RDA Case Statement

GORC International Benchmarking WG

1. ## Charter

The Global Open Research Commons (GORC) is an ambitious vision of a global set of interoperable resources necessary to enable researchers to address societal grand challenges including climate change, pandemics, and poverty. The realized vision of GORC will provide frictionless access to all research artifacts including, but not limited to: data, publications, software and compute resources; and metadata, vocabulary, and identification services to everyone everywhere, at all times.

The GORC is being built by a set of national, pan-national and domain specific organizations such as the European Open Science Cloud, the African Open Science Platform, and the International Virtual Observatory Alliance (see Appendix A for a fuller list). The GORC IG is working on a set of deliverables to support coordination amongst these organizations, including a roadmap for global alignment to help set priorities for Commons development and integration. In support of this roadmap, this WG will develop and collect a set of benchmarks for GORC organizations to measure their user engagement and development internally within the organization, gauge their maturity and compare features across commons.

In the first case, the WG will collect information about how existing commons are measuring success, adoption or use of their services within their organization, such as data downloads, contributed software, and similar KPI and access statistics.

Secondly, we will also develop, validate, collect and curate a set of benchmarks that will allow Commons developers to compare features across science clouds. In the latter case for example, we would consider benchmarks such as evidence or the existence of :

1. A well defined decision making process
2. A consistent and openly available data privacy policy
3. Federated Authentication and Authorization infrastructure
4. Community supported and well documented metadata standard(s)
5. A workflow for adding and maintaining PIDs for managed assets
6. A mechanism for utilizing vocabulary services or publishing to the semantic web
7. A process to inventory research artefacts and services
8. An Open Catalogue of these artefacts and services
9. A proven workflow to connect multiple different research artefact types (e.g. [data and publications](#); data and electronic laboratory notebooks; data and related datasets)
10. A mechanism to capture provenance for research artefacts
11. Mechanisms for community engagement and input; an element or scale for inclusion

We anticipate that the first set of metrics will be quantitative measures used within an organization, while the second set of benchmarks will be comparable across organizations.

2. Value Proposition

This WG is motivated by the broader goal of openly sharing data and related services across technologies, disciplines, and countries to address the grand challenges of society. The deliverables of the WG itself will inform roadmaps for development of the infrastructure necessary to meet that goal, while engagements and relationships formed during the work period will help forge strong partnerships across national, regional and domain focused members which are crucial to its success. Identifying observable and measurable benchmarks in pursuit of the global open science commons will help create a tangible path for development and support strategic planning within and across science commons infrastructures. In the future, best practices for commons development will emerge based on the experience of what actions led to successful outcomes. This work will provide a forum for discussion that will allow members to identify the most important features and the minimal elements required to guide their own development and build a commons that is globally interoperable. Finally, it will

support developers as they seek resources to build the global commons by helping them respond to funding agencies requirements for measurable deliverables.

The proposed WG was discussed at the [RDA 16 virtual plenary](#).[1] Participants discussed the initial work packages and agreed during the meeting this was a worthy goal and an appropriate approach.

3. Engagement with Existing Work

The GORC IG builds on, and incorporates the previous [National Data Services IG](#). The Commons that will be investigated in this WG are likely either to have considered or implemented outputs from other RDA groups, such as the [Domain Repositories IG](#), the [Data Fabric IG](#), and the [Virtual Research Environment IG](#), just to name a few. These groups and many others outside of RDA will have recommendations that speak to functionality and features of various components of Commons; for example the re3data.org schema for collecting information on research data repositories for registration, the EOSC [FAIR WG](#) and [Sustainability WG](#) that seek to define the EOSC as a Minimum Viable Product (MVP). We will review these and other related outputs to see if they have identified benchmarks that we feel will support our goals. This review period will ensure that we do not duplicate existing efforts. [Appendix B](#) of this case statement identifies a few of these existing efforts, both within and without RDA; this list will be expanded and reviewed by the WG members.

4. Work Plan

To create these deliverables, members of the group will:

1. Create a target list of Commons ([Appendix A](#))
2. Review public facing documentation of each Commons to extract benchmarking information (both KPIs and feature lists).

---

[1] P16 [session notes](#) and [presentation](#)

3. Review public facing documentation of recommendations and roadmaps from related communities to extract benchmarking information ([Appendix B](#)). This evaluation phase will include an examination of the outputs from other RDA WGs and position papers available in the wider science infrastructure community, along with experiences gathered by the WG's members.
4. Because benchmarking information may not be easily found in public documents we will conduct outreach to Commons representatives and related organizations to ask for additional feedback and information about benchmarks used by their community. This may include benchmarks already in use, as well as benchmarks that organizations feel would be useful but which are not yet implemented.
5. Synthesize and document the benchmarks into 2 deliverables, described below..

We anticipate that the WG will create sub-working groups or task groups. The WG will decide if they would rather define the task group according to the deliverables, creating a Commons Internal Benchmarking TG and a Commons External Benchmarking TG, or if they would rather subdivide according to a typology of the commons, for example with some members looking at pan-national, national, or domain specific commons, or by some other subdivision of labor.

The WG will proceed according to the following schedule:

Month	Activity
Jan-Mar 2020	Group formation 1. Agreement on the scope of work and deliverables (broad scope) 2. Case statement community review 3. Creation of sub-working groups
Apr-Jun 2021	Begin literature review of public facing documents from Science Commons and related organizations Refine scope: Meeting point to consolidate list of topics to be addressed in the deliverables and assess level of resource available to achieve them Begin to define methodology

	Update at [RDA17](#)
Jul-Sep 2021	Recruit members to WG, Refine method, begin lit review
Oct-Dec 2021	Begin outreach to Science Commons and related organizations

Update at RDA18 |
| Jan-Mar 2022 | First draft: Internal Benchmarks distributed for community review |
| Apr-Jun 2022 | First draft: External Benchmarks distributed for community review

Update at RDA19 |
| Jul-Sep 2022 | Develop adoption plan |
| Oct-Dec 2022 | Final deliverables

Update at RDA20 |

5. Deliverables

This group will create [Supporting Outputs](#) in furtherance of the goals of the [GORC IG](#). Specifically, 2 documents:

D1: a non-redundant set of KPIs and success metrics currently utilized, planned or desired for existing science commons, and

D2: a list of observable international benchmarks of features, structures and functionality that can help define a Commons and that will feed into a roadmap of Commons interoperability.

D3: Adoption Plan: described in section 9 below.

6. Mode and Frequency of Operation

The WG will meet monthly over Zoom, at a time to be determined by the membership. The WG will also communicate asynchronously online using the mailing list functionality provided by RDA and via shared online documents. If and when post-Covid international travel is restored during the 18 month work period of this WG then we will propose and schedule meetings during RDA plenaries and at other conferences where a sufficient number of group members are in attendance.

7. Addressing Consensus and Conflicts

The WG will adhere to the stated [RDA Code of Conduct](#) and will work towards consensus, which will be achieved primarily through mailing list discussions and online meetings, where opposing views will be openly discussed and debated amongst members of the group. If consensus cannot be achieved in this manner, the group co-chairs will make the final decision on how to proceed.

The co-chairs will keep the working group on track by reviewing progress relative to the deliverables. Any new ideas about deliverables or work that the co-chairs deem to be outside the scope of the WG defined here will be referred back to the GORC IG to determine if a new WG should be formed.

8. Community Engagement

The working group case statement will be disseminated to RDA mailing lists and communities of practice related to Commons development that are identified by the GORC IG in an effort to cast a wide net and attract a diverse, multi-disciplinary membership. Similarly, when appropriate, draft outputs will also be published to relevant stakeholders and mailing lists to encourage broad community feedback.

9. Adoption Plan

The WG will create an adoption plan for distributing and maintaining the deliverables. A specific plan will be developed to facilitate adoption or implementation of the WG

Recommendation and other outcomes within the organizations and institutions represented by WG members. This will include possible strategies for adoption more broadly within the global community, and in such a way as to facilitate interoperability of global infrastructures. Pilot adoptions or implementations would ideally start within the 18 month timeframe before the WG is complete. We envision implementation occurring when developers of commons compare themselves with similar organizations. We also envision the adoption plan will speak to howwe include the benchmarks in the larger GORC roadmap being created by the parent IG.

10. Initial Membership

Co-chairs:

1. Karen Payne <ito-director@oceannetworks.ca>
2. Mark Leggott <mark.leggott@rdc-drc.ca>
3. Andrew Treloar <andrew.treloar@ardc.edu.au>

Appendix A: List of Commons

Pan National Commons

1. European Open Science Cloud
2. African Open Science Platform
    a. including H3Africa?
3. Nordic e-Infrastructure Collaboration
4. the Arab States Research and Education Network, ASREN
5. LIBSENSE (LIBSENSE is a community of practice, not an infrastructure. The infrastructure will be built by the RENs, NRENs and universities)

6. [WACREN](#)

7. [LA Referencia](#)

National Commons

European Roadmaps - The European Commission and European Strategy Forum on Research Infrastructures (ESFRI) encourage Member States and Associated Countries to develop national roadmaps for research infrastructures.

1. [German National Research Data Infrastructure](#) (NFDI)

2. [DANS](#)

3. [ATT (Finland)](#)

4. [GAIA-X](#) (non- member state?; [see also](#)) (focused on data sharing in the commercial sectors - without excluding research)

5. UK

   a. [UK Research and Innovation](#)
   b. [JISC](#)

   c. [Digital Curation Centre](#)

Non-European

6. [QNL (Qatar)](#)

7. [China Science and Technology Cloud](#) (CSTCloud); [see also](#)

8. [Australian Research Data Commons](#)

9. [Canadian National Data Services Framework](#) (in development)

10. [National Research Cloud](#) (US; AI focused)

11. [NII Research Data Cloud](#) (Japan)

12. [KISTI](#) (South Korea)

13. [Malaysia Open Science Platform](#)

Domain Commons

1. [International Virtual Observatory Alliance](#) (IVOA) (including [SKA](#)?)
2. [NIH Data Commons](#); [Office of Data Science Strategy](#) (USA)
3. [NIST RDaF](#) (USA)

4. Earth Sciences

    a. [DataOne Federation](#)

    b. [Federation of Earth Science Information Partners](#) (ESIP)
    c. [EarthCube](#)
    d. [GEO](#) / [GEOSS](#) ([GEOSS Requirements](#) lists functionality; GEOSS Common Infrastructure - [GCI](#))
    e. [Near-Earth Space Data Infrastructure for e-Science](#) (ESPAS, prototype)
    f. Polar
        i. The [Arctic Data Committee landscape map](#) of the Polar Community

        ii. [Polar View](#) - The Canadian Polar Data Ecosystem (includes international initiatives, infrastructure and platforms)

        iii. [Polar Commons](#) / Polar International Circle (PIC) [not sure if this is active]

        iv. [PolarTEP](#)

    g. [Infrastructure for the European Network for Earth System Modelling](#) (IS-ENES)

5. [Global Ocean Observing Systems](#) (composed of [Regional Alliances](#))

6. [CGIAR Platform for Big Data in Agriculture](#)

7. [Social Sciences & Humanities Open Cloud](#) (SSHOC)

8. Dissco https://www.dissco.eu/ Research infrastructure for natural collections (a commons for specimens and their digital twins)

9. [ELIXIR Bridging Force IG](#) (in the process of being redefined as "Life Science Data Infrastructures IG")

10. Global Alliance for Genomics and Health (GA4GH)

11. [Datacommons.org](#) - primarily statistics for humanitarian work

Gateway/Virtual Research Environment/Virtual Laboratory communities and other Services

1. [International Coalition on Science Gateways](#)

2. [Data Curation Network](#)

3. [CURE Consortium](#)

4. [OpenAire](#)

5. [RDA VRE IG](#)

Appendix B: Draft List of WG/IG, documents, recommendations, frameworks and roadmaps from related and relevant communities

1. [RDA Outputs and Recommendations Catalogue](#)
2. [RDA Data publishing workflows](#) ([Zenodo](#))
3. [RDA FAIR Data Maturity Model](#)
4. [RDA 9 functional requirements for data discovery](#)
5. [Repository Platforms for Research Data IG](#)
6. [Metadata Standards Catalog WG](#)
7. [Metadata IG](#)
8. [Brokering IG](#)
9. [Data Fabric IG](#)
10. [Repository Platform IG](#)
11. [International Materials Resource Registries WG](#)
12. [RDA Collection of Use Cases](#) [(see also)](#)
13. Existing service catalogues (for example the [eInfra service description template](#) used in the EOSC)
14. [the Open Science Framework](#)
15. [Matrix of use cases and functional requirements for research data repository platforms](#).
16. Activities and recommendations arising from the interdisciplinary [EOSC Enhance](#) program
17. [Scoping the Open Science Infrastructure Landscape in Europe](#)
18. Docs from https://investinopen.org/about/who-we-are/
19. [Monitoring Open Science Implementation in Federal Science-based Departments and Agencies: Metrics and Indicators](#)
20. [Next-generation metrics:Responsible metrics and evaluation for openscience](#). Report of the European Commission Expert Group on Altmetrics ([see also](#))
21. Guidance and recommendations arising from EOSC [FAIR WG](#) and [Sustainability WG](#)

22. Outputs from the International FAIR Convergence Symposium (Dec 2020) (particularly the session Mobilizing the Global Open Science Cloud (GOSC) Initiative: Priority, Progress and Partnership
23. The European Strategy Forum on Research Infrastructures (ESFRI) Landscape Analysis "provides the current context of the most relevant Research Infrastructures that are available to European scientists and to technology developers"
24. NIH Workshop on Data Metrics (Feb 2020)
25. WMO's Global Basic Observing Network (GBON) has internationally agreed metrics to guide investments, "using data exchange as a measure of success, and creating local benefits while delivering on a global public good."
26. Evolving the GEOSS Infrastructure: discussion paper on stakeholders, user scenarios and capabilities
27. There is a national open access policy in Ethiopia that was released last year, one in the first in Africa to my knowledge. Part of AOSP?
28. Briefing Note for CODATA Officers: CAS GOSC (Global Open Science Cloud) Project
29. UNESCO Open Science Recommendation
30. Open Science in the ISC Science Action Plan
31. CODATA: Coordinating Global Open Science Commons Initiatives
32. CODATA: Policies and Interoperability for Global Big Earth Data: a joint CASEarth and CODATA Workshop Session
33. CODATA: Building a global network infrastructure for international cooperation on data-intensive science