GORC Interest Group: Typology

Version: 0.9.1, May 17, 2023

# Context

The [Global Open Research Commons](https://www.rd-alliance.org/groups/global-open-research-commons-ig) is an [Interest Group](https://www.rd-alliance.org/groups/interest-groups) within the [Research Data Alliance](http://rd-alliance.org/). The IG is working to reach a shared understanding of what a “commons” is in the research data space: the functionality, coverage, and characteristics of a digital research environment with a potential global impact. The typology presented here is in the form of a living document that requires revisiting and updating as new technologies and practices are developed in research commons.

### Commons:

A digital ecosystem providing research infrastructure, services, and outputs with potentially global access.

### Strapline:

Digital research resources for the common good

One of the outputs of the IG is a typology of the essential elements in a commons. In setting out to characterize research commons, the IG identified the need to provide a set of definitions for each of the core elements that appeared to be essential to the infrastructure of a commons. This document establishesthe typology with the supporting definitions.

# Scope

In considering a research commons, one needs to define the essential elements of such a digital ecosystem.

## Contents

The purpose and uses of digital research ecosystems may vary according to settings, disciplines, and/or methodologies. Different research commons may therefore prioritize different core infrastructure elements in order to achieve specific sought outcomes.

The scope of GORC is limited to digital objects and the infrastructure that supports research in a digital ecosystem.

## Sources

A wide variety of physical instruments are used in the sciences to capture data from (digitally describe) natural phenomena. These physical instruments fall outside the scope of this typology. Digital objects generated through data processing, whether or not in combination with physical instruments, fall within the scope of this typology.

# The essential elements of a research commons



The three elements in blue constitute the core data processing infrastructure for posing research queries and generating responses. The five elements in white constitute the core supporting human infrastructure required for a research commons to exist and be sustained.

# Definitions

## Governance

Governance here refers to the structures and processes through which decisions in a research commons are made, implemented, and monitored. It encompasses the mechanisms and practices that guide decision-making practices, either human or machine, concerning the exercise of authority, control, and/or management of the common’s resources, policies, and activities. Governance includes the development of mission statements, values, organizational performance metrics, risk management frameworks, policies, and guidelines for financial and operational matters and stakeholder relations.

## Rules of participation (RoP)

Rules of participation (RoP) are guidelines or principles that outline how individuals or entities (human or machine) may engage with a research commones. These rules help establish a framework for collaboration, interaction, and decision-making, ensuring that participants are aware of their rights, responsibilities, and expected behavior. Set of policies defining a minimal set of rights, obligations, and responsibilities governing the activities of those participating in the Commons (this definition and the concept of Rules of Participation are drawn from the European Open Science Cloud). The rights to all participants vary by Commons and are not enumerated here but could include rights"that enable and foster reuse and reproducibility. These should fulfill the following criteria (which were inspired by the EOSC Rules of Participation as listed in [this report from the EOSC Executive Board](https://op.europa.eu/en/publication-detail/-/publication/a96d6233-554e-11eb-b59f-01aa75ed71a1/language-en)):

* The RoP apply to all users of the commons, though specific rights and responsibilities may vary among user category.
* They apply to all aspects of a user’s (human or machine) interactions with the commons, its infrastructure and its digital objects, though the specific rules may vary according to access rights to various data processing infrastructure and digital objects

See too, [Elinor Ostrom's Eight Principles for Managing Commons](https://www.onthecommons.org/magazine/elinor-ostroms-8-principles-managing-commmons).

## Engagement

Methods used to interact with the broad stakeholder community to involve them in activities. These could include requirements for gathering exercises, consultations, usability testing, communications, events, and training.

## Human capacity

The ability of the commons to create a human-centric environment for the stakeholders and community members in all aspects, specifically for users, providers, and internal staff, 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>).

* Two dimensions of human capacity are critical:Users: these need to have the capacity to be able to identify and make use of the services provided by the commons. This may be a combination of general IT skills, and specific skills in the tools/services that are available
* Providers: in order to build, maintain and extend the commons infrastructure, a wide range of human capacity is essential. This can be viewed as “people as infrastructure”, by analogy with “data as infrastructure”.

I would delete your entire definition of human capacity above and suggest the following: Human capacity in research commons refers to the knowledge, skills, expertise, and resources of individuals and/or communities that contribute to the development, maintenance, and utilization of the digital research infrastructure and resources within the commons. It encompasses the collective capabilities and contributions of users, including researchers, data scientists, librarians, technologists, and other stakeholders involved in digital research.

## Sustainability

Sustainability refers to models or agreements made on how to fund or resource activities in a way that can be sustained over the long-term. This may include mixed streams of investment and cost recovery through subscriptions or service payment models to ensure operation of the Commons, as well as the input of effort/time by contributors to, and maintainers of, elements of Commons infrastructure. Reuse of existing components is an effective strategy for more sustainable Commons infrastructures.

I would delete the above and suggest the following:

Sustainability refers to the ability of the commons to maintain its operations, functionality, and value over an extended period. It involves the means to ensure the long-term viability, resilience, and effectiveness of the infrastructure in supporting the digital research activities and meeting the evolving needs of the research community. This includes technical sustainability; financial sustainability, governance and leadership sustainability; relevant policies, legal frameworks, ethical considerations, and data protection regulations to ensure compliance and build trust among users and stakeholders; long-term preservation, curation, and accessibility; and measures to ensure the ongoing active engagement of communities, researchers, and other stakeholders.

## Interoperability

Interoperability indicates how research data objects are structured and formatted to allow for integration and processing across the set(s) of data objects. This may involve using standardized data formats, standardized data models and vocabularies, and providing clear and well-defined data and metadata standards. Interoperability may also be defined as “The ability of data or tools from non-cooperating resources to integrate or work together with minimal effort.” ([Wilkinson, et al. 2016](https://www.nature.com/articles/sdata201618#Sec3)). This ability often supports, but is not limited to, reuse and reproducibility. 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 organizations). ([Janssen et al 2014](https://ieeexplore.ieee.org/document/6926683))

## Standards

A standard is a repeatable, harmonised, agreed and documented way of doing something ([IRENA](https://www.irena.org/inspire/Standards/What-are-Standards)). Within a digital research commons, standards may refer to the following:

* Metadata standards used to describe data according to agreed schema
* Controlled vocabularies and ontologies used to label and assign keywords according to semantically agreed and standardised terms
* Data formats used to structure data
* Service endpoints that allow humans and machines to interact and exploit resources
* Authentication and authorization protocols

## ICT Infrastructure (Compute, Storage, Network, AAI)

By “ICT infrastructure” we mean the physical components that a computer system requires to function and are necessary to conduct research. This includes:

* Compute: the data processing cycles delivered via on-premise hardware or off-premise cloud services
* Storage: data storage delivered via on-premise hardware or off-premise cloud services
* Network: the hardware and software required to connect, compute, and storage, as well as to access Internet resources
* AAI: Authentication and authorization infrastructure (AAI) refers to services and procedures that enable users to access the commons or specific attributes of the commons

## Digital tools and services

Service ([as defined by IVOA](https://www.ivoa.net/documents/WD/ResMetadata/RSM-20030509.html)) A service 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.

Digital tools enable researchers to perform one or more operations on digital objects. Tools may be designed for human or machine use

Because 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, or research commons) that are deployed to support research workflows and the communities of practice engaged in collaborative research. Typically, a research platform’s capabilities include data acquisition and management, processing and visualization, storage and preservation, sharing and discovery. These platforms may provide the full spectrum or a subset of components. Science gateways may be discipline-specific, and may support and enhance scientific collaborations and scholarly communication by facilitating citizen science engagement as well.

I would delete the above and suggest the following:

The terms ‘digital tool’ and ‘digital services’ often overlap in their usage with regard to the digital objects or digital functions to which they refer. Below we define them separately aware that an overlap in usage in specific circumstances and for specific usages is common.

A digital tool refers to a software-based application or program that is specifically designed to support research activities, data management, analysis, collaboration, or other functions within the research commons. Digital tools are instrumental in enabling researchers to leverage the benefits of digital technologies and resources in their work, and they play a crucial role in advancing research outcomes. Digital tools may include data analysis tools, data management tools, research workflow tools, data visualization tools, reference tools, publishing tools, collaboration tools. Digital tools in a digital research commons contribute to leveraging the full potential of digital resources and technologies for users.

A digital service in a research commons refers to a software-based application, platform, or tool provided within a research commons to users in support of research activities within the commons or interconnected to the commons. Digital services are designed to enhance and streamline aspects of the research process, including identity and access management services, data management, collaboration, and knowledge dissemination. These services may include data management services, data analysis services, data visualization services, communication services, publishing and dissemination services, collaboration services. The specific digital services offered in a digital research commons may vary depending on the scope, focus, and goals of the commons. These services aim to empower researchers, enhance their productivity, foster collaboration, and enable the efficient and effective use of digital resources and technologies in the research process.

## Digital Objects

A digital object is an electronic entity that describes a phenomenal entity or set of entities in the physical work or (an)other electronic entity(ies). Digital objects may be data objects, digital object identifiers, digital representations, digital files, databases, metadata records, digital artifacts. The term ‘digital object’ encompasses, thus, various forms and manifestations of digital content, resources, and representations. In the context of research commons, it refers to research on the digital counterparts/representations of the phenomenal physical world within our increasingly digital societies.

# References

The primary references for the typology presented here are the following:

* The [CASRAI glossary](https://casrai.org/rdm-glossary/)
* The [EOSC glossary](https://eosc-portal.eu/glossary)
* A [glossary of terms relating to open scholarship](https://forrt.org/glossary/)
* The draft wikidata open infrastructure elements and definitions