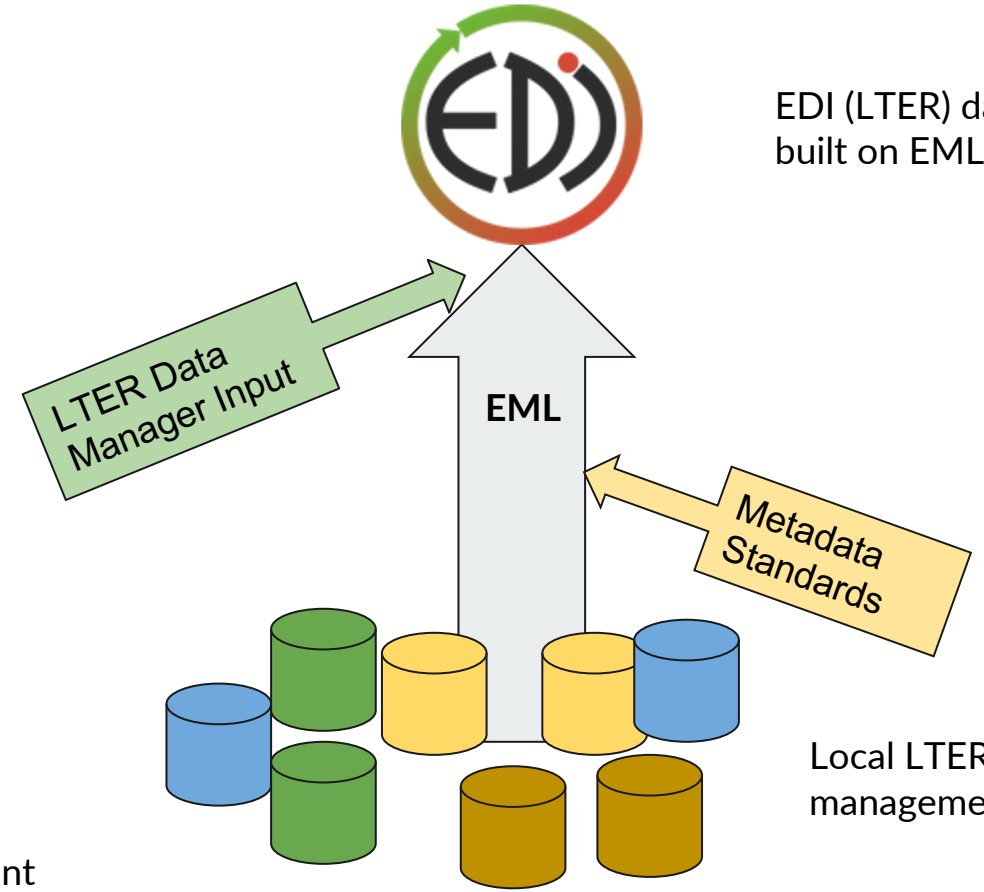
Measuring FAIR

Corinna Gries, Environmental Data Initiative

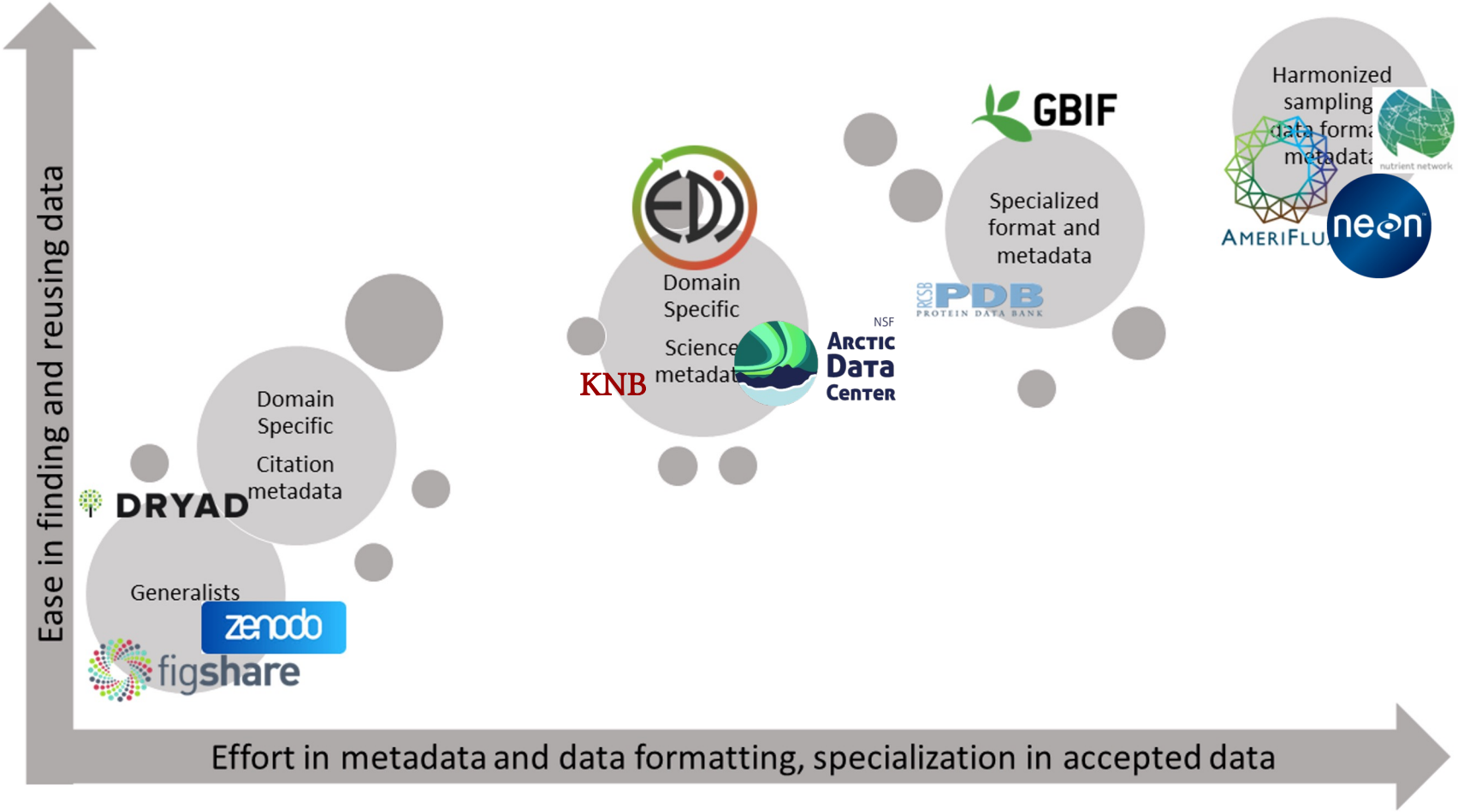
> 40 Years in the Making



EDI (LTER) data repository built on EML

Local LTER site data management systems

NSF Requirement


Ease in finding and reusing data
Effort in metadata and data formatting, specialization in accepted data
Generalists
DRYAD
zenodo
figshare
Domain Specific
Citation metadata
Domain Specific
Science metadata
KNB
NSF
Arctic Data Center
GBIF
Specialized format and metadata
RCSB PDB
PROTEIN DATA BANK
Harmonized sampling
nutrient network
AMERIFLUX
neon

Measuring FAIR

User interviews

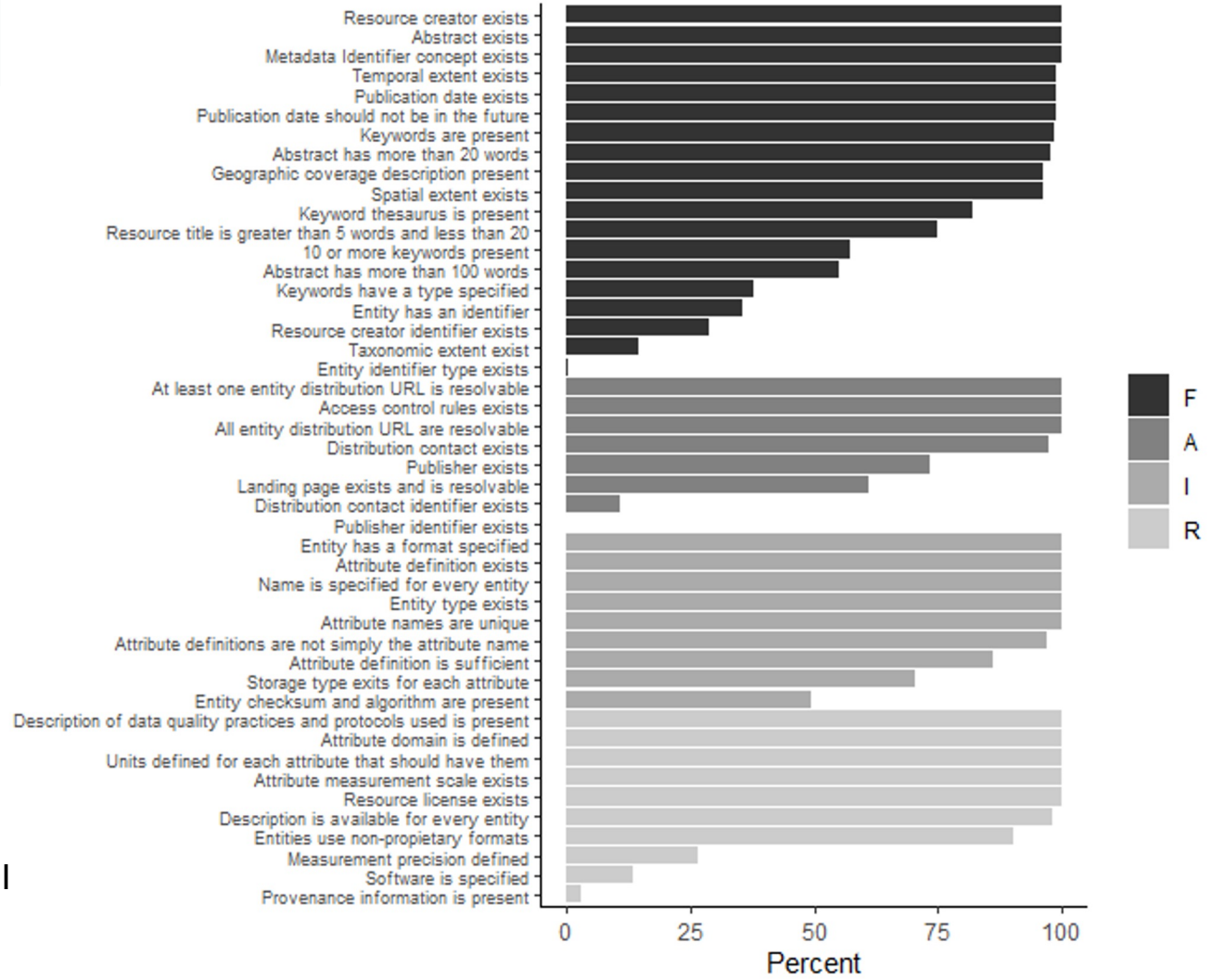
Wilkinson et al. 2016, Research Data Alliance (RDA), DataONE

Data Use: data downloads, data citations, data provenance

Peer review - data papers

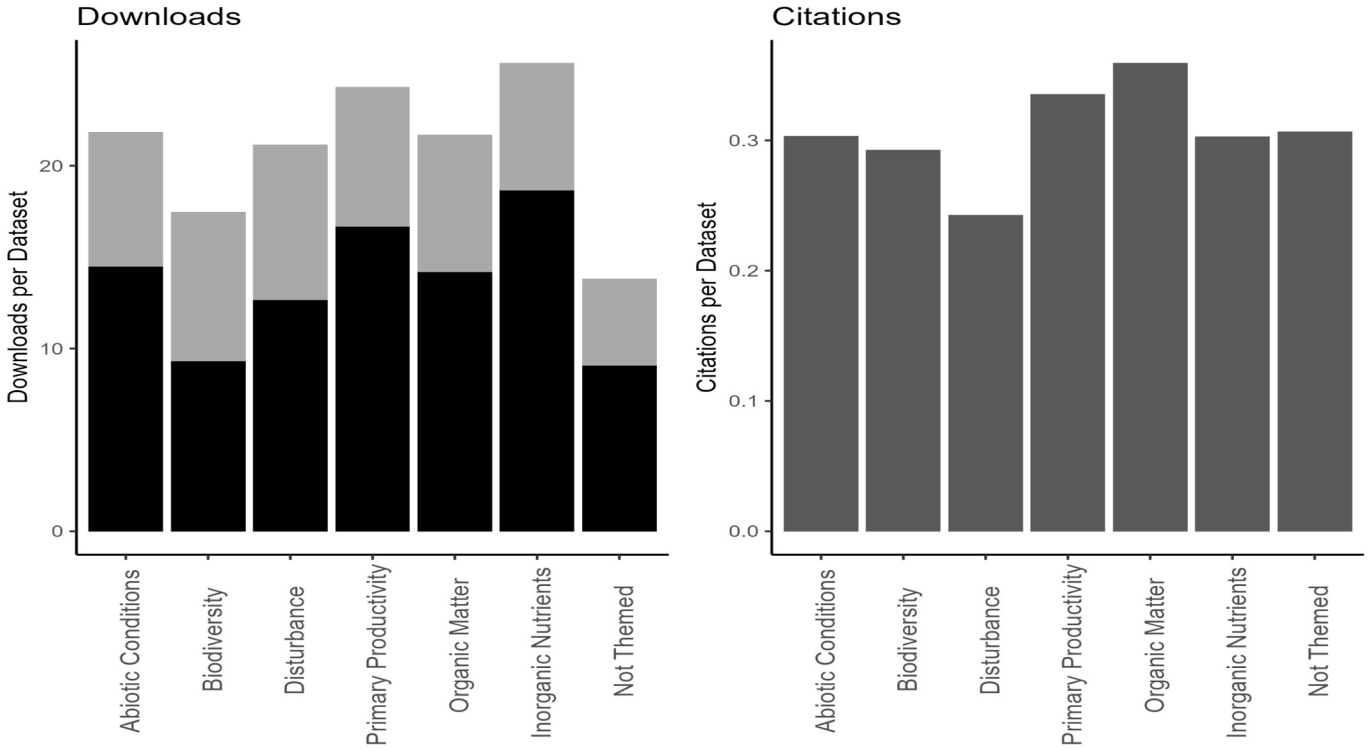
Design patterns

Machine Evaluation FAIR Metadata

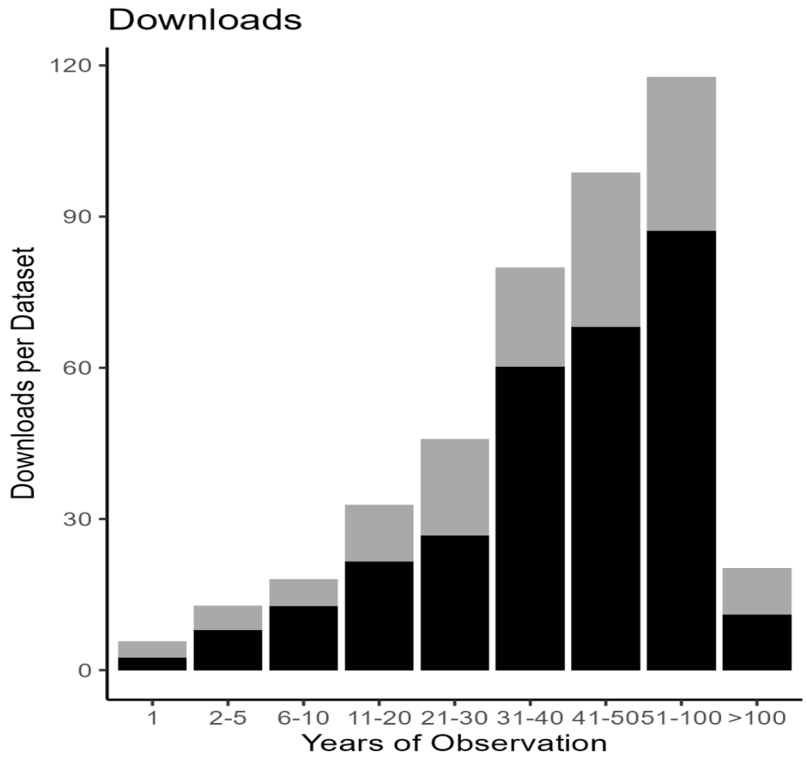


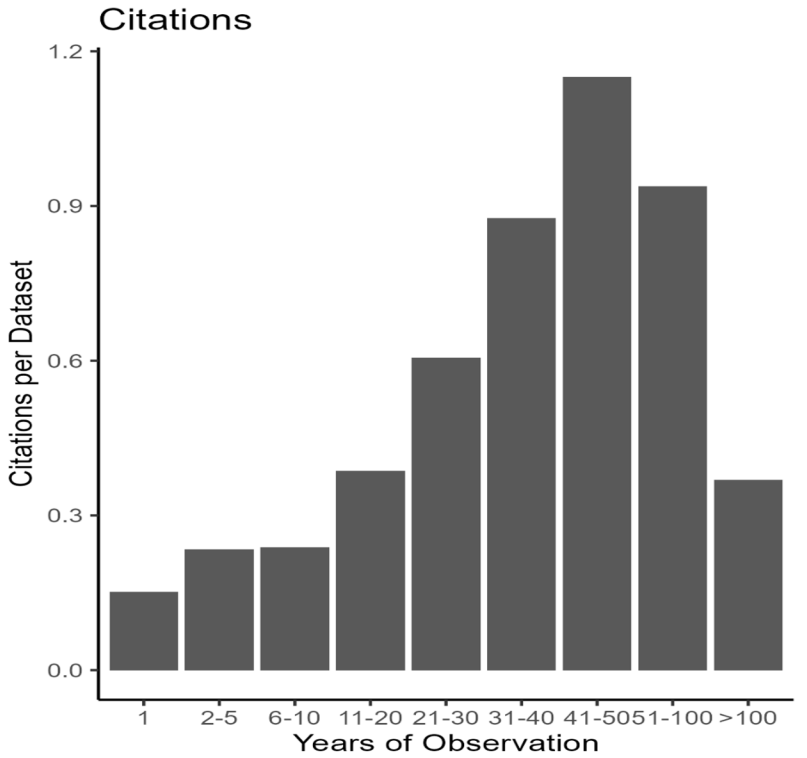
~ 9000 datasets in EDI

Evaluating FAIR through Data Use



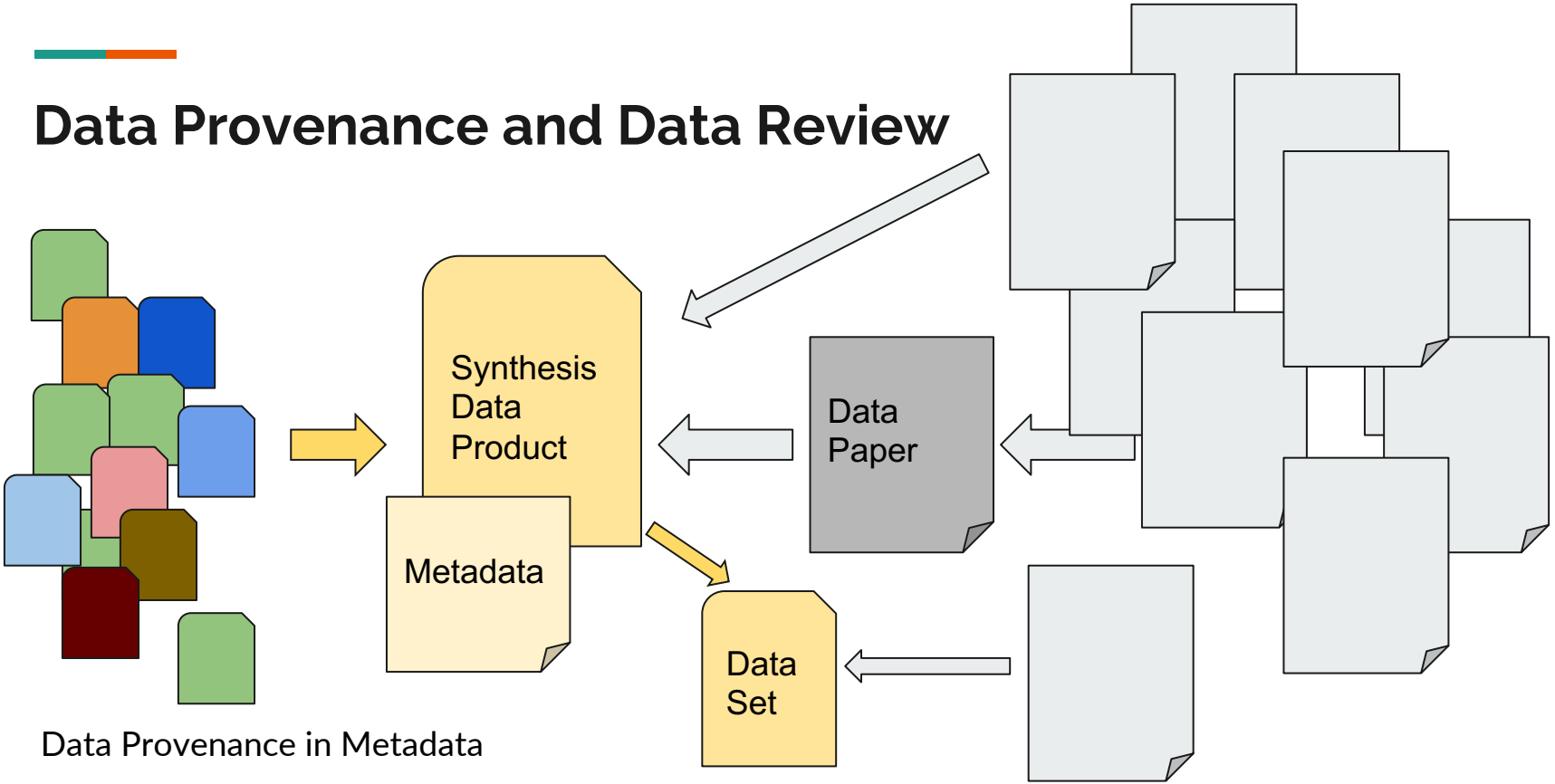
Evaluating FAIR through Data Use





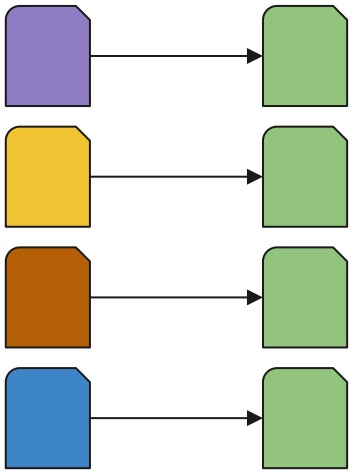
E.g., https://portal.edirepository.org/nis/mapbrowse?packageid=edi.101.3 Soranno et al. 2019

Data Provenance and Data Review



Data Provenance in Metadata

Design Pattern



Considerations:

- Original format had a purpose
- Doesn't lose contextual information
- Automated updates

EDI's ecocomDP
~ 70 datasets
R package

Thank you