WP 1: Project Coordination, Standardisation and Methodology

Deliverable 1.7:

RRING Knowledge database of RRI

Status: Final

22/07/2021

About the RRING project

The overall project aim is to bring RRI into the linked up global world to promote mutual learning and collaboration in RRI. This will be achieved by the formation of the global RRING community network and by the development and mobilisation of a global Open Access RRI knowledge base. RRING will align RRI to the Sustainable Development Goals (SDGs) as a global common denominator. The RRING project acknowledges that each region of the world is advancing its own agenda on RRI. Therefore, RRING will not be producing a Global RRI framework or strategy that is meant to be enforced in a top-down manner. Rather, increased coherence and convergence will be achieved via a bottom-up approach, learning from best practices in RRI globally and from linkages, via the new RRING community, to develop the RRI linked-up world.

Six Objectives of RRING

Objective 1: Promote a linked up global world of RRI by creating the global RRING community network, thereby enabling mutual learning, collaboration, mobilisation of RRI concepts.

Objective 2: Mobilise, promote and disseminate a global open access knowledge base of RRI based on the State of the Art (SoA) and comparative analysis across the key geographies, all stakeholders and sectors. It will cover key platforms, spaces and players, role and influence of stakeholders, drivers and policies for R&I, regulation in public,private sectors and nation states and international organizations.

Objective 3: Align RRI to the UN Sustainable Development Goals (SDGs) to provide a global common denominator for advancement of RRI, and address Grand Challenges globally.

Objective 4: Determine the competitive advantages of RRI and also understand how and where RRI is perceived as a barrier and/ or disadvantage.

Objective 5: Create high level RRI strategy recommendations for the seven geographic zones, trial RRI best practice learning in 2 EU case studies.

Objective 6: Promote inclusive engagement of civil society and researchers.

Document Information

Title	D1.7: RRING Knowledge database of RRI
Distribution	Public
Document Reference	RRING D1.7
Deliverable Leader	Eric A. Jensen ICoRSA
Contributing Authors	Lars Lorenz and Eric A. Jensen ICoRSA

Acknowledgement

Legal Disclaimer

Executive Summary

This report 'D1.7 RRING Knowledge database of RRI' offers a detailed explanation of the Zenodo database, the tool used to upload, store and share RRING Project files within the consortium members. It is explained how it works and why it aligns with task 1.1

Contents

Contents

1. About Zenodo

1.1 About the database

Zenodo, a name derived derived from Zenodotus, the first librarian of the Ancient Library of Alexandria and father of the first recorded use of metadata, is a database built and developed by researchers, and maintained by the European Organization for Nuclear Research (CERN) and stored in their data centre. According to its website, Zenodo code is itself open source and is built on the foundation of the Invenio digital library which is also open source, making it a completely free service. All meta data is openly available under CC0 licence, and all open content is openly accessible through open APIs. Zenodo follows the FAIR principles: to be Findable, to be Accessible, to be Interoperable and to be Reusable.

One of the key advantages of this database is that all files types are permitted and the limit per file is up to 50GB, with unlimited total storage. In this sense, Zenodo does not impose any requirements on format, size, access restrictions or licence. Another feature is that it offers the possibility to safely store closed and restricted content with protected links that can be shared with other peers. All content can also be embargoed and automatically opened when the author/s decide to do so. When made public, each file contains detailed insights on views, downloads and social media sharing. The authors can also edit files after publication and Zenodo supports versioning through DOI versioning.

Finally, CERN, an OpenAIRE partner and pioneer in open source, open access and open data, provide a catch-all repository for EC funded research.

1.2 FAIR compliance

Zenodo follows the FAIR principles as referenced from: Wilkinson, M. D. et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci. Data 3:160018 doi: 10.1038/sdata.2016.18 (2016). The following information was retrieved from Zenodo's website (about.zenodo.org/principles).

1.2.1 To be Findable
- **F1**: (meta)data are assigned a globally unique and persistent identifier
    - A DOI is issued to every published record on Zenodo.
- **F2**: data are described with rich metadata (defined by R1 below)
    - Zenodo's metadata is compliant with DataCite's Metadata Schema minimum and recommended terms, with a few additional enrichments.
- **F3**: metadata clearly and explicitly include the identifier of the data it describes
    - The DOI is a top-level and a mandatory field in the metadata of each record.
- **F4**: (meta)data are registered or indexed in a searchable resource

- Metadata of each record is indexed and searchable directly in Zenodo's search engine immediately after publishing.
- Metadata of each record is sent to DataCite servers during DOI registration and indexed there.

1.2.2 To be Accesible

- **A1**: (meta)data are retrievable by their identifier using a standardized communications protocol
  - Metadata for individual records as well as record collections are harvestable using the OAI-PMH protocol by the record identifier and the collection name.
  - Metadata is also retrievable through the public REST API.
- **A1**.1: the protocol is open, free, and universally implementable
  - See point A1. OAI-PMH and REST are open, free and univesal protocols for information retrieval on the web.
- **A1.2**: the protocol allows for an authentication and authorization procedure, where necessary
  - Metadata are publicly accessible and licensed under public domain. No authorization is ever necessary to retrieve it.
- **A2**: metadata are accessible, even when the data are no longer available
  - Data and metadta will be retained for the lifetime of the repository. This is currently the lifetime of the host laboratory CERN, which currently has an experimental programme defined for the next 20 years at least.
  - Metadata are stored in high-availability database servers at CERN, which are separate to the data itself.

1.2.3 To be Interoperable

- **I1**: (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation.
  - Zenodo uses JSON Schema as internal representation of metadata and offers export to other popular formats such as Dublin Core or MARCXML.
- **I2**: (meta)data use vocabularies that follow FAIR principles
  - For certain terms we refer to open, external vocabularies, e.g.: license (Open Definition), funders (FundRef) and grants (OpenAIRE).
- **I3**: (meta)data include qualified references to other (meta)data
  - Each referrenced external piece of metadata is qualified by a resolvable URL.

1.2.4 To be Reusable

- **R1**: (meta)data are richly described with a plurality of accurate and relevant attributes
  - Each record contains a minimum of DataCite's mandatory terms, with optionally additional DataCite recommended terms and Zenodo's enrichments.
- **R1.1**: (meta)data are released with a clear and accessible data usage license
  - License is one of the mandatory terms in Zenodo's metadata, and is referring to an Open Definition license.

- o Data downloaded by the users is subject to the license specified in the metadata by the uploader.
- **R1.2**: (meta)data are associated with detailed provenance
  - o All data and metadata uploaded is traceable to a registered Zenodo user.
  - o Metadata can optionally describe the original authors of the published work.
- **R1.3**: (meta)data meet domain-relevant community standards
  - o Zenodo is not a domain-specific repository, yet through compliance with DataCite's Metadata Schema, metadata meets one of the broadest cross-domain standards available.

2. About RRING and Zenodo

Zenodo supports "Communities", which are curated collections of Zenodo entries. Each community has its own logo, description, policy and more. All RRING project outputs are stored in the same 'RRING Community' and all external contributors can request their entry to be included in the RRING community. Every contributor has the option to search for files within the community.
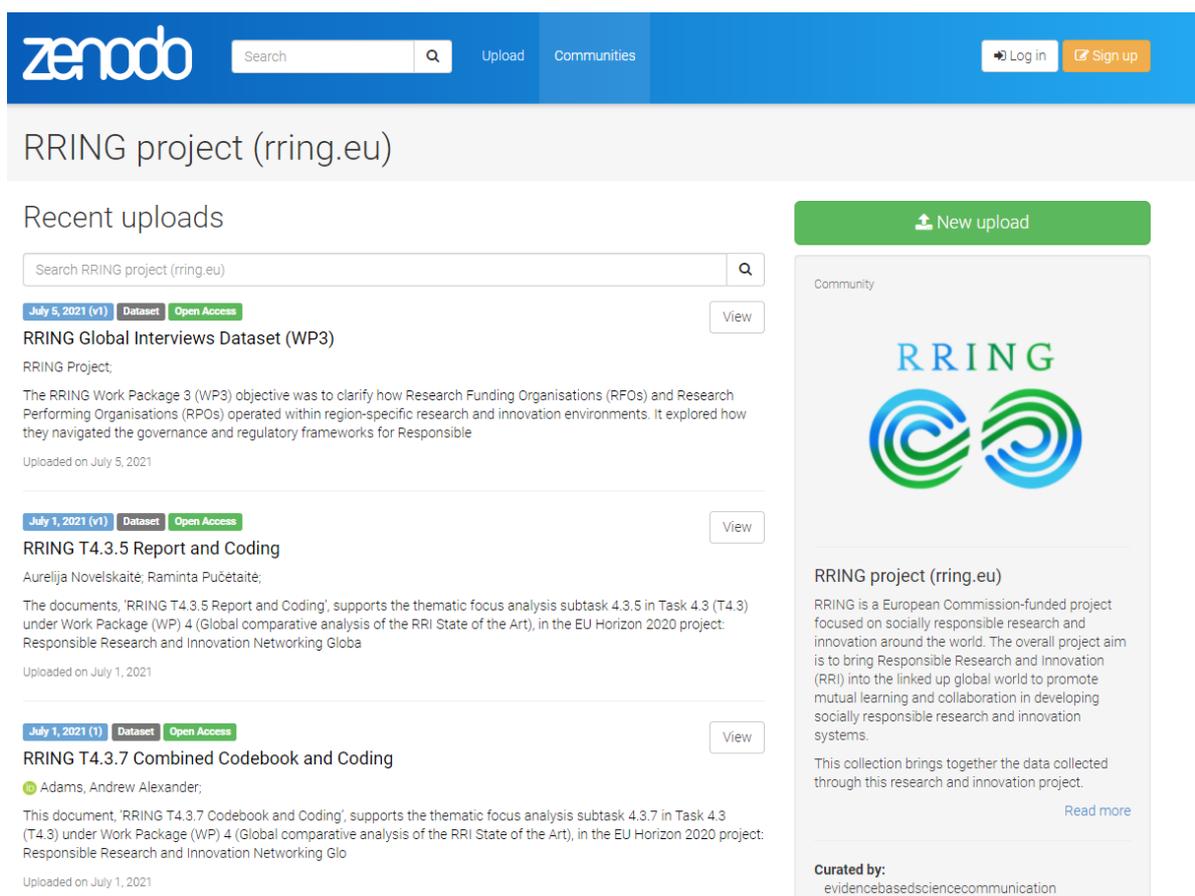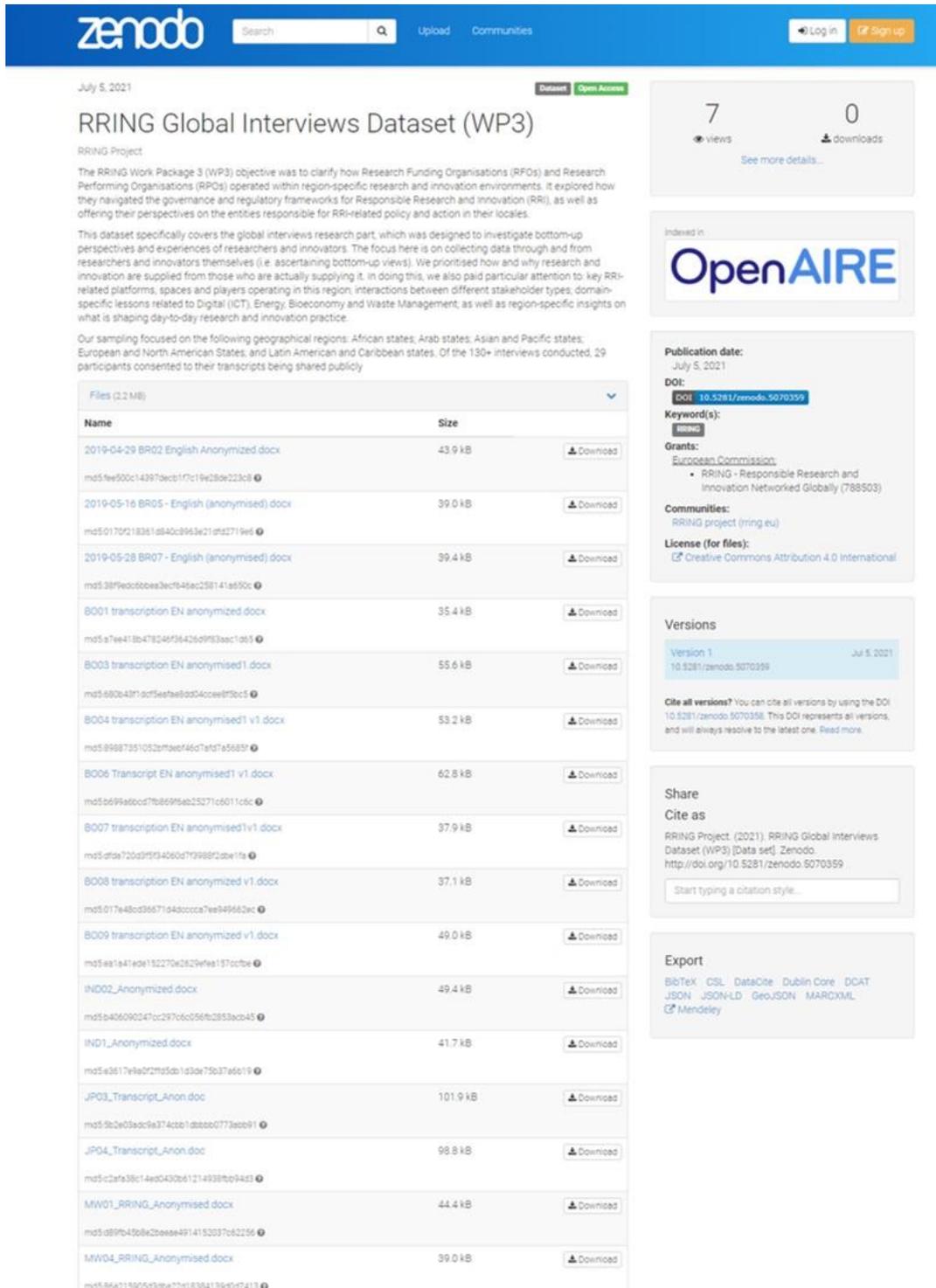

Figure 1. RRING Community on Zenodo

Every Zenodo entry consists of uploaded files and meta info such as title, description, authors, tags, keywords, communities and versioning. Each entry contains detailed insights on views, downloads and social media sharing. All entries are automatically licensed under Creative Commons when published as Open Access.

Figure 2. Entry example: RRING Global Interviews Dataset (WP3)

3. Alignment with Task 1.7

Task 1.7 relates to the selection of tool and the administration of the RRING database. Zenodo was chosen because of the following reasons:

- It is user friendly for stakeholders
  Stakeholders can browse any data without restrictions and without needing to create an account.

- It is easy to populate with data
  Project partners and external collaborators can simply create a free account and submit data to the RRING community.

- It is easy to alter its architecture and code
  While changes to the underlying code are not possible, Zenodo takes all types of data as default.

- It solves potential management issues during project and post project
  Zenodo is completely free and maintained by CERN. The RRING data will therefore persist far past the project's lifespan.

The RRING community can be accesed following this link: zenodo.org/communities/rring

4. Post RRING data sustainability

Data stored on Zenodo will be maintained post RRING by following terms:

- The RRING data is stored in Zenodo, which is a third party website, available for free, and managed externally. Data stored in Zenodo is expected to be stored there in perpetuity , for free, and uncorrupted.

- ICoRSA will supervise the data stored for the first year, making certain that the data remains intact and uncorrupted.

- The RRING network administration will take over the responsibility of the RRING project data stored after year 1. The RRING network will have a Dissemination and Communications person after year 1, who will also be designated as Data Manager

Zenodo ensures longevity in following terms:

- Retention period: Items will be retained for the lifetime of the repository. This is currently the lifetime of the host laboratory CERN, which currently has an experimental programme defined for the next 20 years at least.

- Functional preservation: Zenodo makes no promises of usability and understandability of deposited objects over time.

- File preservation: Data files and metadata are backed up nightly and replicated into multiple copies in the online system.

- Fixity and authenticity: All data files are stored along with a MD5 checksum of the file content. Files are regularly checked against their checksums to assure that file content remains constant.

- Succession plans: In case of closure of the repository, best efforts will be made to integrate all content into suitable alternative institutional and/or subject based repositories.

5. References

Wilkinson, M. D. et al. (2016). *The FAIR Guiding Principles for scientific data management and stewardship*. Sci. Data 3:160018 doi: 10.1038/sdata.2016.18