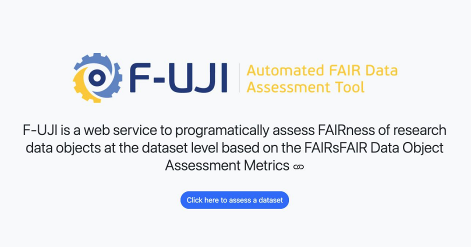
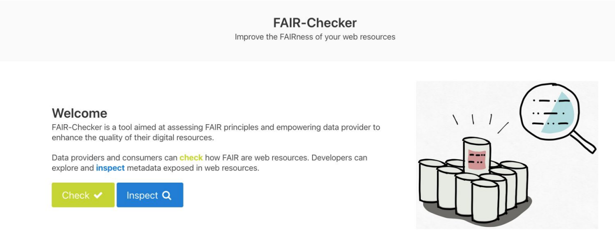
BioDiv-FC: Biodiversity FAIR-Checker

Domain-aware FAIRness evaluation for biodiversity data (genomics • specimens • ecology)

Goal: make FAIR evaluations reflect real biodiversity data practices, object metadata completeness, and turn results into actionable recommendations.



F-UJI is a web service to programatically assess FAIRness of research data objects at the dataset level based on the FAIRsFAIR Data Object Assessment Metrics

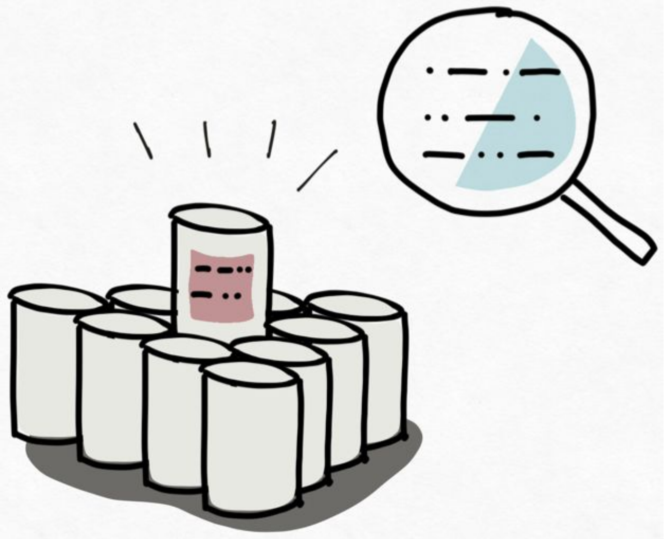

FAIR-Checker

Improve the FAIRness of your web resources

Welcome

FAIR-Checker is a tool aimed at assessing FAIR principles and empowering data provider to enhance the quality of their digital resources.

Data providers and consumers can check how FAIR are web resources. Developers can explore and inspect metadata exposed in web resources.

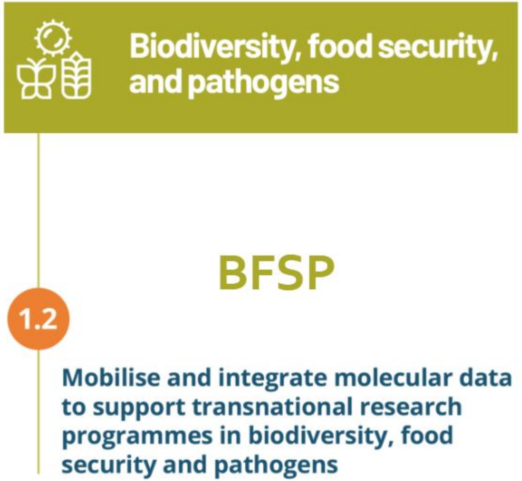


Biodiversity, Food Security, & Pathogens

Science

2024-2028 Programme

BFSP Mission & Scope



Biodiversity, food security, and pathogens

BFSP

1.2

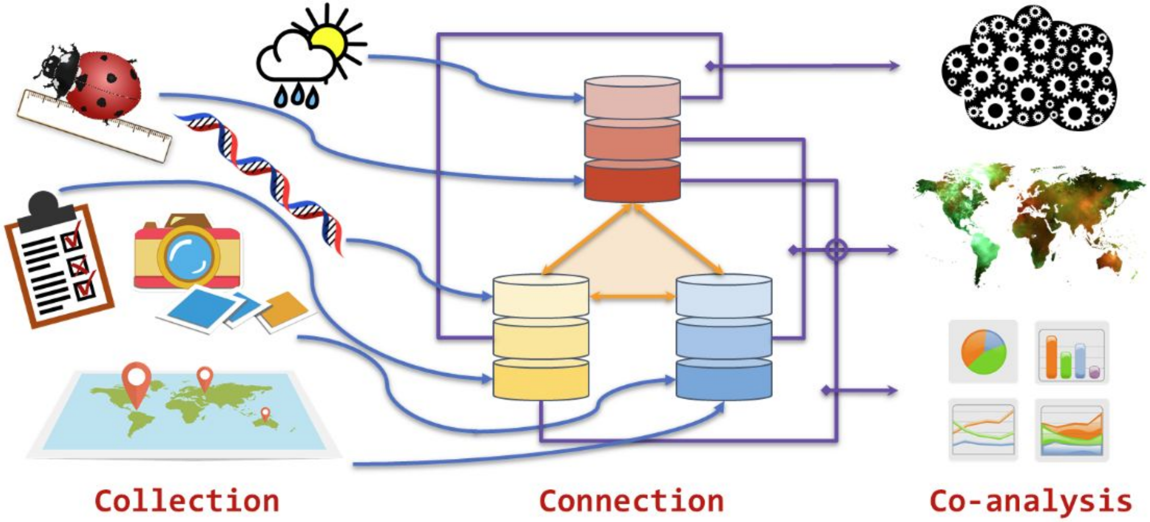
Mobilise and integrate molecular data to support transnational research programmes in biodiversity, food security and pathogens

- Mobilise and integrate genetics, phenomics, observations, and omics data
- Expand Node research in biodiversity, agriculture and aquaculture, and one health approaches to studying pathogens
- Focus on research data management and analysis: tools and services, not production
- Aim to keep activities distinct from Cellular and Molecular Research (CMR) and Human Data and Translational Research (HDTR)

Biodiversity, Food Security, & Pathogens

Science

- **Data brokering solutions**
  - Tools & services supporting data producers to mobilise data
- **Resource interoperability**
  - Technical solutions & common standards connecting data
- **Data discovery & analysis**
  - Software, workflows & curated reference datasets supporting researchers



ELIXIR BFSP Strategy: https://zenodo.org/records/15133254

Generic FAIR checks miss biodiversity specifics

★ Biodiversity outputs are heterogeneous (samples, sequences, specimen records, images, time-series).

★ Generic FAIR tools can under-score well-curated resources when domain expectations are not encoded.

★ The most common blocker is missing or invalid machine-readable metadata (JSON-LD/RDFa/RDF).

★ Low scores reduce discoverability and reuse.

FAIR (in plain terms)

FAIR is about making digital objects easy for both humans and machines to find, access, combine, and reuse using stable identifiers, clear licensing/provenance, and standardised metadata.

Our focus

In FAIR-Checker we evaluate metadata exposure and quality (not the data files and/or their quality). If machines can't reliably read your metadata, your FAIR score and reuse potential drop.

Swiss Institute of Bioinformatics

What we are building

1) Domain-aware assessment baseline

Implement 'core + extensions' metadata requirements across Biodiversity data:

★ Genomics (ERGA/ENA/Barcodes/Bold/GGBN)
★ Digital specimens (DiSSCo)
★ Ecology/time-series (eLTER/LifeWatch)

Accounting for minimal vs recommended fields.

2) Make evaluation actionable

Translate requirements into machine-actionable profiles (Bioschemas + SHACL) and community-specific recommendations so researchers know exactly what to fix and where.

3) Integrate with infrastructure

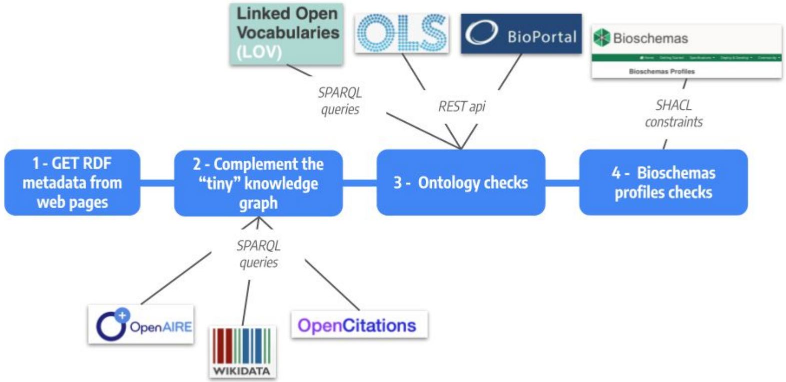
Deliver a biodiversity plugin for FAIR-Checker with:

★ Domain-specific metrics
★ REST API + CLI + dashboard outputs
★ Infrastructure-friendly harvesting

Result: FAIR evaluation becomes a community-owned improvement loop.

SIB
Swiss Institute of
Bioinformatics

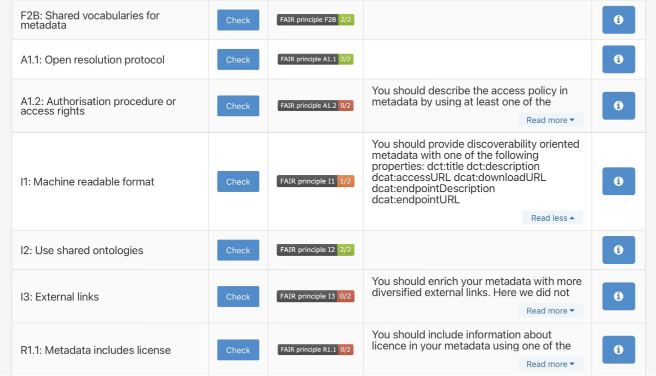
How FAIR-Checker works and what input/output means



Resource identifier (URL/DOI)

https://commons.datacite.org/doi.org/10.3205/09dgnc137

Valid URL/DOI - The input contains the following DOIs that you can also test: 10.3205/09dgnc137



F2B: Shared vocabularies for metadata	Check	FAIR principle F2B 2/2		ℹ
A1.1: Open resolution protocol	Check	FAIR principle A1.1 2/2		ℹ
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 ▾	ℹ
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 ▴	ℹ
I2: Use shared ontologies	Check	FAIR principle I2 2/2		ℹ
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 ▾	ℹ
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 ▾	ℹ

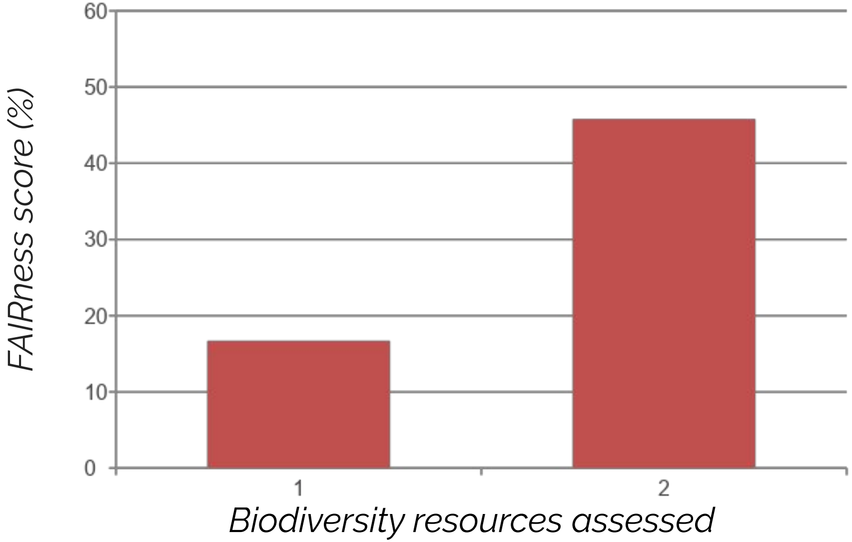
Inputs

- ★ Landing page URL
- ★ DOI URL
- ★ Accession or PID

Outputs

- ★ A score + per-metric results with evidence
- ★ Recommendations

Typical biodiversity resources score low (metadata exposure)



- ★ In our initial benchmark, scores clustered at the low end (≈17–46%).
- ★ Machine-readable metadata is missing, split across places, or technically invalid.
- ★ Even “good” examples often miss licence/provenance links, lowering reuse-related metrics.

What BioDiv-FC changes

We keep generic FAIR checks, but add domain-aware profiles so **good biodiversity practice** is recognised and the tool tells you exactly what to improve next.

How we work with communities

Short contact
object, metadata, priorities

→

Example records
good, typical, tricky

→

Implement profile
minimal vs recommended

→

Community review
feedback + live testing

→

Refined checks
better scoring + advice

Community co-creation of Bioschemas (Types) & Profiles

Step 1 — Create a ‘type’

- ★ **A total set of ‘properties’ (attributes) that help describe something**

- ★ Discussions with Biodiversity community (BGE project) (late 2023, German Biohackathon))
- ★ Survey of broader audience using samples (ELIXIR Biohackathon)
- ★ Formation of Bioschemas Sample Working Group
- ★ **https://bioschemas.org/groups/Samples**

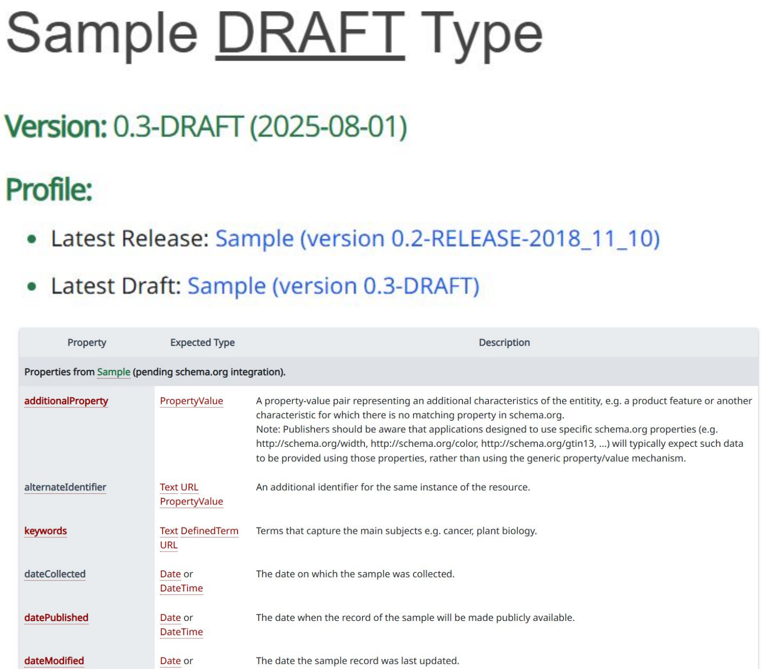
Community co-creation of Bioschemas (Types) & Profiles

Step 1 — Create a 'type'

- ★ **A total set of 'properties' (attributes) that help describe something**

- ★ Discussions with Biodiversity community (BGE project) (late 2023, German Biohackathon))
- ★ Survey of broader audience using samples (ELIXIR Biohackathon)
- ★ Formation of Sample Working Group
- ★ **https://bioschemas.org/groups/Samples**

- ★ Open to broader comment (2 weeks)



Sample DRAFT Type

Version: 0.3-DRAFT (2025-08-01)

Profile:

- Latest Release: Sample (version 0.2-RELEASE-2018_11_10)
- Latest Draft: Sample (version 0.3-DRAFT)

Property	Expected Type	Description
Properties from Sample (pending schema.org integration).		
additionalProperty	PropertyValue	A property-value pair representing an additional characteristics of the entitity, e.g. a product feature or another characteristic for which there is no matching property in schema.org. Note: Publishers should be aware that applications designed to use specific schema.org properties (e.g. http://schema.org/width, http://schema.org/color, http://schema.org/gtin13, ...) will typically expect such data to be provided using those properties, rather than using the generic property/value mechanism.
alternateIdentifier	Text URL PropertyValue	An additional identifier for the same instance of the resource.
keywords	Text DefinedTerm URL	Terms that capture the main subjects e.g. cancer, plant biology.
dateCollected	Date or DateTime	The date on which the sample was collected.
datePublished	Date or DateTime	The date when the record of the sample will be made publicly available.
dateModified	Date or	The date the sample record was last updated.

MANCHESTER 1824
The University of Manchester

Community co-creation of Bioschemas (Types) & Profiles

Step 1 — Create a ‘type’

- ★ **A total set of ‘properties’ (attributes) that help describe something**
- ★ Discussions with Biodiversity community (BGE project) (late 2023, German Biohackathon))
- ★ Survey of broader audience using samples (ELIXIR Biohackathon)
- ★ Formation of Sample Working Group
- ★ **https://bioschemas.org/groups/Samples**
- ★ Open to broader comment (2 weeks)

Sample DRAFT Type

Version: 0.3-DRAFT (2025-08-01)

BioSample (Type)

Version: 0.2-DRAFT (2025-08-15)

Profile:

- Latest Draft: BioSample (version 0.1-DRAFT-2019_11_12)

alternateIdentifier	Text URL PropertyValue	An additional identifier for the same instance of the resource.
keywords	Text DefinedTerm URL	Terms that capture the main subjects e.g. cancer, plant biology.
dateCollected	Date or DateTime	The date on which the sample was collected.
datePublished	Date or DateTime	The date when the record of the sample will be made publicly available.
dateModified	Date or	The date the sample record was last updated.

MANCHESTER 1824
The University of Manchester

Community co-creation of Bioschemas Profiles

Step 2 — Create a ‘profile’

- ★ A prioritised list of ‘properties’ (attributes) with whether they are mandatory, cardinality and expected values (from a list of compliant options)

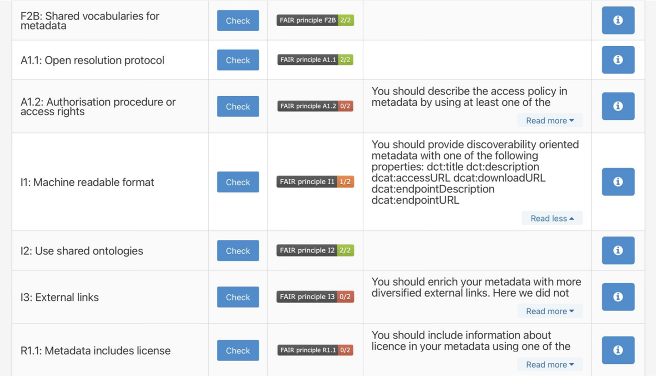
- ★ **Discussions with Biodiversity community - online workshops, surveys, f2f**
- ★ **Inform broader community (bioschemas mailing list, relevant projects where ‘samples’ are used**
- ★ **Prototyping implementation activity (BFSP)**
- ★ **Show and tell, dissemination activities**

- ★ (highlight as a general strategy)

Resource identifier (URL/DOI)

https://commons.datacite.org/doi.org/10.3205/09dgnc137

Valid URL/DOI - The input contains the following DOIs that you can also test: 10.3205/09dgnc137



F2B: Shared vocabularies for metadata	Check	FAIR principle F2B 2/2		ℹ
A1.1: Open resolution protocol	Check	FAIR principle A1.1 2/2		ℹ
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 ▾	ℹ
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 ▴	ℹ
I2: Use shared ontologies	Check	FAIR principle I2 2/2		ℹ
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 ▾	ℹ
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 ▾	ℹ

- ★ A **community-nuanced** score + per-metric results with evidence
- ★ **Tailored** Recommendations - not generic

MANCHESTER 1824
The University of Manchester

What's next

1. We run FAIR-Checker on your exemplars (DONE)
2. Share a short report
3. Publish a draft profile for review
4. Integrate into the BioDiv-FC plugin.

Swiss Institute of Bioinformatics

BioDiv-FC - The teams

SIB Swiss Institute of Bioinformatics

elixir FRANCE

MANCHESTER 1824 The University of Manchester

Coordination + requirements

Robert Waterhouse

Chiara Bortoluzzi

Christian de Guttry

Technical implementation

Alban Gaignard

Fred de Lamotte

Brieuc Quemeneur

Erwan Corre

Community engagement

Nick Juty

Carole Goble