

Enriching Metadata in the Flemish Research Landscape



Dhollander, Evelien 1, a, b, c; Kevin Leonard, 1, c, d; Neyens, Evy 2; Bloemen, Dieuwertje 3; Wuyts, Tom 4; Dengis, Pascale 4; Portier, Marc 5, c

1 Ghent University; 2 ECOOM; 3 KU Leuven; 4 FRIS; 5 VLIZ

a conceptualization; b Writing (first draft); c Writing (review and editing); d visualization (data presentation)

FRDN

Objectives

- Reduce cost for all partners
- Improve quality of research data metadata
- Improve discoverability of research data metadata

Principles

- Enter once, reuse often
- As simple as possible (but no simpler)
- Future proof so growth is possible
- Decentralized solution



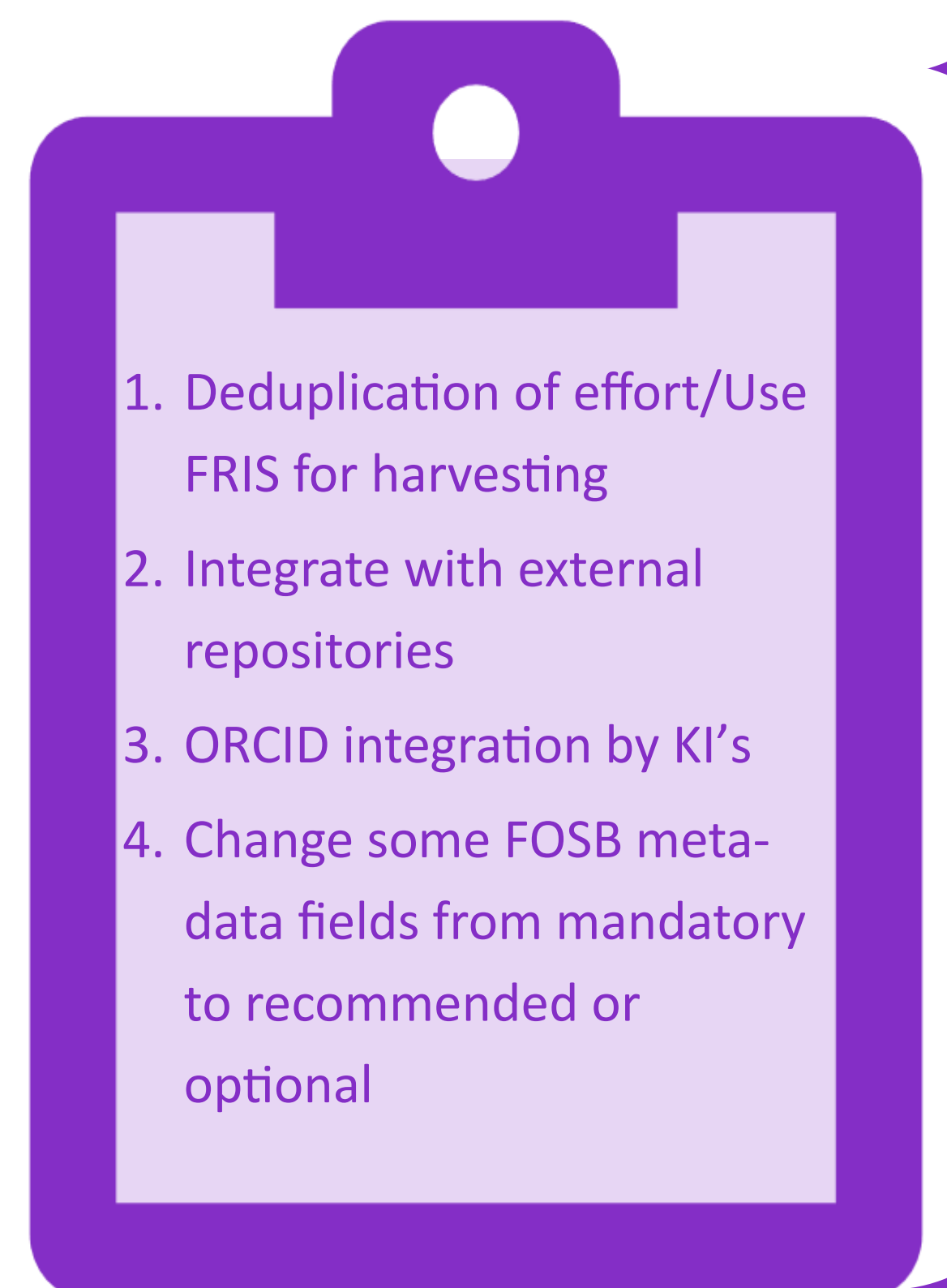
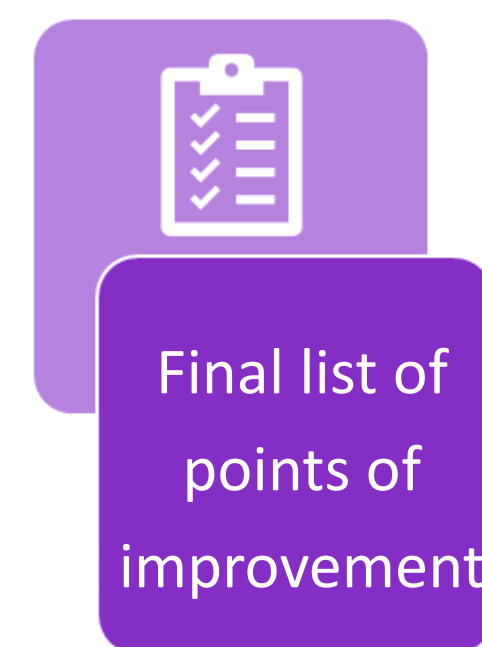
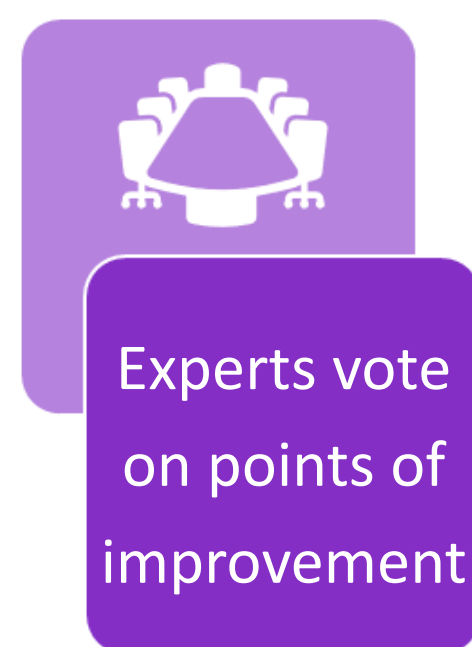
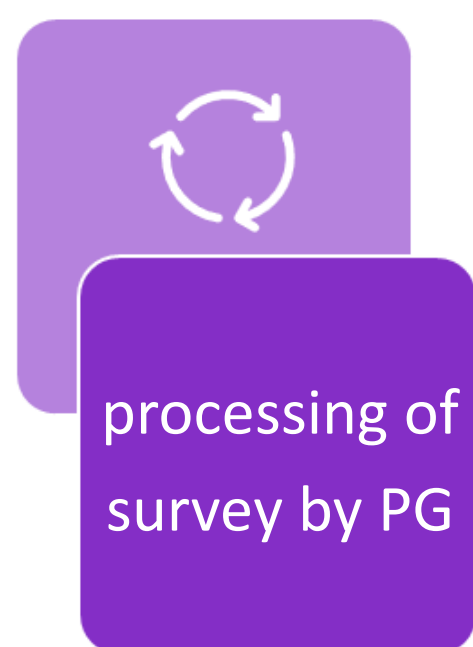
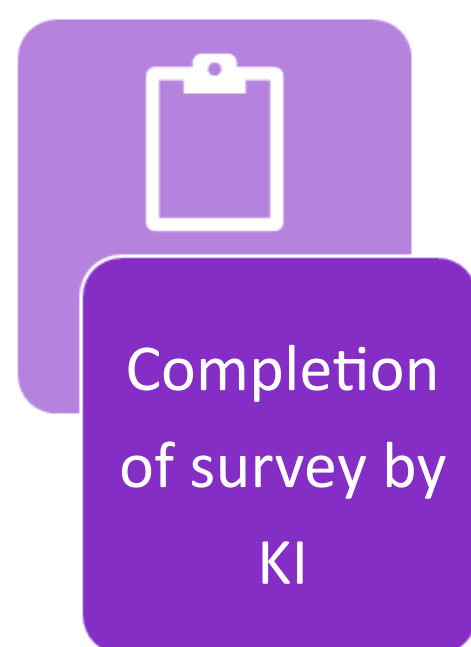
Key deliverable 1

A list of prioritized and externally validated points of improvement within the current ecosystem.

Key deliverable 2

Proof of concept or solution for collecting metadata:

- Define business needs
- Identify limitations of available technologies



PG: project group enriching metadata

KI: knowledge institutions

WG: working group architecture

FRIS: Flemish Research Information Space

CRIS: Current Research Information System

ORCID: Open Researcher and Contributor Identifier

FOSB: Flemish Open Science Board

POC: proof of concept

Assessment of metadata sources

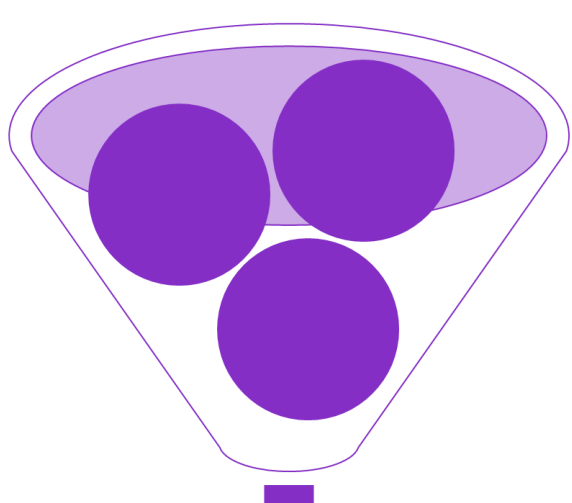
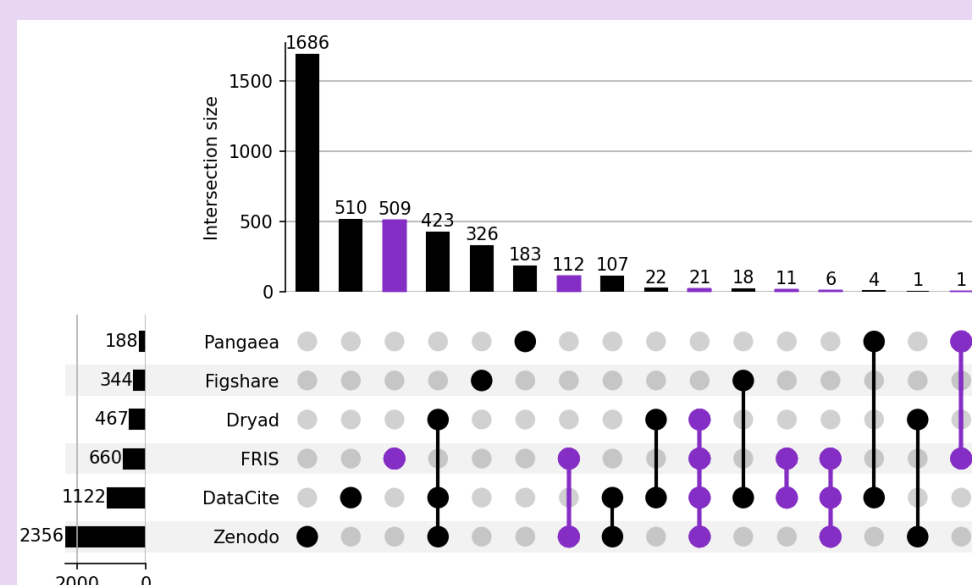
Assess metadata quality of metadata source

Assess ease of use of metadata source

Assess completeness of metadata source

Comparison of different metadata sources

conclusion: DOI aggregator (DataCite) does not return all DOI's of KI's
—> modular approach



POC:

Modular Approach

POC: Modular Approach, simplified flow

FRIS or KI selects source to harvest

Integration built for repository or aggregator?

- Yes —> download code
- No —> create code and share on GitHub

Run code to get csv of persistent identifiers (and metadata)

Integrate found datasets in CRIS or FRIS

Consult institutional dataset finder

Develop query set for repository or aggregator

Deduplicate based on persistent identifier eliminate false positives

