

A comparative study of platforms for research data management: interoperability, metadata capabilities and integration potential



Ricardo Carvalho Amorim, João Castro, João Rocha da Silva, Cristina Ribeiro
WorldCIST'15, Azores

Agenda

- Our team 1 min
- Research data repositories environment 2 min
- Stakeholders
- Evaluation criteria 9min
- Final remarks 3 min

Our team & how we work



Cristina
Ribeiro



João Rocha
da Silva



João Aguiar
Castro



Ricardo
Carvalho
Amorim

- Partnership with researchers from diverse domains;
- Development of solutions for potential data management problems;
- Focus on solutions for the long tail of science;



Our team & how we work



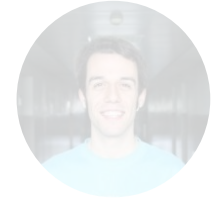
Cristina
Ribeiro



João Rocha
da Silva



João Aguiar
Castro

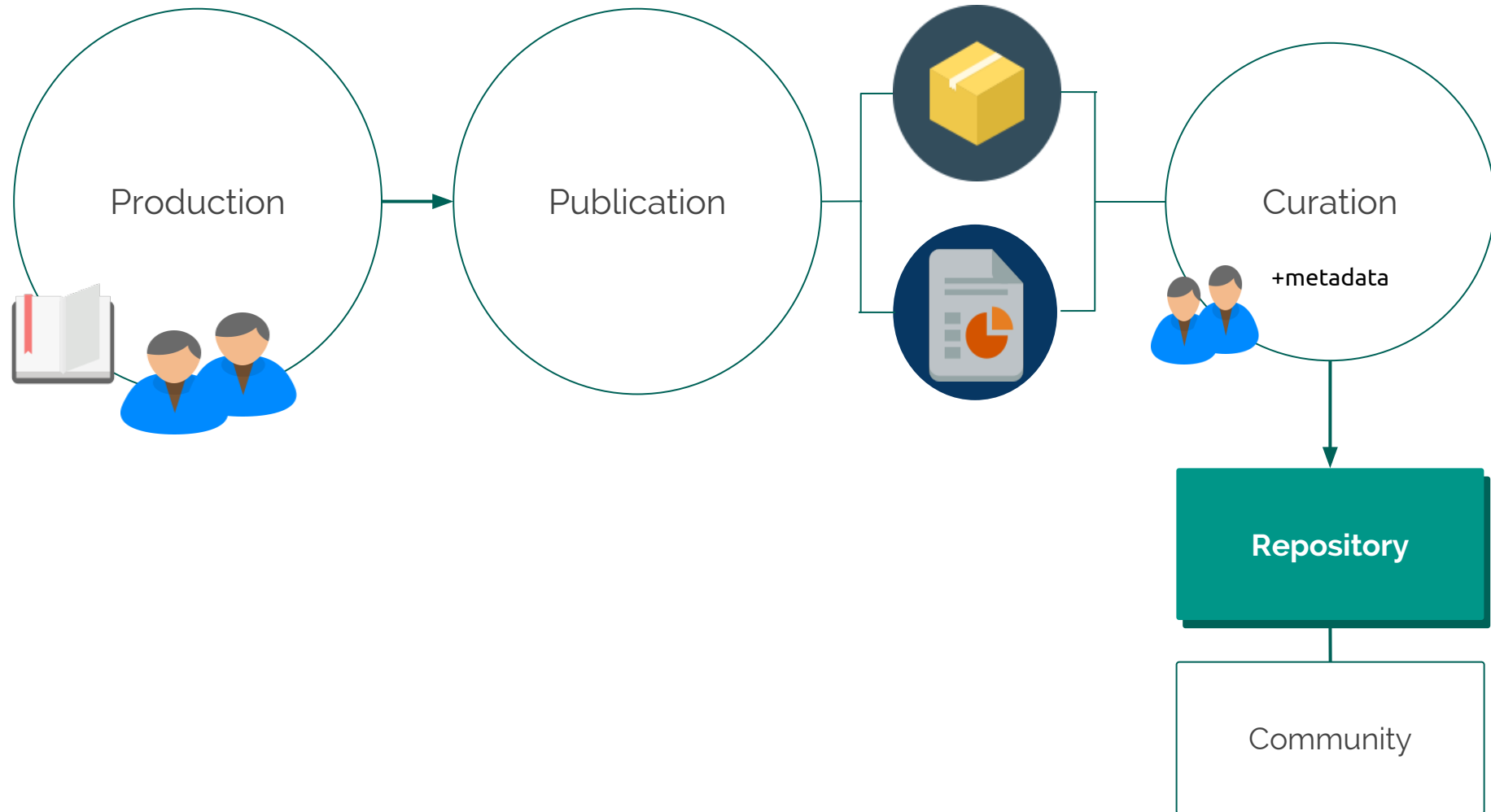


Ricardo
Carvalho
Amorim

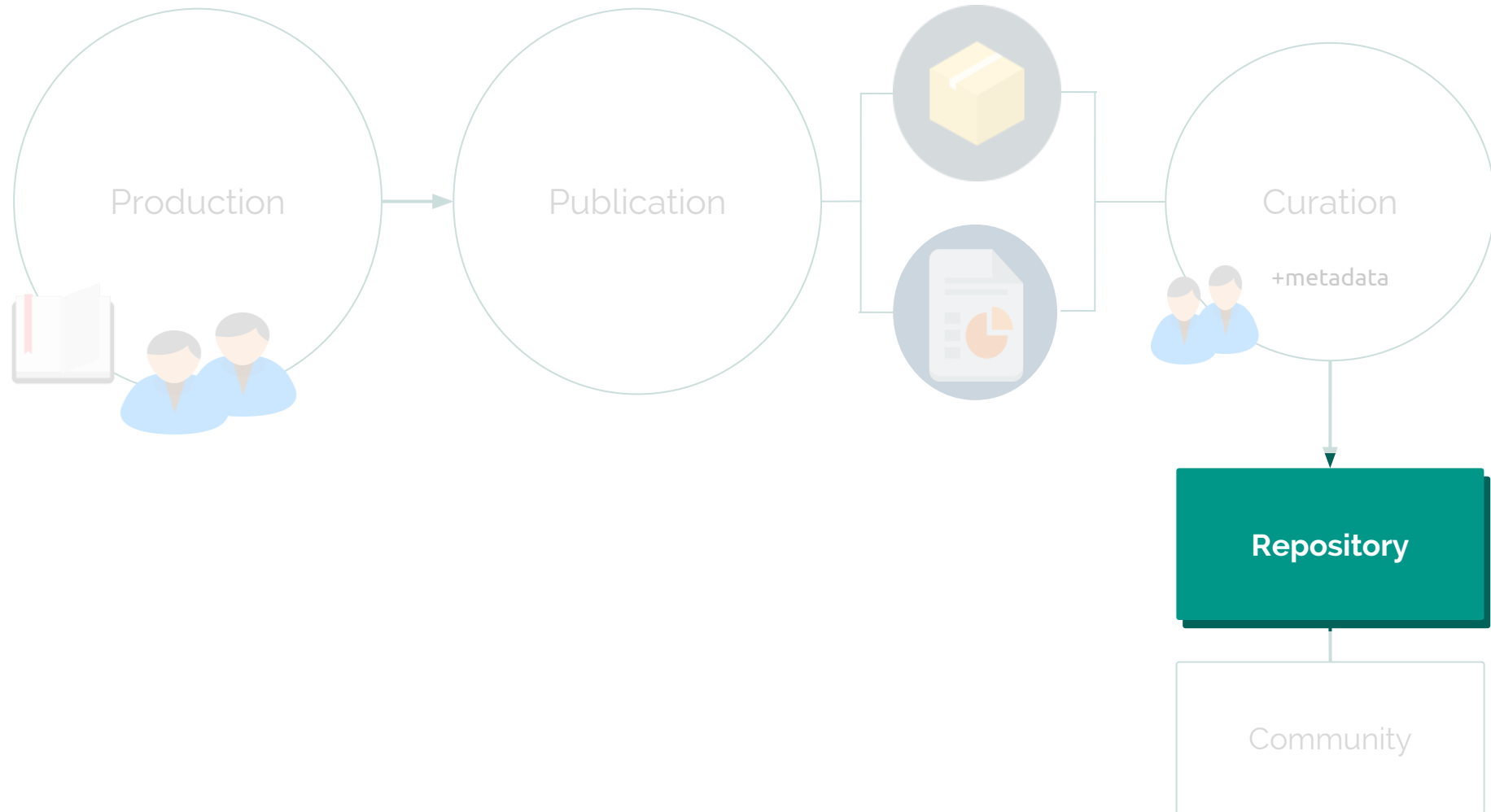
- Partnership with researchers from diverse domains;
- Development of solutions for potential data management problems;
- Focus on solutions for the long tail of science;



Research data repositories



Research data repositories



Research data repositories



Evaluation

Stakeholders & Criteria

Stakeholders

Harvesters

Developers

Institutions

Researchers

Harvesters

Developers

Institutions

Researchers

Established protocols for information dissemination:

- Search engines;
- Repositories directories;
- Data indexers;

Can also be other researchers looking for specific data

Harvesters

Developers

Institutions

Researchers

Work towards integration with other tools

- eg. other sources of data

Maintain the whole infrastructure

Stakeholders

Harvesters

Developers

Institutions

Researchers

Deal with funding issues and budget restrictions;

May have to comply with open access guidelines;

May value full control over their data.

Stakeholders

Harvesters

Developers

Institutions

Researchers

Produce heterogeneous datasets

Often responsible for their management

Experts in their domain

Evaluation criteria

Architecture

Metadata
support

Search
mechanisms

API
completeness

Architecture

Metadata
support

Search
mechanisms

API
completeness

- How is it deployed?
- Is it open source?
- Can it be customized / tailored?

Architecture

Metadata
support

Search
mechanisms

API
completeness

- Are there any metadata **standards** in use?
- Is it able to include **domain-level** metadata?
- Does it validate the gathered descriptions?

Architecture

Metadata
support

Search
mechanisms

API
completeness

- Is there any kind of faceted search?
- Is the repository OAI-PMH compliant?
- Is the metadata available to harvesters?

Architecture

Metadata
support

Search
mechanisms

API
completeness

- Is there a comprehensive API?
- Does this API cover all of the functionality?

Chosen platforms

DSpace

CKAN

Figshare

Zenodo

Results

DSpace

- + Already in use for research publications
- + OAI-PMH compliant
- + Fully customizable metadata schemas
- Can require a development & maintenance team



CKAN

- + Open-source
- + Fully customizable metadata schemas
- + Wide support community

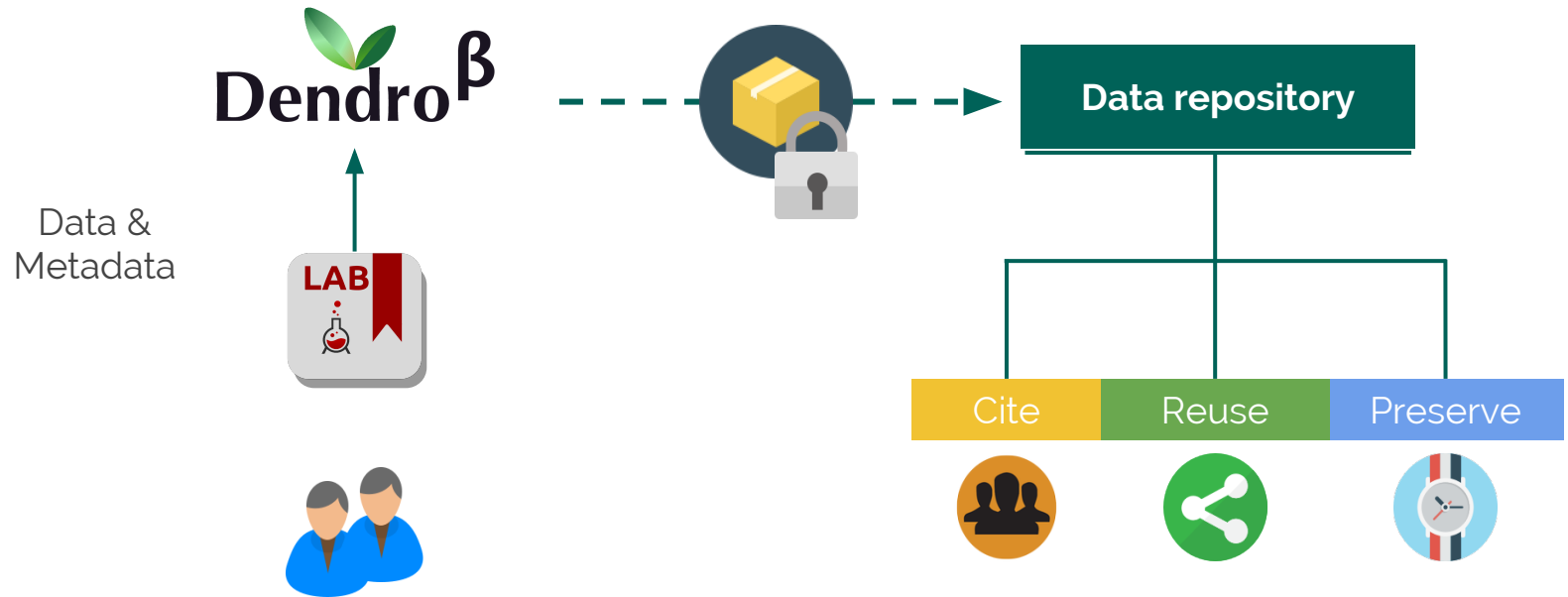
- Not OAI-PMH compliant (natively)
- Non standards-compliant metadata records

Figshare

- + Compatible with known reference manager softwares
- + Credit-based system to record citations and shares
- Very limited data description capabilities

Zenodo

- + Some support for established metadata schemas (DC, MARC, ...)
- + OAI-PMH compliant
- + Enables communities to curate submissions



Amorim, Ricardo Carvalho, et al. "LabTablet: semantic metadata collection on a multi-domain laboratory notebook." MTSR, 2014. Karlsruhe

João Rocha da Silva, Cristina Ribeiro, João Correia Lopes, Ontology-based multi-domain metadata for research data management using triple stores IDEAS 2014, Porto

João Rocha da Silva, João Aguiar Castro, Cristina Ribeiro, João Correia Lopes, Dendro: Collaborative Research Data Management Built on Linked Open Data ESWC 2014, Crete

João Rocha da Silva, João Aguiar Castro, Cristina Ribeiro, João Correia Lopes, The Dendro research data management platform: Applying ontologies to long-term preservation in a collaborative environment iPres 2014, Melbourne