability of shared data, etc. These topics are prominent both in RDA and for CLARIN. Some ideas gathered via RDA can be worked out, you can learn from each other. RDA should not be considered as a “top down” organization. It is also important to contribute, e.g. by active participation in working groups and interest groups and also in sessions on “adoption stories” . “

- In which RDA WGs and IGs are you (and other CLARIN representatives) active?

“ I am involved in the activities of the Data Fabric Interest Group and related initiatives. We contribute to topics like federated identity management, citation of research data and the IG on linguistic data as well as to the groups that work on persistent identifiers. The new IG on social science and humanities research has our attention and we intend to be active in this group. In the past I was co-chair of a WG on dynamic data citation and contributed to the WG on data foundations and terminology. Other people of CLARIN are active in RDA e.g. on groups looking into legal issues in relation to data management. This also is dependent on what is of importance at the moment. The RDA has a lot of groups and it is not always easy to find out what is going on. I try to monitor the RDA activities and communicate this with CLARIN. Next to RDA also ISO as a standardization body is important; several persons involved in CLARIN participate in a number of ISO committees. “

- Which RDA output is relevant for CLARIN?

---

<sup>36</sup> The interview was conducted by René van Horik on May 2nd 2019.

“ What first comes to my mind is the work on data foundations and terminology as it helps us to standardize terms. Also, interoperability aspects in the data fabric group is very relevant, such as the work on PID kernel type information and the Digital Object Architecture. The virtual collection registry is also relevant for CLARIN in which we work together with other stakeholders in the context of Research Data Collections. A last topic that can be mentioned in this respect are repository certifications in relation to the Core Trust Seal. “

- Do you have any suggestions for the RDA to intensify the communication between the RDA and digital humanities scholars and research infrastructures?

“ Examples of practical implementations will help others to use the outcomes of the RDA. So, these aspects should be stressed. RDA Europe does have resources for implementation projects, but the process to prepare these could be less bureaucratic, It could be based on a lump sum and output-driven. This will increase the visibility of the RDA in the research world.“