

FAIRy by Datadabra

A local-first validator, fixer and packager for research datasets that helps reduce bounced submissions and simplifies FAIR data compliance

Goal: portable 'BIOFAIR-ready' preflight you can run locally before handing data to any repository

Seeking pilot collaborators across genomics ↔ environmental data

Why this matters for an integrated data network

- Metadata is messy (weird encodings, inconsistent IDs, non-ISO dates)
- Submissions bounce; curators say “please fix,” leading to multi-day loops
- There’s no portable, local preflight before handoff to curators

This is a standards and training gap: submitters don't get consistent guidance up front and curators have to triage

What FAIRy Already Does

Fix → re-run → tells you what cleared

```
=== FAIRy Preflight ===
Rulepack:      GEO-SEQ-BULK@0.1.0
FAIRy version: 0.1.0
Run at (UTC):  2025-10-29T05:57:07.345059+00:00
FAIL findings: 1
WARN findings: 0
submission_ready: False
Report JSON:   out/report.json

Example finding:
[FAIL] GEO.BIO.CONTEXT_MISSING @ row 1
  why: Each sample needs biological source info (tissue / cell_line / cell_type).
  fix: Provide at least one of tissue, cell_line, or cell_type for each row.

Resolved since last run:
✓CORE.DATE.INVALID_ISO8601
✓CORE.ID.UNMATCHED_SAMPLE
```

- Runs locally (offline)
- PASS/FAIL/WARN with why and how to fix
- Write JSON and Markdown report with timestamp, rulepack version you can share
- On re-run highlights what cleared (fast feedback loop)

Outputs a shareable, data-free readiness sheet (PASS / WARN / FAIL + how to fix) you can hand to a PI or curator . A candidate handoff artifact for a future BIOFAIR node."

Looking for collaborators (BIOFAIR pilot)

Pilot partners wanted: any domain. We'll capture 5–7 must-have fields, automate locally and report aggregate gaps back to the BIOFAIR community

Scan the QR if you're even a little curious!

- Tell me your name / org / data type
- Add “pilot partner” if you're open to testing a tiny rulepack
- I'll follow up



Jennifer Slotnick
FAIRy / Datadabra
hello@datadabra.com
datadabra.com/biofair
linkedin.com/in/jenniferslotnick/

Long-term hope: compare pilots and find the pieces that travel across domains. Start of a BIOFAIR-ready preflight.