RDA TIGER

eosc

Using the RDA Annotator

User guidelines & installation instructions

Table of Contents

Document information

List of contributors

Simon Saldner, KNAW-DANS
ORCID: 0000-0002-1145-7829

Wim Hugo, KNAW-DANS
ORCID: 0000-0002-0255-5101

Version
0.3

Date
April 2025

1. Introduction: why use the RDA Annotator?

RDA members typically consult and reference a large number of web-based resources (websites, publications, and so on) in order to create and formulate their outputs, supplementary materials, and other outputs. In doing so, **a large amount of effort is lost in respect of future reuse** since the finer context of the resources and references being consulted is often not recorded in the final outputs.

The **RDA Annotator** allows users to make detailed annotations to web-based resources using textual descriptions, as well as **RDA vocabularies** and tags that contextualises and categorises the content. Annotations made using the Annotator are then passed to the **RDA Knowledge Base**, a suite of resources where other RDA community members can access them.

This document describes the **installation and use of the RDA Annotator**, a browser extension mainly intended for RDA community members, which lets you annotate materials through a simple menu that captures and annotates any material you select on your browser.

2. The RDA Knowledge Base

The RDA Knowledge Base is a mechanism for the publication and cataloguing of the RDA's body of work, e.g. outputs, supplementary materials, and other useful publications. In addition, it provides prior art, specifications, applications and services, recommendations and good practice, or any other resource that were identified by group participants in RDA as being useful in an RDA-related context. To contextualise such resources for RDA use and re-use, annotation of the resources with RDA-related vocabulary would be a starting point, in addition to the usual free-text annotation.

The RDA Knowledge Base provides a set of interrelated resources:

1. **RDA Publisher:** A mechanism for RDA members or their collaborators to publish resources formally to Zenodo by default,[1] and to supplement the metadata typically required by Zenodo with additional, RDA-specific context, using a newly-developed RDA metadata schema.

2. **RDA Graph:** A comprehensive metadata catalogue containing e.g. RDA outputs, members, groups, as well as external, RDA-Relevant, resources such as vocabularies. The RDA Graph is linked to the Zenodo catalogue, allowing but not requiring resources to be published there.

3. **RDA Annotator**: In addition to containing a catalogue of RDA resources, the RDA Graph can be extended to include any other resource on the web that was annotated by an RDA member, using the RDA Annotator. To the extent that these annotations are linked to RDA vocabulary, groups, and members, they can be found and reused by anyone, including the general public. The annotations

[1] In principle, the system can be configured to publish to other repositories that offer an API service for the purpose.

added via the RDA Annotator are also mirrored to hypothesis.is as free-text annotations, which assists the wider research community with content generated by and on behalf of RDA.

4. **RDA Discovery**[2]: this includes search and discovery tools, and a dashboard interface, allowing users to find and explore resources from the RDA Graph.

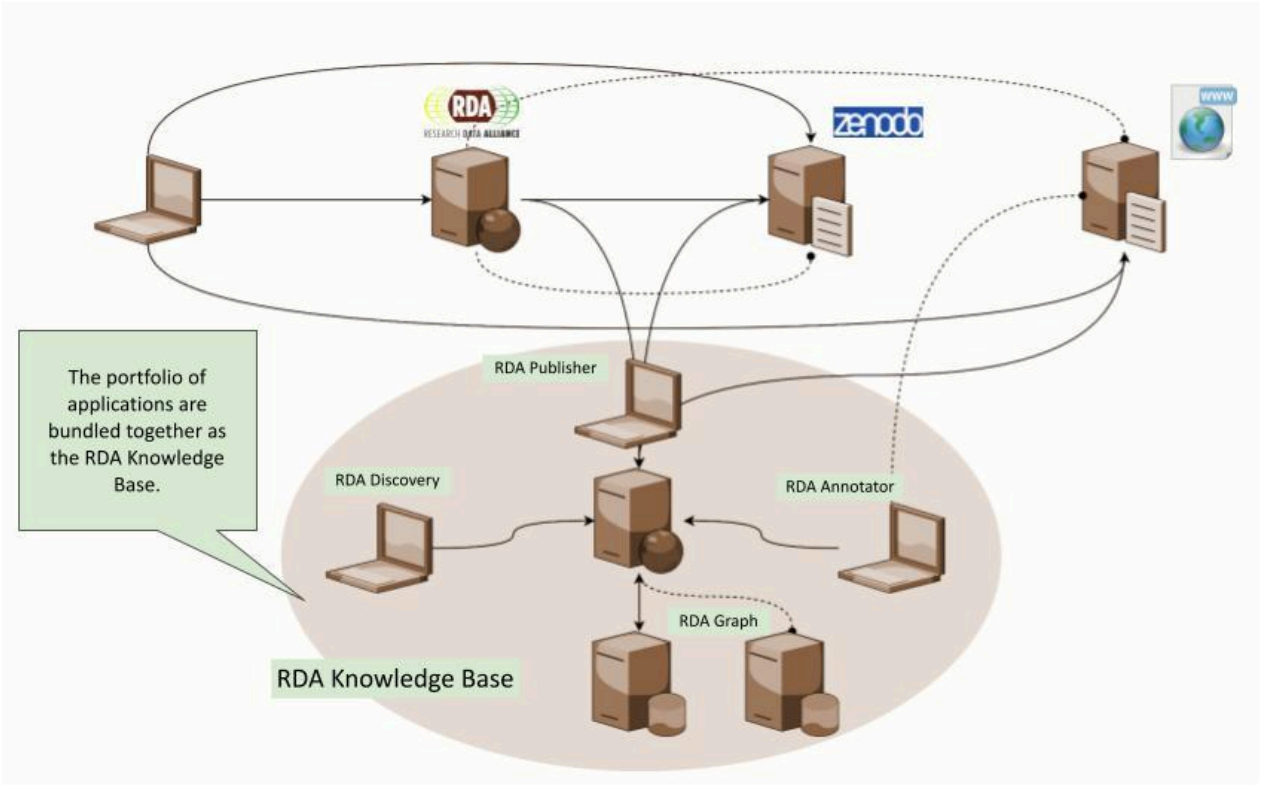


Figure: The RDA Knowledge Base, its components, and links to the RDA Website, Zenodo, and other web-based resources

[2] https://rda.dansdemo.nl/

3. Using the RDA Annotator

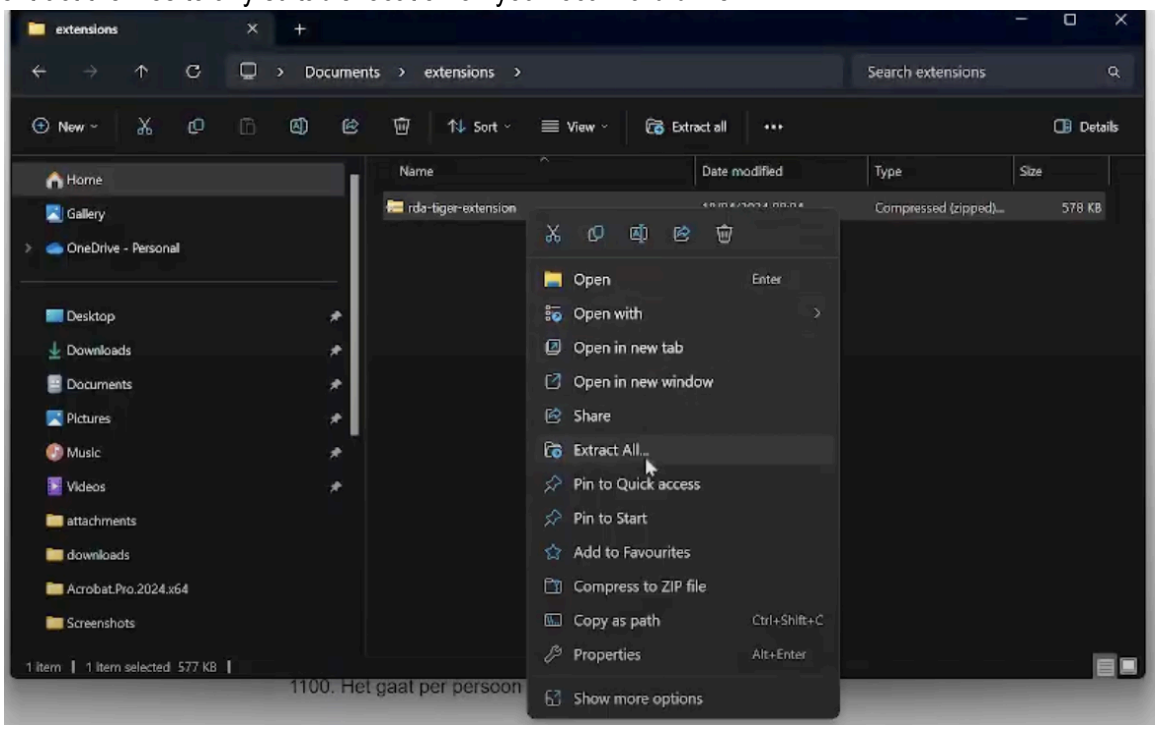
The following 3 sub-sections walk you through how to install and use the Annotator, as well as how to browse existing annotations.

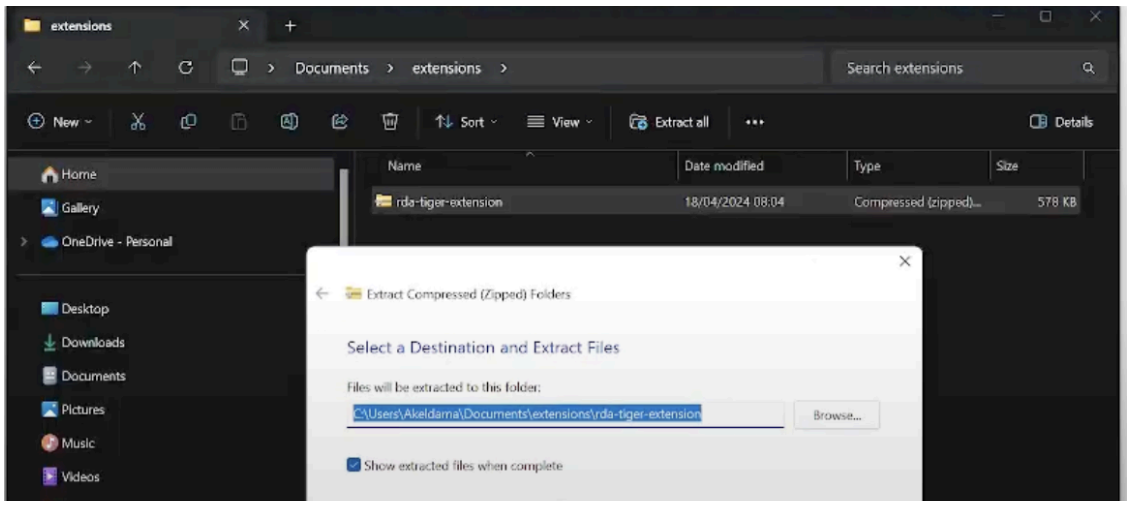
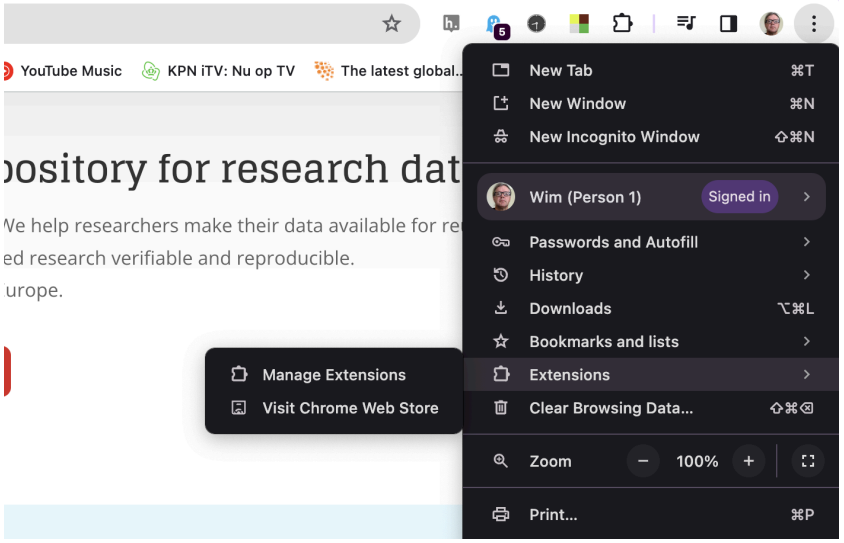
3.1. Installation

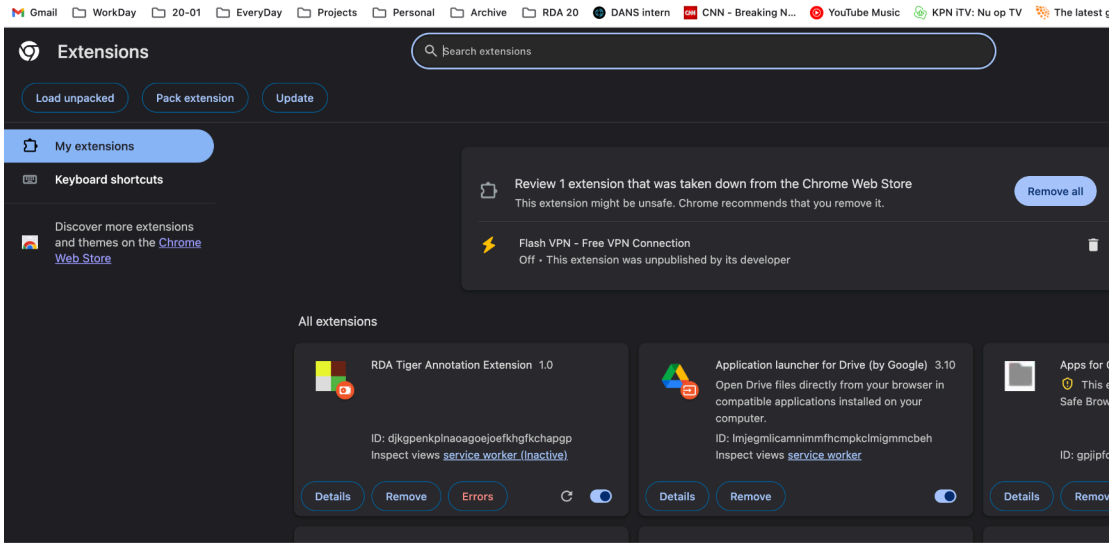
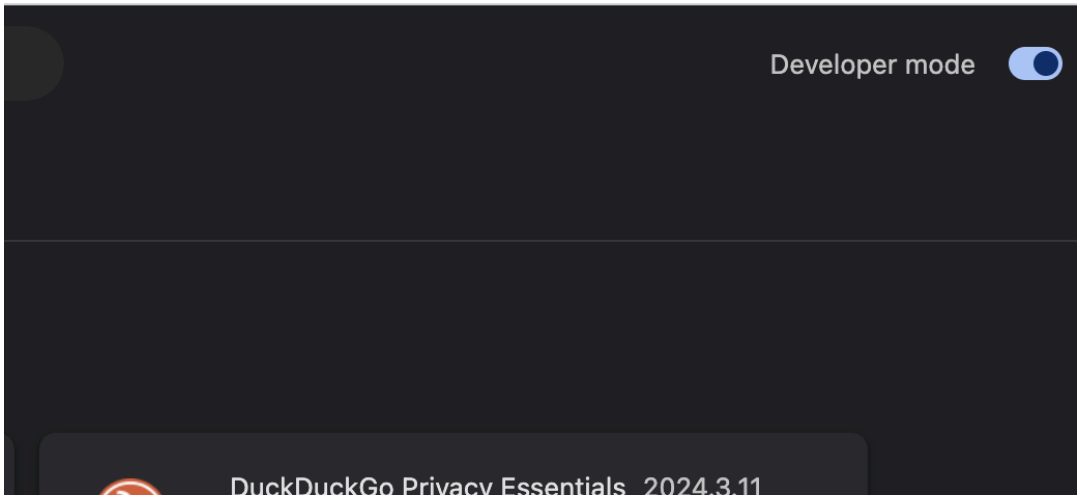
The application consists of a browser extension (also known as a plug-in application) that must be downloaded and activated in your web browser. **The extension currently supports Chromium-based browsers (e.g. Google Chrome, Microsoft Edge, Opera), and support for Firefox browsers will be added soon. The extension can be downloaded at the RDA Annotator GitHub repository. The installation process for Chromium browsers is illustrated in this video.** The table below walks you through the same process as well.

To manage the extension, web browsers generally have a settings page to manage plug-ins (as pictured in step 4) that lets you see whether the extension is activated, and what version it is.

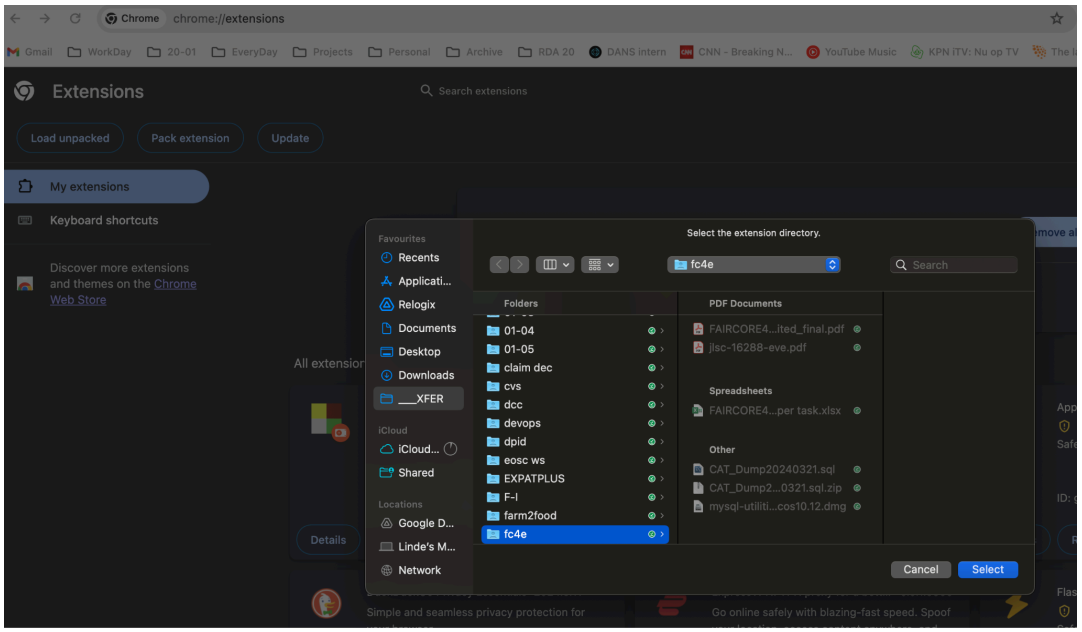
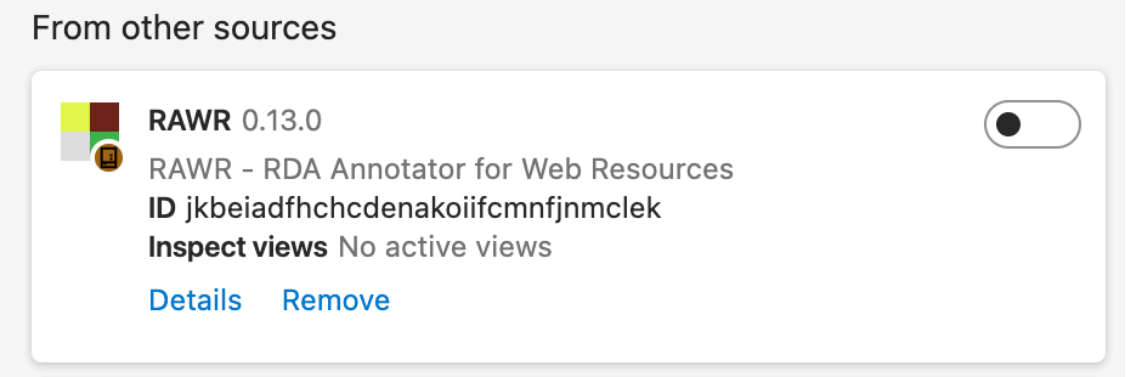
Table 1: Installation

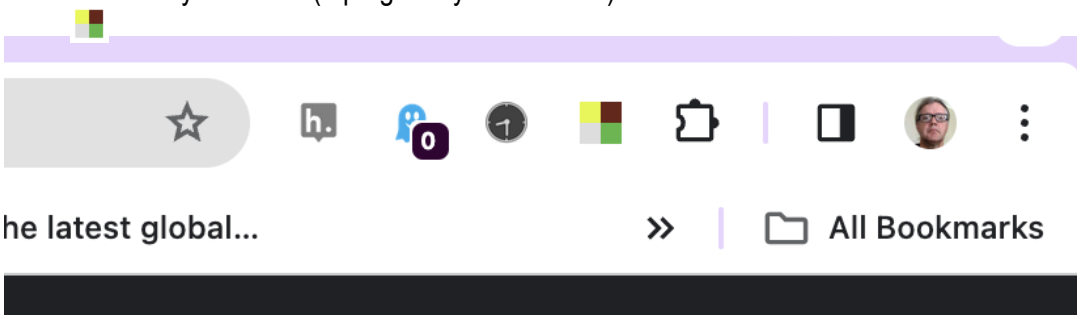
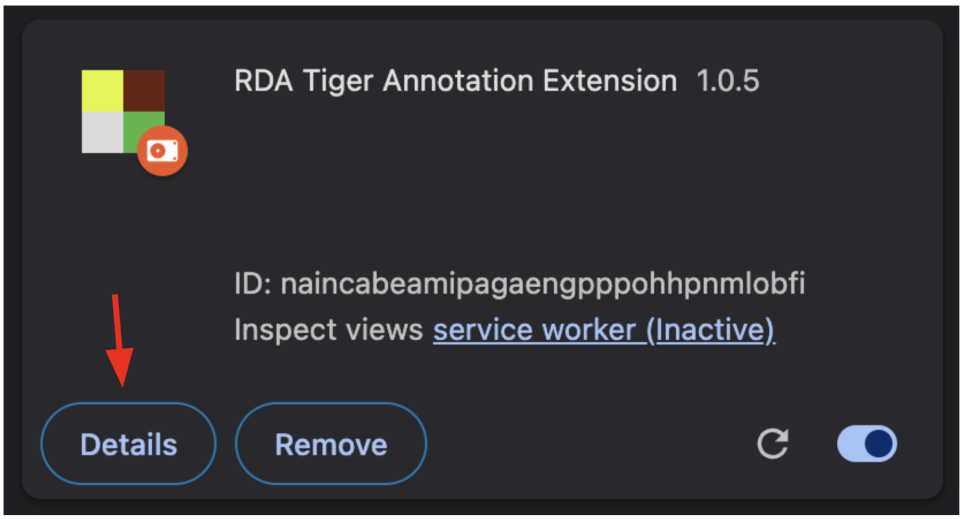
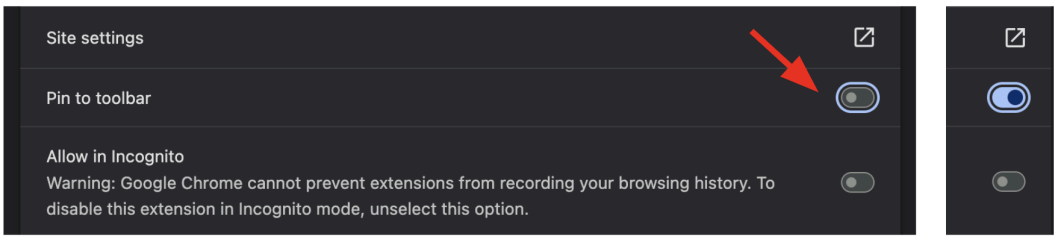
Step	Description
1	First download and expand the zip file containing the browser extension, as shown below, and then extract the files to any suitable location on your local hard drive. 

2	After expanding the zip file, the application can also be moved to a convenient permanent location on your local hard drive (it's not recommended to leave it in the download folder). Depending on the operating system, one can also specify a target for expansion of the zip file. 
3	In the next step, we install the extension in our browser. To add an extension to a **Chromium-browser**, click the settings icon at the very top right of the browser window. The settings menu will open, and in there an entry for 'Extensions' can be found. 

4	For Chromium browsers, click on ‘Manage Extensions’. A page with currently available extensions is shown, and one can add extensions here as well. 
5	For the RDA Annotator application, one will select an application from disk, and not search for it in Google Play.[3] **IMPORTANT: To load extensions from disk, the ‘Developer Mode’ has to be enabled. Once the extension is deployed in Google Play, this step will not be needed.** To enable, click the switch at the top right of the ‘Manage Extensions’ page: 

[3] Once the Plug-In is ready for production use, we will deploy it to Google Play.

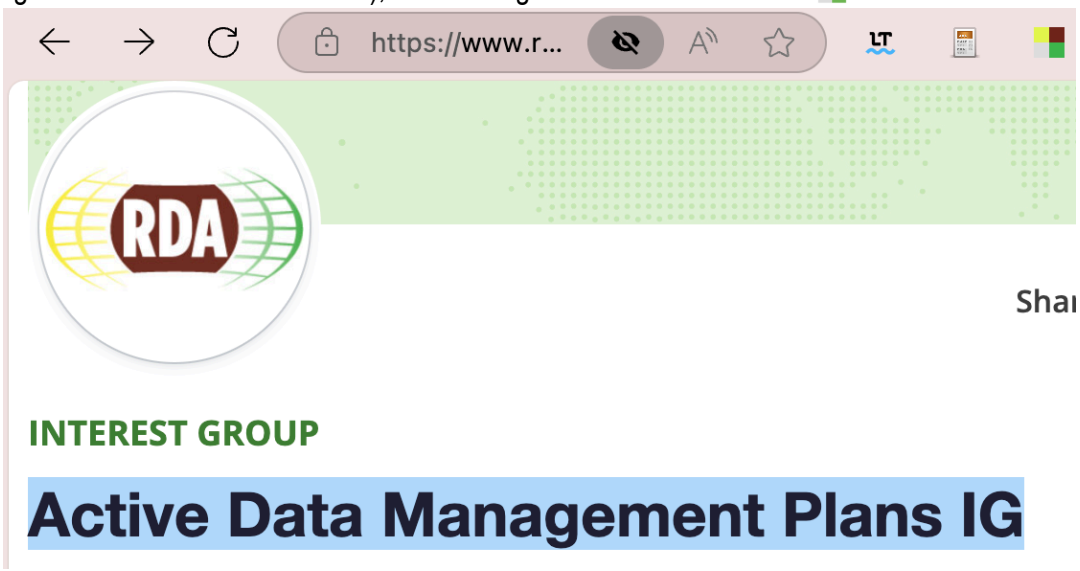
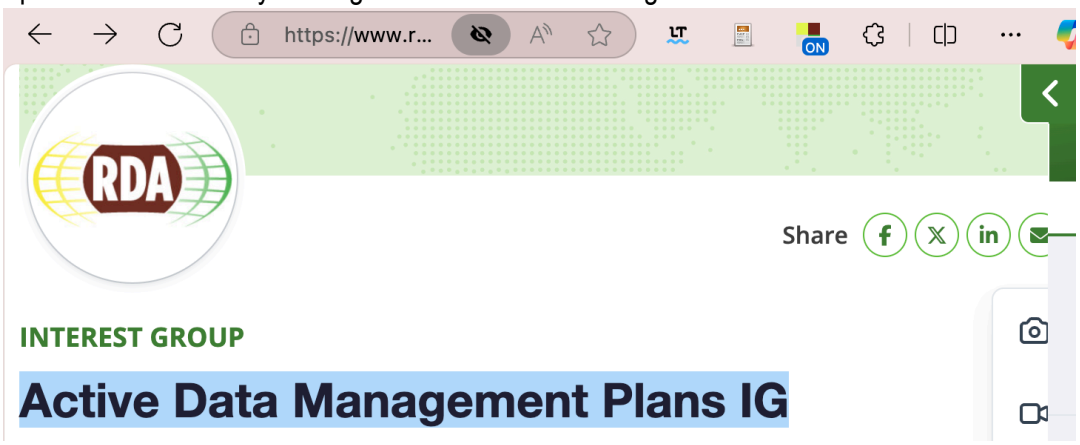
6	To load the extension, click on the ‘Load Unpacked’ button at the top left, and select the location (folder) where the application has been extracted to (steps 2, 3). 
7	Once the Plug-In extension has been registered, it will show up in the list of extensions for the browser (note: depending on the browser, it may be found under a “From other sources” category, as below). Clicking the button at the bottom right of the ‘card’ will enable it in the browser. 

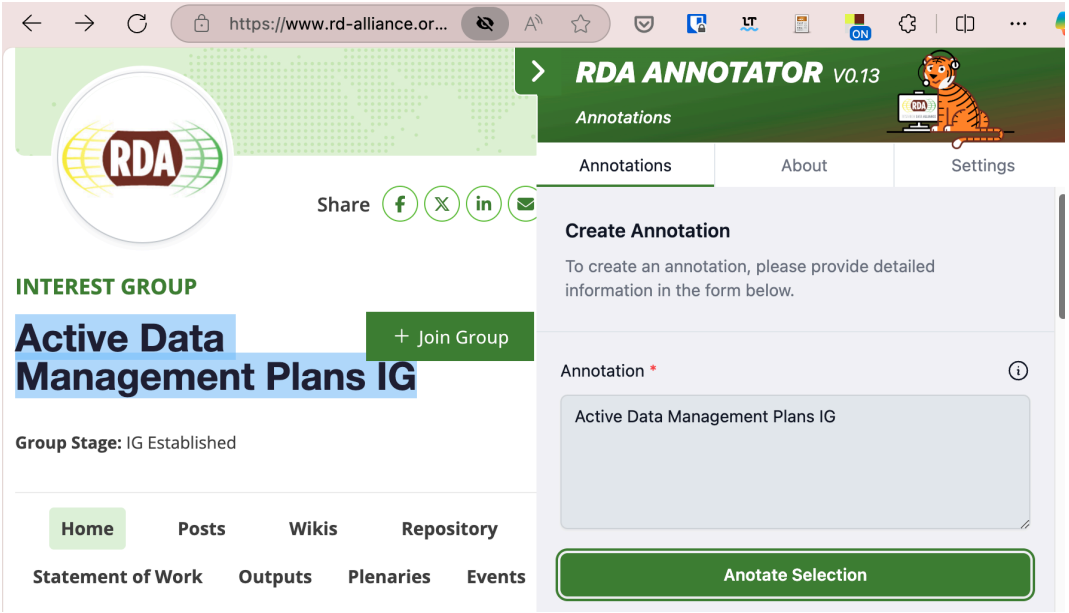
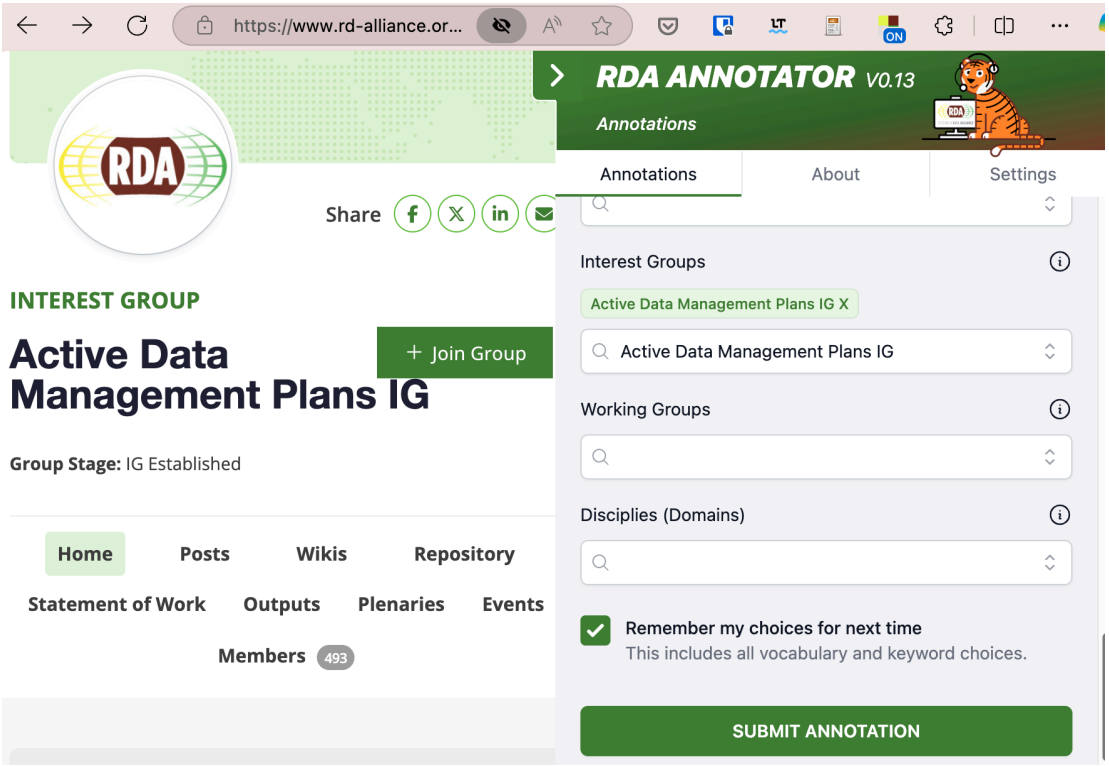
8	Once this is done, the extension can be used, but is is far simpler to pin it to the Browser Toolbar so that it is instantly available (top right of your Browser) - it will show as a small RDA icon: 
9	To pin the extension to the Toolbar, click the 'Details' button in the 'Manage Extensions Page:  

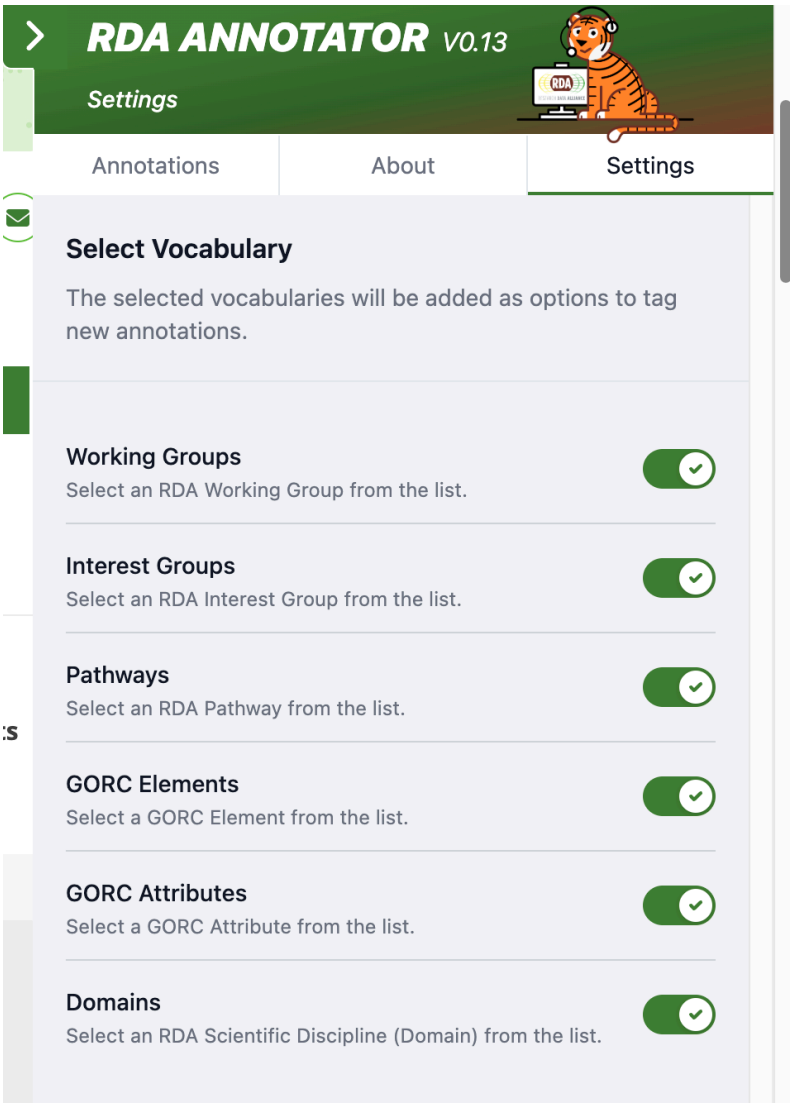
3.2. Creating and Saving Annotations

The following steps describe how to use the RDA Annotator interface to create annotations together with associated tags. For more details and good practices on how to create useful annotations, including how to use the associated vocabularies and tags in a consistent way, see section 4 below.

Table 2: Creating annotations

Step	Description
1	When you want to create an annotation, first turn on the extension by going to the toolbar (usually right of the browser address bar), and clicking the RDA Icon: 
2	Open the side menu by clicking the white arrow to the right side of the browser: 

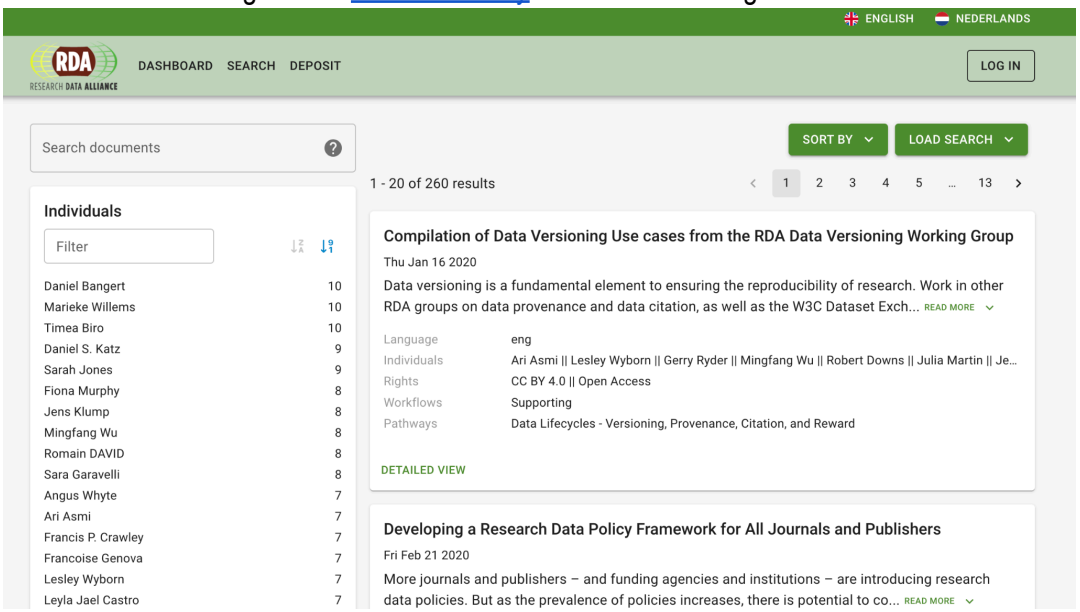
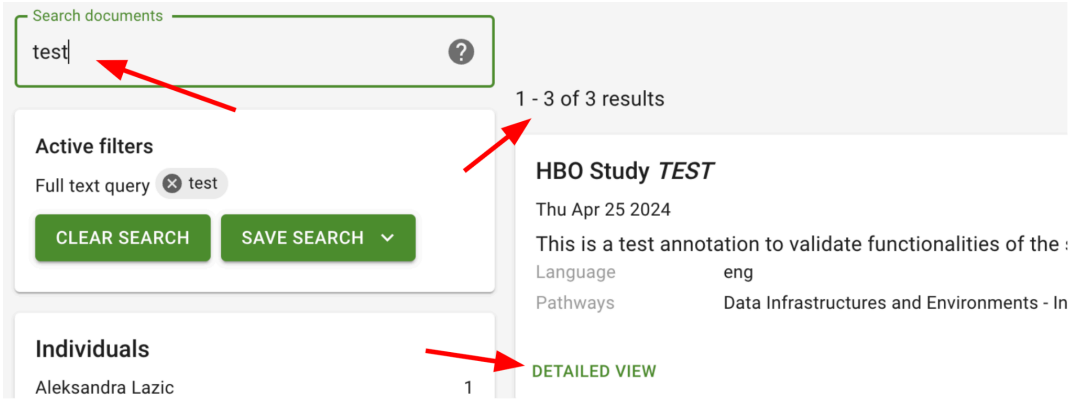
3	Select the text you want to annotate, and click “Annotate Selection” in the side panel. The selected text will now appear in the Annotation window: 
4	Fill out the (required) information, and click “Submit Annotation” at the bottom of the menu. Note that some vocabulary menus allow multiple selections. You can optionally choose to remember the choices made for the next time, which can be useful if you make several annotations on the same page. 

5	The Settings tab allows you to select which vocabularies you are able to choose from each time you make annotations. 

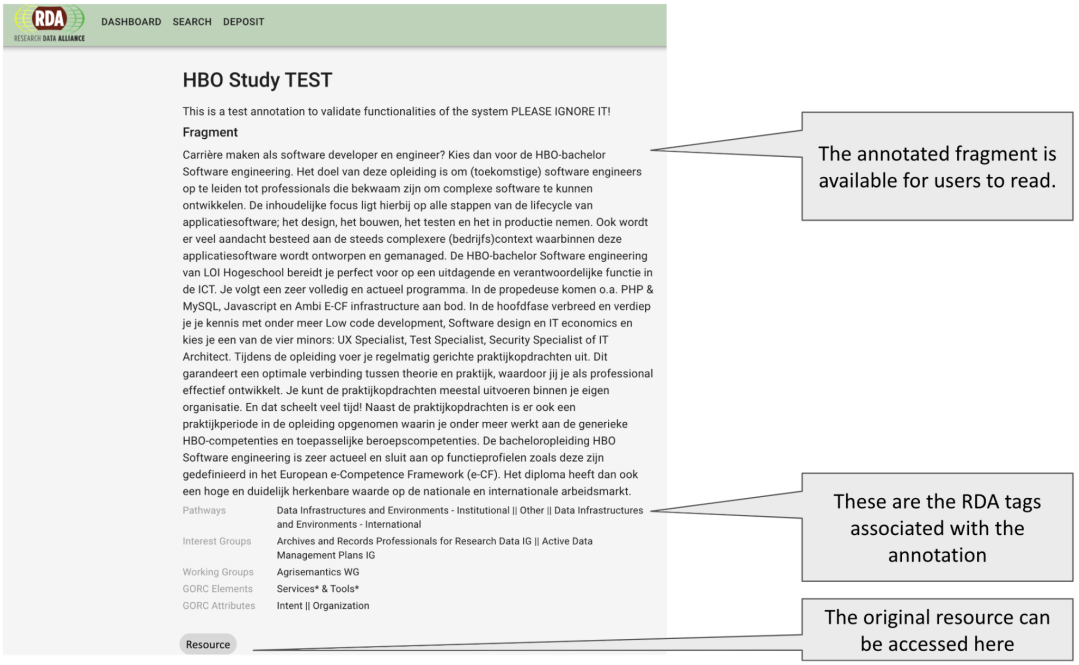
3.3 Browsing Annotations

Saved annotations are stored in the Knowledge Base, and are also made available to the wider research community via hypothes.is (see Annex A for more details). In order to browse your annotations, follow the four steps below:

Table 3: Browsing annotations

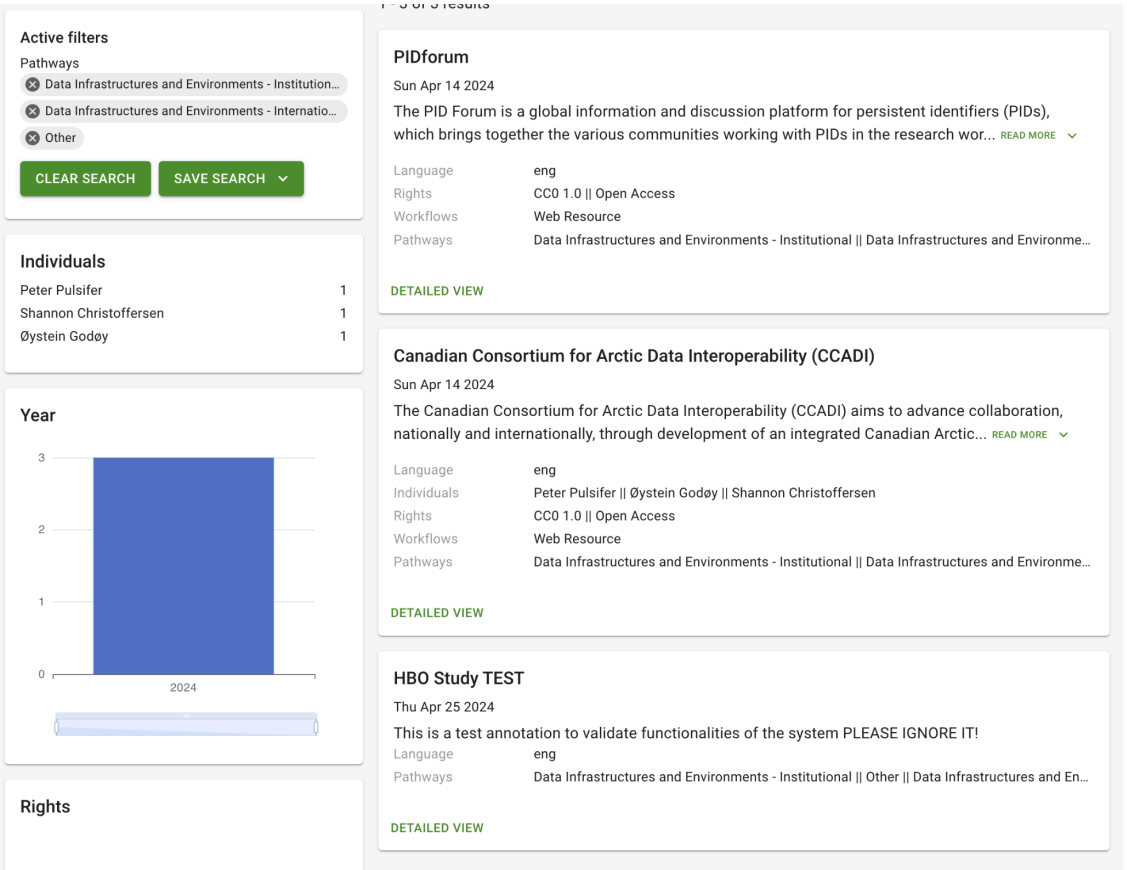
Step	Description
1	The annotations are added to the RDA Knowledge Base, these are catalogued and indexed, and made available through in the RDA Discovery UI of the Knowledge Base. 
2	To find an annotation, one can start by using a free-text search - in our example, we will search for a 'Test' entry:  The right-hand listing of matching resources is updated automatically, and the detailed entry can be seen by clicking 'Detailed View'.

3



4

The resource will also show up using one or more of the tags associated with the annotation in a filter:



4. Guidelines: Making Useful Annotations

Annotations made using this tool can be used by yourself, but **ideally they should also be useful to other RDA members**. The short guidelines in the table below describe best practices when making annotations, including how to use vocabularies and tags consistently. Note that these guidelines should be considered as advice, not as mandatory usage, however applying them can help making annotations more understandable -and therefore useful- to other RDA members.

4.1 Understanding the needs of RDA Members

As mentioned in the introduction, RDA members often go through a lot of materials in the course of their work, much of which does not get recorded in the final output, and therefore cannot be reused by others. Annotations can help RDA members **identify useful web-based content**, saving others from the often arduous process of searching through sites and documents. Naturally, what other members find useful varies greatly: based on their interests, discipline, the intended output, and in which context the work is done. **Members mostly carry out RDA-related work in groups** with the aim of producing specific solutions to data sharing, as described on the RDA website.[4]

RDA Recommendations and Outputs are the technical bridges. They are the technical and social infrastructure solutions developed by RDA Working Groups, Interest Groups or Communities of Practice that enable data sharing, exchange, and interoperability. They have an important impact on solving data sharing problems and adoption in infrastructure environments by individuals, projects and organisations.

In the course of this work, RDA group members typically need information that can:

- Help them understand a problem particular to a community, discipline, technology, or data-sharing problem
- Be used to write recommendations and other outputs able to be understood and adopted by both academic and non-academic audiences, often from various backgrounds
- Be useful in landscaping efforts, i.e. identifying and making an overview of related developments, research, organisations, individuals, etc.

[4] https://www.rd-alliance.org/how-rda-works/

4.2 Use of Annotator Fields and Vocabulary

The following table describes the different fields and vocabulary that can be used to describe annotations. Although not all fields are required, using more fields and fuller, more accurate annotations will help other users identify useful resources. See annex B for examples of annotations using the practices described below.

Table 4: Annotation best practices (* = required)

Field	Description	Best practice
Annotation *	The annotated web content that you selected (autofilled)	This field contains the text snippet that you highlighted on the page. The length of this can vary from a few words to 2-3 sentences, but should not take up several paragraphs. If a longer section of text is of interest, consider annotating a summary, title, or other snippet that describes the gist of the content
Title*	The title of the annotation	A short title that will be displayed prominently e.g. when other users search for annotations, letting understand at a glance roughly what the annotation is about
Description	Description of the annotation for all users	This public description may help yourself or others understand or use the annotation, e.g. by providing additional information and context, links to related resources, opinions about the text in question, etc.
Notes	Notes about the annotation for yourself	
Language *	The language of the annotation (from list)	You can also type in the language instead of searching the list. Only one can be selected. If multiple languages are represented - choose that of the resource/URL, and tag the other language as a keyword

Resource*	The URL of the resource (autofilled)	N/A
Resource Type*	The type of resource the website represents	Choose the type that best matches the resources (e.g. service, publication, news…). If multiple apply, consider adding keywords/tags.
Created at*	Date of annotation creation (autofilled)	N/A
Keywords/ Tags	Classifying the annotation	Applying the annotation or resource more broadly, helping others find and identify the annotation, e.g.: technology, EOSC project, or subject. Multiple choices possible
Pathways	RDA Pathways	Data-related themes or topical perspectives designed by the RDA Technical Advisory Board to help users navigate RDA Plenaries. Multiple choices possible
GORC Elements	Global Open Research Commons International (GORC) Model Elements	The GORC aims to establish a shared understanding of what a “Commons” is in the research data space. GORC Elements form the essential elements of this Commons typology. **Refer to the GORC Model summarised in annex C for more details and definitions**. This field can be used to indicate whether an annotation or associated resource fits within this Commons, e.g. if it contributes as an ICT infrastructure, a Service or Tool, or governance structure. Multiple choices possible.

GORC Attributes	GORC Model Attributes	GORC Attributes essentially refer to the measurable or documentable aspects of Elements, e.g. the Element Interoperability features the Attribute “File Formats”. Since attributes fall under elements, Attributes should ideally be selected with its associated element. **Refer to the GORC Model summarised in annex C for more details and definitions**.
Interest Group	RDA Interest Group (list)	Select all the RDA Interest Groups that relate to or are relevant to the annotation
Working Group	RDA Working Group (list)	As above: select all the RDA Working Groups that relate to or are relevant to the annotation
Disciplines (Domains)	Scientific discipline/domain	As above: select all Scientific disciplines or domains that relate to or are relevant to the annotation

References

[0] Developed in this document

[1] Hugo, Wim, & Saldner, Simon. (2023). RDA TIGER D4.1 – Description of Output Support Services and Maintenance Platform (Version 1). Zenodo. https://doi.org/10.5281/zenodo.8096631

[2] W3C (2012), Media Fragments URI 1.0 (basic), W3C Recommendation 25 September 2012, https://www.w3.org/TR/media-frags/

Appendix

A: What Happens to the RDA Annotator Data?

When using the RDA Annotator, the data is stored in the RDA Knowledge Base but also in hypothes.is.[5] In the Knowledge Base, the data is stored in a richer context, as follows:

1. The tag provided by the annotation will match one of the corresponding concepts or things that was selected by the annotator - see Annexure B. These include other nodes in the RDA Graph, such as Groups, Members, Pathways, and so on. We will use this, rather than a general node such as 'Tag', as the ***object*** of a graph relation.

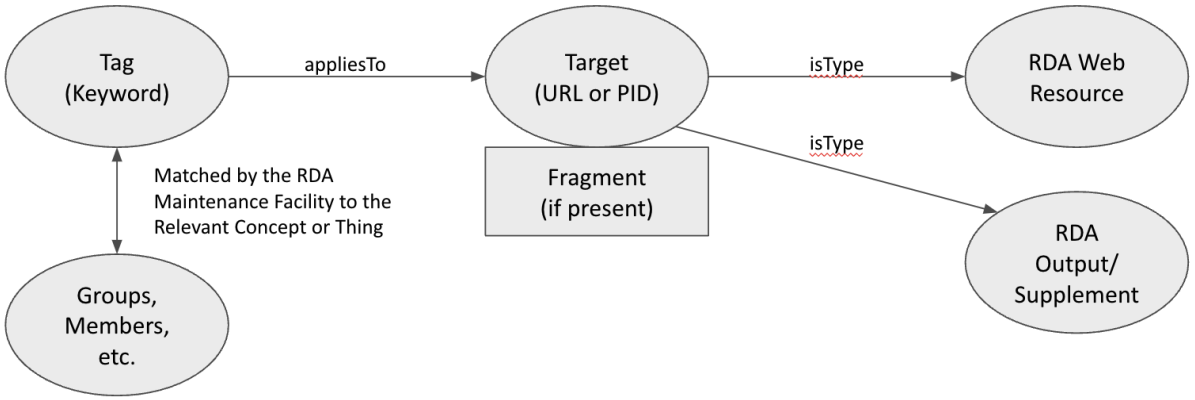
2. The object has a standard relationship with the annotation target (***subject***). In a graph, this relationship (or relation) is called the ***predicate***, and in cases of annotation, it will always be 'appliesTo'.

3. The Target is the PID or URI of the web resource that is being annotated. For HTML pages, images, SVG pages, and similar, it is mostly possible to also identify a fragment unambiguously (such as a region in an image or highlighted text in a web page) [2]. For PDFs, a specification exists for identifying fragments, but this is not consistently implemented by viewer technology that displays PDFs in web pages - in such cases, a fragment cannot be specified.

4. The annotations contributed to the RDA Graph via this channel will always reference an 'RDA Web Resource' (a general category of resources in the web).

5. In some cases, one might annotate existing RDA resources, and if the RDA Maintenance Facility can match it to an existing RDA Output or Supplement via the document URI and Title, the annotation can be linked to that resource.

[5] This allows the owner to view annotations made for RDA together with others made in a different context, and contributes to the larger body of knowledge available to the general research community in the hypothes.is graph.



B: Example annotations

1. AgroPortal. The following annotation was taken from the FAIR-IMPACT Project resources page, describing the AgroPortal website. This example has all annotation fields filled in. The screenshot below demonstrates how this annotation is displayed on the Knowledge Base website:

AgroPortal

Resource description from FAIR-IMPACT website: Vocabularies and ontologies in agri-food are spread out, in different formats, of different size, with different structures and from overlapping domains. There is need for a common platform to receive and host them, align them, and enabling their use in agro-informatics applications. AgroPortal is an ontology repository (more generally a semantic artefact catalogue) for the agri-food domain which features ontology hosting, search, versioning, visualization, comment, and recommendation; enables semantic annotation; stores and exploits ontology alignments; and enables interoperation with the semantic web. The AgroPortal specifically satisfies requirements of the agronomy community in terms of ontology formats (e.g., SKOS vocabularies and trait dictionaries) and supported features (offering detailed metadata and advanced annotation capabilities).

Fragment

AgroPortal A semantic artefact catalogue for agri-food

Pathways	FAIR, CARE, TRUST - Principles || Discipline Focused Data Issues
Interest Groups	Biodiversity Data Integration IG
Working Groups	Agrisemantics WG || Capacity Development for Agriculture Data WG || Smallholder Agriculture Data Collection and Curation Working Group || Rice Data Interoperability WG
GORC Elements	Interoperability* || Services* & Tools*
GORC Attributes	Vocabulary & Ontology
Domains	Agricultural sciences || Agricultural biotechnology

Zenodo

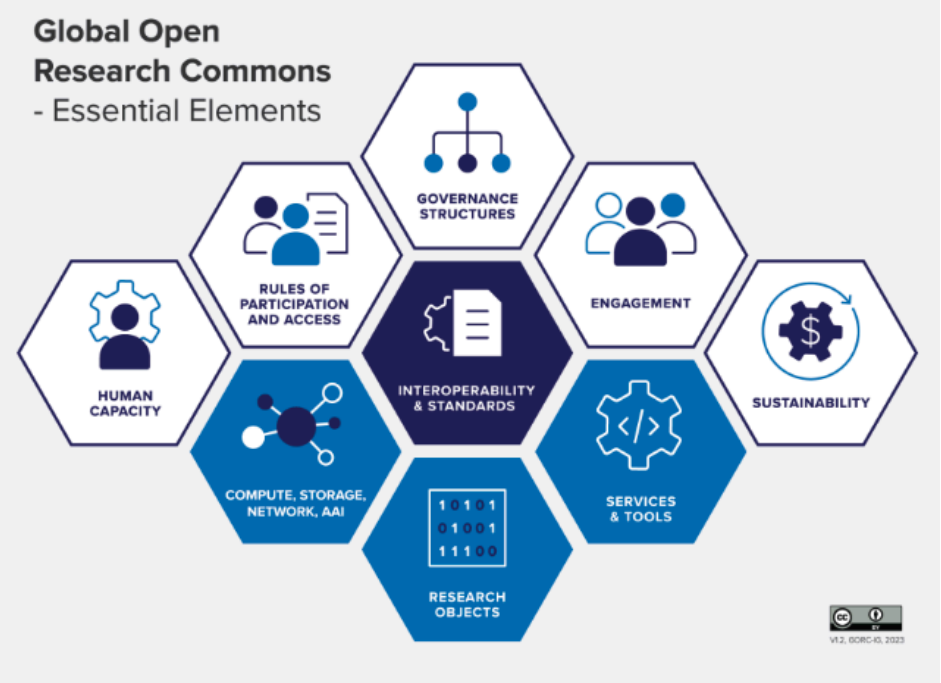
C: Global Open Research Commons International (GORC)

"Open Science Commons" or "Data commons" such as the EOSC, provide a shared virtual space or platform that provides a marketplace for data and services. The GORC Model is developed by an RDA Interest Group and multiple Working Groups aims to establish a shared understanding of what a "Commons" is in the research data space, and what functionality, coverage and characteristics such an initiative requires.

For more details about GORC Elements and Attributes, including a full list of definitions, see the latest GORC Model at DOI: 10.15497/RDA00099. Below are short summaries below extracted from this publication.

GORC Elements

This GORC Model refed to above provides the following useful graphic and summary:

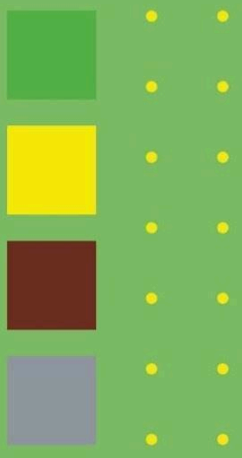


"*The three elements in blue are the underpinning digital elements that constitute the parts of the commons with which people interact. The five elements in white are the social/human elements that are needed to make the commons succeed. The central element in dark blue represents the way in which standards are at the core of a commons.*"

GORC Attributes

This GORC Model described above provides the below definition of a GORC Attribute (p.3). The Model in the above link features a full list and spreadsheet of attributes for each of the Elements listed above.

"An attribute is a *standard, characteristic, functionality or point of reference about an essential element, 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.*"



research data sharing without barriers

rd-alliance.org