The Global Open Research Commons International Model, version 1

A report from the Global Open Research Commons International Model Working Group

*Payne, K., Corrie, B., Crawley, F., Harrower, N., Macneil, R., Maxwell, L., Sansone, S. A., Treloar, A., Woodford, C. J., Åkerström, W. N.

LIST OF TABLES

LIST OF FIGURES

1 Executive summary

In response to the global movement to implement national and cross-national or global commons, a Research Data Alliance (RDA) Interest Group was formed to work towards a community-developed typology for describing research commons. This Interest Group created a Working Group to develop an International Model describing the attributes of Global Open Research Commons. This document supports the release of this RDA Global Open Research Commons (GORC) International Model (IM) v. 1.0, presented as a spreadsheet[1]. This accompanying narrative document provides background information about the initiative, describes its intent and intended audience, the method used to create it, its structure and content. It also provides brief descriptions of communities and activities that have proposed to, or are currently, utilising the model in different contexts, as well as next steps for work in this area. It is important to recognise that the model is aspirational in nature and not prescriptive, drawing on existing good practice and promoting inclusive approaches.

[1] GORC IM WG Commons Model V0.9 (to be updated to V1.0 after RDA Request for comment period): https://docs.google.com/spreadsheets/d/1GLmyczP5Ez32HRK_1DV9H4owlhac8QWdh6SVarKoJKE/edit?usp=sharing

The GORC IM Working Group (WG) consolidated a large range of resources and expert feedback to generate the model, which consists of a number of elements, with associated categories, subcategories, attributes and features, to be considered when undertaking the development of a commons of any kind, at any stage. Although the categories, subcategories, attributes and features are marked as core, desirable or optional, the model does not mandate what should be implemented, or in what way; the decisions on what is relevant, and where resources should be invested will vary depending on the environment and priorities of the implementer. The model is already being used in several contexts that are adapting and testing the model in real world situations. In some cases, the work is being used in the development of commons, while in other cases it is being utilised in other research infrastructure projects.

While the work supports the development of individual commons, it also supports the work necessary to make the commons interoperable. The GORC IM WG outputs provide a firm foundation for the GORC IG as it seeks to create a roadmap for commons integration. They provide a firm, yet flexible, foundation for creating a set of recommendations and a roadmap for building the GORC. The realised vision of GORC will provide frictionless access to all research artefacts including, but not limited to: data, publications, software and compute resources; and metadata, vocabulary, and identification services to everyone, everywhere, at all times. This is the environment that will allow the research community to focus on their enquiries and respond accordingly. It is an audacious goal and we believe that this model will advance our collective efforts in that direction. Interested parties are invited to join the GORC IG[2].

2 Background

2.1 What is a research commons?

Commons are aggregations of resources. A commons in the research context can be defined as a global trusted ecosystem that provides seamless access to high-quality, interoperable research outputs and services and enables data reuse and Open Science more generally[3]. Commons are emerging as an important tool for enabling the reuse of different types of data at the national and global levels. Commons should ideally be developed in keeping with community-developed best practices and underlying values consistent with the FAIR[4], CARE[5], TRUST[6] principles and Open Research more generally.

These commons can focus on all or some of the research artefacts (e.g., data, publications, software and compute resources, vocabularies), as well as the services and tools that generate those artefacts. Many are being funded and built by national, regional, disciplinary and sector actors. Open Research or Data commons provide shared virtual spaces or platforms for access to data and services by researchers and policy makers. Examples include the European Open Science Cloud, the Australian

[2] https://www.rd-alliance.org/groups/global-open-research-commons-ig
[3] https://en.wikipedia.org/wiki/Digital_commons_(economics)
[4] https://www.go-fair.org/fair-principles/
[5] https://www.gida-global.org/care
[6] Lin, D., Crabtree, J., Dillo, I. et al. The TRUST Principles for digital repositories. Sci Data 7, 144 (2020). https://doi.org/10.1038/s41597-020-0486-7

Research Data Commons, the International Virtual Observatory Alliance, and the African Open Science Platform.[7]

2.2 GORC Interest Group: Typology and Definitions

In response to the global movement to implement national and cross-national or global commons, a Research Data Alliance (RDA) Interest Group was formed to work towards a community-developed typology for describing research commons. This Interest Group created a Working Group to develop an International Model describing the attributes of Global Open Research Commons (GORC). This document supports the release of this RDA GORC International Model (IM) v. 1.0. The GORC Interest Group (IG) grew out of a Birds of a Feather meeting held as part of the 11th RDA plenary in Berlin in March 2018[8] and was developed at subsequent related meetings, described in the box on the right. The goal of the IG was to provide a neutral place where people could coordinate the development of a typology to describe what are referred to as "Open Science Commons" or "Data commons".

- At the 11th RDA plenary in Berlin in March 2018 there was a BoF entitled Towards a Global Open Science Commons. This included presentations on the African Open Science Platform, the Australian Research Data Commons, the European Open Science Cloud, the NIH data commons and Canadian activities.
- At International Data Week in Gaborone in November 2018, a SciDataCon session was held on Delivering a Global Open Science Commons. Again a number of presentations were given to profile existing work, after which the group discussion pointed to a number of next steps.
- At the 13th Plenary in Philadelphia in April 2019 there was a BoF entitled Coordinating Global Open Science Commons initiatives. This session focused on group work to advance a typology and identification of areas in which it makes sense to collaborate and coordinate work.
- At the 16th RDA plenary in Costa Rica (virtual) in November 2020 there was a BoF entitled Global Open Research Commons International Benchmarking BoF. This included presentations on the concept and motivation for creating a global commons benchmarking working group (WG) and the review of a draft case statement for the proposed WG. The term "benchmarking" was later removed from the name of the WG.
- August 2020 Draft Case statement submitted internally to GORC IG to review
- Session at RDA 16 Benchmarking BoF session was held at P16 in November 2020. Slides and notes are available.
- January 2021 Case Statement submitted to RDA
- P17 WG submission

As more commons are developed, the need for coordination of these infrastructures on various levels (country, continent, discipline, sector) and focus (for all or some of the research artefacts) is increasing. The Research Data Alliance (RDA) Global Open Research Commons Interest Group (GORC-IG) has developed a set of deliverables to support coordination across commons as organisations work to build the interoperable resources necessary to enable researchers to conduct the cross-commons analyses needed to address societal grand challenges. The GORC's

[7] Coordinating the Global Open Science Commons IG Charter. https://rd-alliance.org/group/coordinating-global-open-science-commons-ig/case-statement/coordinating-global-open-science

[8] https://rd-alliance.org/towards-global-open-science-commons-rda-11th-plenary-bof-meeting

mission is to facilitate access to and reuse of relevant research artefacts including, but not limited to data, publications, software and compute resources; and metadata, vocabulary, and identifier services. The long-term goal is to make commons resources widely available and work towards engaging a diversity of communities and stakeholders to maximise the utility of these commons as an important local and global resource.

As a first step, the GORC-IG generated a definition of a commons, examined a range of existing research commons architectures, and developed a typology of the essential elements in a commons. The typology provides a set of definitions for each of the essential elements of a commons identified by the IG. The typology has arisen from careful discussions within the IG and a process of consultation and refinement at RDA Plenaries over the last 4 years. As this field evolves, the definitions and typology may need to be revisited. This typology was presented to the RDA community as a supporting output in early 2023, revised in July 2023 to respond to community comments, and accepted as a supporting output in August 2023.[9]

2.3 GORC International Model Working Group: Purpose and Intended Audience

The typology and definition were the first deliverables from the GORC IG, and were designed to support the goal of fostering discussion amongst stakeholders. In addition to the typology and providing a forum for conversations about commons, the GORC IG is working towards a roadmap for global alignment and integration of research commons. The GORC International Model WG (GORC IM WG) has been working under the auspices of the GORC IG in support of this roadmap. Specifically, the mission of the GORC IM WG was to "generate a set of pertinent attributes to identify common features across commons" and "review and identify attributes or features currently implemented by a target set of GORC organisations and when possible identify how they measure their user engagement with these features."[10] This model is the realisation of that mission.

Both the GORC IG and IM WG have benefited from the community and support provided by the Research Data Alliance (RDA). RDA has a proven track record of filling a wide range of needs necessary for open and interoperable sharing of research data. Moreover, members of the RDA community have achieved undeniable success in building the social, technical and cross-disciplinary links to enable such sharing on a global scale. Lastly, outputs from RDA groups were a significant source of information for the model presented here. For these reasons, it is an ideal forum to develop and distribute this model.

The first draft case statement for the GORC WG was submitted internally to GORC IG for review in August of 2020. It was followed by a series of case statements that were reviewed by both RDA and the wider research infrastructure community. In total, the WG created 4 case statements, with the

[9] Jones, S., Leggott, M., Lopez Albacete, J., Madalli, D., Pascu, C., Payne, K., Schouppe, M., & Treloar, A. (2023). GORC IG: Typology and Definitions (Version 1.0). Research Data Alliance. https://doi.org/10.15497/RDA00087

[10] IBID.

final version accepted in July of 2021.[11] Since then the members of the WG have been developing the model, which was released on <date>. The method employed by the WG and the model itself are described below.

The GORC IG charter notes that it is intended to function in a similar vein to the RDA funders forum; a space to raise topics of mutual interest, track trends and reach consensus on priorities, and to provide information to inform funding decisions. It is no surprise then, that funders are one of the intended beneficiaries of this model. The model, which can also be thought of as an organisational structure or framework, captures observations of commons elements, attributes, and key performance indicators (KPIs) from real world implementations currently in use or expected in research commons. The model enables funders and other commons' stakeholders to understand and evaluate commons' structure and evolution.

We also anticipate commons developers will find value in the set of curated and validated commons elements and attributes depicted in the model. Commons developers will be able to use the model and the information it provides to guide their development processes, compare and develop plans for improvement within their own infrastructures, and to understand how their commons fits within the global network of commons infrastructures. As commons continue or launch work to develop and deploy services and data, the common typology for describing commons will facilitate the different levels of interoperability needed for global cross-commons analyses.

The model is intended as a guideline rather than a requirement with suggestions for how commons can be better prepared to join or contribute to the vision of the GORC. It is not a prescriptive list and not all elements and attributes will apply to all commons. All items in the model should be given careful consideration by those undertaking the commons development and implementation. Decisions on how to adopt the elements will be commons specific and related to community needs and values, feasibility, and jointly-held cross-commons concerns. The WG created and populated this model to describe commons attributes, but is not endorsing, certifying or otherwise placing a value judgement on existing infrastructures and their features, nor does it intend this to be the only or penultimate model of a commons.

3 Method

The structure of the model is based on the GORC IG Typology and Definitions[12], which defines the essential elements of a commons. The model further refines these essential elements by defining categories and subcategories of the essential elements as well as attributes and features of these entities. The items in the model were identified from a range of sources including:

1. A speaker series and related documentation
    a. 12 speaker series presentations from commons around the globe
    b. Websites and documentation from current commons initiatives

[11] RDA Case Statement GORC International Model WG v.4. https://www.rd-alliance.org/sites/default/files/GORC%20International%20Model%20WG_%20Case%20Statement%20V4%20July%202021.pdf

[12] https://doi.org/10.15497/RDA00087

2. Literature review
    a. Outputs from relevant RDA groups
    b. A thorough literature review of peer reviewed and grey literature, including the GORC IG Typology and Definitions and publications and reports from relevant community groups outside of RDA
3. Community consultation
    a. The wider RDA Community during Plenary Meetings, dedicated workshops, and online asynchronous reviews
    b. WG membership of over 60 individuals globally and in particular discussions amongst task group members charged with refining the model

The method employed by the WG speaks to the core philosophy of this project and of RDA itself. Namely, that we relied heavily on the good work and existing investments that came before the initiation of the GORC IG and WG. The model presented here was derived from significant narrative outputs, feedback from a multi-disciplinary, multi-national team of committed practitioners, and live commons instances across the global research infrastructure community. Doing so allowed us to create products that have inbuilt consensus and reflect current best practices.

3.1 The Speaker Series

Authors of the first WG case statement developed an extensive list of 46 potential commons that we considered reviewing to produce the model typology. Successive iterations of the WG case statement narrowed the scope of those potential commons to 13, of which 12 were featured as speakers. These existing initiatives were foundational to the model developed by the WG. Figure 1 shows the representatives from the speaker series and reflects the international scope of the project.

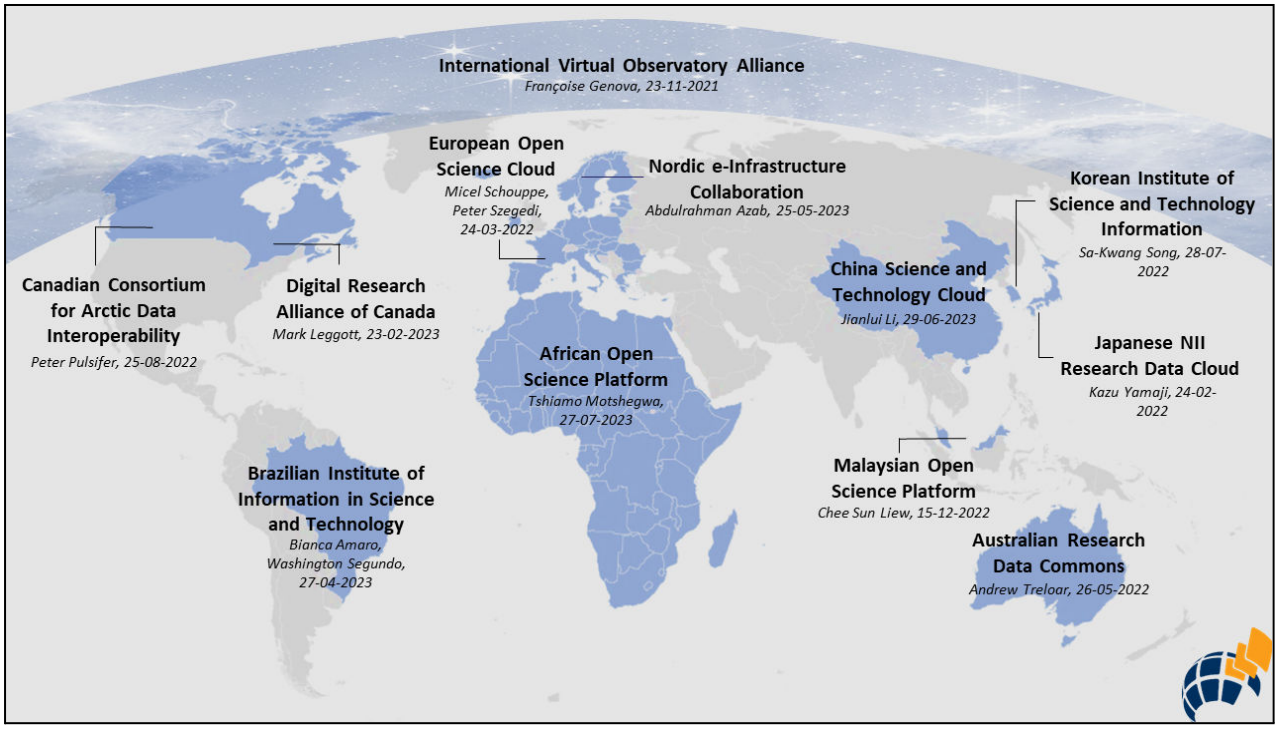

Figure 1: 12 Commons representatives who participated in the WG speaker series between November, 2021 and July, 2023.

In consultation with the RDA community during the 18th RDA Plenary in November 2021, we generated a list of questions that were given to members of the speaker series prior to their presentations. We asked all speakers to respond in writing to these questions which included an open-ended section to capture any relevant concerns for commons that were not covered in the questionnaire. See Appendix A for more information about the questionnaire.

This questionnaire was designed to solicit information about a range of topics such as their perception of whether or not their organisation was a commons, their mission, roadmap and social and organisational constructs and the types of services they provide. This last question was further delineated by asking speakers to mark items from a checklist to identify which, if any, of a set of services they are currently, or planned to provide via their commons. This checklist was based on what was colloquially referred to as "the eInfra standard."[13] The standard is a classification scheme that is designed to represent the entities that are managed within research infrastructure platforms such as services, users, service providers, and key performance indicators (KPIs). The standard is used for building and synchronising electronic service catalogues from research infrastructures. In addition to its use in the EOSC Portal catalogue & Marketplace, the eInfra standard has been used in the OpenAire infrastructure catalogue of services, the European Grid Infrastructure (EGI) Service Catalogue and the Catalogue of Research Infrastructures (CatRIs).

The questionnaire contains two distinct sections: a set of 10 questions with multiple parts, and a table/checklist for eInfra supercategory, category, and subcategory self-identification. Due to the questionnaire being developed before the release of the GORC Typology, a mapping of the 10 questions in the questionnaire to the GORC Typology was created and used to reassess commons that had presented before July 2022 and to assess commons presenting after July 2022, starting with the Korea Institute of Science and Technology Information (KISTI) Korea Research Data Commons (KRDC). The mapping between the questionnaire and the commons essential elements is provided in Appendix A.

Almost all speakers engaged with the eInfra list and identified the services they felt were most relevant to their commons. In the cases where the speakers did not complete the questionnaire, we made a good faith effort to complete it for them based on publicly available information and their presentation and requested their review of the result. Each commons was classified for each attribute or feature as: determined to have, planned to have, or does not have. Publicly available information included the home website for each commons, the home website for their major services if applicable, the websites for their funding agencies and oversight bodies if applicable, and all accessible relevant documentation found in these digital spaces.

The result was an intermediate product referred to as the "einfra matrix," a spreadsheet of the einfra categories and the names of the commons who identified them. The WG makes no claim that all possible relevant material or relevant digital spaces were investigated, only that sufficient sources were found to justify statements found in the model. The einfra matrix is not presented as part of the model, but we anticipate it will be a helpful tool as the IG considers a roadmap for Commons integration globally.

[13] EOSC. (2021). Resource Category, Subcategory (and Supercategory). In EOSC Provider Portal—Resource Profile (v. 3.00). EOSC Portal - A gateway to information and resources in EOSC.

The speakers had the option of creating a slide deck about their commons or just sharing and speaking to their responses in the question document. We recorded all of the speakers and made those recordings, along with any slide decks they provided, publicly available via the RDA website[14]. The written responses to the pre-meeting questionnaire were made available for WG members review.

3.2 Literature Review

A full-time research associate assigned to the WG reviewed the information from the speaker series, and relevant reports and publications in 2 phases. The associate used these sources to create a running list of potential commons attributes and KPIs which was updated by WG members with information that emerged from the speakers series and questionnaire responses and their own work and exploration of commons' related literature. The resulting set of potential attributes were then reviewed by a series of task groups described below. All of the sources for these attributes were retained, and a count of repeated references to potential attributes and KPIs was tracked using a shared spreadsheet.

The running list of attributes was first informed by the examples given in the GORC IG Typology and by initial suggestions given in the WG Case Statement. Attributes observed through the speaker series presentations and investigations were added next. Following this, attributes suggested by other RDA IGs and WGs were reviewed, compared, and added as deemed necessary. Many RDA IG and WG outputs contain recommendations for attributes and features that commons or specific data curators should have or aspire to. These were considered attributes directly or with slight generalisations or modifications. Outputs from the following RDA groups were reviewed and mined for additional attributes (in order of review):

1. Data Discovery Paradigms IG
2. Repository Platforms for Research Data IG
3. FORCE11 FAIRsharing WG
4. RDA FAIRsharing Registry: Connecting data policies, standards and databases WG
5. FAIR Data Maturity Model WG
6. National Data Services IG
7. CURE-FAIR WG
8. FAIR Digital Object Fabric IG
9. FAIR for Research Software WG
10. I-ADOPT WG
11. Research for Metadata Schema WG
12. Data Usage Metrics WG
13. Metadata Principles, from the Metadata IG
14. RDA/TDWG Attribution Metadata Working Group
15. Federated Identity Management IG
16. RDA/CODATA Legal Interoperability IG
17. Data Type Registry WG
18. Data Description Registry Interoperability WG
19. PID Information Types (PIT) WG

[14] https://www.rd-alliance.org/group/gorc-international-model-wg/wiki/gorc-wg-speaker-series

20. DMP Common Standards WG
21. Research Data Collections WG
22. RDA/WDS Publishing Data Workflows WG

Some attributes that were found in the output of one RDA group were occasionally also found in another. In some of these cases the RDA groups that were reviewed first were cited. Only RDA IGs or WGs with completed outputs were reviewed. Active RDA IGs and WGs were consulted to ensure collaboration and to avoid duplicate efforts, but if their outputs postdate this release, they may not be incorporated in this version of the model. We encourage members of these RDA groups to contact us if they believe we have misrepresented their output or have missed crediting their work in the statement of any item in the model.

Research infrastructure documents were reviewed in two phases for potential attributes. The review of resources that occurred between August and October of 2022 is referred to as the "phase 1" review. The October 20, 2022 cut off date was set to bound the narrative review and community input for additional attributes so that the WG could focus on refining and evaluating an initial list. The phase 1 review included 42 sources, which are listed in a Zotero library[15] and in Appendix B.[16] We believe the review of sources was extensive, particularly when coupled with the next phase of the review described below.

Some attributes, features, and KPIs were discussed in more than one source and therefore could be considered as stemming from more than one reference. We created a source crosswalk to identify all of the sources that mentioned each attribute, feature, and KPI, in an attempt to cite our findings correctly and to identify which attributes, features, and KPIs were discussed in the most sources. An overview of this analysis was made available for the phase 1 review[17], and the resulting crosswalk data are in the attributes running list. An overflow document tracking sources and attributes that were encountered after October 20, 2022 was curated for future consideration[18] and comprises the resources reviewed in "phase 2", discussed below.

3.3 Community Consultation

The running list of potential attributes was the starting point for community consultation, which was employed with the goal of refining the set of attributes and features. Consultations were conducted via task groups, RDA meetings, a workshop dedicated to model development and by general request for comments. While the literature reviews undertaken by the research associate were the lion's share of this project in general, the culling and moulding of the potential attributes by the task groups account for a huge amount of the value added by the experts involved in this initiative.

3.3.1 Task Group Methodology

The purpose of the task groups (TGs) was to refine the initial list of attributes, features and KPIs identified during the phase 1 review. A call went out for all interested GORC-WG members to

[15] https://www.zotero.org/groups/2892309/gorc_international_model_wg/library
[16] https://wds-ito.github.io/gorc-wg.github.io//misc/AttributesSources/
[17] https://wds-ito.github.io/gorc-wg.github.io/misc/SourcesOverlap/
[18] https://docs.google.com/document/d/1M4kCEJG8YXyinP2-g9jF40ls8WKXl3WJR0mz-FB3K4I/edit

volunteer for both a first and second choice of TGs. In addition, TG leads were asked to reach out to specific GORC-WG members and solicit their involvement based on their background experience and involvement with the GORC-WG up to that point. TG membership for phase 1 was set by October 27, 2022 and is shown in Appendix C of this document. Each TG adopted a slightly different approach to meeting times, frequency, and synchronous or asynchronous evaluation tactics, remarked on further below.

An evaluation spreadsheet[19] was created for TGs to capture their recommendations, as well as a Google Folder[20] for any meeting rolling notes. The evaluation sheet included an Introduction sheet, a Glossary sheet, a sign-up sheet that was used to create the TGs, and a working sheet for each TG. Each TG sheet contained all of the evaluation items that they needed to review, and columns for recommendations, recommendation notes, importance, and importance notes. Recommendations included what the TG thought the action should be on the item, i.e. Keep, Modify, Merge, Move, Remove, Add. Importance was an internal indicator of criticality for the item to be in a final output, i.e. Essential, Important, Useful. Notes were intended to be used to expand on the recommendation and importance chosen for each item.

There was an uneven distribution of attributes, features and KPIs across the essential elements. For this reason, some TGs evaluated items in multiple essential elements of the GORC typology, while other TGs evaluated items for single essential elements. Figure 2 shows the total number of items (the total number of categories, subcategories, attributes, and features) for each essential element in version 1 of the model, and the assignment of task groups that addressed this unequal distribution. Figure 2 does not include KPIs and metrics. During the phase 1 review KPIs and metrics were assigned and reviewed by the most relevant task groups. During phase 2, a new dedicated TG (TG 6) was created to exclusively review and consolidate them, resulting in 104 KPIs and metrics in version 1 of the model.

[19] https://docs.google.com/spreadsheets/d/1Sav1mYOGGSA3pxi1pvkLXVdkH0vR9Pf7Rfv3tRh8U8E/edit
[20] https://docs.google.com/document/d/1LVt_oh8DXQYWC8ugc_obkzyqKMH-wKSKwOIHUVsnCJ8/edit

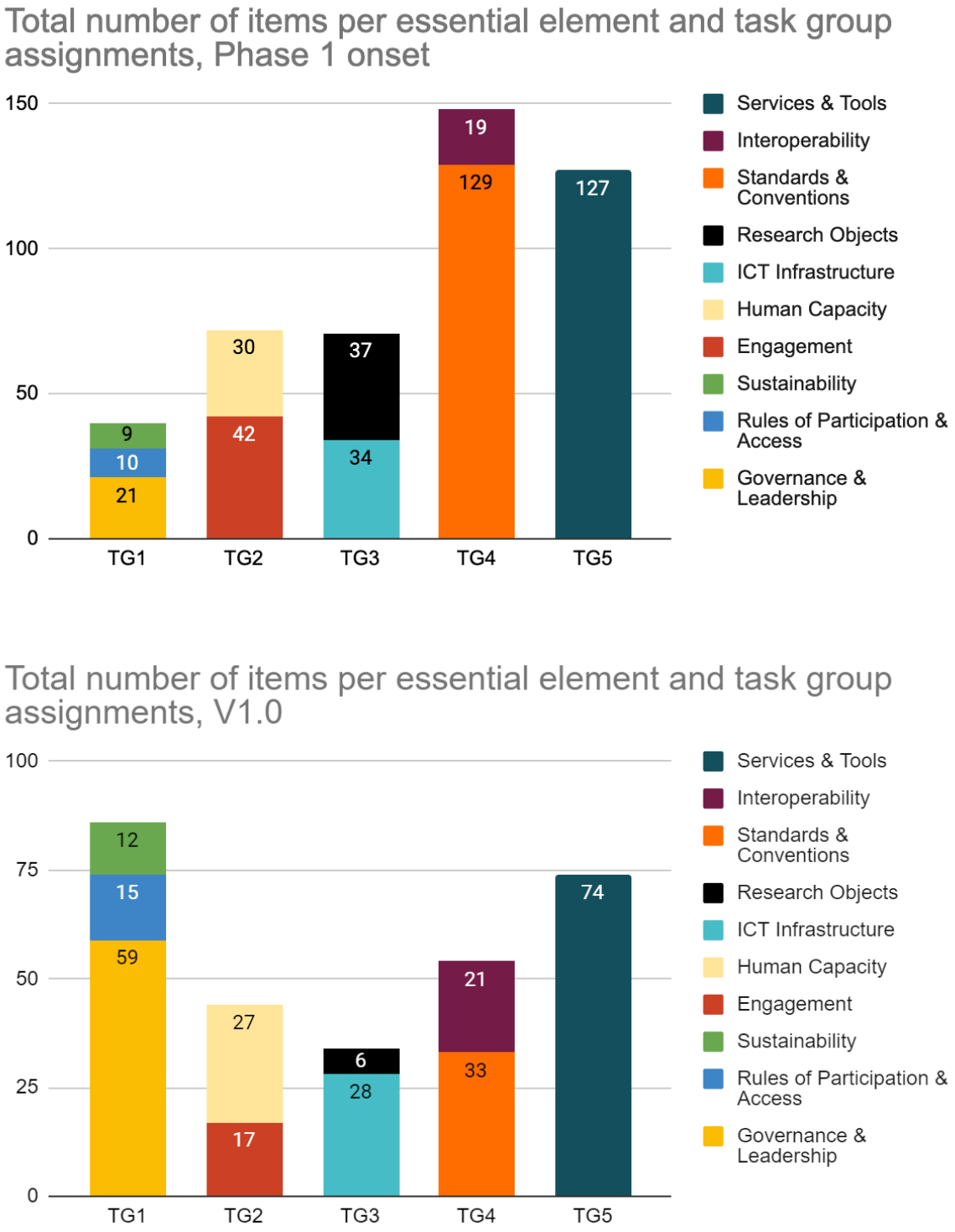

Figure 2: Number of items in the IM and the task group assignments at (top) the Phase 1 onset, or the number of items at the beginning of the first evaluation phase, and (bottom) in V1.0, at present.

The TGs were international and met through online video calls; each TG was charged with creating their own internal communication strategy. For example, TG1, which evaluated items in the essential elements of Governance, Rules of Participation & Access and Sustainability, enacted an asynchronous evaluation process where each member evaluated each of the attributes, features, and KPIs separately before meeting to consolidate their evaluations. In contrast, TG5 members, which

evaluated Services & Tools, met for hourly sessions weekly beginning in early November. During these sessions members discussed each proposed attribute, feature and KPI in turn, clarifying the language and disposition of the attribute until a consensus was reached by all TG members. The slight differences in the implementation of the evaluation across task groups was a reality of working across timezones with colleagues who all hold different responsibilities and were doing this "off the side of their desk." At the end of the day we do not feel this impacted the resulting product, as members of the TGs evaluated all items and made recommendations about them for the model.

TGs worked until March 2023. By and large, the task groups modified and merged items with the occasional addition of items instead of removal. KPIs were the exception, where they were largely either kept as is, modified slightly, or removed instead of merged. As shown in figure 3, before the phase 1 evaluation by the task groups, there were 458 attributes and features, and 125 KPIs for a total of 583 items to be reviewed. After the phase 1 evaluation by the task groups, there were 277 attributes and features, and 66 KPIs for a total of 343 items.

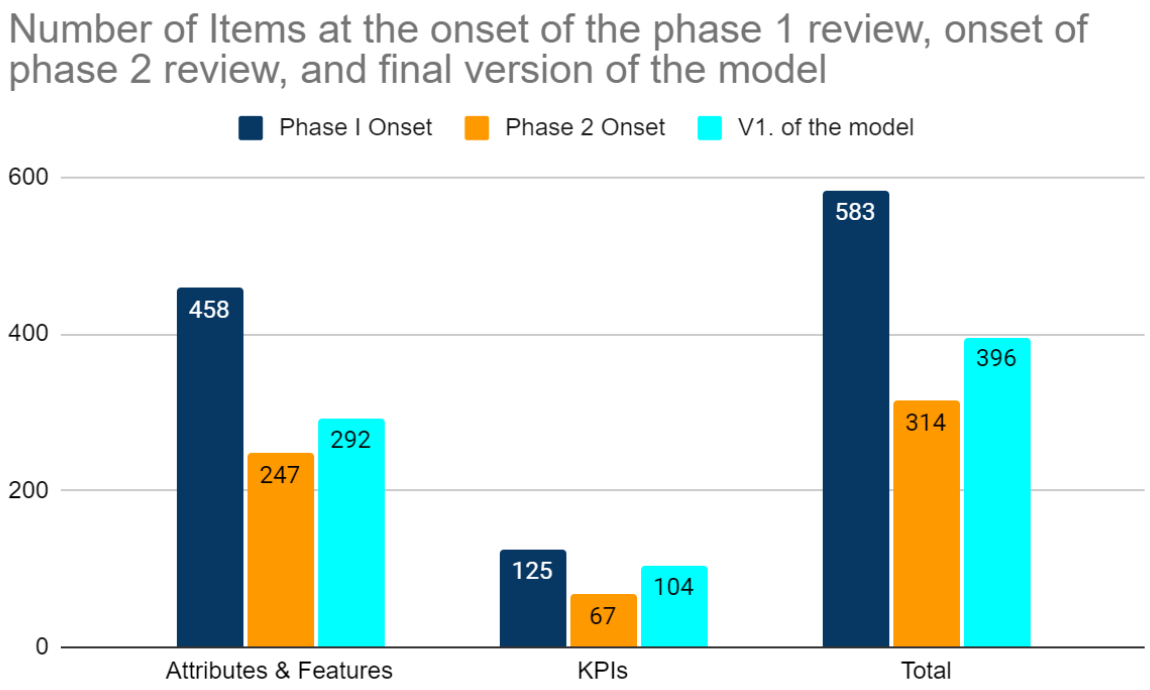

Figure 3: Number of items in the IM at the onset of the phase 1 review, the phase 2 review, and V1.0.

Results of the phase 1 review were showcased during RDA20 in Gothenburg, Sweden[21,22]. With the support of RDA-TIGER[23], the TGs then took part in a full-day in-person workshop[24] co-located with RDA20. The workshop allowed WG members and stakeholders to review each other's work and address working group -wide concerns. The result was version 0.5[25] of the model, which contained

[21] https://www.rd-alliance.org/save-date-global-research-commons-europe-and-beyond
[22] https://www.rd-alliance.org/plenaries/rda-20th-plenary-meeting-gothenburg-hybrid/gorc-international-model-wg-reflections-and
[23] https://www.rd-alliance.org/get-involved/calling-rda-community/rda-tiger
[24] https://forms.gle/Xdx1nHHcJjVGYKgs6
[25] https://docs.google.com/spreadsheets/d/1M6GGO8uPKX-ZYqfg-hBq-2et7QeTr8F3PUrxxyssebI/edit

247 attributes and features, and 67 KPIs for a total of 314 items. Version 1 of the model contains 292 categories, subcategories, attributes and features and 104 KPIs for a total of 396 items.

3.3.2 Final Consultations Phase 2 Review

Version 0.5 of the commons model was shared with the RDA community and interested parties. Prior to the release of version 0.5, as the TGs were engaged in refining the first set of items from the running list identified in the phase 1 review, the dedicated research associate was conducting an additional literature review and collecting new and previously unconsidered attributes, features and KPIs. These items were released to the TGs for review in early June, 2023 and fed into the 'phase 2' evaluation. This phase was necessary in part due to the large number of documents to review and in part because the Speaker Series was ongoing during this time period. A number of the task groups changed their leadership and membership for phase 2, in part to ensure wider input into the process and provide for early cross-checking.

Appendix B lists the narrative sources that were assessed for both the phase 1 and phase 2 review. The number of sources reviewed in phase 2 was far more extensive than those reviewed in phase 1 because of the phase 2 literature review spanning a longer time period. All sources used for the Phase 1 and Phase 2 review are also captured in the GORC-WG Zotero library[26], under the "Key Resource" tag. Sources that do not have any material to refer to (e.g. email threads, meetings that are not recorded, etc.) are referred to as "random encounters". Only those sources tagged as "key resources" in the Zotero library were reviewed and found relevant. The number of sources reviewed is larger than the number that were reviewed and found relevant and each source has an attached note that states whether or not it was reviewed. The resources that were reviewed and found relevant were parsed for items that could be potential inclusion in the model. These items were passed to the TGs for review.

The review of items identified during phase 2 by the TGs resulted in the addition of categories and subcategories to the model. The TGs completed their assessment of the additional items in phase 2 and created version 0.6[27] of the commons model.

3.3.3 Final Consultations

V.0.6 was shared with the RDA community and interested parties on 20 July, 2023. A last round of revisions was conducted by the working group and interested external parties such as representatives from other RDA IGs and WGs before September 13, 2023. The result of these consultations was version 1.0 of the GORC International Model, which was sent to RDA to begin the official community request for comment process in mid September 2023.

[26] https://www.zotero.org/groups/2892309/gorc_international_model_wg/library
[27] https://docs.google.com/spreadsheets/d/1ow2x6alS0SqAK2BaBljrTNhiMXOmQZUt4Smk2WQLeoI/edit?usp=sharing

4 Result

What follows is a narrative summary of the first public release of the model version 1.0, which is presented in detail in the GORC IM spreadsheet[28]. The categories and attributes are grouped by essential elements, as defined by the GORC Typology: Governance & Management Structures, Rules of Participation & Access, Engagement, Human Capacity, Sustainability, Interoperability, Standards & Conventions, Services & Tools, and Research Objects. Categories in this model define sub-elements or classes of the essential elements of a commons, with attributes being the characteristics of essential elements, categories, and subcategories. Features belong to attributes and represent a finer layer of granularity. Attributes and features can be defined for essential elements, categories, and/or subcategories, and are inherited from parent to child in all cases. The model also includes an alphabetical glossary of terms. Every attempt was made to make the definitions more inclusive rather than exclusive.

4.1 Structure

The purpose of the GORC-WG International Model (IM) is to provide a framework and common language to stakeholders around the world who are committed to developing interoperable research services for the public good. The target audience for the model is anyone that is involved in the planning, development, operation, funding or use of a research commons. It is not intended as a prescriptive model, but rather to define and establish a common basis of attributes and features of research commons that users of the model can consider in the context of their own commons as it evolves.

The model is based on the GORC-IG typology outlining the **essential elements** of a commons shown in figure 4. Essential elements are high-level concepts that are essential to the composition of an Open Research Commons.

[28] GORC IM WG Commons Model V0.9 (to be updated to V1.0 after RDA Request for comment period): https://docs.google.com/spreadsheets/d/1GLmyczP5Ez32HRK_1DV9H4owlhac8QWdh6SVarKoJKE/edit?usp=sharing

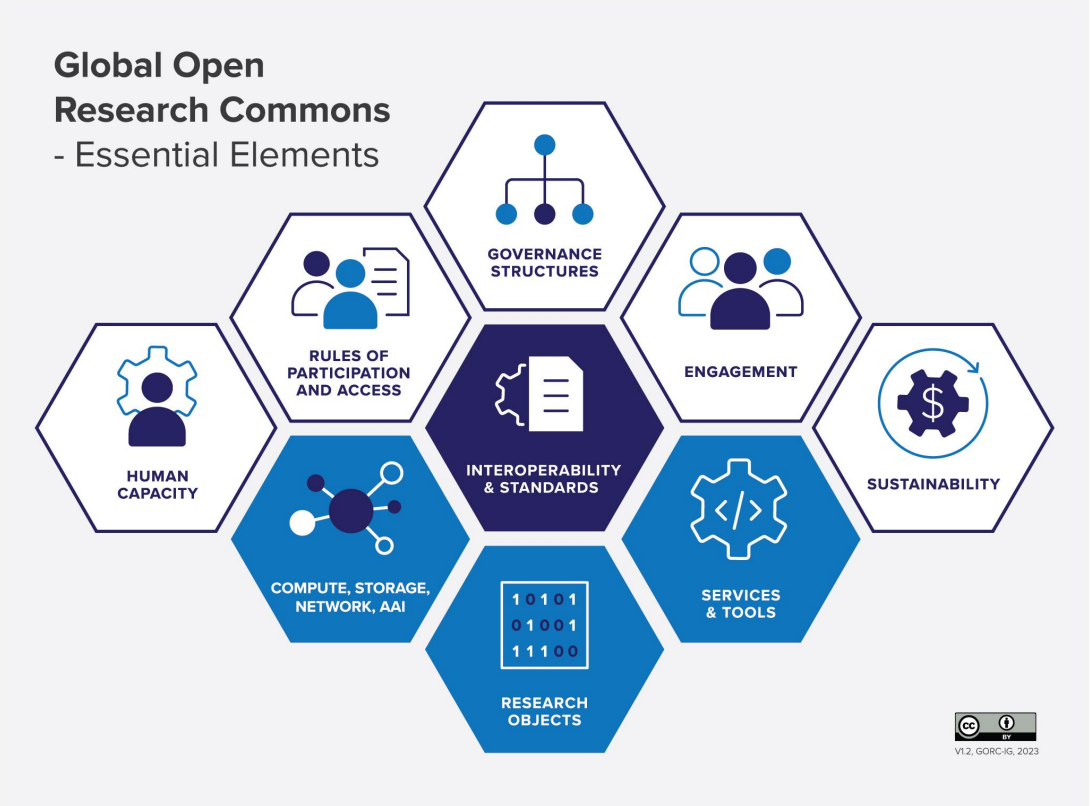

Figure 4: Essential elements of a research commons.

The three elements in blue are the underpinning elements that constitute the parts of the commons with which people interact. The five elements in white are the social/human elements that are needed to make the commons succeed. The central element in dark blue represents the central importance of standards at the core of a commons. The above diagram was used as an organising principle for the model creation. Each essential element in the diagram above is represented as a tab in the model spreadsheet[29], with the exception of *Interoperability and Standards*, which has been split into two separate tabs.

Each row in the model spreadsheet represents a single category, subcategory, attribute or feature that has been identified as something that is worthy of consideration in the planning, development, management, or operation of an Open Research Commons. The model breaks down each essential element into a set of **categories** that provide scope for the broad concepts that each element represents. For example, the Interoperability essential element is broken down into the three categories: *Technical Interoperability, Organisational Interoperability, and Legal Interoperability.* Categories are in turn sometimes broken down into **subcategories** that provide a more granular scoping and definition for the concept being described. For example, *Technical Interoperability* has subcategories of *Syntactic Interoperability, Semantic Interoperability,* and *Other Technical Interoperability.* Not all categories have subcategories, in keeping with the complexity of the concept being considered.

[29] GORC IM WG Commons Model V0.9 (to be updated to V1.0 after RDA Request for comment period): https://docs.google.com/spreadsheets/d/1GLmyczP5Ez32HRK_1DV9H4owlhac8QWdh6SVarKoJKE/edit?usp=sharing

Each category or subcategory has associated with it one or more **attributes.** An attribute is a standard, characteristic, functionality or point of reference about a category or subcategory from which information can be documented, or measurements or comparisons may be made. For example, *Syntactic Interoperability* has attributes that relate to a research commons *planning for interoperability* as well as attributes that consider a research commons implementation of syntactic interoperability in the form of *file and data formats for syntactic interoperability* and *APIs that support syntactic interoperability.* For concepts that are sufficiently complex, attributes can be further subdivided into sets of **features.** Attributes and features can be defined for essential elements, categories, and/or subcategories, and are inherited from parent to child in all cases.

For each row in the model sheet, we have also provided an **extended description**, **examples**, **consideration level (i.e., core, desirable, optional)**, and the **primary sources** that were used to define the category, subcategory, attribute or feature. The model is not intended to be prescriptive or indicate how any one commons should be structured or operate, and so the consideration level is intended as a guideline for prioritising the implementation or refinement of commons' attributes. Items with the *Core* consideration level do not have to be implemented but should be considered by all commons. *Desirable* items may be less critical for all commons to consider and may be more suited to established commons, and *Optional* items may be suited to commons of specific type.

The model also provides a set of **key performance indicators** (KPIs) and **metrics** that can be used to measure uptake, engagement, or use of the essential elements, attributes, and features of a commons and progress in the development of the commons. KPIs can be used to measure performance against indicators of importance to the commons, while metrics can measure a wider range of indicators. This set of KPIs and metrics provide a starting point for commons to consider and create their own set of relevant success indicators and measures.

4.2 Governance & Leadership

Governance is focused on defining that organisation's purpose and the development of the strategies, objectives, values, and policies that frame how that purpose will be pursued by management and internal personnel of the commons. It includes the development of such things as strategic plans including mission statements, values, organisational performance metrics, risk management frameworks; policies and guidelines for financial and operational matters such as commitments to community endorsed principles and frameworks like FAIR and data ethics ; and the creation and maintenance of governance structures, their interactions with stakeholders, and the ways of working with management. Unless they are part of very small teams, there is a boundary between governance and management. Governance is responsible for strategy and direction, while management is the coordination of day to day activities and implementation of policies set by the governing body. Typically the governance processes will be operated via a series of steering groups or boards, involving key stakeholders for the commons such as funders, national services and community representatives. It is likely that the model will be adapted differently for different types of commons, in particular large monolithic commons will have different governance structures than more distributed or federated commons. The model developers also note that the lack of a culture and structure for effective lobbying that drives appropriate regulatory change is challenging.

The FAIR, CARE, and TRUST principles are community-developed, widely implemented principles for data reuse resources. As such they should play a central role in multiple areas of a commons, including governance and have implications for all aspects of the commons.

4.3 Rules of Participation & Access

Rules of participation and access refer to the policies that define the rights, obligations, and accountability for commons' stakeholders. The rules of participation and access define how different stakeholder groups interact with the commons and each other. These rules are designed to ensure that the principles of open research, collaboration, and transparency and other commons-defined values are upheld while promoting responsible and ethical interactions with commons' resources and community members. All commons users and resources are governed by rules of participation and access but the rules and their application may vary by commons and by resource within each commons. In keeping with Elinor Ostrom's framework for the sustainable management of collective resources[30], rules of participation and access should be informed by local values, knowledge, and practices. The rules of participation and access will include policies for access, allocation of resources, privacy, preservation of resources, attribution, and acceptable use. The commons community should develop consequences for research objects not meeting quality standards and for the misuse of commons research objects, services, tools, or infrastructure. This enforcement may be conducted internally, by the commons' community and externally, as needed. The commons should include deterrents or sanctions to promote accountability and prevent the misuse of commons' resources.

4.4 Sustainability

Sustainability includes models and agreements made on how to ensure the viability and operations of the commons. It includes funding and resourcing activities that ensure the commons can be sustained over the long-term. This may include mixed streams of investment and cost recovery through subscriptions, service payment models to ensure operation of the commons, and in-kind contributions in the form of effort/time by contributors to both maintainers of commons infrastructure and interactions with stakeholders. Sustainability should also include defining and developing a strategy for long-term sustainability for all operations and holdings, as well as keeping issues of sustainability in mind when choosing or building commons components. For instance, reuse of existing components is an effective strategy for more sustainable commons infrastructures. Research infrastructure managers and developers have often remarked that brittle policies make it difficult to create a sustainable set of services. In particular, Time-limited funds are used only for time-limited activities, and organisations which define sustainability based on recovering costs can become stagnant.[31] While understandable, in the context of a national commons, funding can be difficult if services are only available to national stakeholders, thus preventing global research. Similarly, new mechanisms should be developed that support multinational funding streams that support international, interoperable services.

[30] https://www.onthecommons.org/magazine/elinor-ostroms-8-principles-managing-commmons/
[31] https://openscholarlyinfrastructure.org/?ref=investinopen.org

Sustainability is approached in the model from three aspects, with each requiring suitable plans, schemes, and implementations to demonstrate sustainability. Sustainability of resourcing and capacity building in the medium and long term includes development of business models, the management of human resources (including retaining accumulated knowledge), and management of any other types of resources required by the commons. Sustainability for Research Objects, Services and Tools includes their stewardship, contextualisation (i.e. metadata, documentation), usability and accessibility over the medium and long term. This requires considerations for ICT infrastructure, human resources and commons operations, as well as the use of transition plans and scalability plans. The third aspect of sustainability is for building and maintaining community trust, which sits more clearly in the domain of the social elements required to sustain the commons, including the social agency to operate and the overall mindset of the commons.

Developers of the model have also noted that infrastructure providers - particularly those in an early phase of development - need to plan for the enterprise to scale and for contingency plans in the event of failure. This will affect their cost and funding structures as well as their use of and ability to migrate between commercial and open services and software. The WG is aware that smaller and newer projects often look towards commercial cloud offerings to help with these issues, but this also has risks of vendor lock-in; extraction may be difficult as the project continues.[32]

4.5 Engagement

Engagement is one of the social/human elements that is needed to make the commons succeed, and it refers to the interaction between the commons and the commons' broadly defined community of stakeholders. Engagement activity should be seen as an iterative cycle, which includes requirements gathering exercises, consultations, usability testing, communications, events and training amongst others. A core attribute of this element is an efficient and effective engagement plan to structure, coordinate and share the relevant level of information (what), to the right target audience (to whom), addressing specific requirement (why), at the right time and with the appropriate frequency (when), and via the most suitable mechanisms (how). An engagement plan is key to ensure the development, maintenance and evolution process is transparent to users and contributors, and that they are aware and involved in the activities, as needed. Also depending on the individual circumstances, the engagement plan should include culturally appropriate materials as well as translations in more than one language. Building community trust, and creating connections to the communities that are being served, is imperative. If that trust is lost, then the mandate for the commons organisation is also lost.

In the model the engagement element is defined by four categories, based on the intended scope/outcome of the engagement process. Community input and feedback, the first category, requires use of methodologies that attract contributors, providing them with credit and incentives, as well as explanation for why a suggestion or requirement was accepted and implemented, or parked, modified, or rejected. Active promotion of and participation in the commons, the second category, requires addressing a multi-level diverse audience, using the means familiar to them: from presentations and training sessions, to ambassadorship programmes, and citizen science events.

---

32 https://a16z.com/2021/05/27/cost-of-cloud-paradox-market-cap-cloud-lifecycle-scale-growth-repatriation-optimization/

Incentivize the participation in and use of the commons, the third category, recommends targeting the intended audience offering consultations and events tailored to provide solutions for their needs, e.g. fund meta-research projects set to contribute to the scholarly ecosystem, technical hackathons to enable co-development. Interoperability with other commons, research institutions, and other potential partners is the fourth category that recommends considering a strategic collaboration and alignments with neighbouring and related efforts, to amplify the impact, or share experience and expertise.

4.6 Human Capacity

The ability of the commons to create a human-friendly environment for all stakeholders and community members in all aspects, specifically for users, providers, and intermediaries, so that the commons can set and achieve objectives, perform functions, solve problems, and continue to develop the means and conditions required to enable this process (adapted from https://www.fao.org/3/y5613e/y5613e08.htm).

The human capacity of a commons should be viewed in the context of a community which includes all human individuals and entities that could be considered stakeholders, users, providers, members of the commons, and intermediaries (i.e., those who do not interact directly with the commons, but use information about or provided by the commons, e.g., policy makers, journals, funders) in the past, present, and future as well as all research communities that the research Commons is a part of in a regional, national, and global context. Depending on the structure of the Commons, stakeholders may include funding and government bodies as well as related commercial entities; these may be in scope through the provision of related services or an association with related research communities.

The model divides human capacity into five main categories:

Internal Capacity includes human resources required to provide services and to plan for growth required by future services, succession and labour turnover planning, and mechanisms for enacting EDII[33] commitments to personnel and the commons. As part of the support for internal capacity, personnel processes need to be documented and shared for a high level of transparency, with regular reviews of working conditions and requirements in place.

The model articulates as an additional category the skills required for planning, managing and assessing service delivery. This includes the development of use cases for services and platforms, the prioritisation of identified needs and derived requirements of users, assessing plans and deployments, and the implementation of continuous improvement mechanisms.

Skill requirements for the commons community focuses on the capacity of individuals interacting with the commons, which includes documentation to facilitate that use and lower the skill requirement to ensure that the commons is an effective choice for its users.

Ease of use for the commons community focuses on the capacity enabled by the commons, including considerations for effective user-centred design employed to make the commons easier to use.

[33] EDII stands for Equity, Diversity, Inclusion, and Indigenization. Similar representations may be EDIA (Equity, Diversity, Inclusion, and Accessibility) and DEI.

Training and education, the fifth category, requires design (structure, content, target audiences and levels), development, and delivery of the material, including mechanism (e.g., modules, summer schools) and means (e.g., online, in person), as well as an evaluation and assessment phase also to measure impact. A specific subcategory of training that commons may consider is an ongoing training program for internal personnel, with the goal of increasing internal capacity.

Finally, the model developers note that increasing internal and external documentation in all of these areas is a mechanism for increasing the sustainability of a commons.

4.7 ICT Infrastructure

By "ICT infrastructure" we mean the hardware and base software components that a computer system requires to function and are necessary to conduct research. These need to be designed to scale with increasing volume, complexity and velocity of projects and expectations. To aid this, a review and update of ICT infrastructure should be scheduled to happen on a regular basis. For all of the infrastructure categories listed below, a knowledge of, and ability to manage, the infrastructure is required by commons operators.

A number of infrastructure categories were identified through the review. Network infrastructure encompasses both the internal network infrastructure (for passing messages within the commons) and external network infrastructure (to facilitate connections to external services and other commons). Compute infrastructure encompasses both the base computing infrastructure (the essential hardware components required for stable and robust minimum viable operation of the commons) and add-on computing infrastructure (components that enable advanced or specialised operations, such as GPUs for ML computing). Compute may be delivered via on-premise hardware or off-premise cloud services. Storage infrastructure encompasses both direct-attached storage in the data centre and network-based storage in the cloud. Base software infrastructure includes the underlying OS elements upon which different applications depend. Authentication and authorization infrastructure refers to services and procedures that enable members of different institutions to access protected information that is distributed on different servers. This includes both base AAI infrastructure and add-on infrastructure.

4.8 Interoperability

Interoperability is the ability of data or tools from non-cooperating resources to integrate or work together with minimal effort[34]and is arguably the most difficult part of implementing the FAIR principles. It can also enable a wider range of cross-commons use-cases. Interoperability enables cross-commons reuse of data and is of central importance to the Commons. Types of interoperability include: Technical interoperability (how artefacts are exchanged), Syntactic interoperability (how to structure information), Semantic interoperability (data are interpreted the same way) and Pragmatic interoperability (agreements between organisations)[35].

[34] Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. (2016).The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 3, 160018. https://doi.org/10.1038/sdata.2016.18

[35] Janssen, M., Estevez, E., & Janowski, T. (2014). Interoperability in Big, Open, and Linked Data--Organizational Maturity, Capabilities, and Data Portfolios. Computer 47,10, pp. 44-49.https://doi.org/10.1109/MC.2014.290

Interoperability in many ways is at the heart of the work of the GORC IG and the WG. The IG is working to coordinate infrastructures" as they work to build interoperable resources necessary to enable researchers to address societal grand challenges". The model divides the issues into technical, organisational and legal interoperability.

Technical interoperability in turn has two main foci: syntactic interoperability and semantic interoperability. Syntactic interoperability means that plans and mechanisms exist to create and maintain Interoperability and compatibility at the syntactic level over time, the commons uses file/data formats that support Syntactic Interoperability, and provides APIs that support Syntactic Interoperability. Semantic interoperability means that plans and mechanisms exist to create and maintain Interoperability and compatibility at the semantic level over time, and that metadata, data, and other Research Objects use standardised community-endorsed vocabularies, and FAIR-compliant community-endorsed vocabularies where possible. In addition to these two main dimensions, commons may offer other Other Technical Interoperability plans, infrastructure, and mechanisms, such as following API search standards, and having a security framework that is shared between Services and Tools, from backend to frontend.

Because commons infrastructure is operated and managed by organisations, it is important to also consider organisational interoperability. This includes ensuring that plans and mechanisms exist to create and maintain organisational Interoperability and compatibility over time. For non-domain specific commons, it also requires ensuring that domain-specific needs are addressed and considered so that the commons remains interoperable with other domain-specific commons and services over time.

The whole point of a commons is to provide access to research objects for use/re-use. This requires attention to a cluster of legal interoperability issues. Access to and reuse of Research Objects should be open and unrestricted as a default rule, or otherwise granted with the fewest limitations possible; ideally a licence for reuse is required for all Research Objects and Tools in the commons. Licences used should be enumerated and harmonised to allow seamless exchange between actors within the commons and outside of the commons. The entities with rights to Research Objects should be specified appropriately via licences and Research Object documentation and identified before dissemination, to ensure no surprises. As an increasing amount of access to commons will be via software, it is important that legal aspects are encoded in a Machine Actionable format that enables automated provision of services and data.

4.9 Standards & Conventions

A standard is a repeatable, harmonised, agreed and documented way of doing something[36]. Standards can be either *de jure* or *de facto*. *De jure* standards, or standards according to law, are endorsed by a formal standards organisation, such as the ISO. The organisation ratifies each standard through its official procedures and gives the standard its stamp of approval.

De facto standards, or conventions, are adopted widely by a community. These standards arise when they become part of the accepted way of doing things within a community. *De facto* standards can become *de jure* standards if they are submitted to a formal standards organisation and approved.

[36] https://inspire.irena.org/Pages/standards/whatarestandards.aspx

Within the data commons context, standards and conventions may cover various things:

- Community supported Research Object standards and conventions
- Community supported Semantic Object standards and conventions
- Applications, Software, Services & Tools standards and convention
- Quality standards, conventions, and/or guidelines
- Standards and conventions for adding and maintaining PIDs for managed assets
- Authentication and Authorization protocols
- Standards and conventions for the commons catalogue of digital objects
- Standards and conventions supporting and describing mechanisms, infrastructure and plans for specific workflows, use cases, and types of interexchange within the commons
- Standards and conventions for regulatory and ethical compliance.

4.10 Services & Tools

Service (as defined by IVOA[37]) is any Commons element that can be invoked by the user to perform some action on their behalf. Services are usually intended for use by machines, and mostly invoked by software. Tools enable researchers to perform one or more operations, typically on data, and often with data as the output. Tools are usually intended for use by humans. In this context we are explicitly excluding physical instruments as tools. Services and tools overlap with users who create processes.

The following categories of services and tools were identified:

1. Research object repositories
2. Discovery service
3. Services and tools for direct research tasks
4. Services and tools that enable workflows and middleware
5. Persistent identifier services
6. Vocabulary and semantic object services
1. Data management services and tools
7. Commons catalogue of all services and tools
8. Security and identification services
9. Helpdesk service

These categories include tools and services used by researchers in their primary research, e.g services and tools for direct research tasks; tools and services used to support or connect tools and services, e.g. tools and services that enable workflows and middleware; tools and services focussed on research data management, e.g. research object repositories and dedicated research data management services and tools; tools and services focussed on discovery; persistent identifier services; vocabulary and semantic object services; and miscellaneous other tools.

The context for these definitions is the emerging and complex intersection of tools, disciplines, services, platforms, hardware, resources, and the people (users, researchers, developers, stakeholders, personnel and communities, etc.) who use and contribute to them.

[37] https://ivoa.net/documents/WD/ResMetadata/RSM-20021011.html

As research infrastructure, services and tools are often made available through research platforms (variously referred to as virtual science labs, virtual research environments (VREs), or Science Gateways,) that are deployed to support both the research workflows and the communities of practice engaged in collaborative research. Typically, a research platform's capabilities include data acquisition and management, processing and visualisation, storage and preservation, sharing and discovery; platforms may provide the full spectrum or a subset of components. Science Gateways may be discipline-specific, and may support and enhance scientific collaboration and scholarly communication by facilitating citizen science engagement as well.Processes, services, and tools all overlap with each other in ways that complicate the discussion. Some particular issues to be considered are these:

- Services can be both internal and external, where some of the layers are invisible to users but identifiable as services that sit between ICT and user-oriented services
- Commons need to have a range of services that reflect and support the processes in the other essential elements
- Every commons shouldn't need to do everything; over time there should arise a range of global services offered to everyone that are needed by every commons. (e.g. citation information could come from cross ref or DataCite)
- Globally the community should be targeting the services and tools that make a commons attractive. It is also the case that different commons will have different histories, funding sources, and business drivers, and that this may lead to parallel service offerings. Looking at specific attributes and services and evaluating them individually might not aggregate up to the big picture.
- Given that the ever changing and fast developing landscape/ecosystem of services and tools used in research is extraordinarily rich, varied and dynamic, the categorisation described below is best understood as a snapshot which attempts to capture the current situation, and will evolve over time with significant changes. Moreover, full blown categorisation of tools and services was not the remit of the GORC WG. This is an expected outcome of the OfR-RDA Mapping the landscape of digital research tools WG[38], and will likely be incorporated into future versions of the model

Specific tools and services may overlap more than one subcategory and categories, especially if it is a larger system with multiple functionalities. Categories and their subcategories are not disjoint.

4.11 Research Objects

Research Objects are the outputs of the research process, but can also be inputs to later processes. Like ICT Infrastructure and Services & Tools, they are the underpinning digital elements that people interact with in the commons. The scope here is limited to digital research objects, and research hardware itself is out of scope; the model developers understand that this is a challenging perspective, since research cannot occur without specially designed hardware. However, it was necessary to contain the scope of the model and the project more broadly. The digital outputs *from* hardware are of course in scope (e.g. calibration data are research data, user manuals are publications/documentation, software endpoints are research software).

[38] https://www.rd-alliance.org/groups/rda-ofr-mapping-landscape-digital-research-tools-wg

The research object approach is primarily motivated by a desire to improve reproducibility of scientific investigations. Central to the proposal is a need to share research artefacts commonly distributed across specialist repositories on the Web including publications, lab notebooks, blog entries, supporting data, software executables, source code, presentation slides, and presentation videos.

The model identifies five main categories of research objects:

- Publications and Research Documentation include any digital, textual, visual, audio, or tactile representations that describe or discuss any aspect of the research project and activities in human or machine readable formats.
- Research Data are a collection of data that is identifiable and has the potential to be curated or published by a single actor and is the result or focus of research activities. Research data can digitally represent a group of observations, a data product from a specific version of a processing algorithm based on observations, output of numerical model(s), or outcomes of laboratory experiments.
- Research Software includes any software component created during the research process or for a research purpose that is implementable or executable by a computer or machine (actionable research documentation that describes protocols, workflows, algorithms etc. is included in the above category of Research Documentation)
- Semantic Objects are a named grouping of descriptive elements that sufficiently describe a distinct identity. Semantic Objects may be in the form of documentation, Research Software, or research data. They may also be referred to as Semantic Artefacts, defined as groups of entities with unique identifiers where entities include subjects, predicates, and objects that can be linked together to form a network that describes a dataset. In this model, Semantic Objects/Artefacts are not considered as collections, since collections are comprised of research objects and not entities.
- Collections are a combination or bundle of research objects that are of the same kind or different kinds that share a relationship and are treated as a digital object which bears a PID. Collections consist of a finite number of digital object identifiers and metadata associated with each referenced identifier.

4.12 KPIs & Metrics

KPIs are qualitative or quantitative measures that the commons, commons' stakeholders, and community of commons can use to measure the uptake, engagement, or use of commons attributes and features. In contrast, metrics are quantitative measures used to assess the evolution or performance of specific processes. While the model includes suggested KPIs and metrics for every essential element, most KPIs and metrics relate to engagement with the commons and to human capacity. Broad themes of KPIs and metrics are apparent and provide some organisation for the KPIs and metrics, namely: commons governance and policy, commons engagement with stakeholders, feedback and satisfaction of stakeholders, commons infrastructure and technology, and Stakeholder Engagement with technology and infrastructure. KPIs and metrics use data generated from the commons (e.g., number of registered users), data generated outside of the commons (e.g., number of publications produced using commons-derived datasets), and data generated by other commons (e.g., the number of commons that implement a given standard to facilitate interoperability across

related data in separate commons). Most KPIs and metrics are quantitative measures that could be externally verified. In a few cases, the model suggests qualitative measures, as in adherence to commons' policies. While some quantitative measures can be uniformly applied across commons (e.g., number of registered users), others, like measures of research impact, will vary in their application across commons. As evidenced in the speaker series, KPIs and metrics can be internal to the commons or openly available. It should be noted that the KPIs and metrics are not nearly as comprehensive as the IM itself and should be considered as observations of success and engagement metrics that we observed in the existing commons. These metrics could be developed further as part of future work, and currently provide a starting point for commons to consider what their own set of KPIs and metrics may be.

5 Conclusion

The purpose of this report was to describe the motivations, goals, methods and outputs of the GORC IM WG, in particular the GORC International Model (IM) V.1.0. While the release of the GORC IM fulfils the goal set out in the WG case statement, there is clearly more work to be done in this sector. This concluding section describes where this work leads next.

The GORC IM lists a number of elements to be considered when undertaking the development of a commons of any kind, at any stage. Although the attributes and features are marked as core, desirable or optional, the model does not mandate what should be implemented, or in what way; the decisions on what is relevant, and where resources should be invested will vary depending on the environment and priorities of the implementer. Our hope is that a large global survey of the current state of the art in commons initiatives, thoughtfully organised, vetted by international experts, and well presented, will provide actionable information for organisations as they make their decisions about what and how to develop their infrastructure in light of the evolution of the commons community. Our work is also intended to allow policy makers to respond to the specific needs of different stakeholders, while facilitating alignment with other domain, national, and international commons initiatives.

In this spirit, we are encouraged by the immediate uptake of this work in several contexts that are adapting and testing the model in real world situations. In some cases, the work is being used in the development of commons, while in other cases it is being utilised in other research infrastructure projects. Specifically, within the RDA context, a new working group supported by Oracle for Research (RDA-OfR)[39,40], Mapping the Landscape of Digital Tools, is utilising our two internal products commons (an analysis matrix and the commons type dimensions[41]) as a reference for the development of an online map of the digital research tool landscape. In this case, it is contributing to the development of a registry, instead of a commons. The IM is also being reviewed as part of a proposal by CODATA to create a WG that will generate a machine-actionable representation of the GORC IM that can be used to describe commons evolution in the health and other domain spaces.[42] The machine-actionable representation will be used to generate a knowledge graph that enables stakeholders, including

[39] https://www.rd-alliance.org/new-rda-working-group-supported-oracle-research
[40] https://www.rd-alliance.org/groups/rda-ofr-mapping-landscape-digital-research-tools-wg
[41] https://wds-ito.github.io/gorc-wg.github.io//python/TypesOfCommons/
[42] https://doi.org/10.12688/openreseurope.15982.1

funders, to measure commons' evolution and identify strategies to maximise opportunities for cross-commons synergies knowledge graph), as well as conducting a discipline specific analysis of the model within the health commons space.

In contrast, the model is also being used in support of commons development. In the development of a domain commons the UKRI-funded BioFAIR initiative[43] in the UK is using the IM to help them organise their work plan for the BioCommons for methods, data and people, while in the national context, the Australian Research Data Commons is adapting the model for local circumstances, in their development of three different thematic national commons.[44]

The GORC IM WG was designed and developed under the auspices of the GORC IG, and it is now appropriate to turn back and remark on the ongoing initiatives of the IG. First and foremost, we understand that the GORC IM V.1.0 is a snapshot in time and that in the future iterations it will need to be revised and updated. We encourage representatives of commons and allied institutions to review our analysis and investigation, and to contact the IG with comments, corrections or updates. For example, future iterations could include reviews of additional commons identified in earlier versions of the WG case statements, or ideas about how to manage stand alone services that are valuable but not yet part of a commons initiative. Future versions should also include the expanding range of open science initiatives promoted by the United States administration and implemented across federal agencies.[45]

The model itself could also be revamped; while it is presented here as a spreadsheet and accompanying narrative document, there have been ongoing discussions about the feasibility and utility of presenting the model in alternate formats such as an RDF triplestore, a geographical map or a mural[46]. The IG may decide to commission new WGs to address updates and create new products.

While the work supports the development of individual commons, it also supports the work necessary to make the commons interoperable. The GORC IM WG outputs provide a firm foundation for the GORC IG as it seeks to create a roadmap for commons integration. It provides a firm, yet flexible, foundation for creating a set of recommendations and a roadmap for building the GORC. It is also worth noting that the WG in general, and the speaker series in particular, was a rich source of information and inspiration about different organising principles and infrastructure evolutions in various commons. While the model is more generalised than the source information, we believe that these real world stories of commons development at a more detailed level will be helpful in the roadmap development. This roadmap could be built in concert with a group such as the Open Science Commons Executives' Roundtable (OSCER).

The unprecedented challenges faced by the world today demand an equally unprecedented response from the scientific and broader research community. In order for this response to occur, especially at the accelerated pace necessitated by the multiple global crises facing humanity, our scientists will need to have access to the next generation of scientific platforms. Rather than think of them as

---

[43] https://biofair.uk/

[44] https://docs.google.com/document/d/1klhB4uMaEViJxgKYg_hcab9Ce0pm6fWbD58NoGjaqGY/edit#heading=h.lsi0b3gdler1

[45] https://www.whitehouse.gov/ostp/

[46] https://app.mural.co/t/wdsitoworkspace3502/m/wdsitoworkspace3502/1690201881742/4fd99ef0453073af14948edea1e0997bc1285008?sender=5f159514-8453-4b08-b35f-3e1359370f2e

singular ‘platforms’ this work envisions an interoperable set of platforms that build on both the advances of the internet and the consensus and strengths of the research community. The realised vision of GORC will provide frictionless access to all research artefacts including, but not limited to: data, publications, software and compute resources; and metadata, vocabulary, and identification services to everyone, everywhere, at all times. This is the environment that will allow the research community to focus on their enquiries and respond accordingly. It is an audacious goal and we believe that this model will advance our collective efforts in that direction. Interested parties are invited to join the GORC IG[47].

Acknowledgements

The authors gratefully acknowledge the contributions of all members of the GORC IG and WG, and individuals who participated in workshops and provided feedback on products generated over the last 2 years. We recognize that these members volunteer their time and expertise on top of their other obligations, including their full time positions in key organisations that are building the future of global science; we remain inspired by their commitment. We also appreciate the contributions and support from RDA, especially in areas of communications and event planning. We reserve a special thank you to members of the speaker series who provided insights into their work and their communities. We also gratefully recognize the enormous amount of previous work in open science and research infrastructure that the model builds on. This work would not have been possible without the funding provided by The Digital Research Alliance of Canada, who provided resources to hire Dr C.J. Woodford as a full time research associate to spearhead this work, and without whom this would not have been possible.

[47] https://www.rd-alliance.org/groups/global-open-research-commons-ig

Appendix A: Mapping of Speaker Questions to GORC-IG Typology

Essential Elements	Questions (mapped to essential elements)
Governance & Leadership	8. b) Different takes/models/approaches to governance by different commons - could you describe the governance approach taken by your commons?
Rules of Participation & Access	3. a) How are you managing integration? Both in terms of (i) adding new services to your commons - how do standalone services become part of a commons? and (ii) in terms of integrating between services - Do you have plans for interoperable workflows between components of your commons?
	3. b) Are there any scoping boundaries in terms of data providers or data user participation?
	8. a) Are there "Rules of Participation & Access" for the commons, is it implicit or explicit?
	8. f) How do you promote/ensure adoption of standards and services by users (service providers, end users)?
Engagement	5. How do you define success for your commons?
	5. a) Do you have a set of benchmarks or milestones you are aiming for?
	5. b) Are you tracking any KPI or engagement metrics in your Commons?
	6. a) If you are a domain specific commons, are you aware of any other domains that have been inspired by or have adopted your standards? If you are geographically bounded, are you aware of any other geographic regions that have been inspired by or have adopted your standards or consume your services?
	8. d) How is social trust obtained within the commons? (trust in contributions, trust in brokers, trust in aggregators, trust in consumers of data, and data infrastructures)?
	8. e) What are the organisational norms, community norms etc. that incentivise or dis-incentivise participating in the commons? Is this actively managed or emergent?
	8. f) How do you promote/ensure adoption of standards and services by users (service providers, end users)?
	4. Which eInfra tags do you see applying to your commons example?
Human Capacity	8. c) what are the arrangements for partnerships between providers-users, provider-provider?
Sustainability	5. How do you define success for your commons?

	9. Do you have a roadmap or strategic plan for the future of your commons?
	9. a) What are your immediate goals? Long term goals?
Interoperability	3. a) How are you managing integration? Both in terms of (i) adding new services to your commons - how do standalone services become part of a commons? and (ii) in terms of integrating between services - Do you have plans for interoperable workflows between components of your commons?
	6. Are there any interfaces between your commons and other commons out in the world, or conversations at your commons about connecting with other groups or resources? - things like setting standards, establishing/operating a catalogue, providing/organising compliance activities, setting strategy for the commons, outreach/promotion of commons, etc.
Standards & Conventions	4 a) What standards are supported by your commons?
	6. b) Does your commons rely on services provided by organisations outside of your commons cohort? For example do you rely on PID systems like ORCID or DOIs? Exterior vocabulary services? Do you use OpenAire services to connect publications to date?
	7. How are you addressing cybersecurity? Is the responsibility for security centralised or distributed throughout your partners? How is it governed?
	7. a) Do you have a specific security framework or standard you require, have implemented, or plan to implement?
	4. Which eInfra tags do you see applying to your commons example?
ICT infrastructure	3. Is your research commons an example of an individual service or platform, or is it an aggregation of multiple disparate organisations/commons?
	4. Which eInfra tags do you see applying to your commons example?
Services & tools	3. Is your research commons an example of an individual service or platform, or is it an aggregation of multiple disparate organisations/commons?
	3. c) Do you have a catalogue of services? Do you have a map of the data and services ecosystem for your commons?
	4. Which eInfra tags do you see applying to your commons example?
Research Objects	3. c) Do you have a catalogue of services? Do you have a map of the data and services ecosystem for your commons?
	4. Which eInfra tags do you see applying to your commons example?

Table 1 : A mapping of questions asked of each speaker series participant to the GORC-IG Typology of essential elements. The questions asked to each speaker were created before the release of the GORC-IG Typology of essential elements.

Appendix B: Sources Reviewed

Phase 1 Review

NB: This list does not include the outputs from RDA groups that were also reviewed and listed in the body of this report.

1. the US National Institute of Standards and Technology Research Data Framework (NIST RDaF)[48],
2. the Big Data Maturity Matrix from the Big Data UN Global Working Group[49],
3. the FAIR Maturity Indicators created by the FAIR Metrics Group[50],
4. the International Image Interoperability Framework[51], the OpenAire Research Graph[52],
5. the OpenAire Monitor Indicators[53],
6. the re3data Quality Assurance Survey[54],
7. the XSEDE final report on operations[55],
8. the Scientific Information Policy SPD-41 from the Science Mission Directorate in the United States[56], and select papers and journal policies[57] were reviewed for any attributes, features, and KPIs that were not already captured. Lastly, WG member observations in various meetings, workshops, and webinars were also summarised and included in the potential list of attributes.

Phase 2 Review

1. Speaker series occurring after October 2022: MOSP, Digital Research Alliance of Canada, IBICT, NeiC, CSTCloud and the AOSP
2. New content released by speaker series members, such as ARDC's data management framework[58]
3. Supplemental materials to phase 1 review items, specifically: Data Quality Assurance at Research Data Repositories[59], Introducing the FAIR Principles for research software[60], Research Metadata Schema WG Output 2[61], updates from the Global Open Science,

---

[48] https://www.nist.gov/programs-projects/research-data-framework-rdaf, see https://www.nist.gov/system/files/documents/2021/05/12/RDaF%20Preliminary%20Framework%20Core%2005-12-21_0.pdf for the listing of specific categories and subcategories

[49] https://unstats.un.org/bigdata/task-teams/training/Big%20Data%20Maturity%20Matrix%20v1.0.pdf

[50] https://fairdata.services:7171/FAIR_Evaluator/metrics and https://github.com/FAIRMetrics/Metrics

[51] https://iiif.io/

[52] https://graph.openaire.eu/

[53] https://monitor.openaire.eu/ and https://monitor.openaire.eu/indicators/themes

[54] https://zenodo.org/record/6457849#.Yz2nn3bMLIU

[55] https://www.ideals.illinois.edu/items/124108

[56] https://science.nasa.gov/science-red/s3fs-public/atoms/files/Scientific%20Information%20policy%20SPD-41.pdf

[57] 10 Simple Rules for Improving Research Data Discoverability (https://doi.org/10.1371/journal.pcbi.1009768), Supplemental Information to the NIH Policy for Data Management and Sharing: Selecting a Repository for Data Resulting from NIH-Supported Research (https://grants.nih.gov/grants/guide/notice-files/NOT-OD-21-016.html)

[58] https://ardc.edu.au/resource/research-data-management-framework-for-institutions/

[59] https://datascience.codata.org/article/10.5334/dsj-2022-018/

[60] https://www.nature.com/articles/s41597-022-01710-x

[61] https://docs.google.com/document/d/1VTA-rFhSn6dceNZpp0Nq3mRRUoAjYMPaqQtEW-ls9Qs/edit

Cloud[62,63], version 1.5 of the NIST Research Data Framework[64]

4. A Brief Overview of Open Data Initiatives in South America[65]
5. Beyond Accuracy: What Data Quality Means to Data Consumers: Journal of Management Information Systems[66]
6. Global Community Guidelines for Documenting, Sharing, and Reusing Quality Information of Individual Digital Datasets[67]
7. Quality of Open Research Data: Values, Convergences and Governance [68]
8. The landscape of open science in Malaysia, 2022[69]
9. Science and Society, OSTP memo[70]
10. An Update on Research Security: Streamlining Disclosure Standards to Enhance Clarity, Transparency, and Equity | OSTP[71]
11. Playing catch-up in building an open research commons[72]
12. SQAaS[73]
13. A set of common software quality assurance baseline criteria for research projects[74]
14. A Set of Common Service Quality Assurance Baseline Criteria for Research Projects[75]
15. Research Software Current State Assessment[76]
16. Six Recommendations for implementation of FAIR practice by the FAIR in practice task force of the European open science cloud FAIR working group[77]
17. A Persistent Identifier (PID) policy for the European Open Science Cloud (EOSC)[78]
18. Recommendations on FAIR metrics for EOSC[79]
19. FAIRsFAIR Data Object Assessment Metrics v0.5[80]
20. EOSC interoperability framework[81]
21. EOSC Authentication and Authorization Infrastructure (AAI)[82]
22. Turning FAIR into reality[83]

---

[62] https://codata.org/initiatives/decadal-programme2/global-open-science-cloud/
[63] https://docs.google.com/document/d/1g6jw03gPRXvuswmjJnQ5nI1c9bfXQm1_kjb1Z9rQpMU/edit?pli=1
[64] https://www.nist.gov/programs-projects/research-data-framework-rdaf
[65] https://zenodo.org/record/7418976#.Y8aKf3bMLIV
[66] https://doi.org/10.1080/07421222.1996.11518099
[67] https://datascience.codata.org/articles/10.5334/dsj-2022-008/
[68] https://www.mdpi.com/2078-2489/11/4/175
[69] https://www.rd-alliance.org/system/files/documents/The%20Landscape%20of%20Open%20Science%20in%20Malaysia_2022.pdf
[70] https://www.whitehouse.gov/ostp/ostps-teams/science-and-society/
[71] https://www.whitehouse.gov/ostp/news-updates/2022/08/31/an-update-on-research-securitystreamlining-disclosure-standards-to-enhance-clarity-transparency-and-equity/
[72] https://www.science.org/doi/10.1126/science.abo5947
[73] https://sqaaas.eosc-synergy.eu/#/
[74] https://digital.csic.es/handle/10261/160086
[75] https://digital.csic.es/handle/10261/214441
[76] https://alliancecan.ca/sites/default/files/2022-03/rs_current_state_report_1.pdf
[77] https://data.europa.eu/doi/10.2777/986252
[78] https://data.europa.eu/doi/10.2777/926037
[79] https://data.europa.eu/doi/10.2777/70791
[80] https://doi.org/10.5281/zenodo.6461229
[81] https://data.europa.eu/doi/10.2777/620649
[82] https://op.europa.eu/en/publication-detail/-/publication/d1bc3702-61e5-11eb-aeb5-01aa75ed71a1/language-en
[83] https://data.europa.eu/doi/10.2777/1524

23. Twenty-Year Review of GBIF[84]
24. Report on a maturity model towards FAIR data in FAIR repositories (D4.6)[85]
25. D1.2 Data Management Plan[86]
26. D2.1 Report on fair requirements for persistence and interoperability[87]
27. D2.3 Set of FAIR data repositories features[88]
28. D2.4 2nd report on FAIR requirements for persistence and interoperability[89]
29. D2.8 FAIR Semantics Recommendations Third Iteration)[90]
30. 2.7 Framework for assessing FAIR services[91]
31. D2.9 Second reference implementation of the data repositories features and client application[92]
32. D2.10: 3rd report on FAIR requirements for persistence and interoperability[93]
33. D3.1 FAIR policy landscape analysis[94]
34. D3.2 FAIR Data practise analysis[95]
35. D3.3 Policy Enhancement recommendations[96]
36. D3.4 Recommendations on practice to support FAIR data principles[97]
37. D3.5 Description of FAIRsFAIR's Transition Support Programme for Repositories[98]
38. D3.6 Proposal on integration of metadata catalogues to support cross-disciplinary FAIR uptake[99]
39. D3.8 Final Report on Policy and Practice recommendations and support[100]
40. D4.1 Draft Recommendations on Requirements for Fair Datasets in Certified Repositories[101]
41. D4.3 Report on the certification support and guidance for repositories and reviewers[102]
42. D4.5 Report on FAIR Data Assessment Toolset and Badging Scheme[103]
43. D4.7 Tools for finding and selecting certified repositories for researchers and other stakeholders[104]
44. D5.1 FAIRsFAIR Communication, Marketing and Engagement Plan[105]

[84] https://doi.org/10.35035/ctzm-hz97
[85] https://zenodo.org/record/6699520
[86] https://doi.org/10.5281/zenodo.6656144
[87] https://doi.org/10.5281/zenodo.5535719
[88] https://doi.org/10.5281/zenodo.5361952
[89] https://doi.org/10.5281/zenodo.5356517
[90] https://zenodo.org/record/6675295#.ZEGtUHbMLIU
[91] https://zenodo.org/record/6656431#.ZEGmHnbMLIU
[92] https://doi.org/10.5281/zenodo.6699160
[93] https://doi.org/10.5281/zenodo.6685820
[94] https://doi.org/10.5281/zenodo.5537032
[95] https://zenodo.org/record/5362079#.ZEmIiHbMLIU
[96] https://doi.org/10.5281/zenodo.5362183
[97] https://doi.org/10.5281/zenodo.5357329
[98] https://doi.org/10.5281/zenodo.5362210
[99] https://doi.org/10.5281/zenodo.5357560
[100] https://doi.org/10.5281/zenodo.6699333
[101] https://doi.org/10.5281/zenodo.5362222
[102] https://doi.org/10.5281/zenodo.6656437
[103] https://doi.org/10.5281/zenodo.6656444
[104] https://doi.org/10.5281/zenodo.6700832
[105] https://doi.org/10.5281/zenodo.6656060

45. D5.3 Report on the First Synchronisation Force Workshop[106]
46. D5.4 FAIRSFAIR Communication, Marketing & Engagement Plan (Final Version)[107]
47. D5.5 Report 2 of the Synchronisation Force[108]
48. D5.6 Report 3 of the Synchronisation Force[109]
49. D5.7 Recommendations for a FAIR EOSC - White Paper FAIRsFAIR Synchronisation Force 2021[110]
50. D5.8 Pan-European Uptake FinalReport[111]
51. D6.1 Overview of needs for competence centres[112]
52. D6.2 Initial Core Competence Centre Structures[113]
53. D6.4 Final Report on Competence Centre with Knowledge Base[114]
54. D6.5 Report on three annual schools in core data skills for researchers[115]
55. D6.6 Data Steward Instructor Training[116]
56. D6.7 Report on schools run through franchising with local organisers[117]
57. D7.5 Good Practices in FAIR Competence Education[118]
58. M2.7: Assessment report on 'FAIRness of services'[119]
59. M2.10 Report on basic framework on FAIRness of services[120]
60. M2.15 Assessment report on 'FAIRness of software'[121]
61. M4.2 Draft Maturity Model Based on Extensions and-or Additions to CoreTrustSeal Requirements[122]
62. M4.3 CoreTrustSeal+FAIRenabling, Capability and Maturity[123]
63. M4.7 Improved Description of Data Repositories[124]
64. M4.9 Report on Fair Data Assessment Mechanisms to Develop Pragmatic Concepts for Fairness Evaluation at the Dataset Level[125]
65. Sendai Framework[126]
66. FAIR Implementation Profile (FIP) Ontology[127]

[106] https://doi.org/10.5281/zenodo.5361052
[107] https://doi.org/10.5281/zenodo.5244705
[108] https://doi.org/10.5281/zenodo.5361417
[109] https://doi.org/10.5281/zenodo.5595863
[110] https://doi.org/10.5281/zenodo.6674042
[111] https://doi.org/10.5281/zenodo.5786729
[112] https://doi.org/10.5281/zenodo.5361524
[113] https://doi.org/10.5281/zenodo.3732888
[114] https://doi.org/10.5281/zenodo.6700985
[115] https://doi.org/10.5281/zenodo.6701243
[116] https://doi.org/10.5281/zenodo.6701372
[117] https://doi.org/10.5281/zenodo.6701167
[118] https://doi.org/10.5281/zenodo.6657165
[119] https://doi.org/10.5281/zenodo.5470375
[120] https://doi.org/10.5281/zenodo.5473015
[121] https://doi.org/10.5281/zenodo.5472911
[122] https://doi.org/10.5281/zenodo.5471568
[123] https://doi.org/10.5281/zenodo.5346822
[124] https://doi.org/10.5281/zenodo.5471811
[125] https://doi.org/10.5281/zenodo.5471977
[126] https://www.undrr.org/implementing-sendai-framework/what-sendai-framework
[127] https://peta-pico.github.io/FAIR-nanopubs/fip/index-en.html

67. National Action Plan for Open Research (2022-2030)[128]
68. Trends in preprint, data, and code sharing, 2019-2022[129]
69. Briefing Paper on EOSC: Federating Research Infrastructures in Europe for Fair Access to Data[130]
70. An Analysis of Scientific Practice towards FAIR Digital Objects[131]
71. The EOSC Executive Board Working Group (WG) Architecture Task Force (TF) SIRS[132]
72. Landscape analysis | ESFRI Roadmap 2021[133]
73. ERIC Work Packages & Tasks[134]
74. Report on proposed approach and dashboard for common ERIC KPIs[135]
75. Coalition Publica strategic plan[136]
76. ISC by the numbers in 2021[137]
77. CoreTrustSeal Requirements 2023-2025[138]
78. Erudit 2021-2022 annual report[139]
79. NSF Public Access Initiative[140], NSF Public Access Plan 2.0[141]
80. Data characteristics affecting levels of openness[142]
81. Amsterdam Declaration on Funding Research Software Sustainability[143]
82. Brno Declaration on Fostering a Global Ecosystem of RIs[144]
83. PARIS Principles[145]
84. Legal Interoperability and the FAIR Data Principles[146]
85. Deliverables & Milestones | SSHOPENCLOUD[147]
86. Digital Objects – FAIR Digital Objects: Which Services Are Required?[148]
87. OpenAIRE Guidelines[149]
88. Challenges of Research Data Management for High Performance Computing[150]
89. Four tips for metadata management[151]

[128] https://doi.org/10.7486/DRI.ff36jz222
[129] https://theplosblog.plos.org/2023/04/open-science-indicators/
[130] https://www.scienceeurope.org/our-resources/briefing-paper-on-eosc/
[131] https://b2share.eudat.eu/records/e14269d07ce84027a7f79ee06b994ef9
[132] https://op.europa.eu/s/oK7d
[133] https://roadmap2021.esfri.eu/landscape-analysis/
[134] https://www.eric-forum.eu/work-packages-task-leaders/
[135] https://www.eric-forum.eu/wp-content/uploads/D4.1-Report-on-proposed-approach-and-dashboard-for-common-ERIC-KPIs.pdf
[136] https://www.erudit.org/public/documents/EN_CO2020_stratplan.pdf
[137] https://council.science/annual-report-2021/2021-by-the-numbers-graphic/
[138] https://zenodo.org/record/7051012#.Y39xi3bMLIV
[139] https://apropos.erudit.org/2021-2022-annual-report/?lang=en
[140] https://beta.nsf.gov/public-access
[141] https://nsf-gov-resources.nsf.gov/2023-06/NSF23104.pdf?VersionId=cSTD31SSPUEkM_Vm25HSlgZBDeiPvzdQ
[142] https://docs.google.com/document/d/1Oqbxj5rneWkuKyrYmfwPqvvM9j8iaLgB/edit
[143] https://zenodo.org/record/7330542#.Y5ts83bMJdg
[144] https://www.esfri.eu/latest-esfri-news/brno-declaration-ris
[145] http://www.infocomm-journal.com/bdr/EN/10.11959/j.issn.2096-0271.2023013
[146] https://zenodo.org/record/4471312#.Y86hSnbMLIU
[147] https://sshopencloud.eu/publications/deliverables
[148] https://datascience.codata.org/articles/10.5334/dsj-2020-015/
[149] https://guidelines.openaire.eu/en/latest/index.html
[150] https://link.springer.com/chapter/10.1007/978-3-319-67008-9_12
[151] https://osf.io/gsfc6

90. Storage 2020: A Vision for the Future of HPC Storage[152]
91. Research Software Workshop: guidelines and metrics for metadata curation (co-located with RDAP20) | FAIR-IMPACT[153]
92. The four pillars of governance best practice for New Zealand Directors[154]
93. Metadata Principles and their Use[155]
94. Metadata Standards Catalog[156]
95. RDA Brokering Framework: Preliminary Recommendations[157]
96. Data Citation of Evolving Data[158]
97. RDA Recommendation on PID Kernel Information FINAL[159]
98. Research Data Repository Interoperability WG Final Recommendations[160]
99. The final version of the RDA COVID-19 Recommendations and Guidelines for Data Sharing[161]
100. A Collection of Crosswalks from Fifteen Research Data Schemas to Schema.org[162]
101. Addressing the Gaps: Recommendations for Supporting the Long Tail of Research Data[163]

Table2: Sources reviewed for this project.

Appendix C: Task Group members

The following tables capture a snapshot of the status of each TG during phase 1 and phase 2. The TG wiki[164] contains more detailed information and step-by-step actions in the timeline. Note that "TG items" refers to the total number of categories, subcategories, attributes, features, and KPIs and metrics, where KPIs and metrics were split and evaluated by each TG in phase 1 and were moved into a separate space before the RDA20 workshop. For phase 1, dates next to individuals' names indicate when they joined the TG after October 2022. "Start" refers to the items in the original running list[165]

[152] https://escholarship.org/uc/item/744479dp
[153] https://fair-impact.eu/events/fairimpact-events/research-software-workshop-guidelines-and-metrics-metadata-curation
[154] https://www.iod.org.nz/resources-and-insights/4-pillars-landing-page/#
[155] https://www.rd-alliance.org/group/metadata-ig/outcomes/metadata-principles
[156] https://rdamsc.bath.ac.uk/
[157] https://docs.google.com/document/d/1cMpleTatYUckijSL8zqjAwvtKAn6oUAoHcNKrCCgXjI/edit?pli=1
[158] https://www.rd-alliance.org/system/files/RDA-DC-Recommendations_151020.pdf
[159] https://www.rd-alliance.org/system/files/RDA%20Recommendation%20on%20PID%20Kernel%20Information_final.pdf
[160] https://www.rd-alliance.org/group/research-data-repository-interoperability-wg/outcomes/research-data-repository-0
[161] https://www.rd-alliance.org/group/rda-covid-19-rda-covid19-omics-rda-covid-19-epidemiology-rda-covid19-clinical-rda-covid19-2
[162] https://www.rd-alliance.org/group/research-metadata-schemas-wg/outcomes/collection-crosswalks-fifteen-research-data-schemas
[163] https://www.rd-alliance.org/group/long-tail-research-data-ig/outcomes/addressing-gaps-recommendations-supporting-long-tail-0
[164] https://www.rd-alliance.org/group/gorc-international-model-wg/wiki/gorc-wg-task-groups
[165] https://docs.google.com/spreadsheets/d/1SiWnZRwkmbGuAX1sS33EWLqyRJDQWgkXj8l1EwdQsi0/edit?usp=sharing

for each TG, “Pre-RDA20 workshop” refers to the phase 1 draft[166], “Version 0.5” refers to version 0.5 of the model[167] released in April 2023, “Version 0.6” refers to version 0.6 of the model[168] released in July 2023, and “Final (Version 1.0)” refers to the current version of the model, version 1.0[169], that this report accompanies.

Phase 1 (October 2022 - March 2023)

TG	TG Lead	TG Members	TG Scope	TG items
TG1	Françoise Genova	Kheeran Dharmawardena CJ Woodford (12/22)	Governance Rules of Participation & Access Sustainability	Start: 42 Pre-RDA20 workshop:33 Version 0.5: 49
TG2	Sarah Jones	Elisha Wood-Charlson Amy Nurmberger	Engagement Human Capacity	Start: 149 Pre-RDA20 workshop:41 Version 0.5: 40
TG3	CJ Woodford (12/22)	Qian Zhang (12/22)	ICT Infrastructure Research Objects	Start: 93 Pre-RDA20 workshop: 20 Version 0.5: 23
TG4	Andrew Treloar (10./22-11/22, 02/23) Rory Macneil (11/22-02/23)	Rory Macneil (10/22-11/22, 02/23) Mikiko Tanifuji (12/22) Wolmar Nyberg Åkerström (12/22)	Standards Interoperability	Start: 158 Pre-RDA20 workshop: 111 Version 0.5: 65
TG5	Karen Payne	Javier Lopez Albacete Hans Pfeiffenberger	Services & Tools	Start: 141 Pre-RDA20 workshop: 75 Version 0.5: 70
-			KPIs & Metrics	Pre-RDA20 workshop: 66

[166] https://docs.google.com/spreadsheets/d/1x5TlGeTkCTlN37rh1uxJkPCsbzg28XK_huE5H7VQqmM/edit?usp=sharing
[167] https://docs.google.com/spreadsheets/d/1M6GGO8uPKX-ZYqfg-hBq-2et7QeTr8F3PUrxxyssebI/edit?usp=sharing
[168] https://docs.google.com/spreadsheets/d/1ow2x6alS0SqAK2BaBljrTNhiMXOmQZUt4Smk2WQLeoI/edit?usp=sharing
[169] GORC IM WG Commons Model V0.9 (to be updated to V1.0 after RDA Request for comment period): https://docs.google.com/spreadsheets/d/1GLmyczP5Ez32HRK_1DV9H4owlhac8QWdh6SVarKoJKE/edit?usp=sharing

				Version 0.5: 67

Table 3: A snapshot of the GORC-WG task group members and workload for Phase 1.

Phase 2 (April 2023 - September 2023)

TG	TG Lead(s)	TG Members	TG Scope	TG items
TG1	Javier Lopez Albacete Nick Jones	Hans Pfeiffenberger CJ Woodford	Governance & Leadership Rules of Participation & Access Sustainability	Version 0.5: 49 Version 0.6: 89 Final (Version 1.0): 86
TG2	Kheeran Dharmawardena Noel Chibhira	Amy Nurnberger CJ Woodford	Engagement Human Capacity	Version 0.5: 40 Version 0.6: 43 Final (Version 1.0): 44
TG3	CJ Woodford	Qian Zhang Devika Madalli	ICT Infrastructure Research Objects	Version 0.5: 23 Version 0.6: 30 Final (Version 1.0): 34
TG4	Françoise Genova	Mikiko Tanifuji Wolmar Nyberg Åkerström Laurents Sesink	Standards & Conventions Interoperability	Version 0.5: 65 Version 0.6: 55 Final (Version 1.0): 54
TG5	Andrew Treloar	Rory Macneil CJ Woodford	Services & Tools	Version 0.5: 70 Version 0.6: 75 Final (Version 1.0): 74
TG6	CJ Woodford		KPIs & Metrics	Version 0.5: 67 Version 0.6: 98 Final (Version 1.0): 104

Table 4: A snapshot of the GORC-WG task group members and workload for Phase 2.