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. 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 the roadmap, this WG will generate a set of pertinent attributes to identify common features across commons. We will not coordinate or certify the use of specific attributes by research commons. Rather, we will review and identify attributes or features currently implemented by a target set of GORC organizations and when possible identify how they measure their user engagement with these features.

In the first case we will collect and curate a set of attributes that will allow Commons developers to compare features across science clouds.
The GORC IG has defined the Commons as “A global trusted ecosystem that provides seamless access to high quality interoperable research outputs and services.” In short: “Digital research resources for the common good.” This WG is motivated to identify attributes or features that a user would expect to find in this system, especially those attributes that lead to interoperability across different Commons infrastructures. In light of that, in this context we define attributes: **a standard, feature, functionality or point of reference from which information can be documented, or measurements or comparison may be made.**

For example, we would consider attributes 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
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

The attributes listed above will be an initial starting point of what we would expect to find in a mature research commons. The initial list will be augmented and refined by members of the WG during a landscape level review of the target commons, as described in section 5 below.

When relevant, the WG will collect information about how existing commons are measuring success, adoption or use of their attributes and services within their organization, such as number or amount of data downloads, contributed software, and
similar key performance indicators (KPI) and access statistics. In our model, the attributes will be the existence of a feature or service and are comparable across organizations, although we will not be comparing the organizations themselves. The KPIs are quantitative measures used within an organization to measure the uptake, engagement or use of an attribute, feature or service. We seek to identify how organizations are measuring engagement but do not seek to quantify or identify the value of the KPIs in the Commons we review. Finally, while it is not the primary goal, whenever possible we will collect information about exemplars of attributes or features as we review the commons. This would include a finer level of detail about the implementation or deployment of the attribute such as access protocols and encoding standards.

For example, the attribute: “A mechanism for utilizing (or accessing) vocabulary services” is exemplified by the existence of the NERC Vocabulary Server (NVS) in the EOSC, and by Research Vocabularies Australia (RVA) in the Australian National Data Service (ANDS). NERC uses the Simple Knowledge Organization System (SKOS) to represent concepts in the vocabulary service and provides access via both SPARQL and SOAP endpoints. ANDS RVA also serves SKOS-encoded vocabularies and provides a SPARQL endpoint, but also a RESTful API and the option to bulk download complete vocabularies in a single file for local processing. The attribute in this case is “evidence of the ability to use a vocabulary service,” satisfied by exemplars in both RVA and NVS. A relevant KPI that may be tracked by ANDS could be “the number of download requests for complete vocabularies.” The WG will collectively decide what constitutes an attribute, and will place a priority on attributes that increase or enable cross domain and cross-commons interoperability. For example, the ANDS RVA service also has the ability for users to self-register and create, edit or upload vocabularies, a function not available in the NVS. In this case, the WG will decide if the ability to create and edit, not just access, a vocabulary service should constitute a separate attribute.

2. A Model, not a Certification

The WG will create and populate a model or framework to describe Commons attributes, but we are not interested in endorsing, certifying or otherwise placing a value
judgement on Commons and their features, nor do we intend this to be the only or penultimate model of a Commons. The intent of the model is to provide an opportunity for Commons developers to self assess their options for their roadmaps based on attributes that exist or are desired by other Commons entities. We will collect information about each of the attributes we see “in the wild”, but the attributes are not intended to be prescriptive regarding implementation, nor are they ranked. We would like to state explicitly that not all commons-developing organisations are expected to develop all these features; what they decide to implement will depend on their mission and community requirements. The goal is for self identified commons that have different histories and areas of focus to be able to use these attributes at their discretion as they identify their development priorities. Organizations should be able to add their own information about services and attributes that intersect with the model if that is useful for them. We anticipate this will be an iterative process, especially at the onset, and has been described by one of our members as the “fruit salad approach.” We recognize that Appendix A, where we have provided the target list of Commons that we will review to extract attributes, contains a mix of research commons, organizations, and specific national projects/initiatives. The list is intended as a starting point for the search for attributes; the members of the WG have agreed to iteratively work through this list in good faith and seek out attributes in support of global research services interoperability. Our approach is to have an initial diversity of research infrastructure types to help refine the focus to where it needs to be based on the group discussions. There are additional organizations that some may feel belong in this target list, and we have identified and included those organizations in the potential list of documents to review in Appendix B. This reflects the WG’s desire to narrow the scope of the project, without removing relevant initiatives from the conversation entirely, especially those that showcase cross disciplinary or cross domain interoperability. WG members will cull through the list as part of the work plan described in section 5.

3. 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 attributes 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 attributes and features and the minimal elements required to guide their own development and build a commons that is globally interoperable. Building interoperable commons will support many research efforts including work focused on societal grand challenges and UN Sustainable Development Goals (SDGs). 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 both the RDA 16 and RDA17 virtual plenaries. Participants discussed the initial work packages and agreed during the meeting this was a worthy goal and the approach described in the Work Plan in section 5 below was deemed appropriate.

4. Engagement with Existing Work

This WG will review all appropriate IG and WG outputs to determine intersection with this work, and engage with the WG/IGs as appropriate. Some of the efforts are reasonably well known now: the GORC IG builds on, and incorporates the previous National Data Services IG, which was embarking on a similar exercise when the GORC started; the Domain Repositories IG, specifically the repository-specific discovery metrics. The RDA 9 functional requirements for data discovery will be very informative of the attributes associated with data repositories and the WG on Data Usage Metrics will inform KPIs for data services. FAIR plays an essential role for enabling seamless

1 P16 session notes and presentation; P17 session notes and presentation
access to data and other digital objects, the RDA FAIR Data Maturity Model will be a priority for the WG to review.

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 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 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 attributes 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.

5. Work Plan

This WG is predominantly an extensive review of Commons and Commons-adjacent white and peer-reviewed papers, conference proceedings, strategic and implementation plans, guidelines and related literature, coupled with discussion and review by WG members. Figure 1 shows the proposed workflow to be tested and refined by the group.
Figure 1. Proposed workflow to be tested and refined by the GORC International Model WG.

**Step 1: Identify target list of commons**

Appendix A is the current target list of Commons as agreed upon by WG members at the 17 July 2021 working meeting. This list represents a significant reduction in Commons listed in the first 2 versions of this charter. The target Commons reveal representation in Europe, Africa, ME, Asia-Pacific and North America, with global representation captured in the 2 domain commons.

**Step 2: Define the form of the deliverable**

This step begins by creating and identifying the most effective online collaboration mechanism for the group. The overall goal is to create online spaces that allow members to propose, refine and capture attributes, KPIs and implementations. As part of this process the WG co-chairs have begun developing options for the group to accept, tweak or reject as appropriate community building spaces. These include the dedicated online RDA space, an online form for submitting attributes and generating

---

2 https://www.rd-alliance.org/group/global-open-research-commons-international-benchmarking-wg-global-open-research-commons-ig
slides used in WG discussions, a Zotero instance of references, a database to house identified attributes, and an online collaborative document. We can also propose to the group: a github structure, a wiki and commercial or opensource project management systems such as Trello, Jira, Microsoft Project, MyCollab or OpenProject. After the collaboration mechanism has been clarified we can begin to refine the project deliverable.

We have identified two potential final forms of the deliverable for this working group. One option would be to create an online narrative document that describes attributes, KPIs and implementations, with extensive footnotes, hyperlinks and references. This would be the easiest form to implement and has the advantage of supporting easy asynchronous editing by WG members at their convenience. At a minimum this group will create a narrative document that describes our process and outcome.

In addition, we could create a relational database of attributes that holds information about the source material used to review the Commons, along with their related KPIs and implementations. Ideally these would be described using the same language found in other parts of the commons, for example classified according to language found in service catalogues like the e-infra catalogue standards and the GORC IG Typology of Commons. Figure 2 shows an early draft of a UML that demonstrates the concept of attributes as a database.

In addition to the form, the WG will need to decide what properties will be captured in the description of each attribute, KPI and implementation. This will be the subject of working group meetings in mid to late 2021. To the extent possible we will align the classification of attributes with the topology of commons being developed by the GORC IG.
Figure 2. Draft UML for a database used to capture attributes.

**Step 3: Generate Initial List of Attributes and KPIs**
In this step we will seed the database or document with an initial set of attributes that WG members would expect to find in a research commons, similar to the examples provided earlier in this case statement. This seed list will be based on the expertise and experiences of WG members, and could be aided by existing Commons service catalogues, such as the eInfra Service Catalogue standard. Service catalogues and registries from mature commons will be very helpful in this WG, but do not account for other non-digital features necessary to develop a commons, such as the existence of an openly available data privacy policy.

**Step 4: Create Sub/Task Groups**

We anticipate that WG members will want to subdivide the group into task groups according to their areas of expertise (networking, research data management), domain (earth sciences, health and life sciences), geographic location (Asia-Pacific, Africa), familiarity with particular Commons services and developers or by some other subdivision of labor. As they see fit, WG will create task groups, each responsible for reviewing a subset of the target list and ancillary documents.

There are multiple ways for the WG to create task groups. The WG will decide if they would rather define the task group according to the deliverables, creating a Commons Internal TG and a Commons External 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,

**Step 5: Literature Review/Collect and propose attributes**

We have begun collecting relevant literature pulled from Appendix B and beyond in a Zenodo library. WG members will self-select documentation from this collection to extract information and propose attributes, KPIs implementations and report back to the larger WG. 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.

**Step 6: Discuss proposed attributes in WG. Approve or remove from inclusion.**
During each regular WG meeting, members will collectively discuss and ultimately accept or reject proposed attributes. As part of their preparatory work for this project, the Chairs of the WG will use the collaboration platform approved by the group in step 2. For example, the Chairs can create an online form that will allow members to submit attributes, KPIs and implementations, which are captured in a spreadsheet and used to automatically create a slide for each submission. Either the slides, or alternatively, the growing, live document describing attributes can be collectively debated, edited and wordsmithed during the calls. Once a submission is approved it will be incorporated into the final product.

**Step 7: Conduct outreach**

Because attribute 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 attributes used by their community. This may include attributes, services or features already in use, as well as attributes that organizations feel would be useful but which are not yet implemented.

**Step 8: Create adoption plan in conjunction with IG and user consultation**

In the final step the WG will synthesize the attributes, draft an adoption plan and create the 2 deliverables described in section 6 below.

6. **Deliverables**

This group will create Supporting Outputs in furtherance of the goals of the GORC IG. Specifically:

D1: a list of observable international attributes of features, services structures and functionality that can help define a Commons and that will feed into a roadmap of Commons interoperability. The attribute criteria needs to remain simple, understandable and not skewed towards the particular reality of some of the commons so as not to appear as irrelevant or unattainable to Commons developers. It will include a description of implementations observed or planned in Commons examined in this work. This will either be a narrative document or database and will include a
non-redundant set of KPIs and success metrics currently utilized, planned or desired for existing science commons, and classified by functional layers defined by the GORC IG; how do we define a minimal interoperability. This is the primary output of the WG: an organizational structure and framework that can capture observations of attributes, KPIs and implementations currently in use or expected in research commons.

D2: Contributions to the GORC IG Roadmap for Commons integration. The contribution will integrate the attributes into the GORC IG typology of commons, and relate any integration or intersections with the CODATA GOSC work, including use cases. We recognize that RDA working groups are expected to produce endorsable and adoptable “Recommendations.” In this case, our recommendations will be captured in the larger IG roadmap document. We felt that this body of work was sufficiently large to warrant a separate WG from the IG and feel that it is appropriate for the IG document to be the vehicle for the recommendations as it will lay out a plan for Commons integration globally.

7. Schedule

The WG will proceed according to the following schedule:

<table>
<thead>
<tr>
<th>Month</th>
<th>Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan-Mar 2020</td>
<td>Group formation</td>
</tr>
<tr>
<td></td>
<td>1. Agreement on the scope of work and deliverables (broad scope)</td>
</tr>
<tr>
<td></td>
<td>2. Case statement community review</td>
</tr>
<tr>
<td>Apr-Jun 2021</td>
<td><strong>RDA17</strong></td>
</tr>
<tr>
<td></td>
<td>Refine scope: Agree to target list of commons and organizational approach</td>
</tr>
<tr>
<td></td>
<td>Report on progress to International Symposium on Global Open Science Cloud (June 2021)</td>
</tr>
<tr>
<td></td>
<td>Recruit additional members to WG</td>
</tr>
<tr>
<td>Jul-Sep 2021</td>
<td>Begin to refine methodology and online collaboration process, especially the form for data collection and initial set of commons</td>
</tr>
<tr>
<td></td>
<td>Identify subgroups</td>
</tr>
<tr>
<td></td>
<td>Begin literature review of public facing documents from Science Commons and</td>
</tr>
<tr>
<td>Period</td>
<td>Activities</td>
</tr>
<tr>
<td>-------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
</tbody>
</table>
| Oct-Dec 2021| Continue lit review and discussions  
Begin outreach to Science Commons and related organizations  
Update at RDA18  
| Jan-Mar 2022| Continue lit review and discussions |
| Apr-Jun 2022| First draft: External attributes distributed for community review  
Update at RDA19 |
| Jul-Sep 2022| First draft: Internal attributes distributed for community review |
| Oct-Dec 2022| Develop adoption plan  
Update at RDA20 |
| Jan - Mar 2023| Final deliverables |

8. **Mode and Frequency of Operation**

The WG will meet monthly over Zoom, at a time to be determined by the membership, likely every 4 Thursday of the month at 13:00UTC. 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.

9. **Addressing Consensus and Conflicts**

The WG will adhere to the stated [RDA Code of Conduct](https://www.rda.datacloud.org/) 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.

10. 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. The GORC International Model effort is also being facilitated by the RDA Secretariat, providing a strong intersection with the EOSC community - this will provide an additional level of community engagement. Similarly, the CODATA GOSC work, and the associated coordination of both efforts by the Data Together group, will provide additional engagement and outreach to the WDS and GO FAIR communities. When appropriate, draft outputs will also be published to relevant stakeholders and mailing lists to encourage broad community feedback, this will include both the GORC WG and GORC IG membership. When appropriate we will ask members of the WG to reach out to their own networks.

11. 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>

Current members represent Europe, the U.S., Canada, Australia, and the UK. There was a significant rise in membership after the P17 and we have recently submitted a call for an
additional co-chair from outside of the global north. We will continue recruitment via promotions of P18, via members of the GORC IG and CODATA GOSC WG. We hope the CODATA-led GOSC Symposium being planned for September 2021, will also generate additional memberships.

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

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. GAIA-X (non- member state?; see also) (focused on data sharing in the commercial sectors - without excluding research)
3. UK JISC Open Research Framework

Non-European

4. China Science and Technology Cloud (CSTCloud); see also
5. Australian Research Data Commons
6. NDRIQ (Canada)
7. NII Research Data Cloud (Japan)
8. KISTI (South Korea)
Domain Commons

9. International Virtual Observatory Alliance (IVOA) (including SKA?)
10. GEO / GEOSS (GEOSS Requirements) lists functionality; GEOSS Common Infrastructure - GCI

Appendix B: Draft List of WG/IG, organizations, documents, recommendations, frameworks and roadmaps from related and relevant communities to be reviewed during research phase

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. Vocabulary Services IG
11. Repository Platform IG
12. International Materials Resource Registries WG
13. RDA Collection of Use Cases (see also)
14. Existing service catalogues (for example the eInfra service description template used in the EOSC)
15. the Open Science Framework
16. Matrix of use cases and functional requirements for research data repository platforms.
17. Activities and recommendations arising from the interdisciplinary EOSC Enhance program
18. Scoping the Open Science Infrastructure Landscape in Europe
19. Docs from https://investinopen.org/about/who-we-are/
20. Monitoring Open Science Implementation in Federal Science-based Departments and Agencies: Metrics and Indicators
22. Guidance and recommendations arising from EOSC FAIR WG and Sustainability WG
23. Outputs from the International FAIR Convergence Symposium (Dec 2020) (particularly the session Mobilizing the Global Open Science Cloud (GOSC) Initiative: Priority, Progress and Partnership
24. 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”
25. NIH Workshop on Data Metrics (Feb 2020)
26. 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."

27. Evolving the GEOSS Infrastructure: discussion paper on stakeholders, user scenarios and capabilities

28. 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?

29. Briefing Note for CODATA Officers: CAS GOSC (Global Open Science Cloud) Project
30. UNESCO Open Science Recommendation
31. Open Science in the ISC Science Action Plan
32. CODATA: Coordinating Global Open Science Commons Initiatives
33. CODATA: Policies and Interoperability for Global Big Earth Data: a joint CASEarth and CODATA Workshop Session
34. CODATA: Building a global network infrastructure for international cooperation on data-intensive science

35. Outputs from European Plate Observing System (EPOS) under ERI (European Research Infrastructure Consortium) upcoming work package: "Strategy for engagement across solid Earth research infrastructures on a global scale" in the section Key initiative and infrastructure [architecture]

36. A Research Data Infrastructure for Materials Science
37. CeNAT (Costa Rica)
38. Canada's Roadmap for Open Science
39. Are there any ontologies for metrics and measurements we should be aware of?
40. RDA Adoption stories for implementations in the wild
41. Relevant to KPIs: RDA17 Session: From principles to metrics to evaluation, increasing TRUST in data repositories
42. RDA Brokering Framework: Preliminary Recommendations
43. EOSC Interoperability Framework (v1.0)
44. An Analysis of Scientific Practice towards FAIR Digital Objects
45. E-infra catalogues include KPIs - https://einfracentral.eu/node/151
46. Is the new Digital Earth project in CapeTown part of the AOSP? https://www.digitalearthafrica.org/
47. DANS
48. Earth Sciences
   a. DataOne Federation
   b. Federation of Earth Science Information Partners (ESIP)
   c. EarthCube
49. Near-Earth Space Data Infrastructure for e-Science (ESPAS, prototype)
50. Polar
   a. The Arctic Data Committee landscape map of the Polar Community
   b. Polar View - The Canadian Polar Data Ecosystem (includes international initiatives, infrastructure and platforms)
   c. Polar Commons / Polar International Circle (PIC) [not sure if this is active]
   d. PolarTEP
51. Infrastructure for the European Network for Earth System Modelling (IS-ENES)
52. Global Ocean Observing Systems (composed of Regional Alliances)
53. Global Climate Observing System
54. CGIAR Platform for Big Data in Agriculture
55. Health and Life Sciences
   a. ELIXIR Bridging Force IGI (in the process of being redefined as “Life Science Data Infrastructures IG”)
   b. NIH Data Commons; Office of Data Science Strategy (USA)
   c. AIRR Data Commons
d. Global Alliance for Genomics and Health (GA4GH)
56. Social Sciences & Humanities Open Cloud (SSHOC)
57. Dissco https://www.dissco.eu/ Research infrastructure for natural collections (a commons for specimens and their digital twins)
58. Datacommons.org - primarily statistics for humanitarian work
59. ATT (Finland) initiative has ended. Currently these activities has organised by national open science coordination: https://www.avointiede.fi/en
60. The EOSC Executive Board Working Group (WG) Architecture Task Force (TF) SIRS (https://op.europa.eu/s/oK7d)
61. RDA for EOSC Commons Compendium