FAIR assessment :
Le défi de la comparaison des scores FAIR

Frédéric de Lamotte - 0000-0003-4234-1172

Nancy Novembre 2025

Plan

- Qui suis-je ?
- Contexte
- Fonctionnalités
- Démo
- Analyse des usages
- Perspectives

INRAE

Un établissement public de recherche parmi les tout premiers leaders mondiaux en sciences agricoles et alimentaires, en sciences du végétal et de l'animal.

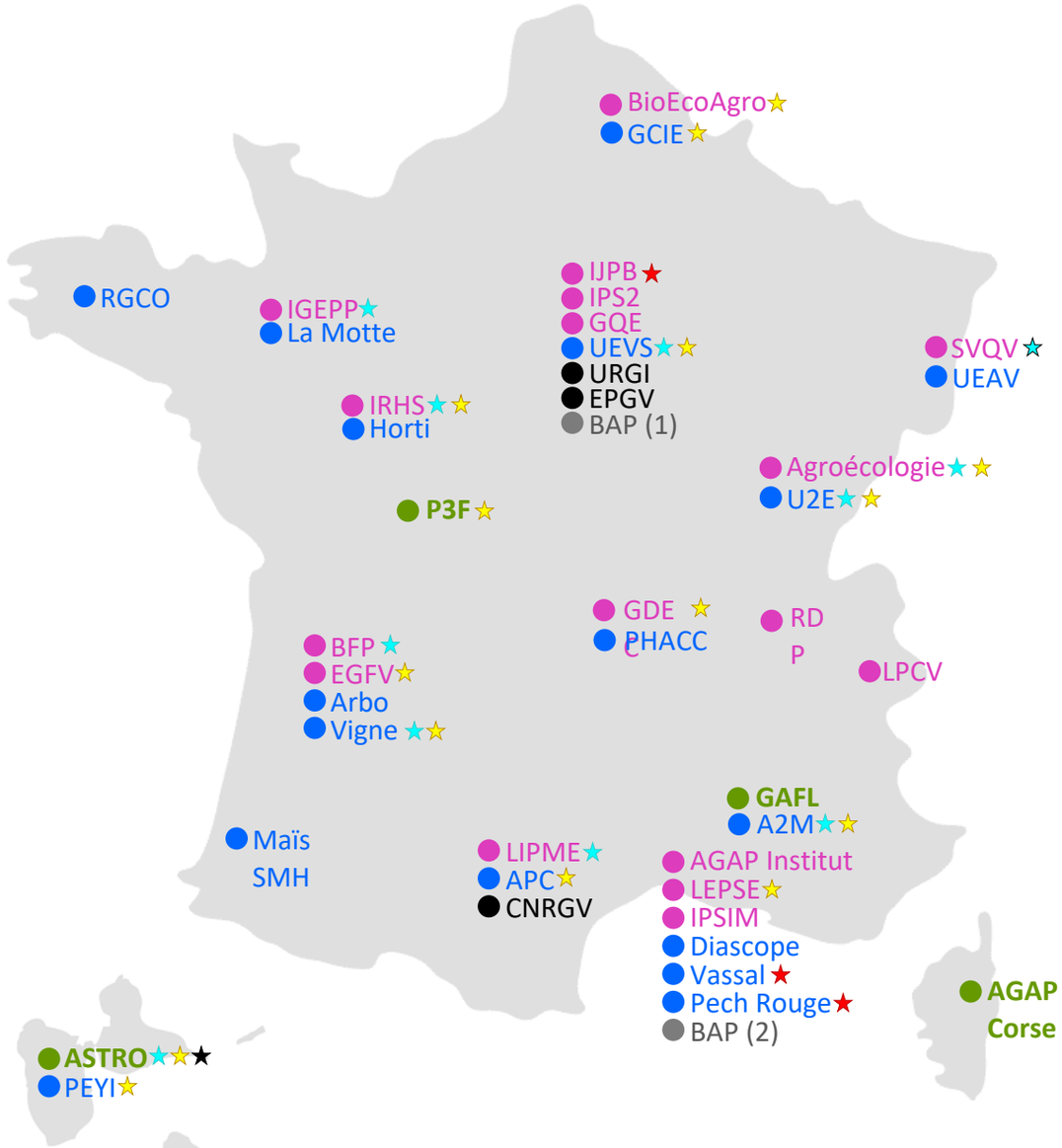
- 12 000 collaborateurs et collaboratrices
- 272 unités de recherche et de service
- 18 centres implantés sur toute la France
- + 1 milliard d'€ de budget

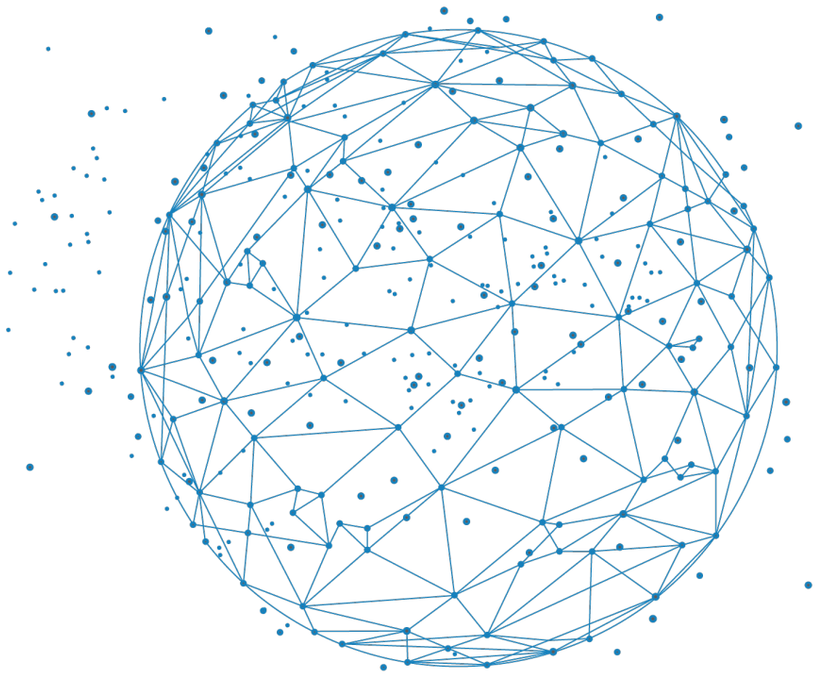
Référent Données Stratégique
Département BAP
Biologie et Amélioration des Plantes

1200 personnes

43 unités dont :

- 17 Unités mixtes de recherche (UMR)
- 4 Unités propres de recherche (UR)
- 17 Unités expérimentales (UE)
- 3 Unités de service
- 2 Unités d'appui au département BAP





L’Institut Français de Bioinformatique (IFB/ELIXIR-FR)

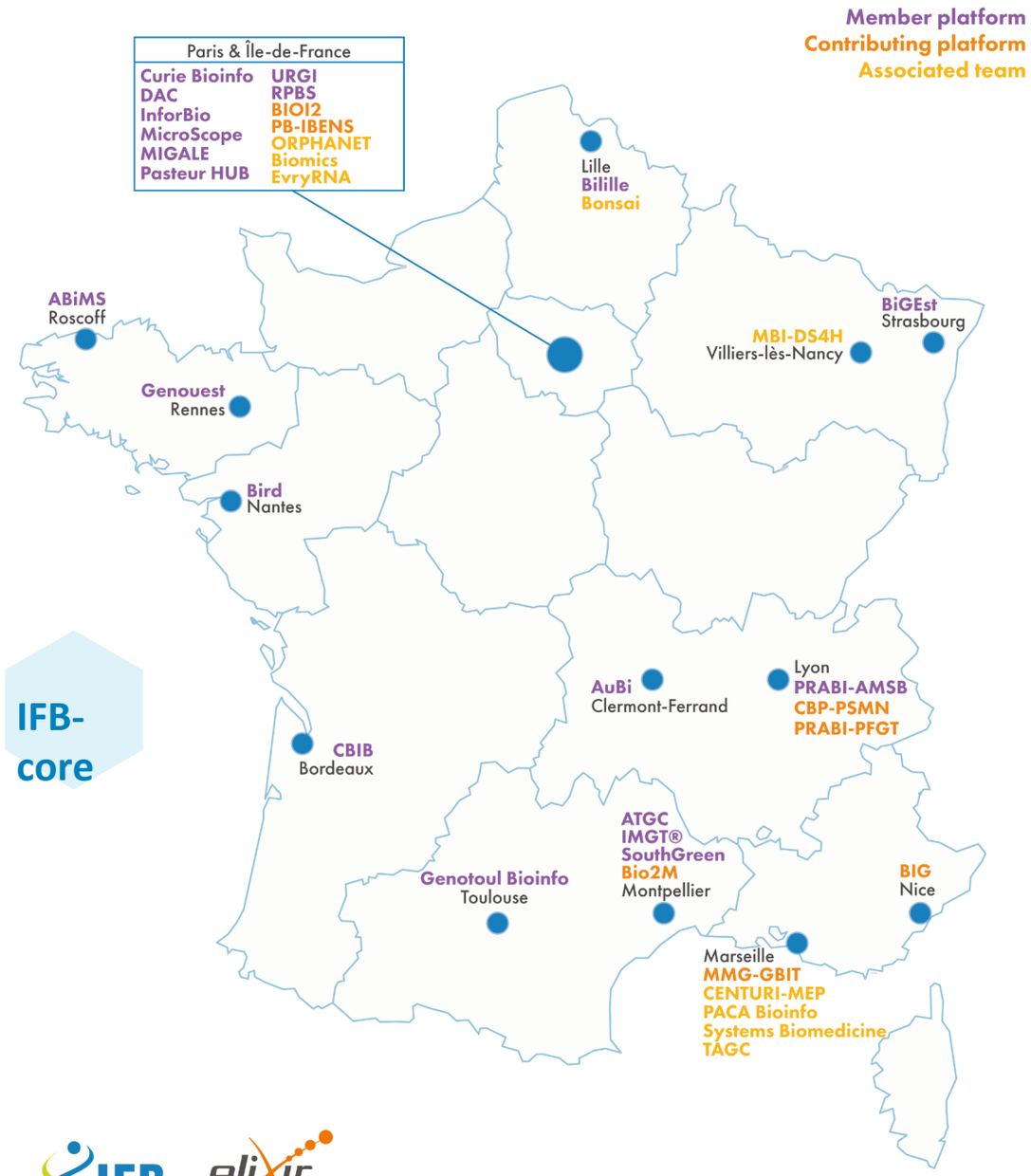
Une infrastructure nationale de bioinformatique pour les communautés des sciences de la vie et le nœud français d'ELIXIR

IFB INSTITUT FRANÇAIS DE BIOINFORMATIQUE

elixir FRANCE

cnrs INRAE Inserm cea FRANCE 2030

L’IFB, alias ELIXIR-FR, une **fédération de plateformes et d’équipes de bioinformatique**



v. Avril 2025

Une infrastructure distribuée sur toute la France métropolitaine

- **20** plateformes membres
- **7** plateformes contributrices
- **8** équipes associées
- **1** unité de coordination (IFB-core)

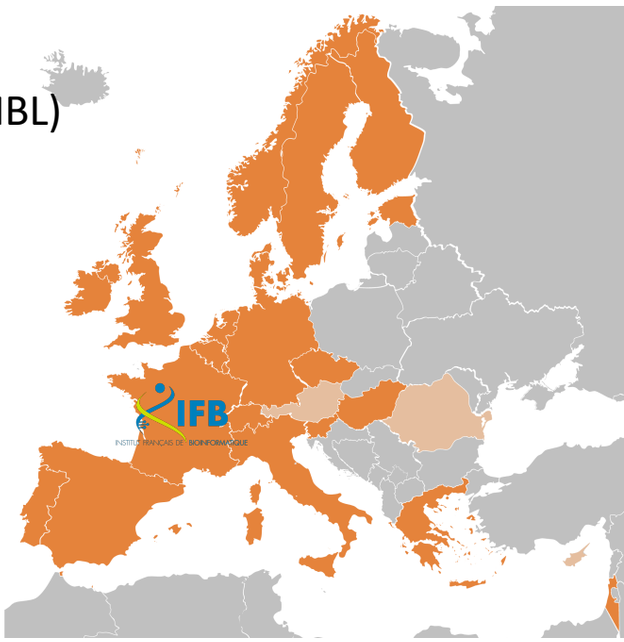
Une intégration dans l'Europe

- Nœud français d’ELIXIR Europe
- Point d'entrée pour les projets européens

ELIXIR, ESFRI de bioinformatique

- elixir-europe.org
- 24 noeuds nationaux
- 1 noeud international (EMBL)
- ELIXIR-Hub: coordination

IFB = ELIXIR FRANCE



Pilotage Science Ouverte & Interopérabilité

Ouvrir la science !

Collège des Données de la recherche

Dr. Véronique STOLL & Dr. Frédéric de Lamotte (pilotes du Collège des Données de la recherche)

https://orcid.org/0000-0002-4118-3574 & https://orcid.org/0000-0003-4234-1172

Missions

- Issues de l'axe 2 du PNSO : structuration, partage, ouverture des données
- Besoins exprimés par les communautés scientifiques, remontées de terrain et sollicitations ponctuelles du MESR (prix SO Données, Etude sur les métiers…)

1. Renforcer la connaissance des pratiques en matière de collecte, de gestion et d'ouverture des données de recherche
2. Encourager l'application des principes FAIR (Facile à trouver, Accessibles, Interopérables, Réutilisables) dans la gestion des données de recherche
3. Promouvoir et développer les pratiques d'ouverture et de réutilisation des données de recherche
4. Soutenir l'adoption d'une politique de données sur l'ensemble du cycle de vie

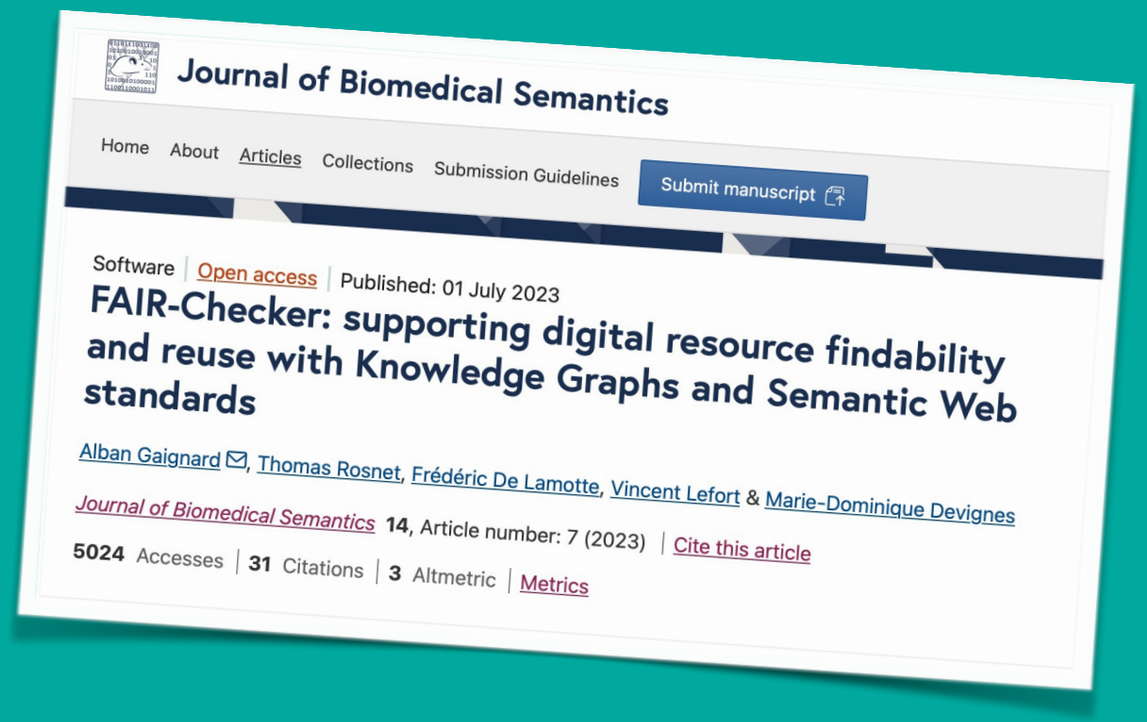
Context

Alban Gaignard

Thomas Rosnet

Marie-Dominique Devignes

Frédéric de Lamotte



Comment pouvons-nous évaluer si les données de recherche sont suffisamment FAIR ?

[1] Gaignard, A., Rosnet, T., de Lamotte, F., Lefort, V., & Devignes, M. (2023). *FAIR-Checker: supporting digital resource findability and reuse with Knowledge Grap*

FAIR : Essentiel pour des sciences ouvertes et reproductibles

L’évaluation du FAIR nécessitent des outils

Australian Research Data Commons

Et ce n’est pas simple

Comment mettre en œuvre ces principes ...

... et aller au-delà des checklists ?

https://www.go-fair.org/fair-principles

https://www.nature.com/articles/sdata201618

2 points de vue

Producteurs de données

Make

- Dataset
- Training
- Software tools
- etc …

discoverable !

Où publier ?
Quel entrepôt ?
Quelles métadonnées ?
Est-ce assez FAIR ?

Gestionnaires d'infra

Make

- Dataverse
- Bio.tools
- Pangea, eBrains, Pubmed …

more FAIR !

Améliorer la qualité des métadonnées ?

Normes spécifiques à la communauté ?

Quelle Techno ?

Why a (nother) tool ?

Idée de départ

- **On doit pouvoir s'appuyer sur des techno comme "Linked Data" et Semantic Web**
  … mais des compétences techniques sont nécessaires.

Objectifs

- Fournir une interface web permettant aux fournisseurs de ressources d'évaluer les indicateurs FAIR et de progresser dans la FAIRification
- Fournir des outils supplémentaires (RDF, SPARQL, SHACL) aux développeurs pour améliorer la qualité des métadonnées

FAIR Checker

Un bon schéma vaut mieux qu'un long discours … ou pas !

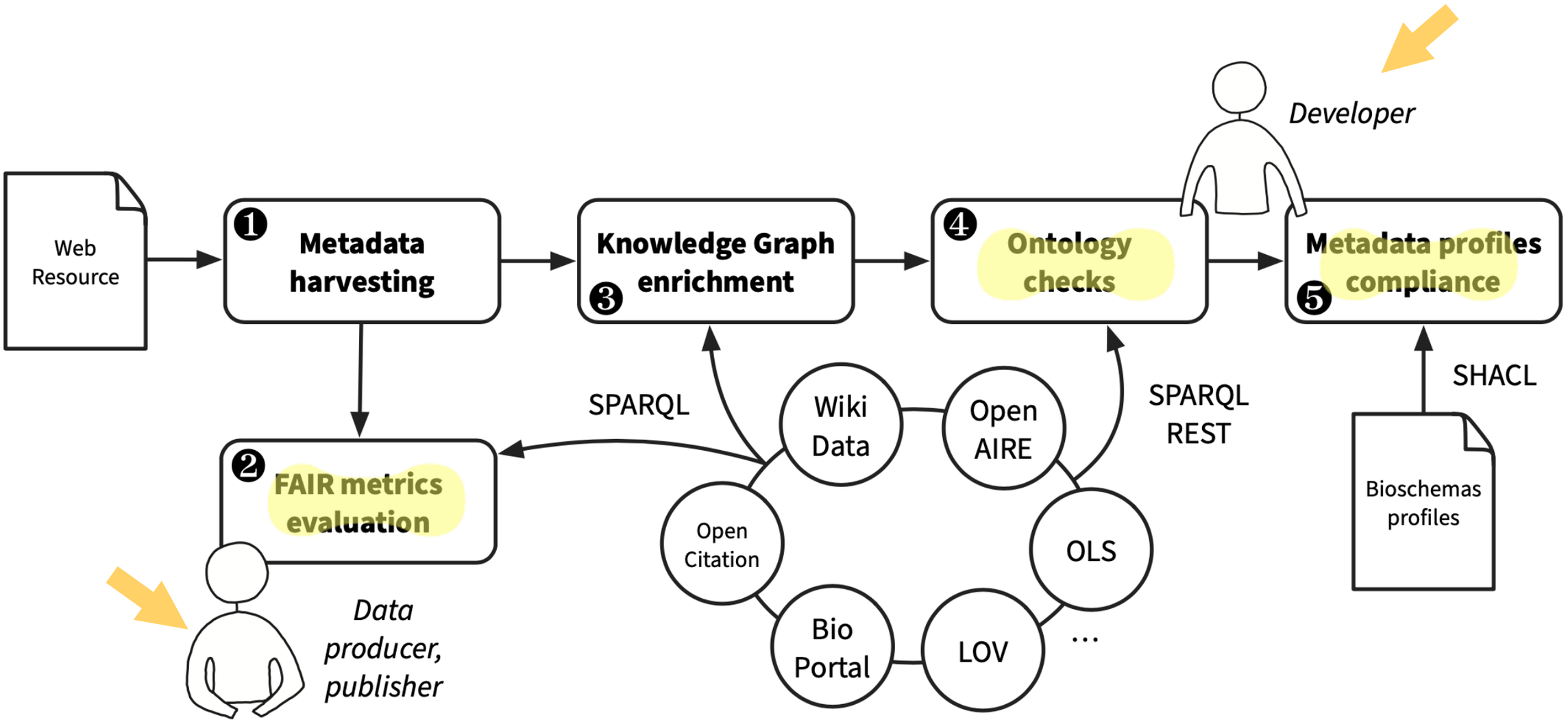
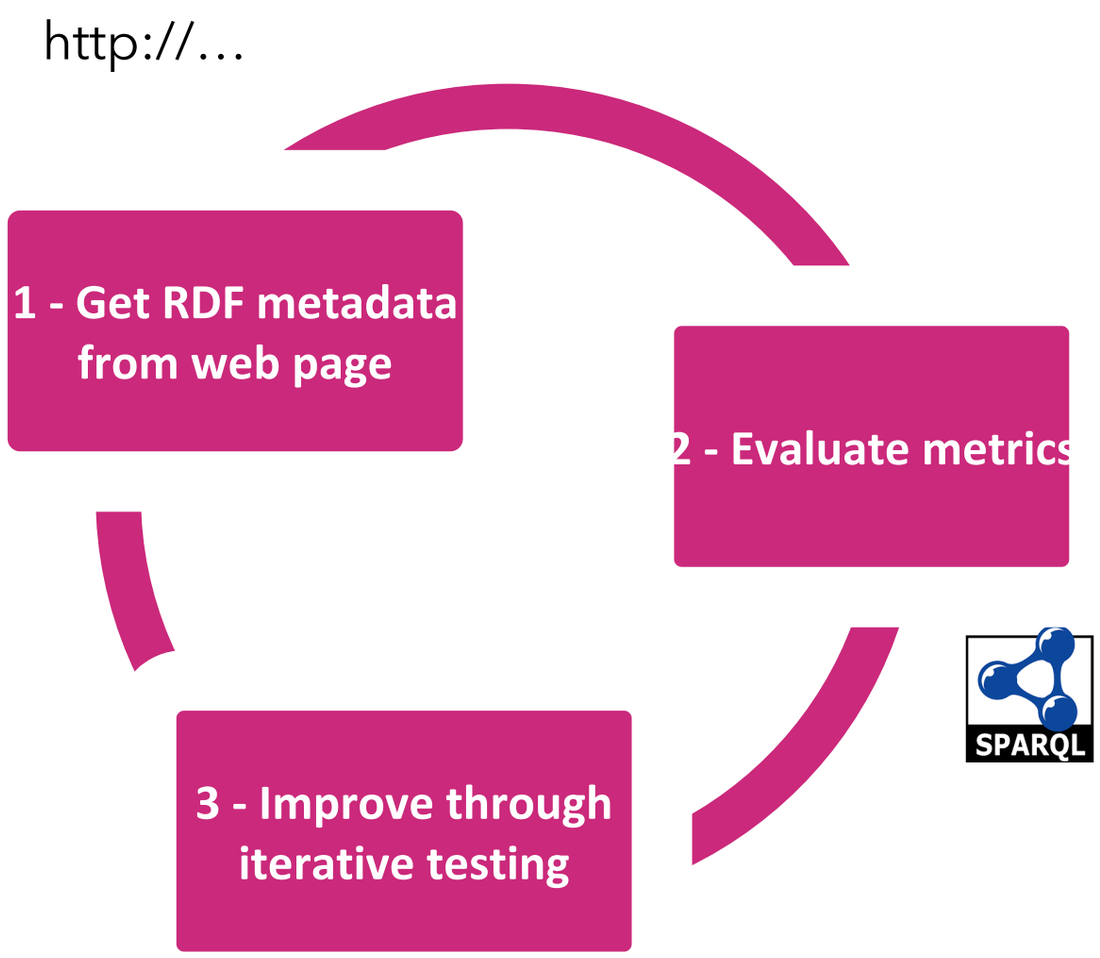


Figure 1 Gathering, enriching and analyzing semantic web annotations in line with FAIR principles.

① Checking metadata

Checking metadata



Semantic web technologies

- RDF
- SPARQL

FAIR cookbook

RDMkit

https://fairplus.github.io/the-fair-cookbook/content/home.html

https://rdmkit.elixir-europe.org

What is evaluated ?

Findability F1B, F2	Accessibility A1.2	Reuse (licenses) R1.1	Reuse (provenance) R1.2
dct:identifier	odrl:hasPolicy	schema:license	prov:wasGeneratedBy
schema:identifier	dct:rights	dct:license	prov:wasDerivedFrom
dct:title	dct:accessRights	doap:license	prov:wasAttributedTo
dct:description		dbo:license	prov:used
dcat:accessURL		cc:license	prov:wasInformedBy
dcat:downloadURL		xhv:license	prov:wasAssociatedWith
dcat:endpointDescription		sto:license	prov:startedAtTime
dcat:endpointURL		nie:license	prov:endedAtTime
			dct:hasVersion
			dct:isVersionOf
			dct:creator
			dct:contributor
			dct:publisher
			pav:hasVersion
			pav:version
			pav:hasCurrentVersion
			pav:createdBy
			pav:authoredBy
			pav:retrievedFrom
			pav:importedFrom
			pav:createdWith
			pav:retrievedBy
			pav:importedBy
			pav:curatedBy
			pav:createdAt
			pav:previousVersion
			schema:creator
			schema:author
			schema:publisher
			schema:provider
			schema:funder

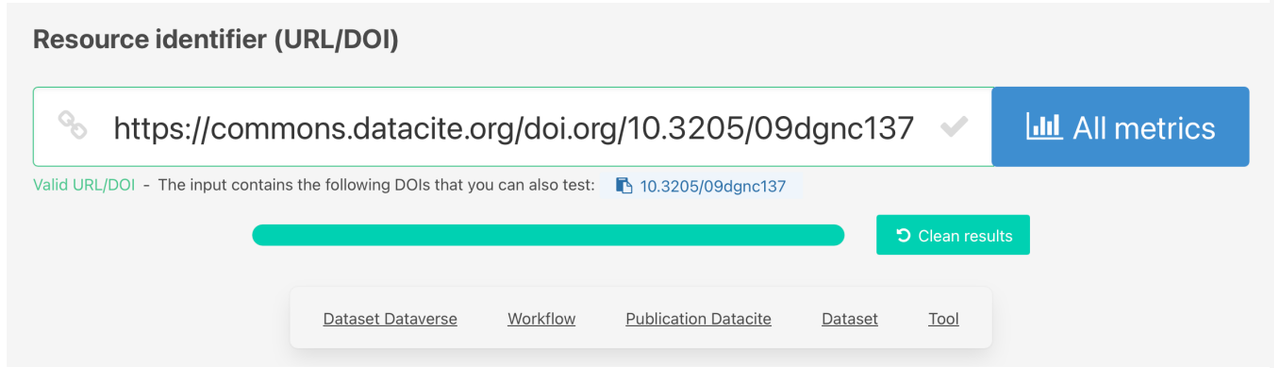
DC-Terms
DCAT
Schema.org

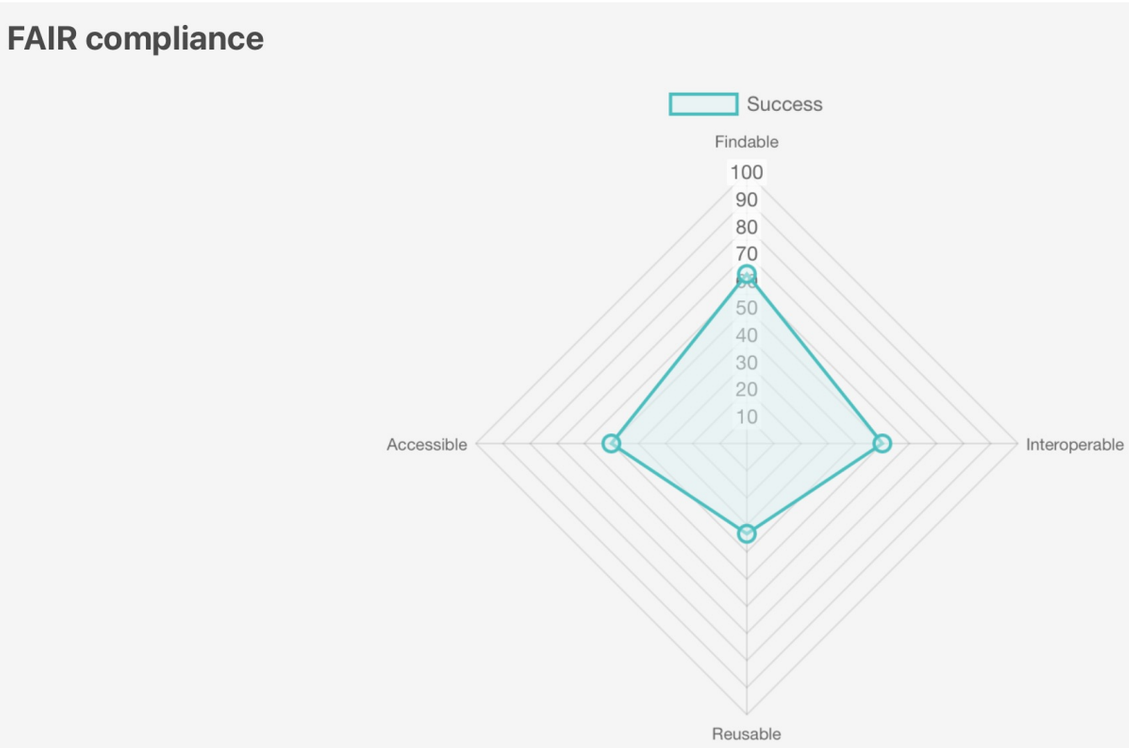
+ **ODRL**

+ **DOAP,**
DBO, CC ...

+ **PROV-O, PAV**

Table 2 Summary of the selected ontology properties relevant to assess three specific FAIR principles in *FAIR-Checker*





1. Provide resource URL
2. Retrieve metadata
3. Evaluate relevant metadata fields (ontology classes and properties)
4. Propose **relevant recommendations** from FAIR Cookbook and RDMkit
5. Improve metadata and **try** again …

F2B: Shared vocabularies for metadata	Check	FAIR principle F2B 2/2		
A1.1: Open resolution protocol	Check	FAIR principle A1.1 2/2		
A1.2: Authorisation procedure or access rights	Check	FAIR principle A1.2 0/2	You should describe the access policy in metadata by using at least one of the Read more	
I1: Machine readable format	Check	FAIR principle I1 1/2	You should provide discoverability oriented metadata with one of the following properties: dct:title dct:description dcat:accessURL dcat:downloadURL dcat:endpointDescription dcat:endpointURL Read less	
I2: Use shared ontologies	Check	FAIR principle I2 2/2		
I3: External links	Check	FAIR principle I3 0/2	You should enrich your metadata with more diversified external links. Here we did not Read more	
R1.1: Metadata includes license	Check	FAIR principle R1.1 0/2	You should include information about licence in your metadata using one of the Read more	

Resource identifi

[redacted]

Valid URL/DOI - The input c

FAIR compliance

Detailed r

Principle

F1A: Uniqu

F1B: Persis

F2A: Struc

F2B: Share

A1.1: Open

A1.2: Autho

I1: Machine

I2: Use sha

I3: Externa

R1.1: Metac

R1.2: Meta

R1.3: Comr

Did not find y

For additiona

Detailed results

Export

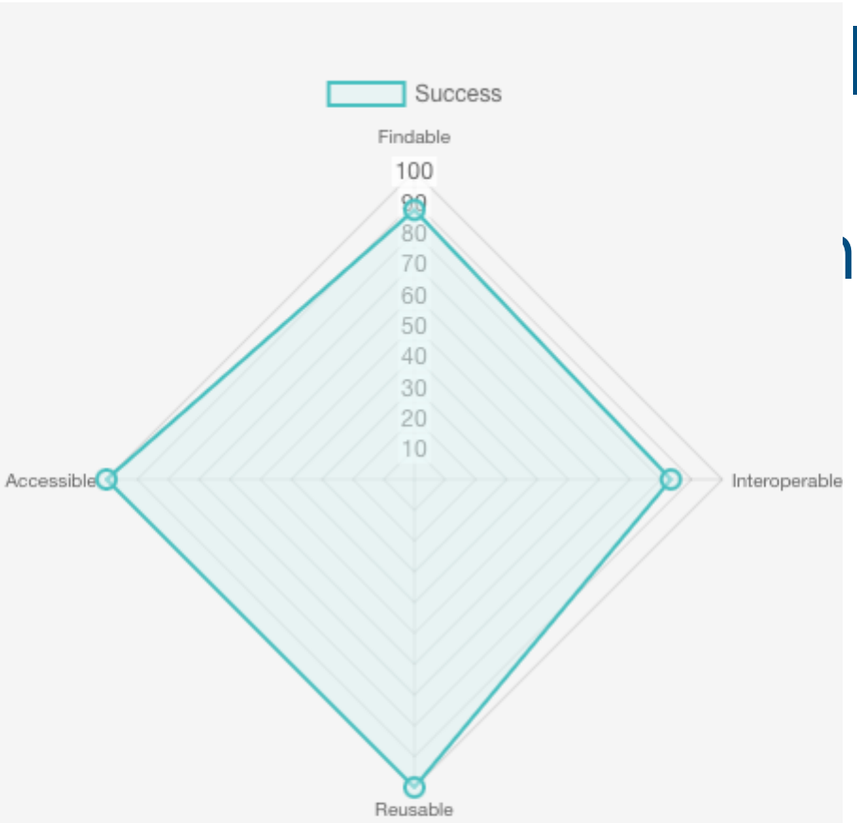
Principle	Test	Result	Recommandation	Details
F1A: Unique IDs	Check	FAIR principle F1A 2/2		
F1B: Persistent IDs	Check	FAIR principle F1B 0/2	To ensure that the used identification scheme is persistent, you should build your resource ID with a namespace that can Read more	
F2A: Structured metadata	Check	FAIR principle F2A 1/2	You should provide discoverability oriented metadata with one of the following properties: dct:title dct:description Read more	
F2B: Shared vocabularies for metadata	Check	FAIR principle F2B 2/2		
A1.1: Open resolution protocol	Check	FAIR principle A1.1 2/2		
A1.2: Authorisation procedure or access rights	Check	FAIR principle A1.2 0/2	You should describe the access policy in metadata by using at least one of the following properties: odrl:hasPolicy Read more	
I1: Machine readable format	Check	FAIR principle I1 1/2	You should provide discoverability oriented metadata with one of the following properties: dct:title dct:description Read more	
I2: Use shared ontologies	Check	FAIR principle I2 2/2		
I3: External links	Check	FAIR principle I3 0/2	You should enrich your metadata with more diversified external links. Here we did not detect more than two distinct Read more	
R1.1: Metadata includes license	Check	FAIR principle R1.1 0/2	You should include information about license in your metadata using one of the properties below: Read more	
R1.2: Metadata includes provenance	Check	FAIR principle R1.2 0/2	You should include information about provenance in your metadata using one of the properties below: Read more	
R1.3: Community standards	Check	FAIR principle R1.3 2/2		

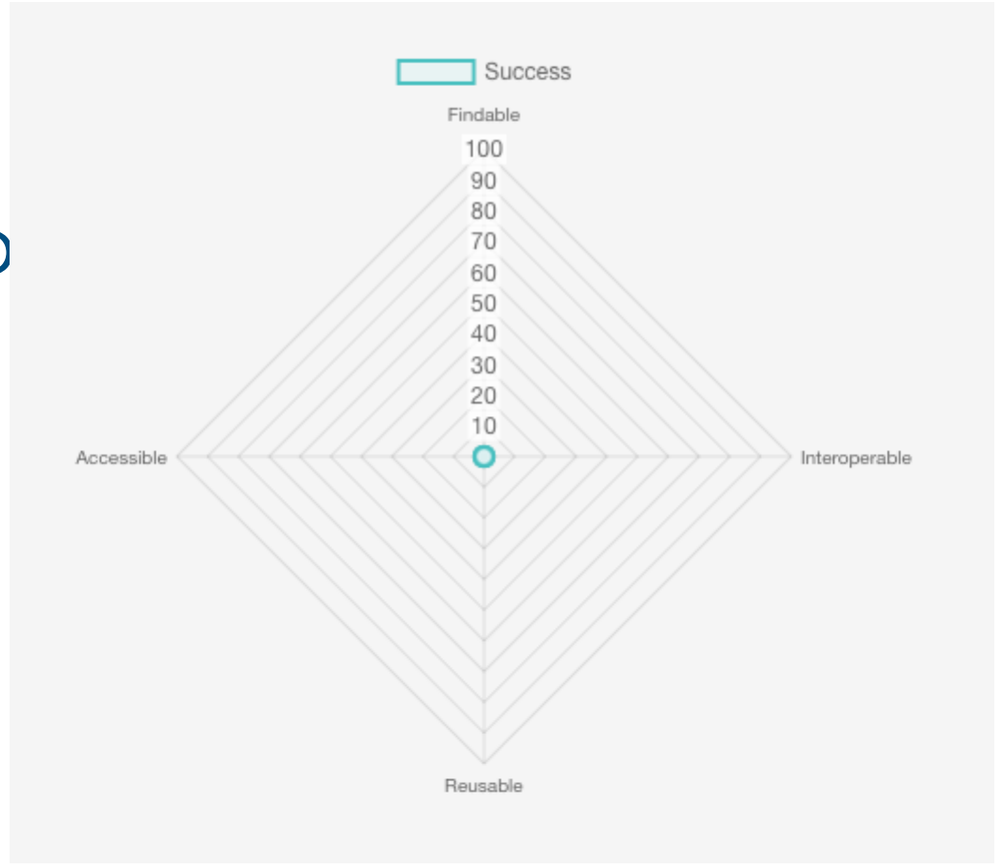
Did not find your metadata term ? Please submit a request and let's discuss with the community ! Ask for a new term

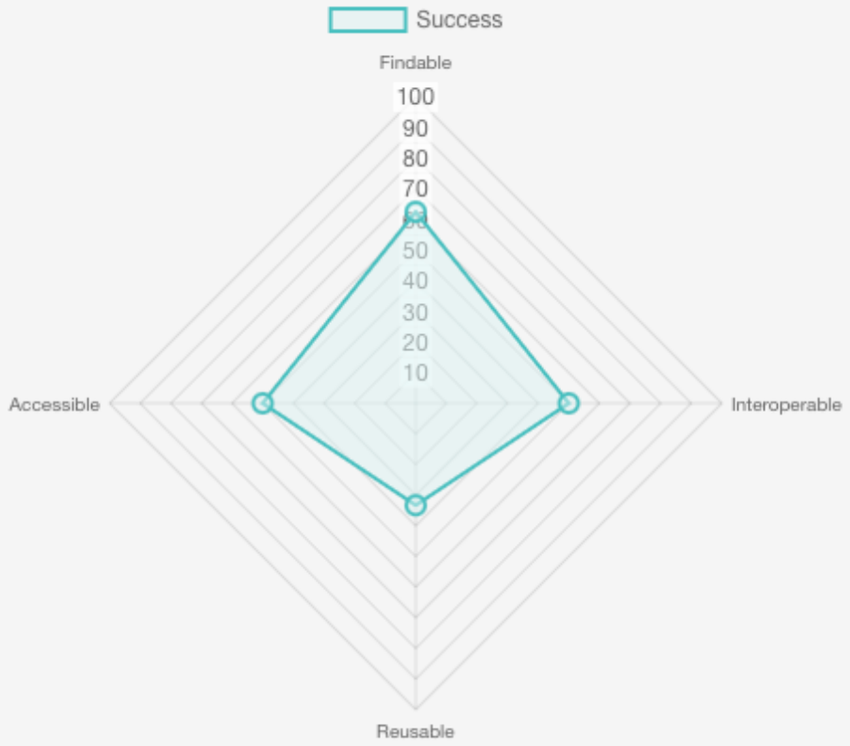
For additional tips and recommendations, please look at the FAIR Cookbook: FAIR Cookbook

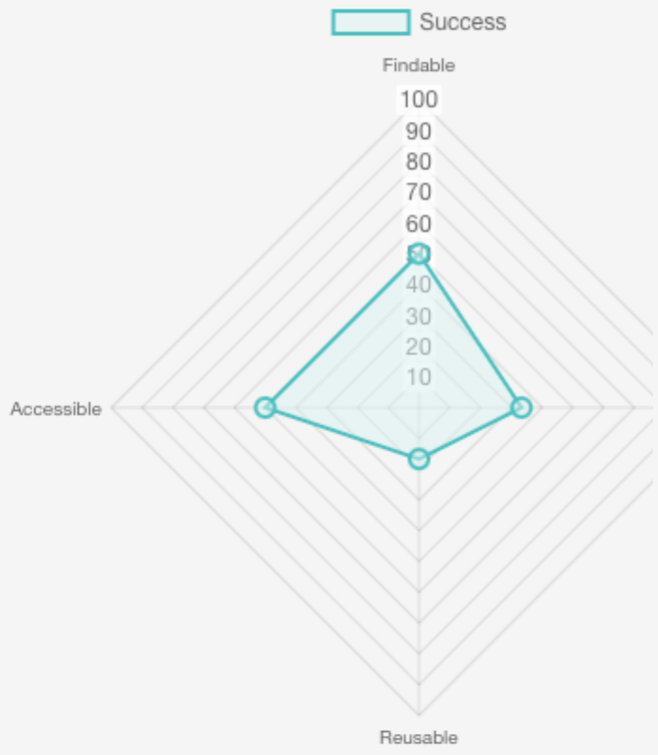
[l

(me o


Success
Findable
100
90
80
70
60
50
40
30
20
10
Accessible
Interoperable
Reusable


Success
Findable
100
90
80
70
60
50
40
30
20
10
Accessible
Interoperable
Reusable


Success
Findable
100
90
80
70
60
50
40
30
20
10
Accessible
Interoperable
Reusable


Success
Findable
100
90
80
70
60
50
40
30
20
10
Accessible
Reusable


Success
Findable
100
90
80
70
60
50
40
30
20
10
Accessible
Interoperable
Reusable

② Inspecting metadata

Approach

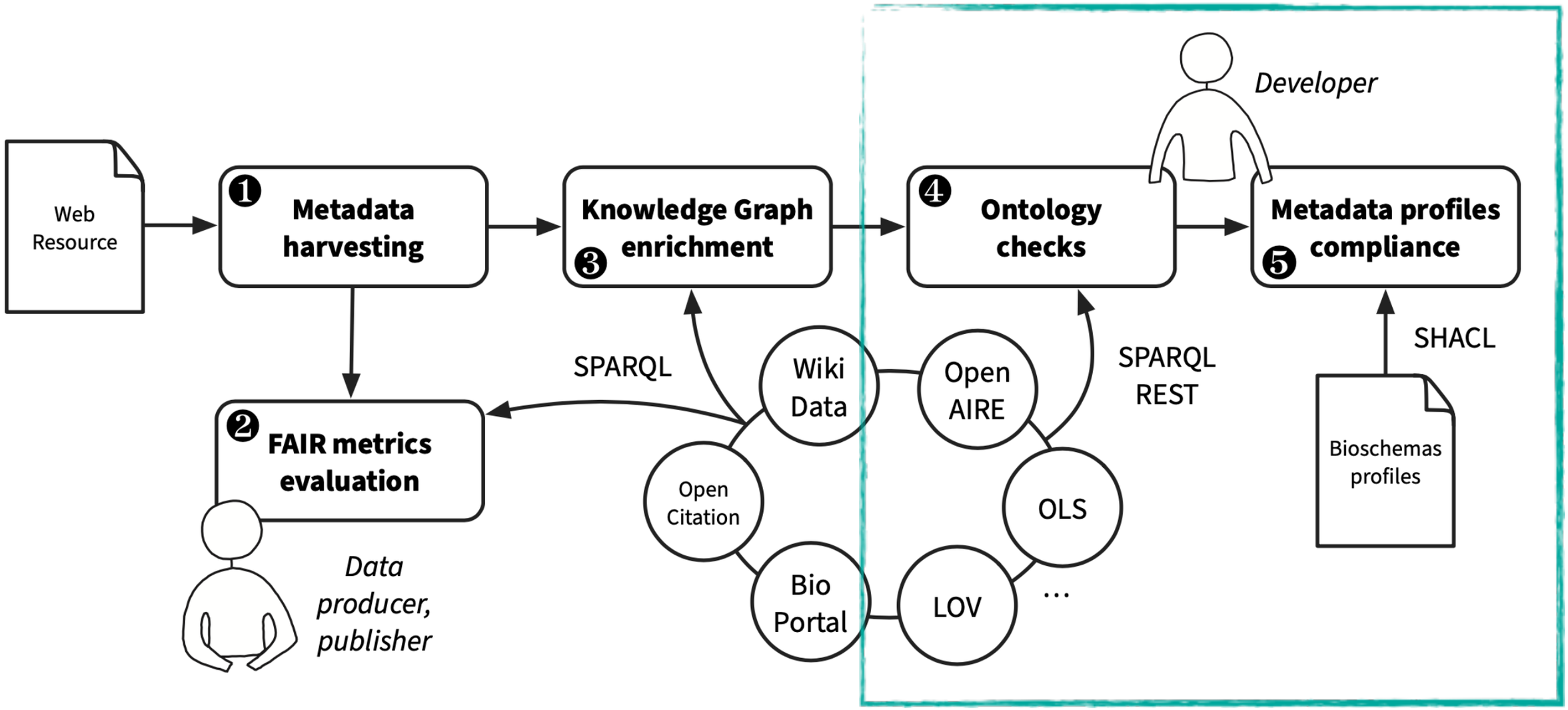
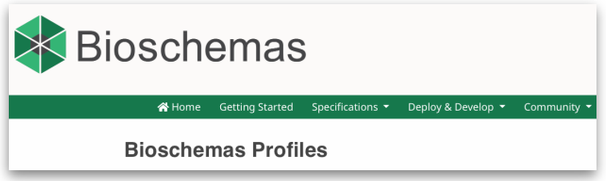


Figure 1 Gathering, enriching and analyzing semantic web annotations in line with FAIR principles.

Metadata completeness

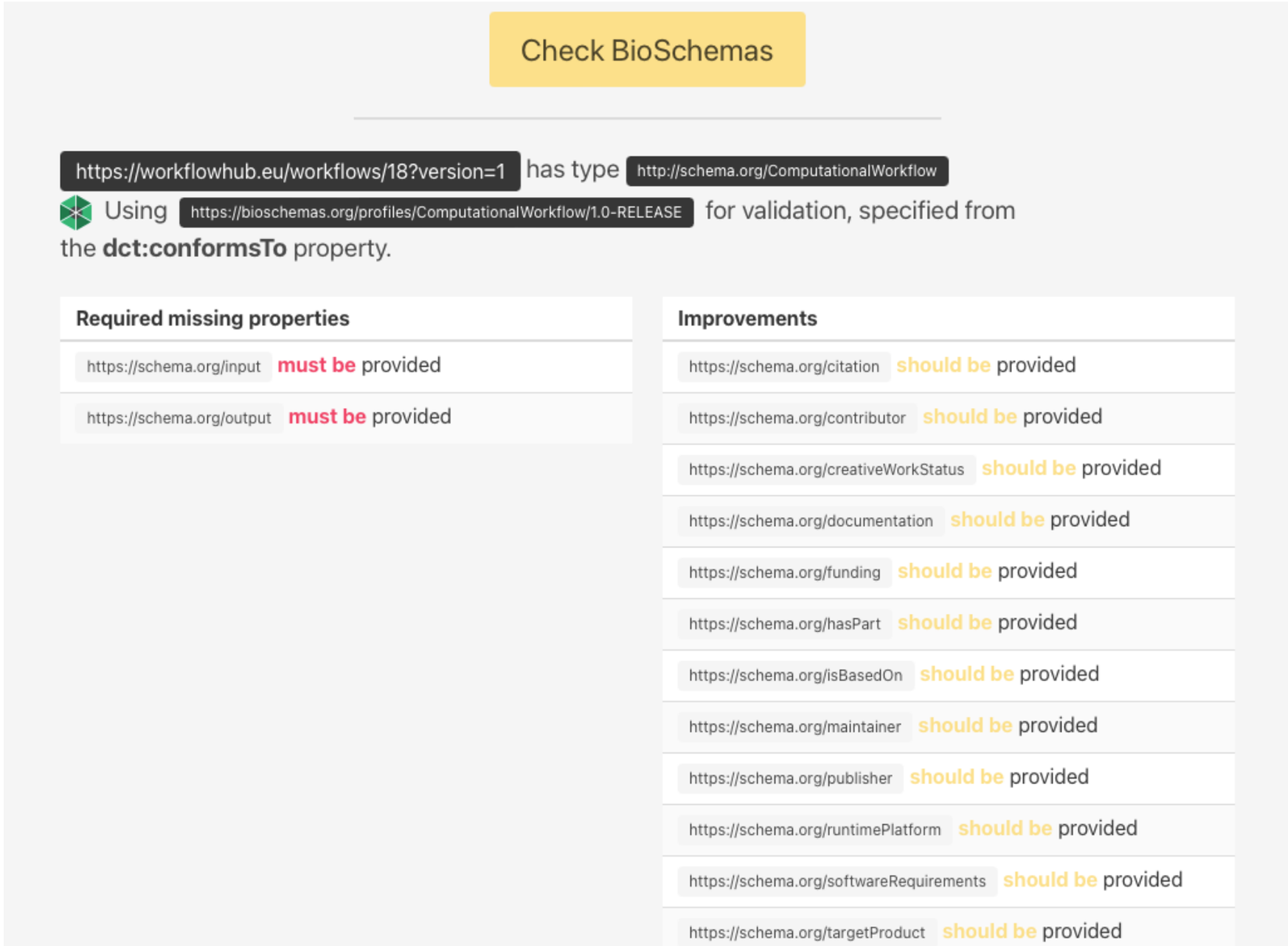


[R1.3: (Meta)data meet domain-relevant community standards]

Validation of Bioschemas profiles:

→ **rank missing** metadata

→ developer **focus** on **minimal** metadata first



Reuse of ontologies

BioPortal OLS **Linked Open Vocabularies (LOV)**

F4: (Meta)data are registered or indexed in a searchable resource

I2: (Meta)data use vocabularies that follow the FAIR principles

R1.3: (Meta)data meet domain-relevant community standards

Step 3: Metadata quality checks

Controlled vocabularies | Bioschemas

We now have a Knowledge Graph grounded to ontology concepts (classes) and relations (properties). Are these classes and properties already known in reference ontology registries such as LOV, OLS or BioPortal ?

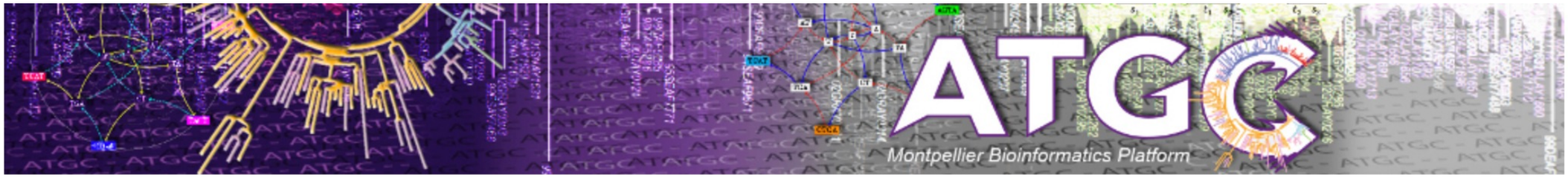
Check Vocabularies

Congratulations ! All Classes and Properties are referenced in one or more of the registries checked !

Classes			
http://schema.org/DataDownload	OLS	LOV	BioPortal
http://schema.org/Organization	OLS	LOV	BioPortal
http://schema.org/Person	OLS	LOV	BioPortal
https://schema.org/Dataset	OLS	LOV	BioPortal

Properties			
http://ogp.me/ns#description	OLS	LOV	BioPortal
http://ogp.me/ns#site_name	OLS	LOV	BioPortal
http://ogp.me/ns#title	OLS	LOV	BioPortal
http://ogp.me/ns#url	OLS	LOV	BioPortal
http://schema.org/affiliation	OLS	LOV	BioPortal
http://schema.org/author	OLS	LOV	BioPortal
http://schema.org/contentSize	OLS	LOV	BioPortal
http://schema.org/contentUrl	OLS	LOV	BioPortal
http://schema.org/creator	OLS	LOV	BioPortal
http://schema.org/dateCreated	OLS	LOV	BioPortal
http://schema.org/dateModified	OLS	LOV	BioPortal
http://schema.org/datePublished	OLS	LOV	BioPortal

Zenodo dataset "Comparison of spatial transcriptomics technologies used for tumor cryosections" https://zenodo.org/records/10863259



PhyML 3.0: new algorithms, methods and utilities

Please cite:
"New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0."
Guindon S., Dufayard J.F., Lefort V., Anisimova M., Hordijk W., Gascuel O.
Systematic Biology, 59(3):307-21, 2010.

Get the latest PhyML releases from GitHub.

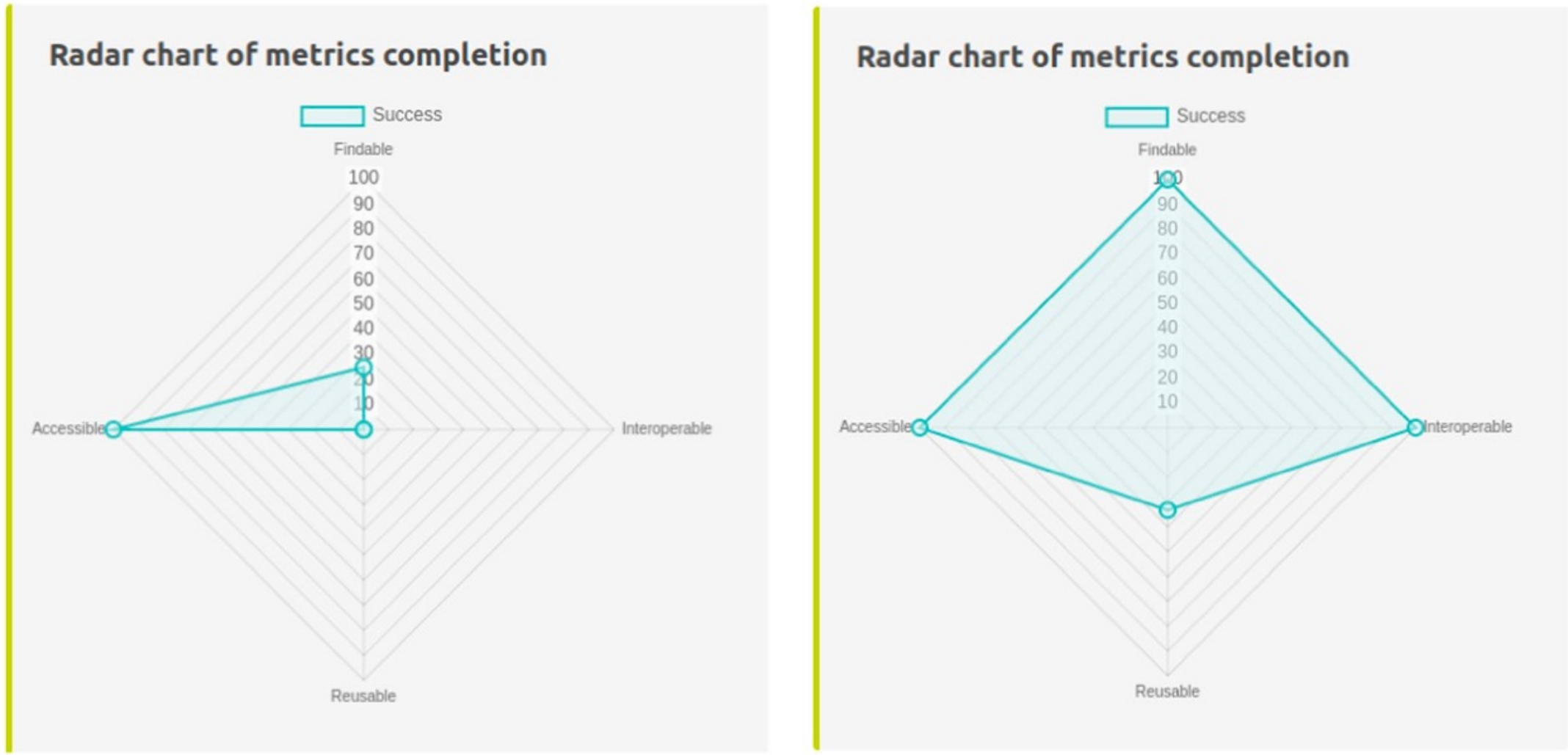


Fig. 6 Comparison of the "Check" results on the PhyML landing page before (left) and after (right) the addition of JSON-LD metadata

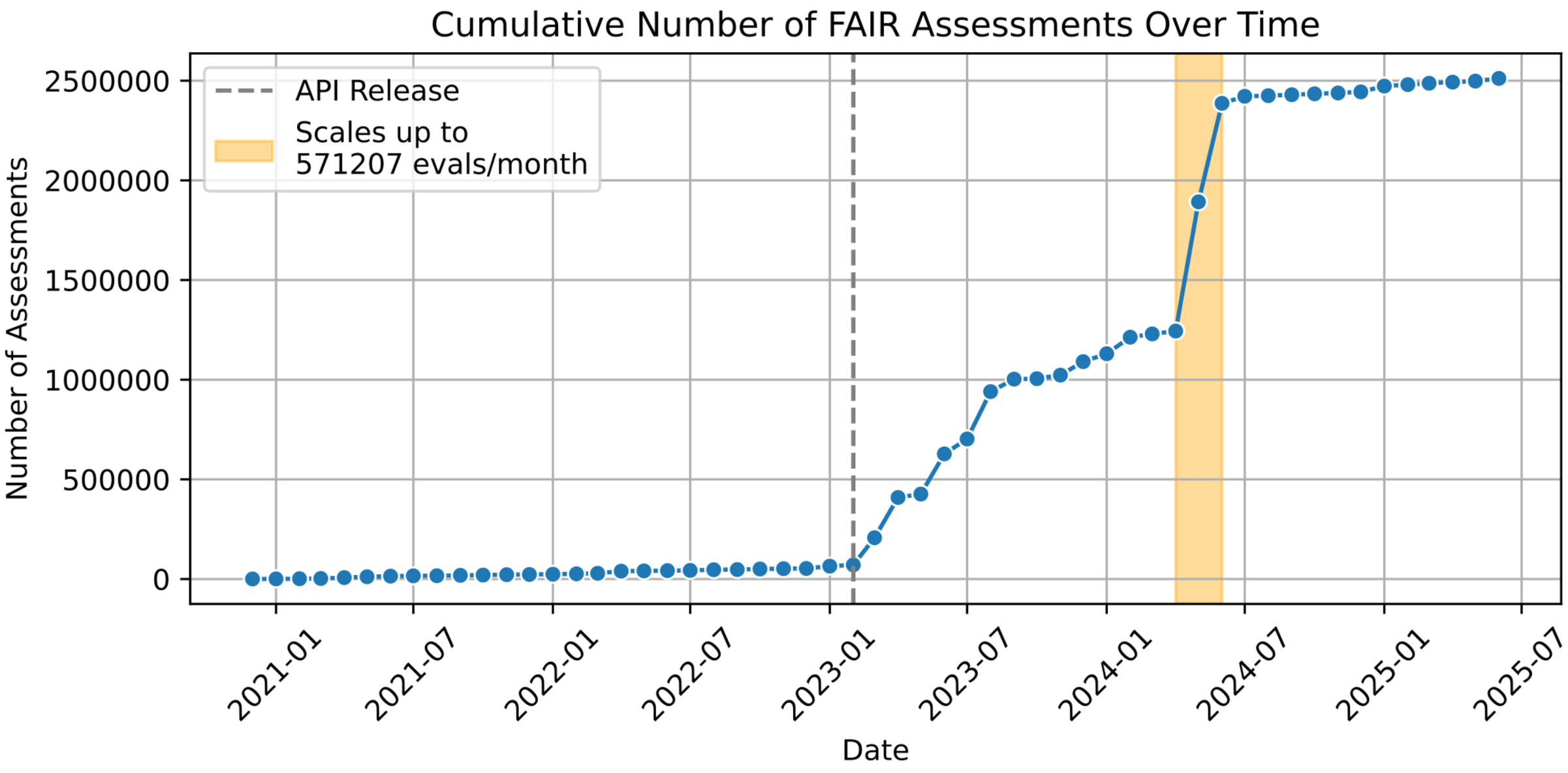
Instant Bonaldi

https://fair-checker.france-bioinformatique.fr/

FAIR-Checker

est-il utilisé?

Usage statistics

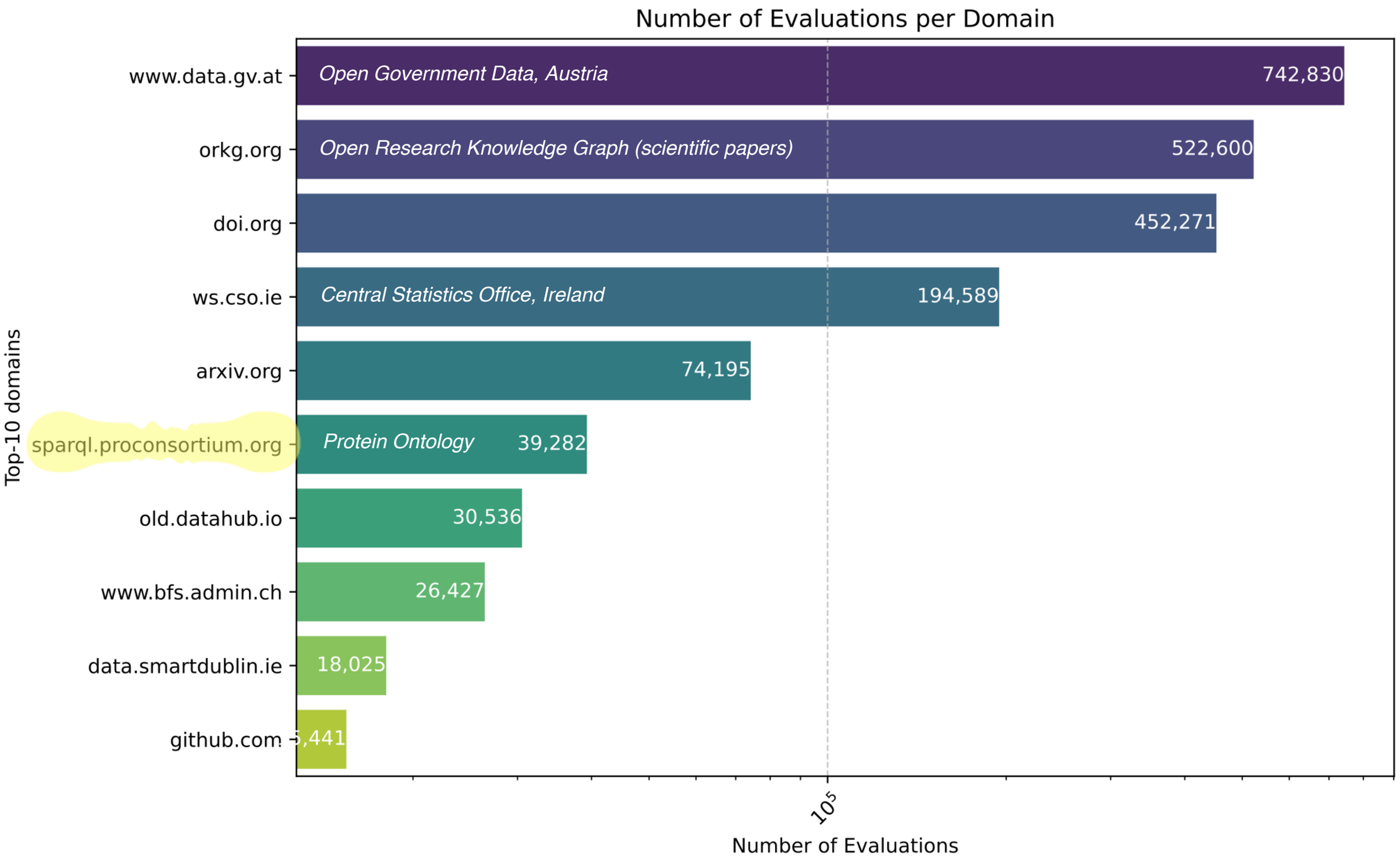


102,180 unique target URLs

47,537 unique target URLs evaluated **more than 2 times**

437 unique target URLs evaluated **more than 10 times**

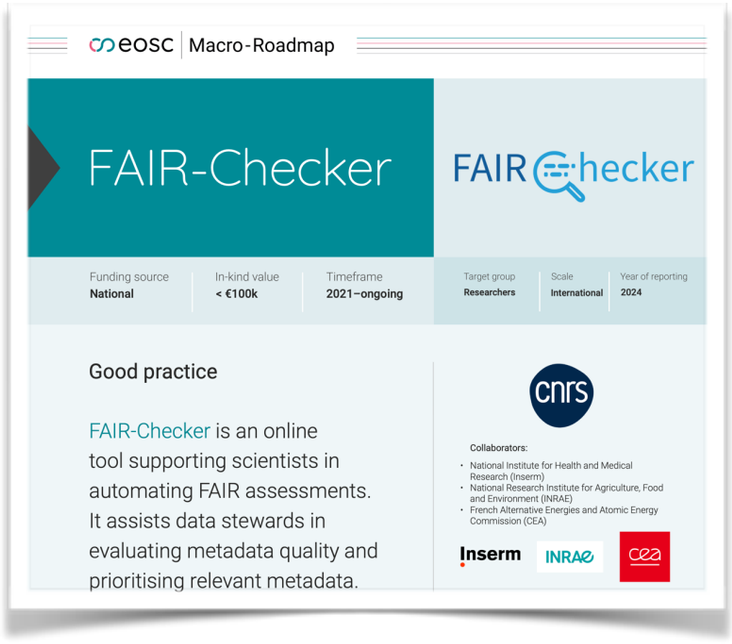
Principaux utilisateurs



Perspectives

Inscription au « Software Delivery Plan » d'Elixir

https://fair-checker.france-bioinformatique.fr



Future works

- Support "FAIR-Signposting" for **better metadata consumption**
- Bioschemas **profile recommender**
- Allow users to provide **missing metadata**
- Suggest semantic metadata based on **AI generation** pipeline
- Retrospective **usage study**

Remerciements

Marie-Dominique Devignes

Thomas Rosnet

Alban Gaignard

alban.gaignard@univ-nantes.fr

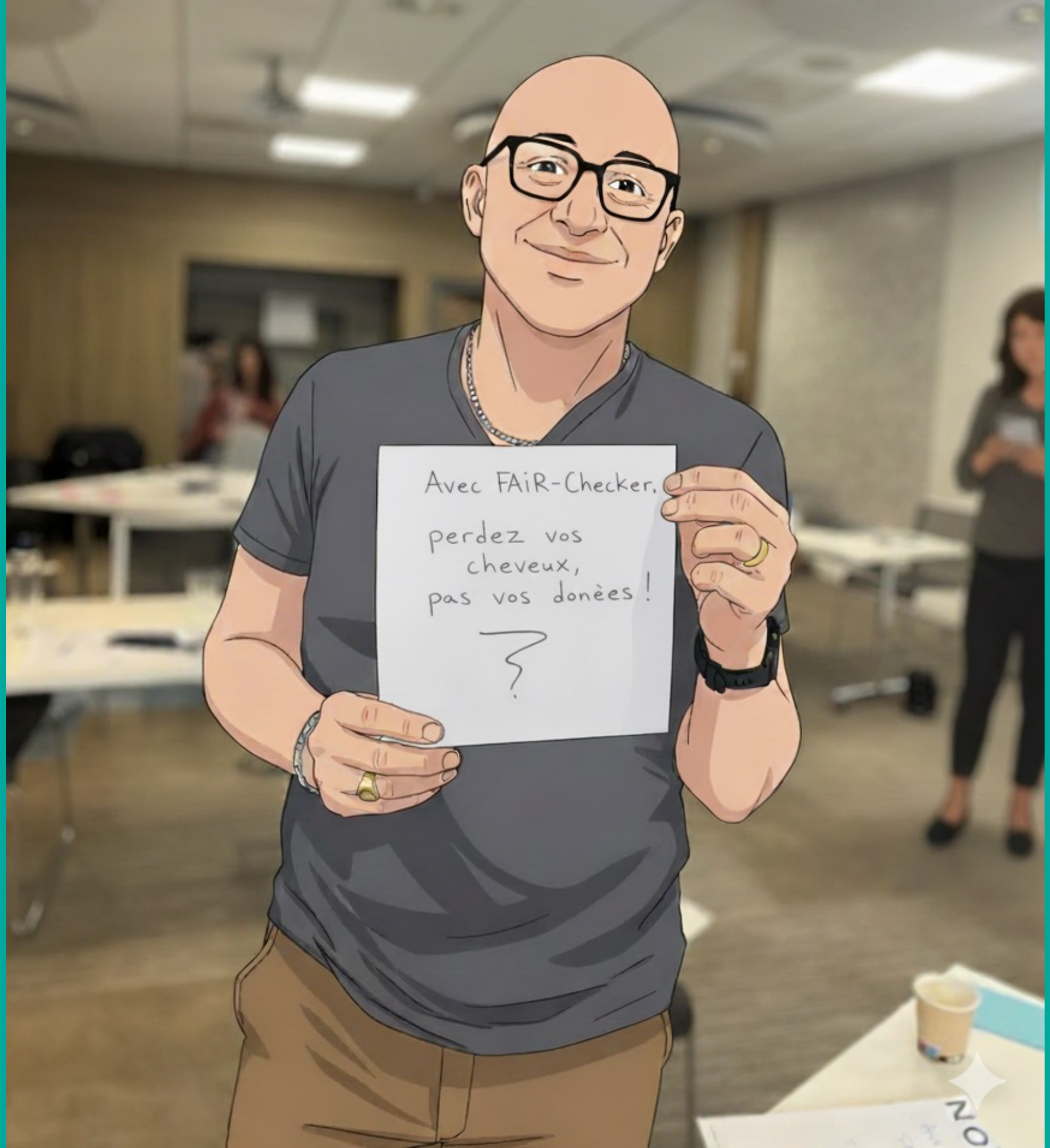
Philippe Lamare

Sylvie Cazalens

Jennien Andersen

Funding:
ANR AAPG2019 DeKaloG
French institute for Bioinformatics (Elixir-FR)

https://fair-checker.france-bioinformatique.fr


Avec FAiR-Checker,
perdez vos
cheveux,
pas vos donèes !
?