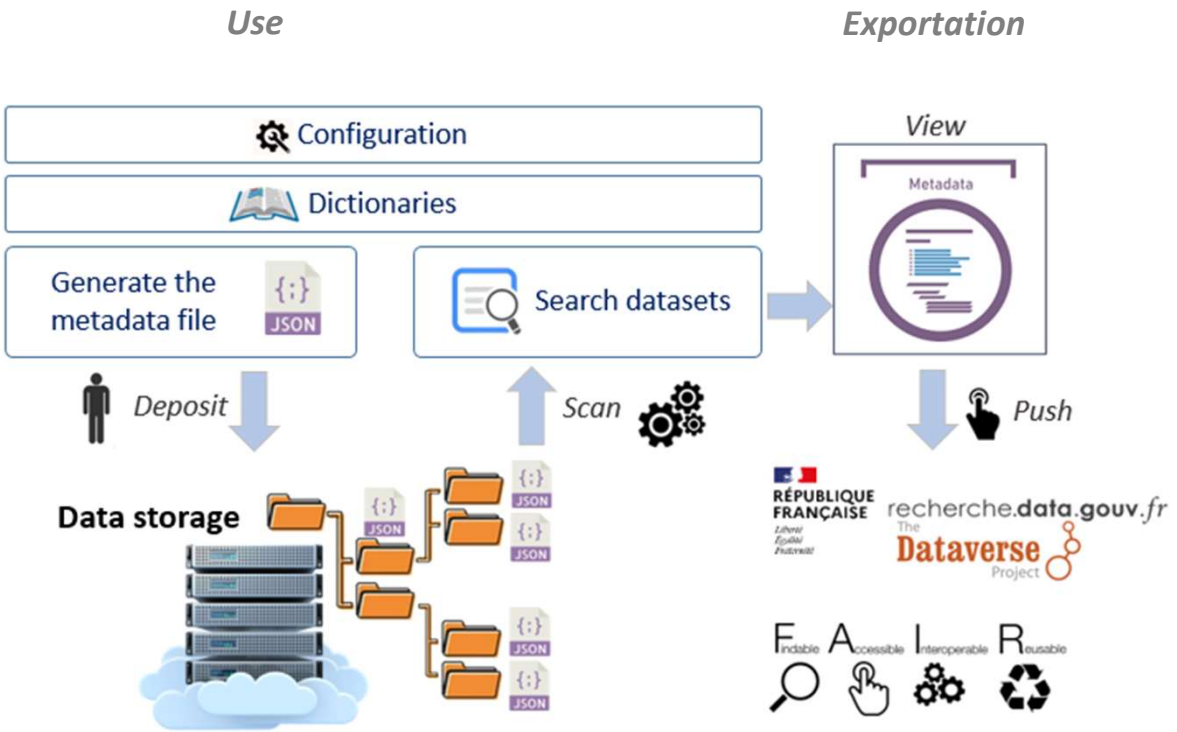
Maggot
Metadata aggregation on data storage

A tool dedicated to metadata



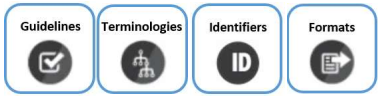
Assistance in entering and managing metadata

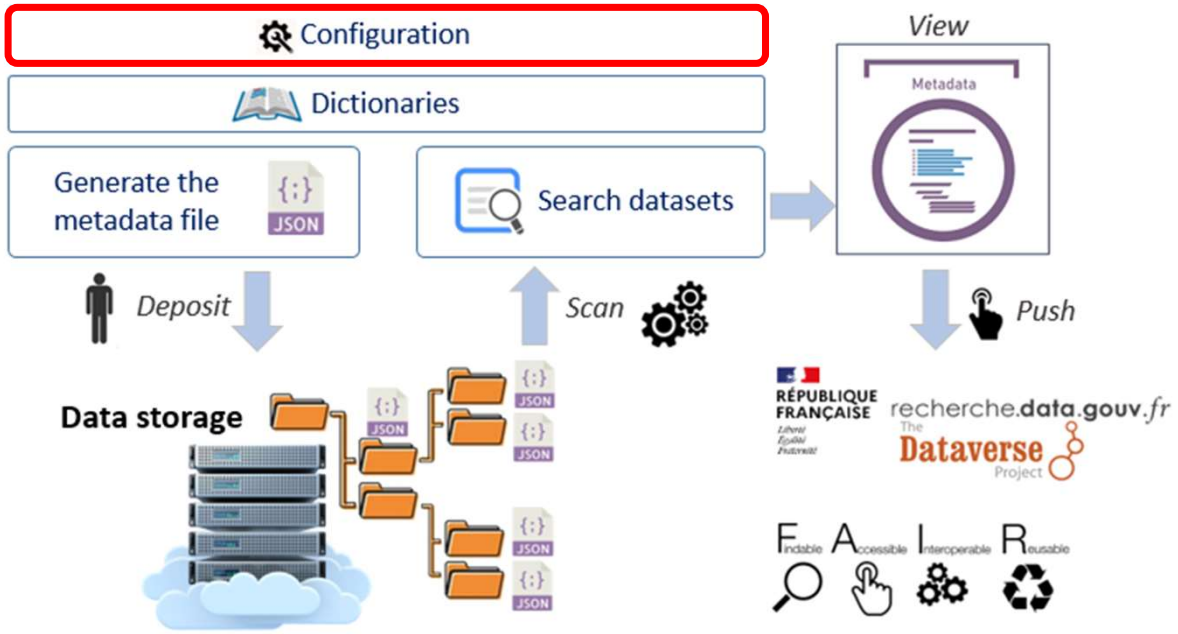
Filing Automation
avoids re-entering metadata during distribution

A tool dedicated to metadata

Maggot
Metadata aggregation on data storage

Standards & Terminologies
The 4 **FAIR** pillars





Assistance in entering and managing metadata

Filing Automation
avoids re-entering metadata during distribution

Configuration flexibility

Configuration

Definition of all terminology and functionality of the input interface → Terminology definition file

Dictionaries allowing you to enrich metadata without having to re-enter → Dictionary definition files

Definition of correspondences between vocabulary and ontology → Mapping definition file

Definition of correspondences between terminology et semantics → Mapping definition file to JSON-LD

Standards & Terminologies
The 4 **FAIR** pillars

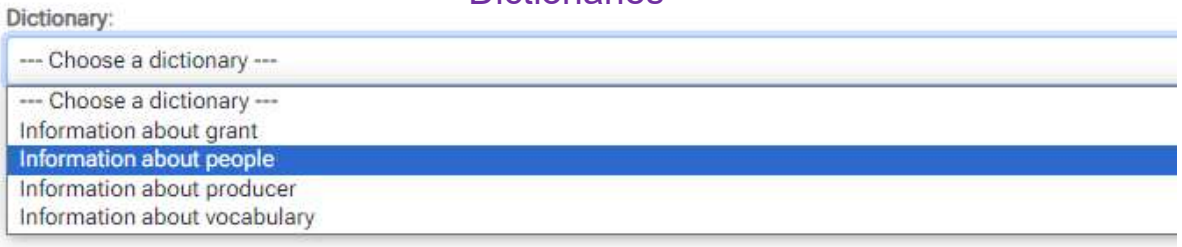
https://pmb-bordeaux.fr/maggot/config/view

View all current definition files

Terminology definition file	conf/config_terms.txt
Documentation definition file	conf/config_doc.txt
Dataverse definition file	conf/dataverse.txt
Mapping definition file	conf/mapping.txt
Schema.org JSON-LD definition file	conf/schema-jsonld.txt

https://pmb-bordeaux.fr/maggot/dictionary

Dictionaries



Maggot
Metadata aggregation on data storage

Definition of metadata

Terminology definition file

Completely done using a single configuration file (TSV format)

Field	Section	Required	Search	ShortView	Type	features	Label	Predefined terms
title	definition	Y	N	1	textbox	width=350px	Short name	
fulltitle	definition	Y	Y	2	textbox		Full title	
subject	definition	Y	Y		checkbox	open=0	Subject	Agricultural Sciences, Arts and Humanities, Astronomy and Astrophysics, Business and Management, Chemistry, Computer and Information Science, Earth and Environmental Sciences, Engineering, Law, Mathematical Sciences, Medicine Health and Life Sciences, Other, Physics, Social Sciences
description	definition	Y	Y		areabox	rows=6, cols=30	Description of the dataset / Organization of directories	
note	definition	N	Y		areabox	rows=4, cols=30	Notes	
status	status	N	Y	3	dropbox	width=350px	Status of the dataset	In progress, Processed, Unprocessed
access_rights	status	N	Y	4	dropbox	width=350px	Access rights to data	Mixte, Private, Public
language	status	N	Y		checkbox	open=0	Language	Czech, Danish, Dutch, English, Finnish, French, German, Greek, Hungarian, Icelandic, Italian, Lithuanian, Norwegian, Romanian, Slovenian, Spanish, Swedish
license	status	N	Y		textbox	autocomplete=license	License	
contacts	management	Y	Y		multiselect	autocomplete=people, min=1	Contacts	
authors	management	Y	Y		multiselect	autocomplete=people, min=1	Authors	
collectors	management	N	Y		multiselect	autocomplete=people, min=1	Data collectors	
curators	management	N	Y		multiselect	autocomplete=people, min=1	Data curators	
members	management	N	Y		multiselect	autocomplete=people, min=1	Project members	
leader	management	N	Y		multiselect	autocomplete=people, min=1	Project leader	
wpleader	management	N	Y		multiselect	autocomplete=people, min=1	WP leader	
depositor	management	N	Y		textbox		Depositor	
producer	management	N	Y		multiselect	autocomplete=producer, min=1	Producer	
grantNumbers	management	N	Y		multiselect	autocomplete=grant, min=1	Grant Information	
kindOfData	descriptors	Y	Y		checkbox	open=0	Kind of Data	Audiovisual, Collection, Dataset, Event, Image, Interactive Resource, Model, Other, Physical Object, Service, Software, Sound, Text, Workflow
keywords	descriptors	N	Y		multiselect	autocomplete=bioportal, onto=EFO: JERM: EDAM: MS: NMR: NCIT: OBI: PO: PTO: AGRO: ECOCORE: IOBC: NCBITAXON	Keywords	

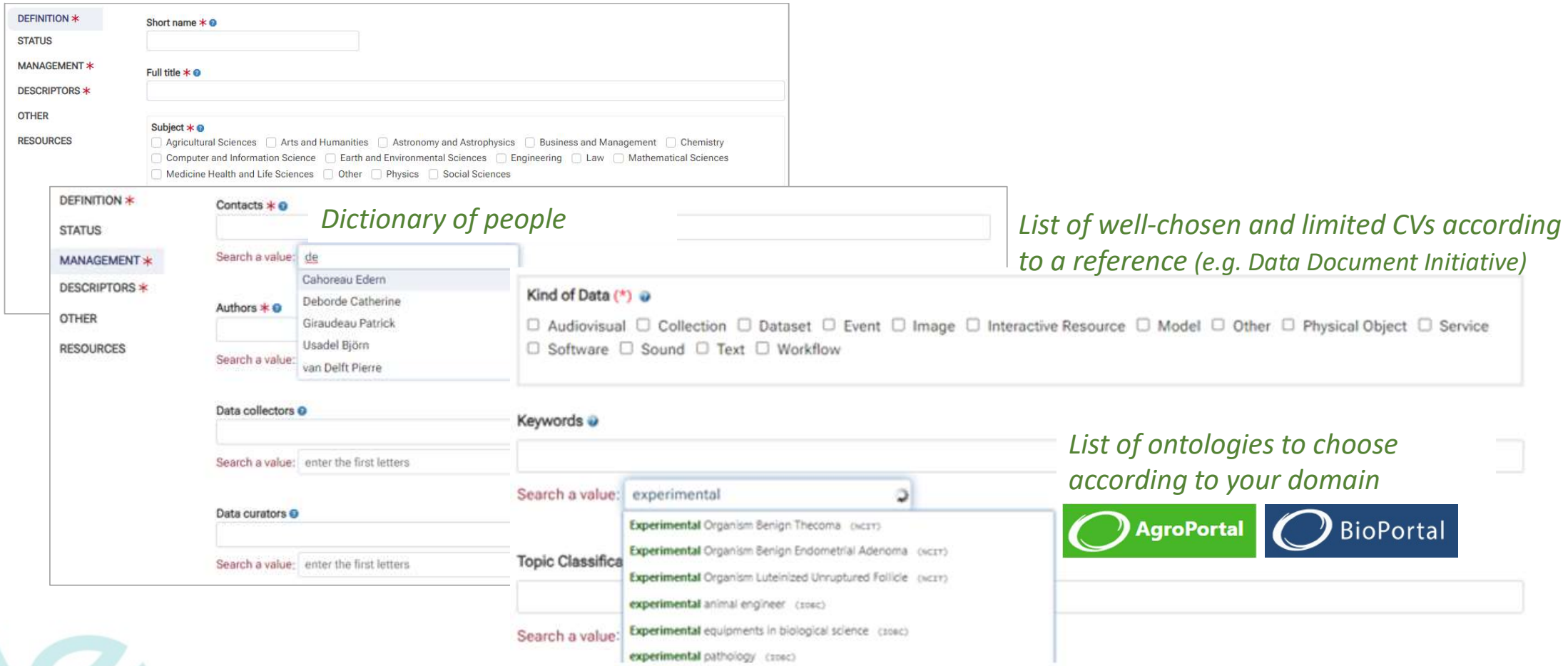
Built the web interface according to the **specifications given upstream**

Maggot
Metadata aggregation on data storage

Definition of metadata

Terminology definition file

Built the web interface according to the ***specifications given upstream***



Dictionary of people

List of well-chosen and limited CVs according to a reference (e.g. Data Document Initiative)

List of ontologies to choose according to your domain

AgroPortal BioPortal

Maggot
Metadata aggregation on data storage

Mapping to JSON-LD

Mapping definition file to JSON-LD

#datasetField	Field	Type	Level	Array	Multiple	Property	Value type	Value source	Index	Prefix
	@type		0	FALSE	FALSE		schema:Dataset			
	@id		0	FALSE	FALSE					https://pmb-bordeaux.fr/maggot/metadata/
		schema:version	0	FALSE	FALSE		1.0			
	title	schema:identifier	0	FALSE	FALSE		#VALUE			maggot:
	fulltitle	schema:name	0	FALSE	FALSE		#VALUE			
	description	schema:description	0	FALSE	FALSE		#VALUE			
	note	schema:comment	0	FALSE	FALSE	#VALUE				
	license	schema:license	0	FALSE	FALSE		#VALUE			
	contacts	schema:creator	0	TRUE	TRUE					
	contacts	schema:creator	1	FALSE	FALSE	@type	schema:Person			
	contacts	schema:creator	1	FALSE	FALSE	name	#VALUE			
	contacts	schema:creator	1	FALSE	FALSE	identifier	#DICO	cvlist:people	3	https://orcid.org/
	contacts	schema:creator	1	FALSE	FALSE	email	#DICO	cvlist:people	4	
	contacts	schema:creator	1	FALSE	TRUE	affiliation				
	contacts	schema:creator	2	FALSE	FALSE	@type	schema:Organization			
	contacts	schema:creator	2	FALSE	FALSE	name	#DICO	cvlist:people	2	
	authors	schema:author	0	TRUE	TRUE					
	authors	schema:author	1	FALSE	FALSE	@type	schema:Person			
	authors	schema:author	1	FALSE	FALSE	name	#VALUE			
	authors	schema:author	1	FALSE	FALSE	identifier	#DICO	cvlist:people	3	https://orcid.org/
	authors	schema:author	1	FALSE	TRUE	affiliation				
	authors	schema:author	2	FALSE	FALSE	@type	schema:Organization			
	authors	schema:author	2	FALSE	FALSE	name	#DICO	cvlist:people	2	
	collectors	contributor	0	TRUE	TRUE					
	collectors	contributor	1	FALSE	FALSE	@type	schema:Person			
	collectors	contributor	1	FALSE	FALSE	contributorType	DataCollector			http://purl.org/datacite/v4.4/
	collectors	contributor	1	FALSE	FALSE	name	#VALUE			
	collectors	contributor	1	FALSE	FALSE	identifier	#DICO	cvlist:people	3	https://orcid.org/
	collectors	contributor	1	FALSE	TRUE	affiliation				
	collectors	contributor	2	FALSE	FALSE	@type	schema:Organization			
	collectors	contributor	2	FALSE	FALSE	name	#DICO	cvlist:people	2	
	curators	contributor	1	TRUE	TRUE					
	curators	contributor	1	FALSE	FALSE	@type	schema:Person			

Maggot
Metadata aggregation on data storage

Linked Open Data

Metadata schema (standard) *(container)*

DDI® *Data Document Initiative*

Findable Accessible Interoperable Reusable

Interoperability

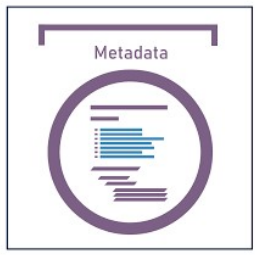
Vocabulary *(content)*

- Dictionaries
- Control Vocabulary

AgroPortal BioPortal

Thesaurus SKOSMOS

(VOINRAE, LOTERRE, ONTOSTACK, …)



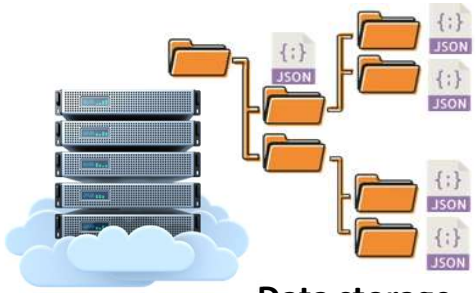
External Resources *(related data)*

JSON *Metadata*

TSV **Mapping definition file to JSON-LD**

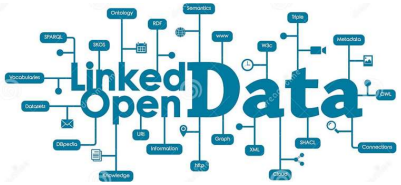
Mapping

JSON-LD



Data storage

Your local (meta)data repository

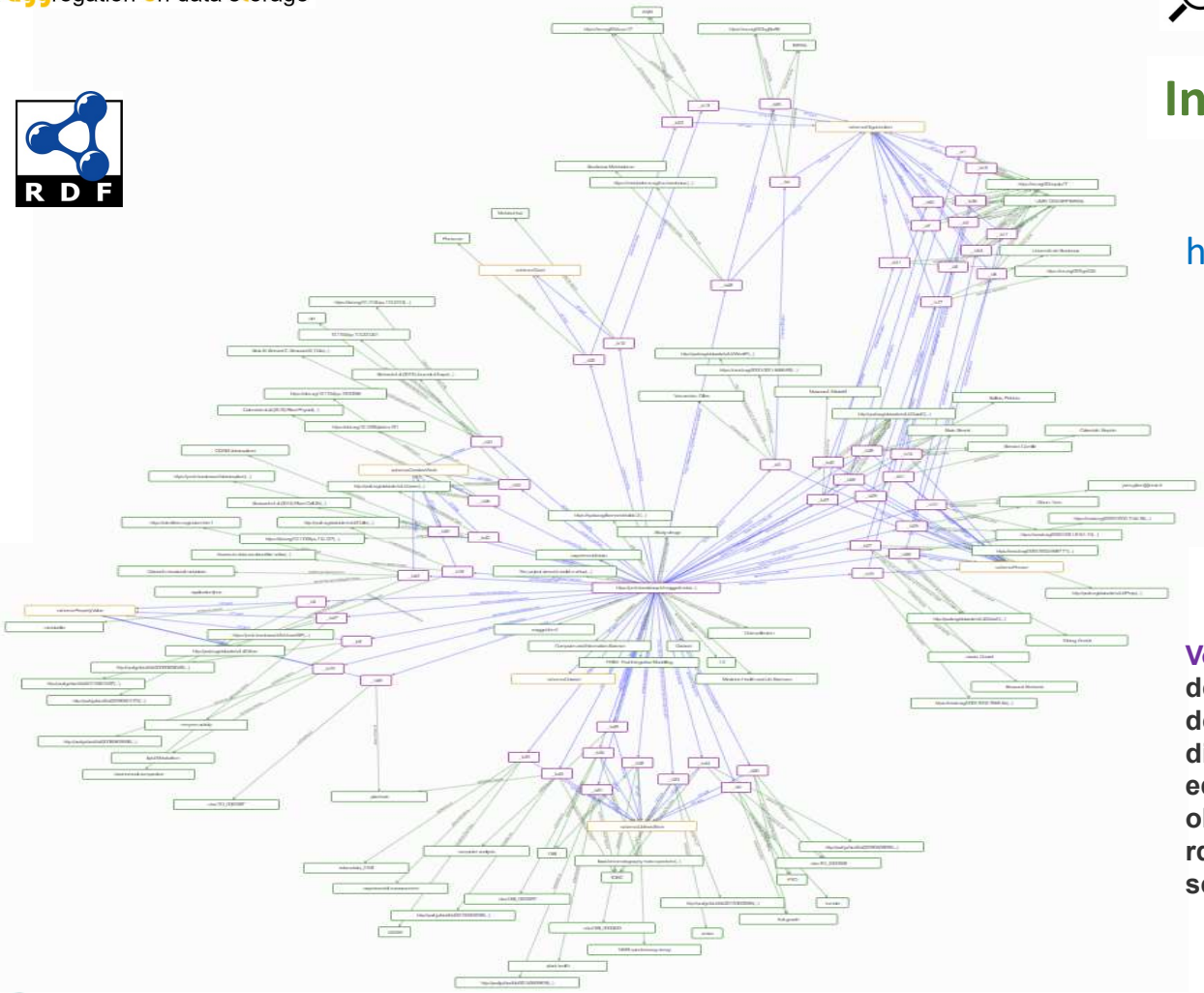


5 ★ Open Data

https://5stardata.info/fr/

Maggot
Metadata aggregation on data storage

Linked Open Data



https://pmb-bordeaux.fr/maggot/metadata/frim1?format=jsonld

Findable Accessible Interoperable Reusable

Interoperability

https://issemantic.net/rdf-visualizer

Legend:
- Type/Class
- Instance/Individual
- Value/Literal
- Object property
- Data property
- ..... Quoted triple
- _:b Blank node

Vocabularies/prefixes:
dcite: http://purl.org/spar/datacite/
dcterms: http://purl.org/dc/terms/
disco: http://rdf-vocabulary.ddialliance.org/discovery#
edam: http://edamontology.org/
obo: http://purl.obolibrary.org/obo/
rdf: http://www.w3.org/1999/02/22-rdf-syntax-ns#
schema: http://schema.org/

Maggot

Metadata aggregation on data storage

A milestone on the way up

« Climb the LOD mountain »
gently and step by step.

Metadata Profile

DDI
Data Document Initiative

BioPortal
AgroPortal
SKOSMOS
API

TSV **Terminology definition file**

TSV **Mapping definition file to JSON-LD**

TSV **Mapping definition file**

TSV **Dictionary definition files**

Metadata Capture

Metadata

JSON
JSON Schema

Good practices

doi ORCID iD

Interoperability

Mapping

JSON-LD

RDF

Linked Open Data

Data storage

Upstream

Downstream

Roles

The **notion of role** makes it possible **to take into consideration the contribution of each person according to their « job skills ».**

Standards & Terminologies
The 4 **FAIR** pillars

DDI
Data Document Initiative

AgroPortal BioPortal
Thesaurus SKOSMOS
(VOINRAE, LOTERRE, ONTOSTACK, …)

Roles

Upstream

Fix the frame based on standards and vocabularies recognized by the international community in the field.

Guidelines Terminologies

« Data manager »
+
« Data scientist »

Metadata Capture

« Input Mask »

Allowing metadata to be captured according to a data description standard

Good Practices

Identifiers Terminologies

« Data manager »

Downstream

Transform data for interoperability

greatly facilitated by the task upstream

JSON-LD

Guidelines Terminologies Identifiers Formats

«Data scientist »

Role : Kind of action or behavior, in the place one occupies. A person can have several roles; similarly several people can have the same role.