

# The development of the matrix of use cases and functional requirements for research data repository platforms

Last updated: 2016-09-26 by Stefan Kramer

## Introduction

The [Repository Platforms for Research Data Interest Group](#) (hereafter: RPRD IG) of the Research Data Alliance (RDA) was formally launched at the [5th RDA Plenary Meeting](#) in San Diego, California, in March, 2015. Originally proposed as a Working Group (WG), but changed to an Interest Group following feedback from the RDA Technical Advisory Board, the RPRD IG had the charter, within the 18-month timeframe expected of an RDA WG, to “gather and analyze research data use cases in the context of repository platform requirements,” with the “primary deliverable [being] a matrix relating use cases with functional requirements for repository platforms” for research data. The primary target audience for the matrix was envisioned to be developers and service providers of repository software, with secondary audiences including repository implementers and managers, funding agencies, grant applicants, and researchers generating and using scientific data.

## Methodology

Members of the RPRD IG identified other RDA groups with expected outcomes or activities, as well as documents, relevant to its own work; these can be found in its wiki under [RPRD Related Work](#). RPRD IG members developed a Microsoft Word template for the collection of use cases, and identified communities of practice - and their online groups or mailing lists - from which to solicit use case submissions. The template, and the resulting collected use cases, can be found under [Collection of use cases](#); a listing of which communities were contacted when and by whom, under [RPRD Use Case Dissemination](#); and an example of a use case solicitation message in Appendix 1 of this document.

Specific RPRD IG members then volunteered to transfer the stated functional requirements and their respective importances<sup>1</sup> from the use case documents to a matrix in the form of a Google Sheet. That format/platform had been chosen because it lends itself well to convenient group collaboration and was familiar to IG members. The RPRD IG co-chairs then de-duplicated and combined the collected functional requirements, since many of them overlapped across use cases, and computed a “functional requirement score” for each, based on the importance values from the use cases, but inverted so that a *higher* value indicates *more* (combined) importance. At the top of the “Functional requirement score” column are summary statistics for that measure.

---

<sup>1</sup> which were expressed on a scale from 1 = very important to 5 = not at all important

## *Research Data Alliance Repository Platforms for Research Data IG*

The respective use case authors were then invited to review their original use cases vis-a-vis the matrix..

RPRD IG members discussed and coordinated progress on the aforementioned work at in-person meetings at RDA Plenary Meetings 5-8 (spring 2015-autumn 2016), and in regular virtual meetings (shown under [RPRD Meetings](#)).

## Conclusions

Due to the non-scientific collection method for use cases, and the relatively low resulting collected number (N=11), the resulting work cannot claim to be representative of all use cases and functional requirements for research data and repositories that may be used to store and deliver them. Furthermore, some of the collected use cases are of a generic nature (e.g., research data related to doctoral dissertations), some describe already implemented repositories (e.g., VecNet – Vector-borne Disease Network), others describe repositories in planning stages (e.g., at the Center for Information Services and High Performance Computing of the Technische Universität Dresden). Nonetheless, it is hoped that the resulting matrix may be informative to those developing, improving, selecting, deploying, customizing, or using a research data repository.

Future work of the RPRD IG, and/or of WGs spawned by it, may include: a rating of existing research data repository platforms based on the gathered functional requirements; a survey of developers, managers, and users of research data repository platforms, based on the gathered functional requirements; and the development of a generic API for interoperability of research data repositories.

---

Appendix 1: example use case solicitation email

Date: Tue, 18 Aug 2015 14:30:22 -0400 (Eastern Daylight Time)  
From: Stefan Kramer <skramer@american.edu>  
To: iasst-l@lists.columbia.edu  
Subject: Use case gathering for research data repositories by Research Data Alliance IG

[Apologies if you receive this email through multiple channels]

Dear fellow IASSISTers,

On behalf of the Research Data Alliance's "Repository Platforms for Research Data" Interest Group, I would like to invite you to submit use cases for research data repositories. The charter of this group is to "gather and analyze research data use cases in the context of repository platform requirements. The primary deliverable will be a matrix relating use cases with functional requirements for repository platforms." More details on the group and its work can be found at <http://www.rd-alliance.org/groups/repository-platforms-research-data.html>.

At <https://www.rd-alliance.org/use-cases.html-3> you will find:

- 1.) A Microsoft Word template for describing a use case:  
[Template\\_UseCase\\_RepoPlat.docx](#)
- 2.) Another Word file ([Dissertation\\_UseCase\\_RepoPlat.docx](#)) with an example of a completed use case, namely about a doctoral dissertation and its associated research data, which addresses concerns of the dissertation creator, their university, and potentially a funding agency.
- 3.) Already submitted use cases - as of this writing, there's one, on "Institutional Life-cycle Research Data Management" from the Donders Institute for Brain, Cognition and Behaviour in Nijmegen, The Netherlands.

If you'd like to participate in this information gathering, please download the template mentioned in 1.), complete it with your use case and its associated requirements for a research data repository, and then either upload it to the collection at <https://www.rd-alliance.org/use-cases.html-3> (which requires logging in to the RDA site), or email it to me and I will upload it on your behalf. The initial "deadline" for that is Sep. 15.

On Sep. 25, the "Repository Platforms for Research Data" IG will have a working meeting (<https://rd-alliance.org/ig-repository-platforms-research-data-p6-meeting-session.html>) at the 6th Plenary of the Research Data Alliance (<https://www.rd-alliance.org/plenary-meetings/rda-sixth-plenary-meeting.html>),

*Research Data Alliance Repository Platforms for Research Data IG*

which will include review of the use cases submitted by then. On Sep. 24, there will also be a joint meeting of the RDA interest groups "Libraries for Research Data," "Long tail of research data" and "Repository Platforms for Research Data"

(<https://rd-alliance.org/ig-libraries-research-data-p6-joint-session.html>).

Please don't hesitate to contact me and my fellow co-chairs with questions or comments.

Stefan Kramer  
Associate Librarian &  
Associate Director for Research Data Services  
American University Library  
4400 Massachusetts Ave. NW, Washington, DC 20016

Co-chairs of the RDA Repository Platforms for Research Data IG:

David Wilcox <[dwilcox@duraspace.org](mailto:dwilcox@duraspace.org)>

Stefan Kramer <[skramer@american.edu](mailto:skramer@american.edu)>

Ralph Müller-Pfefferkorn <[ralph.mueller-pfefferkorn@tu-dresden.de](mailto:ralph.mueller-pfefferkorn@tu-dresden.de)>