A Routemap to Information Nodes and Gateways (RING)
for Agricultural Research for Development (ARD)

Handbook

What is the RING and why it was created
Using the RING portal
Getting machine-readable data from the RING
Basic concepts and technologies for interoperability
# Table of Contents

About this handbook................................................................................................................................. 3

Part 1. The RING........................................................................................................................................ 4
  1.1 What is the RING............................................................................................................................... 4
     1.1.1 Functions of the RING .............................................................................................................. 4
     1.1.2 Why the RING? ......................................................................................................................... 4
     1.1.3 Who uses the RING? .................................................................................................................. 5
     1.1.4 Expected impact of the RING .................................................................................................. 5
     1.1.5 The way forward: leveraging the RING Registry to build advanced services ...................... 5
     1.1.6 What are “information services” and “datasets” in the RING ................................................. 6
  1.2 What you can do with the RING ...................................................................................................... 9
     1.2.1 Search for information services / datasets ........................................................................... 10
     1.2.2 Learn how to exploit / create interoperable services ............................................................. 12
     1.2.3 Publish on the RING: make your sources known and re-usable ........................................... 13
     1.2.4 Have metadata about your datasets re-published in the RING ............................................. 21

Part 2. The machine-readable RING ...................................................................................................... 22
  2.1 RDF store and SPARQL endpoint .................................................................................................... 22
     2.1.1 URIs ......................................................................................................................................... 22
     2.1.2 SPARQL endpoint ................................................................................................................... 22
     2.1.3 Vocabularies ........................................................................................................................... 22
     2.1.4 Sample SPARQL queries ........................................................................................................ 23
     2.1.5 Examples of filters that can be used in SPARQL queries ..................................................... 26
  2.2 REST API .......................................................................................................................................... 28
     2.2.1 Ws_datasets ............................................................................................................................ 28
     2.2.2 Ws_datasetsbyuri ..................................................................................................................... 29

Part 3. Interoperability: standards and technologies ............................................................................. 30
  3.1 Interoperability .................................................................................................................................. 30
  3.2 Description and indexing standards ............................................................................................... 31
     Schemas / metadata models ................................................................................................................ 31
     KOS and Authority Lists .................................................................................................................... 32
     Notation, rules, encoding ................................................................................................................... 32
  3.3 Protocols .......................................................................................................................................... 33
  3.4 “Architectures” .................................................................................................................................. 33
     RSS .................................................................................................................................................... 33
     OAI-PMH .......................................................................................................................................... 34
     RDF, the Semantic Web and Linked Data .......................................................................................... 34
  3.5 Existing standards adopted / recommended in agriculture ............................................................. 37
About this handbook

This handbook provides information on the CIARD RING portal and instructions on how to use it. In the second part, the handbook gives an overview of standards and technologies for interoperability that are essential for understanding the classifications of services in the RING and for correctly registering new services. This is the fifth revision and was last updated on 19 July 2017.
1.1 What is the RING

The RING is a directory / metadata registry of agri-food information services and datasets.

The services and datasets registered in the RING are described in detail and categorized according to criteria that are relevant to the use of the service and its interoperability (such as formats, metadata standards / models adopted, subject vocabularies / code lists used, access protocols etc. - see part 2 on interoperability).

The RING facilitates the discovery of sources of agriculture-related information across the world.

While we collect information on any data service, including web catalogs and search engines, the focus of the RING is real datasets. Datasets can be registered and described using an online form, or metadata (including links to the files) can be harvested from other catalogs (especially if managed with mainstream tools like CKAN or Dataverse). In the latter case, datasets can then be further manually curated by the catalog owners.

1.1.1 Functions of the RING

- to provide a map of accessible information sources with instructions on how they can be used effectively;
- to provide a dataset sharing platform for the agri-food sector;
- to federate metadata from existing sources whenever possible and alternatively allow for manual submission and curation;
- to provide examples of services that show good practices on implementing “interoperability”;  
- to clarify the level and mode of interoperability of information sources;

1.1.2 Why the RING?

Consumers of agricultural information and data have complex information needs that require up-to-date information from different sources. They need gateways that give integrated access to several sources. Managers of information services, who should provide such gateways, have difficulties in identifying relevant sources and in re-using information from them to re-purpose it for their end-users.

The RING was created with the objective of making data sources more easily “discoverable” and to allow other services to re-use and re-package these data to make it accessible in different ways (different browsing and search options, new visualizations, different formats, different channels for different users, new combinations of data...).
1.1.3 Who uses the RING?

The RING is designed mainly for agri-food data managers and IT professionals. The main objective is to help them provide better information services. However, the RING was created with the needs of the end-users of agricultural information in mind. Consumers of agricultural information will benefit from the RING infrastructure to the extent that the RING will be exploited by information service managers in order to: a) create better interoperable sources and b) leverage existing interoperable sources to provide better integrated information systems for their users. Consumers can also use the RING as a “bookmark” list of agricultural data services.

1.1.4 Expected impact of the RING

The potential impact of the RING is not so much in the collected information itself as in what can be built out of it. Providing structured machine-readable metadata on the formats, protocols, standards and vocabularies used in each registered source will facilitate the building of applications like:

- services that offer a common browsing or searching interface to data from different sources;
- aggregation and harvesting services;
- services that re-package information and make it available in different forms (conversions, visualizations) or through different channels (text messaging, radio etc.);
- services that combine data from different sources and software tools that process data;
- services that interface the different knowledge organization systems (KOS) used by different sources;
- applications providing value-added services like digests, bibliographies, best practices, surveys etc.

1.1.5 The way forward: leveraging the RING Registry to build advanced services

Once the descriptions collected about datasets and data services are detailed and structured enough, and once the number of registered services is sufficiently large, some advanced services can be built directly on the content of the RING website. Examples of such services are:

- A global harvester of all registered OAI-PMH providers
- A viewer/navigator for registered RDF stores
- Thematic aggregators that harvest from registered RSS
- An application that reads all datasets in a specific format containing a specific type of data
- A weather data visualization tool that can read datasets in specific formats
The few advanced services that may be built in the RING are mainly for demonstrative purposes, while the actual added-value services giving access to relevant data should be implemented by the various actors that have a mandate to serve specific stakeholder groups / communities and improve their access to information and knowledge.

Fig. 2. The objective of the RING: facilitating the building of gateways like the one above

1.1.6 What are “information services” and “datasets” in the RING

In the RING, the two main sections for searching information sources are “Datasets” and “All data services”. Under All data service you will find all the services that have been registered to the RING, while under Datasets you will find only data services that make at least one collection of data available for access or download in one or more formats. The following definitions explain this distinction better.

“Data services” is a generic term for any type of data service on the web, from a simple website to a search engine to an application programmable interface to a data dump. A more technical definition of what is considered "data service" in the RING is: any platform that provides data services from one server instance (website, mail server, web services endpoint, XML archive) to any client (browsers, email clients, news readers, special protocol clients).

Any service that is registered in the RING will be listed in the “All data services” section.
“Datasets” is a more specific term that has been defined in several ways, all of which further specify or extend the basic concept of “a collection of data”. The way datasets are conceived in the RING follows the definition given by the W3C Government Linked Data Working Group to the concept of “dataset”: a dataset is “a collection of data, published or curated by a single source, and available for access or download in one or more formats”. According to the same definition, the “instances” of the dataset “available for access or download in one or more formats” are called “distributions”: a distribution is “a specific available form of a dataset. Each dataset might be available in different forms, these forms might represent different formats of the dataset or different endpoints. Examples of distributions include a downloadable CSV file, an API or an RSS feed”.

Therefore, datasets in the RING are a subset of the more generic data services and comprise only the services that make a collection of data available for access or download in one or more formats (distributions). The word “access” here has a specific technical meaning indicating machine-access at a certain address through a certain protocol, not just access through a web user interface (therefore, an online catalog search is not a dataset). In the same way “in one or more formats” here means in one or more machine-processable formats (therefore, a downloadable Word or PDF file with a list of bibliographic citations is not a dataset).

For example, a website that has a search engine where the user can interactively search and browse a collection of data is not a dataset, while the following can all be considered datasets:

- an RSS feed reachable at a URL;
- an XML dump downloadable via FTP or reachable at a URL;
- a SPARQL engine that responds to a query with an RDF response;
- an OAI-PMH target that responds to a verb call with an XML response;
- a CSV or NetCDF file available at a URL;
- any web service or API endpoint whose response is in a machine-readable format.

The template for registering a dataset and a data service is the same: the only difference is that when registering a dataset the “Access” metadata are mandatory. When users want to register a dataset, they must start registering a “Dataset / data service” and then, under the “Access to data” tab, fill in the information about the available distributions of the datasets: only records that have the “access to data” tab filled are included in the Dataset search.

Following the above definitions, any information

---

W3C Government Linked Data Working Group: [http://www.w3.org/TR/vocab-dcat/#class--dataset](http://www.w3.org/TR/vocab-dcat/#class--dataset)
DISC, Data Information Specialists Committee – UK: [http://www/disc-uk.org/qanda.html](http://www/disc-uk.org/qanda.html)
A service that is registered in the RING is listed among the generic “data services”, while only those services for which at least one “distribution of data” is available for access or download in one or more formats are listed among the datasets.

For the purposes of data sharing and re-use and the building of better information and data services, registering a service with at least one real accessible “dataset” goes a much longer way than registering just a website or an interactive search engine: data in a dataset are re-usable, data behind a search engine are not.
1.2 What you can do with the RING

Information managers may go to the RING to:

- Look for information sources (datasets in any format, RSS feeds, OA archives, RDF stores ...) that can be harvested or imported into their systems
- Look for information services (data catalogs, bibliographic databases, search engines, harvesters, aggregators...) to which they can contribute their contents
- Learn more about what others are doing, look at examples and tutorials on how to build better interoperable sources and better added-value services
- Register the datasets and services for which they are responsible in order to make them known and facilitate their discovery by other information managers
- If they have a data catalog that they want to see harvested in the RING, they can send the URL of their API (any REST API, an OAI-PMH target, a SPARQL engine) or their metadata dump and indicate in which format and under which metadata model their metadata exports are available.

Developers may go to the RING to:

- Look for the technical details about datasets that can be harvested or imported into their systems
- Get the technical details about the RDF layer of the RING and learn how to query the RING data from their applications

In the following sections you will find instructions and examples on how to use the RING for different purposes.
1.2.1 Search for information services / datasets

Services are indexed in the RING according to different criteria.

**Content criteria such as**
- thematic coverage
- geographic coverage
- content type
- target audience

**Technical criteria such as**
- schema / model adopted
- KOS\(^2\) used
- format
- protocols\(^3\) implemented

The technical criteria apply mostly to services that expose a dataset in some form, because they give indications on how the dataset can be accessed (through which protocol, in which format etc.). More information on these technical criteria can be found in Part 2 on Interoperability.

Thanks to the above indexing, services / datasets can be browsed and searched using different criteria. However, it must be noted that not all service providers always submit comprehensive information about their services, which results in some services not showing up under a category they would belong to because they were not indexed under that category by their owners. Furthermore, datasets federated from external catalogs may miss some key metadata: some attempt at semantic enrichment through keyword mapping is made, but still the quality of metadata for different records is not the same.

The main search interfaces in the RING are the “faceted” search pages for “All data services” and for “Datasets” in the main horizontal menu.

---

\(^2\) Knowledge Organization Systems: controlled lists (flat, hierarchical, or organized as an ontology) of terms or concepts, used to organize data according to subject areas, domains, geographic coverage etc. Examples are classifications, subject headings, thesauri, ontologies... See part 2 on Interoperability

\(^3\) See part 2 on Interoperability
1.2.1.1 Example: search for datasets that you can harvest and parse

Example: an information system manager is building a web application combining data on crops from different sources and his application can parse TSV files. He needs accessible datasets so he will go to the Dataset search page.

In the Dataset search page, in the “Content filter” column he can select the type of data “Agronomic data > Plants / germplasm > Crops”. In the “Technical filters” column on the right he can select the technical filter that limits the search to files in the TSV format.

Each of these filters will narrow the search until the results will be only datasets with crop data in TSV format.

1.2.1.2 Example: search for services that may want your contents

Example: an information manager manages a document repository containing resources on the subject of plant protection; the repository implements the OAI-PMH interface and
he/she is looking for service providers that may be interested in harvesting documents from the repository.
In the search page under “All data services”, he/she can limit the search to the domain “Plant Production and Protection” and to the type of service he is looking for: OAI harvester.

1.2.2 Learn how to exploit / create interoperable services

Beside a section providing tutorials and glossaries on technologies and standards for interoperability, the RING provides useful learning material also in the metadata available for each service, where the owners of the service are invited to provide instructions on how to use their service (query it, export from it, or contribute to it) and clear examples.

1.2.2.1 Instructions and examples provided by service managers

Although these fields are not mandatory, the owners of services registered in the RING are invited to provide instructions on how to use their service (query it, export from it, or contribute to it) and clear examples.
This feature is especially useful when a service offers a technologically complex interface (web services, an OAI-PMH interface, a SPARQL engine).

Example(s):

Fig. 4 Specific example provided for the AgriFeeds service in order to get a specific feed

1.2.2 Tutorials

Under “How to”, tutorials (http://ring.ciard.net/tutorials) and glossaries are becoming available.

OAI Harvesting

8 December 2010 - 1:00am — FERREIRA

The Open Archives Protocol for Metadata Harvesting (OAI-PMH) has been widely adopted as an approach to allow harvesting of metadata. Many of the CIARD RING services are using institutional or thematic repositories that expose metadata using an OAI provider. This tutorial provides a brief overview of the requests that can be made, a description of the response, and some example code which demonstrates use of an OAI Harvester.

- OAI Harvesting: Overview
- OAI Harvesting: Identify
- OAI Harvesting: ListMetadataFormat
- OAI Harvesting: ListSets
- OAI Harvesting: ListIdentifiers
- OAI Harvesting: ListRecords
- OAI Harvesting: GetRecord
- OAI Harvesting: Summary

Fig. 5 A tutorial on OAI-PMH harvesting
1.2.3 Publish on the RING: make your sources known and re-usable

1.2.3.1 Register an account

Users need to create an account in order to register new services and new organizations. Once a user creates an account, whenever he adds a new service he becomes the author of that record and will be responsible for the information contained in that record.

To create an account, click on Register in the upper right corner.

Mandatory fields are indicated by an asterisk.
The user registration form

If you have created an account but you do not remember your password, you can ask the system to send you a new temporary password and then reset your password to whatever you like.

Fig. 6 The user registration form

This control helps us prevent automated registrations by spamming systems

This information is not mandatory, but we encourage users to tell us more about their involvement with CIARD and their expertise, so that this platform can also work as a resource for finding experts.

When you’ve filled in the form, click on Create new account

Fig. 7 The link to reset your password if you forgot it

On the next page, type your username if you remember it, or your email address. You will receive a temporary login link in your mailbox.

When you receive the temporary login email, click on the link in the email body: you will be directed to a page where you can login. After clicking on the Login button, you will be directed to your Account page where you can reset your password.
1.2.3.2 Manage your submissions

When you are logged in, you have an additional navigation menu in the upper left part of the website. This menu allows you to register new services, register new organizations, review your submissions, review your account and log out.

Fig. 8 The user menu visible in the upper left part of the web page when you are logged in

If you have already submitted records (services or organizations), you can see a list of them and edit any of them by clicking on the “Your submissions” link in the left column under “Members Area”.

Fig. 9 The page where you can administer your submissions.

Clicking on the title of one of your submissions, you can see the full record. Clicking on Edit, you can edit the record: clicking on it will open the editing form, which is identical to the input form described below in chapter 1.2.3 (Register an information source / service)
1.2.3 Register an information service / dataset

After logging in, you can click on “Add a service / dataset” in the user menu in the upper left part of the web page to register a new information service or dataset. Before registering a new service, please check if the service has already been registered by searching for it using the fulltext search box in the “All info services” page: http://ring.ciard.net/services-datasets

What to register in the RING
Please refer to chapter 1.1.6 What are "information services" and "datasets" in the RING for an explanation of what type of services you can register in the RING and how the RING distinguishes between generic information services and datasets / data distributions.

Who can register services in the RING
Anybody can register new services in the RING, provided that they know the necessary administrative and technical information about the service and they are responsible for the service or authorized by a responsible person. User registration is required in order to trace the source of information and allow users to edit their submission in the future. (For instructions on registration, see previous chapter)

How to register an information service

1. Click on "Add a service / dataset" in the user menu in the upper left part of the web page.

2. Start filling in the mandatory information in the form. The form consists in a number of “tabs”: the tabs group fields in different sections in order to make the registration process easier.

Only a few fields are mandatory, however please fill in as many as possible in order to make your service easily searchable in the directory.

Upon registration of the service, please fill in at least the basic mandatory fields under the first two tabs below: "Basic " and "Geo ", then you can come back later and fill in the remaining fields. However please note that without detailed information on the type of service, type of data, technical accessibility, your service will not be easily found. Even if at different stages, please fill in as many fields as possible in order to make your service easily searchable in the directory.

Before filling in the sections on Standards and Access to data, please consider that you may need the assistance of the technical staff responsible for the service.
Instructions for individual fields

1. Name of the service

When filling the "name" field, please avoid using an article at the beginning. (Remember to first search the system to see if the service has been already registered)

3. Basic information

Use this button to add more organizations if the service is already in the system: if it is, the full name will appear. Otherwise, click on the plus button on the right to add your organization.
4. Geographic location and scope

This is the country where the service is maintained.

This is the geographic coverage of the service: it can be global, regional (in which case select the region) or national (in which case select the country).

5. Thematic information

Start typing some words that identify the topic covered by your service: the system will give you a list of AGROVOC terms that contain the letters you are typing. Using AGROVOC keywords is highly recommended.
6. Type and content

Start typing to see if the data type has already been used. Please try also different terms and whenever possible use a term that has already been used: this will improve the searching of services.

7. Standards

You can select more than one item by pressing Ctrl and clicking on the various items with the mouse.
8. Dataset data distribution(s)

This is very important. Please provide here only URLs that return data in machine-readable formats (XML, CSV, RDF/XML, Json...)

Click here to expand the technical section and provide more technical information. See the following figure.

For each service, you can register as many dataset / data distributions as you like. For instance, for a document repository, you may have an OAI-PMH target and also an RSS feed: click on Add another item to register another distribution for your service.

These three fields here are essential to allow external services to get data from your dataset: please get a technical person who knows your system to provide this information. For more information on these technical parameters, see Part 2 on interoperability.
9. Submit the service

While you fill in the data under the various tabs, you always have the **Save button** visible at the bottom of the form. You can save your record at any time, provided that you have filled in the mandatory information (the system will warn you if mandatory information is missing). Again, it is highly recommended to go through all the tabs and provide as much information as possible beyond the mandatory fields.

### 1.2.4 Have metadata about your datasets re-published in the RING

If you have a data catalog that you want to see harvested in the RING, you can send us the URL of your API (any REST API, an OAI-PMH target, a SPARQL engine) or your metadata dump and indicate in which format and under which metadata model their metadata are available.

The procedures for importing from the CKAN and Dataverse data repository tools are already implemented, so if you use one of these tools just give us the base URL of the catalog and we will import the metadata.

Otherwise, the procedure for federating external data catalogs is different for each catalog. Write to **valeria.pesce@fao.org** to agree on the procedure.
Part 2. The machine-readable RING

2.1 RDF store and SPARQL endpoint

The RING database is also an accessible RDF store. An RDF store is a way of storing data using a machine-readable "grammar" (the Resource Description Framework) and documented semantics (RDF vocabularies).

### 2.1.1. URIs

On each service page, there is a link to its RDF representation. The RDF vocabulary that has been chosen to describe information services, and especially datasets, is DCAT, which was conceived to describe datasets and the forms in which they are exposed, their "distributions".

The URI for each service is built as follows: `http://ring.ciard.net/node/{service-ID}`.

For example, the URI of the AgriFeeds service in this RDF store is:

http://ring.ciard.net/node/2417

Here you can find a list of the URIs of all the RING entities (services/datasets, organizations, KOSs etc.): `http://ring.ciard.net/entity-uris`.

Here you can find a list of the URIs of all RING concepts (countries, topics, regions, protocols etc.): `http://ring.ciard.net/concept-uris`.

### 2.1.2 SPARQL endpoint

The whole RDF store is accessible through a SPARQL engine, which means that any system can run remote queries and get the resulting triples.

The endpoint of the SPARQL engine is:

http://ring.ciard.net/sparql1

### 2.1.3. Vocabularies

The vocabularies used in the RDF store are:

- RDF: http://www.w3.org/1999/02/22-rdf-syntax-ns#
- RDFS: http://www.w3.org/2000/01/rdf-schema#
- DC: http://purl.org/dc/terms/
- SKOS: http://www.w3.org/2004/02/skos/core#
- DCAT: http://www.w3.org/ns/dcat#
- ADMS: http://www.w3.org/ns/adms#
- DCAT-EXT: http://vocabularies.aginfra.eu/dcatext#
- DOAP: http://usefulinc.com/ns/doap#
- FOAF: http://xmlns.com/foaf/0.1/
- VCARD: http://www.w3.org/2006/vcard/ns#
- SCHEMA: http://schema.org/
- VOID: http://rdfs.org/ns/void#
- WS: http://vocabularies.aginfra.eu/ws#

The core vocabularies for information services and datasets among the above are:

- The W3C Data Catalog Vocabulary (DCAT)
The Asset Description Metadata Set (ADMS) extension of DCAT

The DCAT Application profile\(^4\) for data portals in Europe (DCAT-AP) is a specification based on the Data Catalogue vocabulary (DCAT) for describing public sector datasets in Europe. It combines the two vocabularies above, plus classes and properties from Dublin Core, SKOS and Vcard, in an Application profile.

Figure: the DCAT Application Profile RDF model diagram.

### 2.1.4 Sample SPARQL queries

Here are some examples of SPARQL queries that you can execute for testing the endpoint: the response is in RDF/XML (you can change this by adding the output parameter to the SPARQL call, see the SPARQL documentation). The queries below are DESCRIBE queries, so the resulting RDF contains ALL the triples containing the requested entities (dataset, distribution, owner, contact, topic, standard, format, protocol) as subject. The queries can be simplified by removing any of these entities when no properties of these entities are needed in the query results.

**Get all datasets available through OAI-PMH**

*Note: http://ring.ciard.net/taxonomy_term/108 is the URI of the OAI-PMH protocol in the RING: see all URIs of protocols here: http://ring.ciard.net/concept-uris/rdf?vid_1=6.*

Query:

\(^4\)The elaboration of the DCAT-AP was a joint initiative of DG CONNECT, the EU Publications Office and the ISA Programme. See [https://joinup.ec.europa.eu/asset/dcat_application_profile/asset_release/dcat-application-profile-data-portals-europe-final](https://joinup.ec.europa.eu/asset/dcat_application_profile/asset_release/dcat-application-profile-data-portals-europe-final)
Get all datasets available through OAI-PMH using the AGRIS AP metadata set

Note: http://ring.ciard.net/taxonomy_term/108 is the URI of the OAI-PMH protocol in the RING: see all URIs of protocols here: http://ring.ciard.net/concept-uris/rdf?vid_1=6; http://ring.ciad.net/node/19244 is the URI of the AGRIS AP metadata set in the RING: see all URIs of metadata sets here: http://ring.ciad.net/entity-uris?type_1%5B%5D=metadata_set.

Query:

Get all entities of a certain type in the RING (e.g. all standards)

Note: useful for instance to get values – URI and names - for a select list that will be used to then add filters to a query sent to the RING.

Example for entities of type “Schema” and “KOS”: rdf:type is dc:Standard:

Example for entities of type “Organization”: rdf:type is foaf:Organization:
Get all terms in a taxonomy in the RING (e.g. countries or data types)

Note: useful for instance to get values – URIs and names - for a select list that will be used to then add filters to a query sent to the RING.

For all taxonomy terms, the rdf:type is skos:Concept.

Example for data types: these concepts are in the ConceptScheme 4 (Data types):

Example for countries: these concepts are in the ConceptScheme 13 (Countries):

List of taxonomies / concept schemes:

Types of data: http://ring.ciard.net/taxonomy_vocabulary/4
Countries: http://ring.ciad.net/taxonomy_vocabulary/13
Formats / notation: http://ring.ciad.net/taxonomy_vocabulary/7
Licensing / access: http://ring.ciad.net/taxonomy_vocabulary/8
Protocols: http://ring.ciad.net/taxonomy_vocabulary/6
Languages: http://ring.ciad.net/taxonomy_vocabulary/15
Regions: http://ring.ciad.net/taxonomy_vocabulary/14
GFAR regions: http://ring.ciad.net/taxonomy_vocabulary/17
Types of institutions: http://ring.ciad.net/taxonomy_vocabulary/10
AGROVOC terms in the RING: http://ring.ciad.net/taxonomy_vocabulary/18

Get AGROVOC concepts used to tag datasets

Note: The results of this query will contain also the original Agrovoc URI. Useful for instance to get values – original URIs and names - for a select list that will be used to then add the selected AGROVOC URI to a query sent to the RING.

Query:

2.1.5 Examples of filters that can be used in SPARQL queries

You can add as many filters as you need to your query. Filters in SPARQL queries are normally found in the WHERE clause and they are represented by those “triples” (triples in queries are series of three “strings” delimited by dots) containing a variable (starting with ?), a property (a prefix plus the property name, e.g. dc:publisher) and a value (either a string or a URI) (the other triples, containing variable + property + variable, are not filters).

The examples in the previous chapter include one or two filters (protocol the first one and protocol + metadata set the second one). You can use the same base query structure above and replace the filters or add other filters. Remember to separate filters with dots.

**Base query**
(with no filters, it would return all datasets)

NOTICE: add your filters BEFORE the first OPTIONAL clause.

```

**Examples of useful filters.**

Some filters have the ?dataset as subject, others have the distribution (?distro) as subject: if you use the base query above, you don't need to worry about defining these variables, they are already defined by the query.

Some filters consist of just a triple, others need two triples because they have to relate an internal URI to an external URI using owl:sameAs.

**Protocol**

Add a filter to get only datasets that contain a specific data type:

```
?distro adms:representationTechnique <URI of data type>
```

E.g. for "OAI-PMH":

```
```

All the URIs for the various protocols can be found here:

```
http://ring.ciard.net/concept-uris/rdf?vid_1=6
```
Schema
Add a filter to get only datasets that comply with a specific schema or data standard:
?distro dc:conformsTo <URI of data type>

E.g. for “AGRIS AP”:
?distro dc:conformsTo <http://ring.ciad.net/node/19244>

All the URIs for the various metadata sets can be found here:
http://ring.ciad.net/entity-uris?type_1%5B%5D=metadata_set

Data type
Add a filter to get only datasets that contain a specific data type:
?dataset dc:coverage <URI of data type>

E.g. for crops data:
?dataset dc:coverage <http://ring.ciad.net/taxonomy/term/2262>

All the URIs for the various data types can be found here:
http://ring.ciad.net/concept-uris/rdf?vid_1=4

Topic (using AGROVOC URI)

?dataset dcat:theme ?topic . ?topic owl:sameAs <AGROVOC URI>

E.g. for “Livestock” (AGROVOC URI: http://aims.fao.org/aos/agrovoc/c_4397)

All the URIs for AGROVOC terms used in the RING and the corresponding original AGROVOC URIs (which you need for the query above) can be found here:
http://ring.ciad.net/concept-linked-uris?vid_1=18

Country (using RING URI or Geopolitical ontology URI)

Using RING URIs:
?dataset dc:spatial <URI of country>

Using Geopolitical Ontology URIs:
?dataset dc:spatial ?country . ?country owl:sameAs <Geopolitical Ontology URI of country>

E.g. for “China”:
- Using the RING URI:
- Using the Geopolitical Ontology URI:

All the URIs for countries used in the RING and the corresponding Geopolitical Ontology URIs can be found here: http://ring.ciad.net/concept-linked-uris?vid_1=13

Data owner
The RING Handbook

?dataset dc:publisher <URI of organization>

E.g. for “Global Forum on Agricultural Research”
?dataset dc:publisher <http://ring.ciard.net/node/2753>

All the URIs for organizations registered in the RING can be found here: http://ring.ciard.net/entity-uris?type_1%5B%5D=organization

2.2 REST API
(Still under development)
The CIARD RING exposes a simple REST API to get metadata about the datasets. It allows you to apply several filters so that you get only the results you need.

This API currently has two methods:

2.2.1 ws_datasets

API call: http://ring.ciard.net/api/ws-datasets?
This API will return an XML response with the resulting records.
The XML for each record is as follows:

```
<item>
  <dc_title>GFAR document repository</dc_title>
  <dc_conformsTo>AGROVOC Multilingual agricultural thesaurus, GFAR Categories</dc_conformsTo>
  <notation>N-Triples</notation>
  <dc_subject>agriculture, Agricultural education, extension activities</dc_subject>
  <dc_identifier>2754</dc_identifier>
</item>
```

Parameters:
- freetext={free-text search terms}
- owner={ownerID}
  [look up owners IDs at: http://ring.ciard.net/entity-uris?type_1%5B%5D=organization]
- metadataset={metadataset ID}
  [look up metadata sets IDs at: http://ring.ciard.net/entity-uris?type_1%5B%5D=metadata_set]
- kos={KOS ID}
  [look up KOS IDs at: http://ring.ciard.net/entity-uris?type_1%5B%5D=kos]
- protocol={protocol ID}
  [look up protocols IDs at: http://ring.ciard.net/concept-uris/rdf?vid_1=6]
- notation={notation ID}
  [look up notations IDs at: http://ring.ciard.net/concept-uris/rdf?vid_1=7]
- fao-topics={domain ID}
  [look up domains IDs at: http://ring.ciard.net/concept-uris/rdf?vid_1=2]
- Agrovoc={AGROVOC term local ID in the RING}
  [look up AGROVOC local IDs at: http://ring.ciard.net/concept-uris/rdf?vid_1=18]
  (see below for the AGROVOC URI filter option)
- country={country local ID in the RING}
  [look up countries IDs at: http://ring.ciard.net/concept-uris/rdf?vid_1=13]
- region={region local ID in the RING}
  [look up regions IDs at: http://ring.ciard.net/concept-uris/rdf?vid_1=14]
• network={network ID}
  [look up networks IDs at: http://ring.ciard.net/entity-uris?type_1%5B%5D=network]
• datatype={data type ID}
  [look up data types IDs at: http://ring.ciard.net/concept-uris/rdf?vid_1=4]

Examples:

Get all information services and datasets registered by GFAR:
http://ring.ciard.net/api/ws-datasets?owner=2753

Get all information services and datasets about “Animal Production and Health”
http://ring.ciard.net/api/ws-datasets?domain=2

Get all information services and datasets managed in Kenya:
http://ring.ciard.net/api/ws-datasets?country=331

Get all information services and datasets that use AGROVOC terms:
http://ring.ciard.net/api/ws-datasets?kos=19103

### 2.2.2 Ws_datasetsbyuri

http://ring.ciard.net/api/ws-datasetsbyuri?

Parameter:
  • agrovocuri={AGROVOC term’s URI}

Example:
### Part 3. Interoperability: standards and technologies

#### 3.1 Interoperability

“Interoperability is a feature both of data sets and of information services that give access to data sets. When a data set or a service is interoperable it means that data coming from it can be easily “operated” also by other systems. The easier it is for other systems to retrieve, process, re-use and re-package data from a source, and the less coordination and tweaking of tools is required to achieve this, the more interoperable that source is.

Interoperability\(^5\) ensures that distributed data can be exchanged and re-used by and between partners without the need to centralize data or standardise software. Some examples of scenarios where data sets need to be interoperable:

- transfer data from one catalog to another;
- harmonize different data and metadata sets;
- aggregate different data and metadata sets;
- virtual research environments;
- creating documents from distributed data sets;
- reasoning on distributed datasets;
- creating new information services using distributed data sets.”\(^6\)

Information sources need to become more easily “discoverable” and to allow other services to semantically query, re-use and re-package their information; in other words they must become truly interoperable and allow for easy automatic retrieval of information, while work on mapping between vocabularies or advanced natural language processing must be done to improve the semantic accessibility of information.

Standards and technologies that improve interoperability are described below. These standards and technologies are used in the RING as indexing criteria in order to make it easy for information professionals to quickly identify sources and services that adopt them.

In the following chapters, standards and technologies are roughly classified under Description and indexing standards (metadata sets, KOS, rules and encoding), Protocols and Architectures, but since there is no agreed way of really defining and classifying these techniques and users may find a granular and exact classification too rigid and artificial, in the RING the organization is slightly different:

- **Schemas / models / metadata sets**
  These include metadata sets that have been formalized in any way (schemas, definition files, namespaces) and the reference list comes from the Agrisemantics Map of Agri-food Data Standards\(^7\) (former VEST Registry in the Agricultural Information Management Standards [AIMS] portal).

---


\(^6\) From “Building the CIARD architecture for data and information sharing. Background Note for the e-Consultation April 4 - 15”

\(^7\) [http://vest.agrisemantics.org](http://vest.agrisemantics.org)
• **KOS**

  “Knowledge Organization Systems”: classifications, subject headings, thesauri, ontologies used to index data. The reference list comes again from from the Agrisemantics Map of Agri-food Data Standards.

• **Technology / architecture**

  This classification covers what in the following chapter is described under both Protocols and Architectures. This is an open list (a free-tagging list), which at the moment includes OAI-PMH, RESTful web services, RSS, SOAP, SRW/SRU, XML-RPC.

• **Format / notation**

  This refers to what in the following chapter is described under “Notation / serialization”. Since there is no reference list for such “formats”, this started as an open list (a free-tagging list), but is now being better curated as a standard list and is mapped to existing standards whenever possible (e.g. IANA types or W3C formats).

### 3.2 Description and indexing standards

<table>
<thead>
<tr>
<th>Schemas / metadata models</th>
</tr>
</thead>
</table>

**Metadata**

“Data about data”: the elements that describe an entity of a specific type, e.g. for a person:

<table>
<thead>
<tr>
<th>First name: John</th>
</tr>
</thead>
<tbody>
<tr>
<td>Last name: Doe</td>
</tr>
<tr>
<td>Country: United States of America</td>
</tr>
</tbody>
</table>

Agreeing on a metadata set means agreeing on a common set of elements to exchange information of a certain type.

**Metadata vocabularies**

Formalization of a metadata set in a series of agreed “property names” for metadata elements, e.g. for a person:

<table>
<thead>
<tr>
<th>given_name: Valeria</th>
</tr>
</thead>
<tbody>
<tr>
<td>family_name: Pesce</td>
</tr>
<tr>
<td>country: Italy</td>
</tr>
</tbody>
</table>

Vocabularies allow machines to share metadata using the same “labels” for metadata properties.

**Namespaces**

Metadata elements only have a specific meaning within the vocabulary where they were created; these vocabularies are defined in “namespaces” and elements must associated with a namespace in order to have some meaning:

- dc:date indicates the “date” element in the Dublin Core namespace (shortened in the dc: prefix: mappings between namespace URIs and prefixes must be given to machines)
- foaf:given_name indicates the “given_name” element in the AgMES namespace (shortened in the ags: prefix)
Namespaces are needed in order to avoid duplication of element names and misinterpretation

<table>
<thead>
<tr>
<th>KOS and Authority Lists</th>
</tr>
</thead>
</table>
| Knowledge Organization Systems are controlled lists (flat, hierarchical, or organized as an ontology) of terms or concepts, used to organize data according to subject areas, domains, geographic coverage etc. Examples are classifications, subject headings, thesauri, ontologies...
| Authority lists (which include KOS) are more in general controlled lists of “entities” (e.g. journal titles, author names, geographic names/codes, corporate body names...) used to identify an entity univocally. |

<table>
<thead>
<tr>
<th>Notation, rules, encoding</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Notation, “serialization”, format</strong></td>
</tr>
<tr>
<td>Metadata can be “serialized” in different notations of formats:</td>
</tr>
<tr>
<td>Examples:</td>
</tr>
<tr>
<td>• CSV comma separated values (-&gt;Excel)</td>
</tr>
</tbody>
</table>
| • XML

```
<firstname>Justin</firstname>
<lastname>Chisenga</lastname>
<country>Ghana</country>
```

<table>
<thead>
<tr>
<th>first_name, last_name, country</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valeria, Pesce, Italy</td>
</tr>
<tr>
<td>Justin, Chisenga, Ghana</td>
</tr>
</tbody>
</table>

• Microformats
• Json
• ...

**“Rules”: structure, data types and encoding**

Vocabularies can be defined in specific “definition files” (DTDs, XML schemas, RDF schemas...) that provide machine-readable **rules for structure, data types and encoding**, e.g. the nesting of the “ags:locationCountry” element inside the “ags:location” element, or ISO encoding for countries and languages, specific date formats etc.

```
<ags:location>
  <ags:locationCountry>GHA</ags:locationCountry>
</ags:location>
```

(With the advent of RDF, the structure has become standardized and “structure” doesn’t indicate the nesting anymore but the class / property model of an RDF schema.)

**Notes on metadata sets and interoperability**

Recently, the advent of RDF and the common use of technologies for converting metadata (stylesheets, parsers, converters) has shifted the focus from format/structure (notation, vocabulary) to values (the actual values exposed through metadata): the use of the same data format and structure has become less essential for exchanging information, while the use of agreed encoding standards and authority lists for the
values has become more and more important: for integrating a source in a service, it is essential that the necessary data are there and that they are encoded in an agreed machine-readable way.

### 3.3 Protocols

“A communications protocol is a formal description of digital message formats and the rules for exchanging those messages in or between computing systems and in telecommunications. Protocols may include signaling, authentication and error detection and correction capabilities. A protocol describes the syntax, semantics, and synchronization of communication and may be implemented in hardware or software, or both”

In order not to create too many classifications in the RING and risk confusing users between technologies, protocols and architectures, the RING only has one classification for protocols and what we call “architectures” below.

Creating a reference list for protocols and architectures is difficult also because some protocols leverage in turn other existing protocols and some architectures also define protocols: for instance, in the case of web-based services like those registered in the RING, the basic communication protocol is always HTTP, which is used by all the others (OAI-PMH, RESTful web services, SOAP, SRW/SRU, XML-RPC, SPARQL) and for instance SRW is a sub-specification of SOAP and SRU and RSS are forms of RESTful web services).

### 3.4 “Architectures”

Architectures in this context can be defined as “technological frameworks” designed to work thanks to a combination of standards and technologies of different types, e.g. combining a protocol with one or more metadata sets and one or more notations.

---

**RSS**

RSS stands for Really Simple Syndication. An RSS feed is a file that exposes syndicated contents from a website (or any source) in a way that can be read by RSS readers. RSS can be considered an “architecture” in that it is built from several standards and technologies:

- it uses a metadata set defined in the RSS namespace (actually, two namespaces, RSS 1.0, which implements RDF, and RSS 2.0): in the RING this metadata set is included in the “Metadata sets” reference list;
- it is serialized as XML;
- in version 1.0, it implements RDF;
- it is a form of RESTful webservice.

Basic RSS 2.0 feed record:

```xml
<item>
  <description>What does Web 2.0 mean to your company and products? What are the risks and opportunities? What are the proven strategies for successfully capitalizing on these changes?</description>
  <pubDate>Sun, 01 November 2006 00:00:00 GMT</pubDate>
  <guid>ISBN:0-596-52769-1</guid>
  <author>Tim O'Reilly</author>
  <link>http://radar.oreilly.com/research/web2-report.html</link>
  <category>technology</category> <category>web development</category>
</item>
```

---

[8](http://en.wikipedia.org/wiki/Communications_protocol)
Since the RSS specification defines a basic metadata set (the core RSS metadata set) but also allows to extend this with any additional metadata set, knowing that a source is available as RSS feed doesn’t tell us everything about its metadata. This is why some sources that are indexed as RSS feeds in the RING are also indexed against different metadata sets.

Example of RSS feed extended with Dublin Core metadata:

```
<rss version="2.0" xmlns:dc="http://purl.org/dc/elements/1.1/">
  <channel>
    <title>O'Reilly publications</title>
    <link>http://www.oreilly.com/</link>
    <item>
      <title>Web 2.0 Principles and Best Practices.</title>
      <link>http://www.marchal.com/en/</link>
      <dc:creator>Marchal</dc:creator>
      <dc:date>2006/11/01</dc:date>
      <dc:rights>Copyright 2006 O'Reilly</dc:rights>
    </item>
  </channel>
</rss>
```

**OAI-PMH**

In the context of the OA (Open Access) Initiative, the technical protocol called OAI-PMH (Protocol for Metadata Harvesting) is the agreed protocol to harvest metadata from repositories.

The OAI-PMH architecture is based on OAI providers (or data providers) and OAI harvesters (or service providers).

An **OAI provider** maintains one or more repositories (web servers) that support the OAI protocol as a means of exposing metadata. Implementing the OAI-PMH protocol means providing HTTP response pages to the six OAI “verbs” (Identify, ListSets, ListMetadataFormats, ListIdentifiers, ListRecords, and GetRecord). Responses are serialized as XML and can use one or more metadata sets, with the minimum mandatory metadata set being Dublin Core.

An **OAI harvester** is a service that can import metadata from a remote OAI provider sending HTTP requests containing the OAI verbs.

More details on the OAI-PMH architecture are available on the RING portal in the OAI harvesting tutorial: [http://ring.ciard.net/oai-harvesting](http://ring.ciard.net/oai-harvesting).

**RDF, the Semantic Web and Linked Data**

The “Resource Description Framework” (RDF)\(^9\) is more than an architecture, it is a conceptual framework supported by some accessory technologies that together help build what is called the “semantic web”.

The conceptual framework is based on a very simple assumption: “triples” constituted by a subject, a predicate and an object can represent and describe everything. In the RDF model, subject and object are entities, or instances of “classes”, while predicates are “properties”.

Reducing metadata to this essential structure eliminates a certain degree of incoherence and arbitrariness that is often in the definition of metadata structures, adding more rigor to the model that has to be followed in describing things. Besides, the basic concept that subject and object should be represented as often as possible by URIs gives the possibility to identify entities univocally in different records, thus allowing to interlink records across sources and expand searches navigating through URIs.

### 3.3 Examples of source declaration

Let us consider the following RDF graph stating that a resource has a title ("RDF Source") and a creator and that this creator is of type Person and has a name ("Fabien Gandon") and a mailbox ("mailto:fgandon@inria.fr").
Using foaf:knows to relate people together

The "friend of a friend" aspect of FOAF arises from the ability to say that one Person "knows" another Person. "Knows" is used in a deliberately vague sense here that includes "is friends with", "corresponds with", etc. Additional elaborations on the basic FOAF data can further clarify these relationships.

```xml
  <foaf:Person>
    <foaf:name>Leigh Dodds</foaf:name>
    <foaf:firstName>Leigh</foaf:firstName>
    <foaf:surname>Dodds</foaf:surname>
    <foaf:mbox_sha1sum>71b88e95b85f07518d89ebb4b4a451005bc3ca5</foaf:mbox_sha1sum>
    <foaf:knows>
      <foaf:Person>
        <foaf:name>Dan Brickley</foaf:name>
        <foaf: mbox_sha1sum>342451215bde6219989027579566a36a6</foaf:mbox_sha1sum>
        <foaf:knows>
          <foaf:Person>
            <foaf:name>Angela Beesley</foaf:name>
            <foaf: mbox_sha1sum>90e9137e726dc6cu352</foaf:mbox_sha1sum>
          </foaf:Person>
        </foaf:knows>
      </foaf:Person>
    </foaf:knows>
  </foaf:Person>
</rdf:RDF>
```

Which says "there is a person called Leigh Dodds with a mailbox whose hashed value is ... who knows a person called Dan Brickley whose mailbox is ... and that Dan Brickley is further described in foaf.rdf"
3.5 Existing standards adopted / recommended in agriculture

For an extensive list of existing data standards (from data formats to schemas to vocabularies of all types) used in the agri-food sector, look at the Agrisemantics Map of Agri-food Data Standards:

http://vest.agrisemantics.org

This Map was developed in the GODAN Action project, building on existing directories and repositories like the VEST Registry of AIMS (now merged with this one) and the AgroPortal (http://agroportal.lirmm.fr), a specialized repository of RDF vocabularies.