

Sustaining and modernizing open source repository platforms (in the context of Brazil)



Washington Segundo

Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict)

COAR Annual Meeting 2023

17th of May 2023

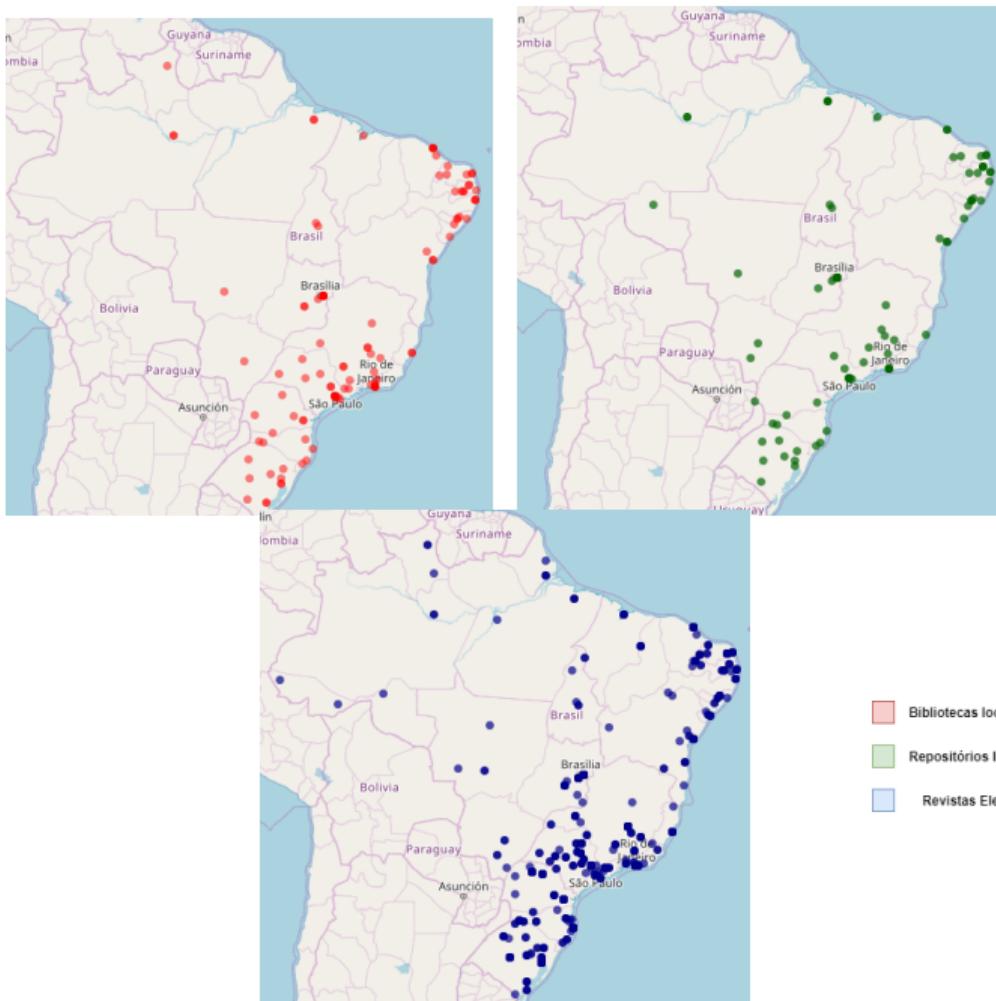


A Hive



Brazilian Portal of Publications and Scientific Data in Open Access (approx. 4,335,000 items) <<https://oasisbr.ibict.br>>





Updating versions 4–6 to 7



Requirements



Linux



- The update procedure sets all the environment automatically, and it can be performed/repeated without affecting previous executions and/or the currently operational DSpace
- The updater is non-invasive to the DSpace in production
- It works with copies of the database (dump) and the directories "config," "assetstore," and "solr"
- It only exposes HTTP ports
- For each installation, it creates random passwords (MD5) for the database

- A ".properties" file provided by the updater, specifying the directory of the "dspace.dir" copy, as well as the addresses for the backend and frontend deployment
- Migrates all the data registered in the old DSpace to the new one
- Updates the submission forms to the new standard
- Migrates the Solr statistics

Source code (BETA version)

<https://github.com/projetos-codic-ibict/migracao-dspace>

Research data repository network



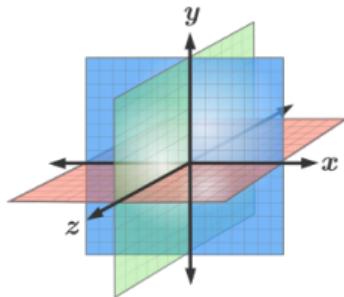
Research data repository network



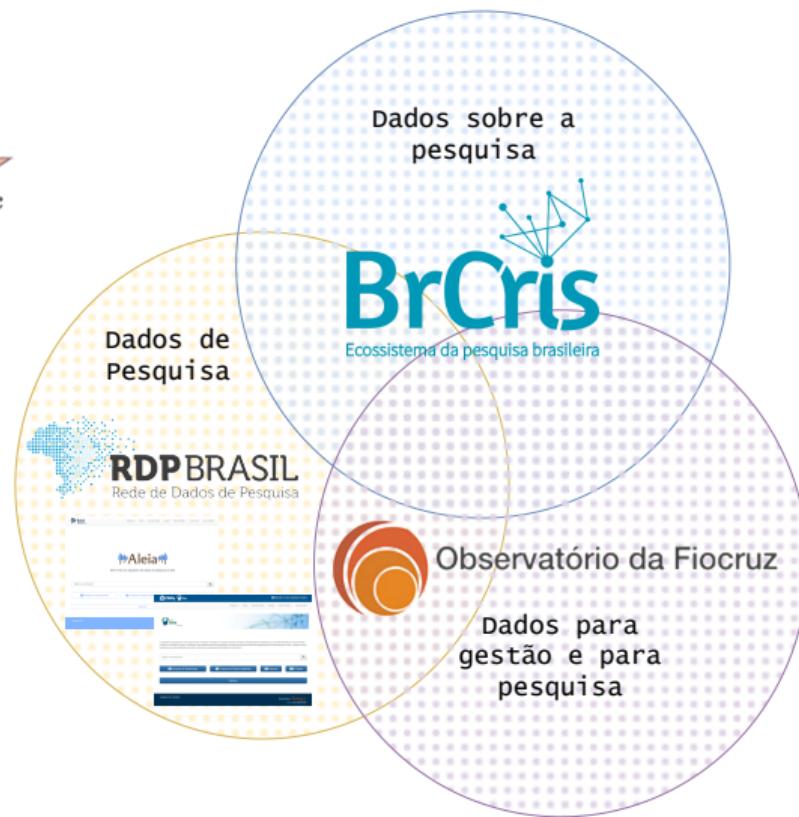
LA Referencia
Red de repositorios de acceso abierto a la ciencia

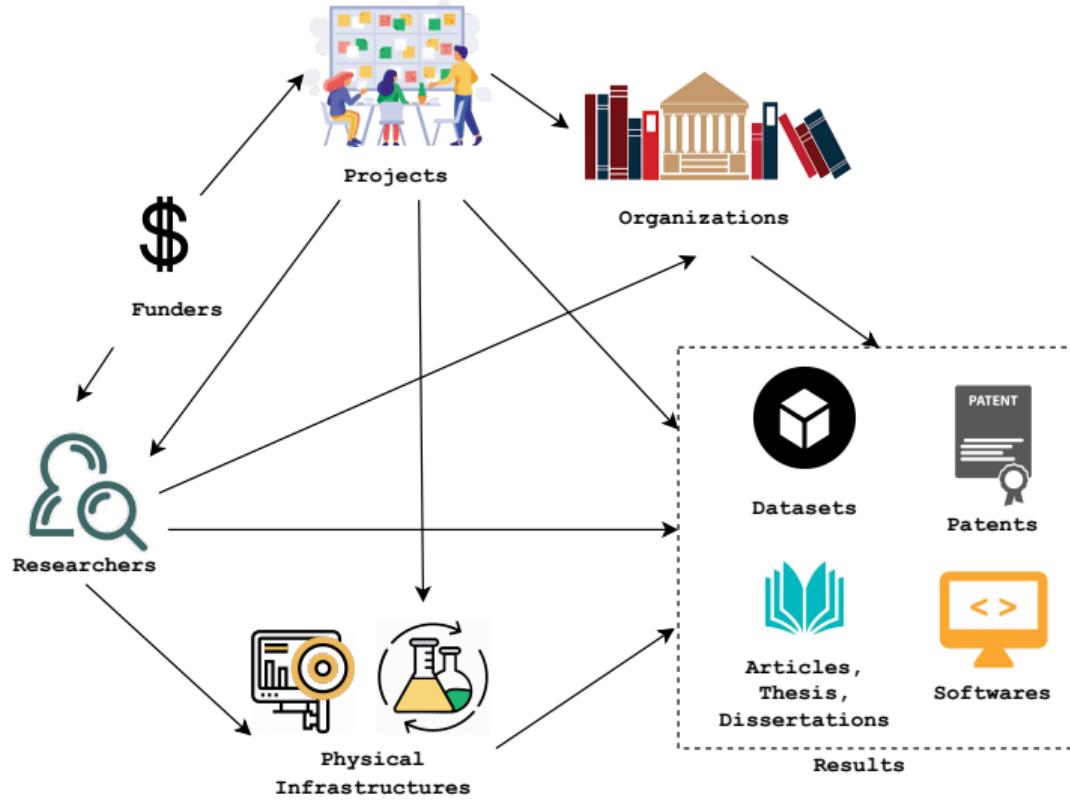


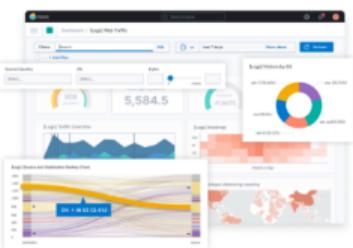
3Data



Data
from,
about and
for the
research



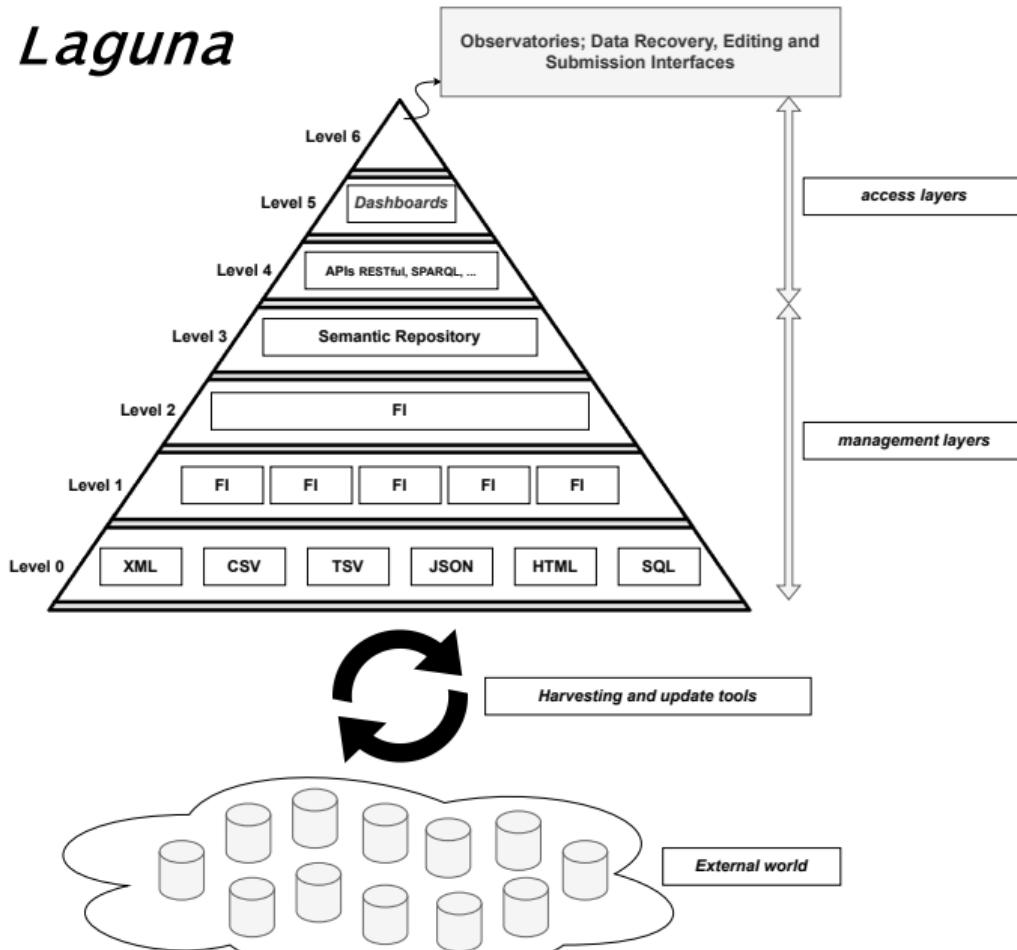




ISSN



Laguna





ISSN → 1975 ORCID → 2018 DataCite → 2020 ROR → 2021

CoNCiênciA



Consortium lead

Organization 1
(Also Consortium Lead role)

Repository



Repository



Organization 2



Reposito



Organization 3

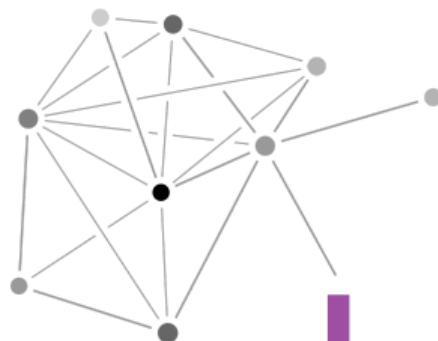


Repository



FOICRUZ





dARK



ISSN

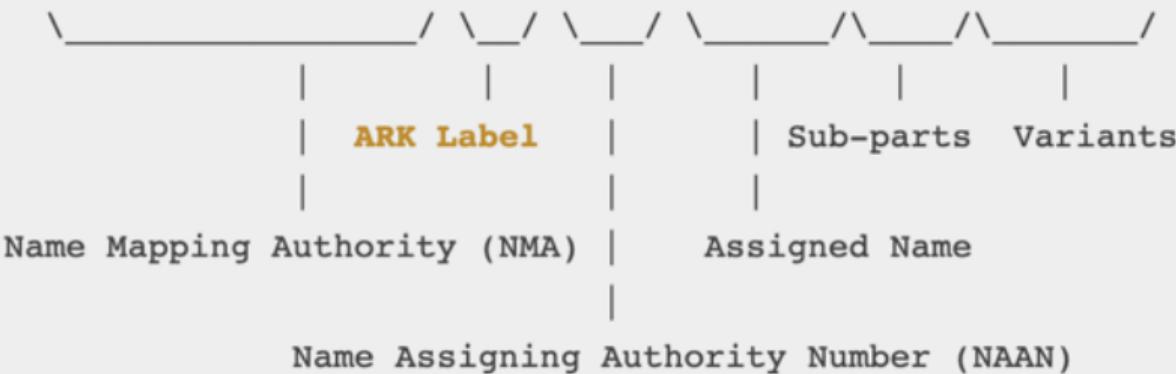
Goals

- **Avoid** disconnected multiple PID assignments.
Example: **a journal article (deposited in two different repositories) with two different DOIs that are not known**
- **Facilitate** PID assignments to multiple types of digital objects (**articles, thesis, data sets, DMP, instruments, ...**) in many diffusion channels (**repositories, journals, preprint servers, ...**)
- **Generate** services open to **publication, diffusion and evaluation systems**
- **Preserve** metadata through **blockchain** mechanisms

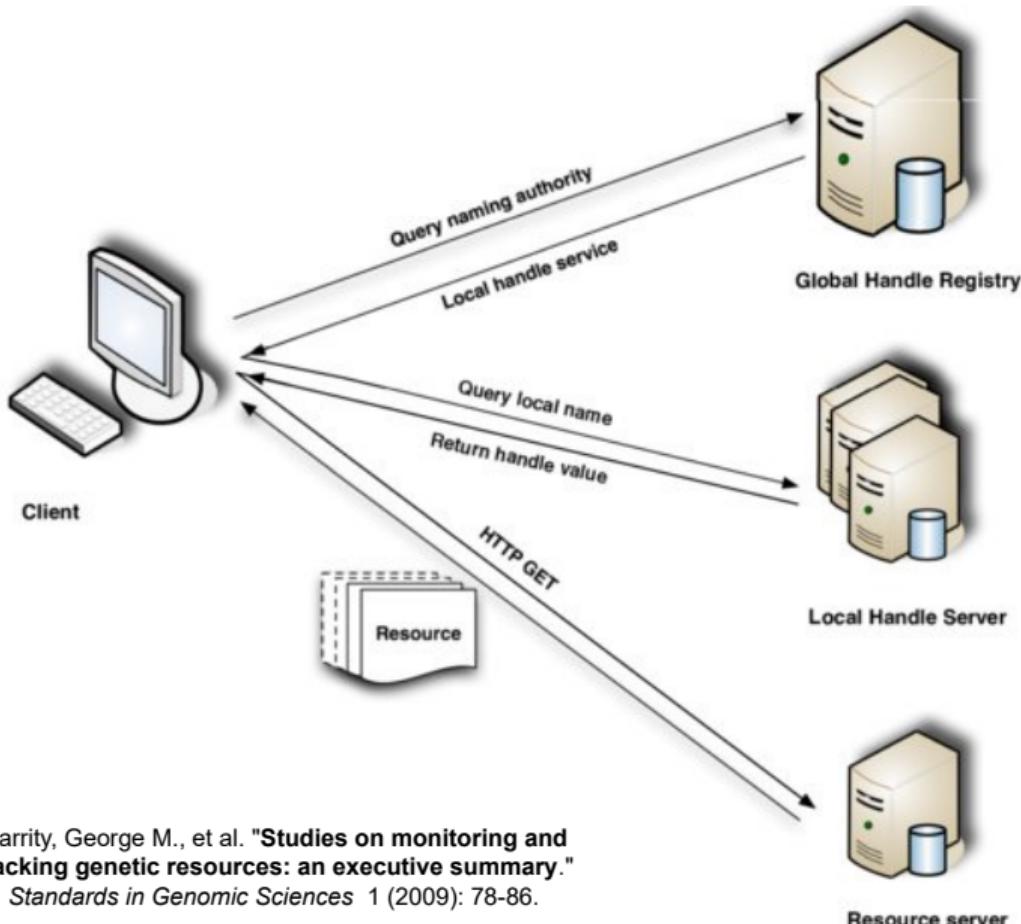


ARK Alliance

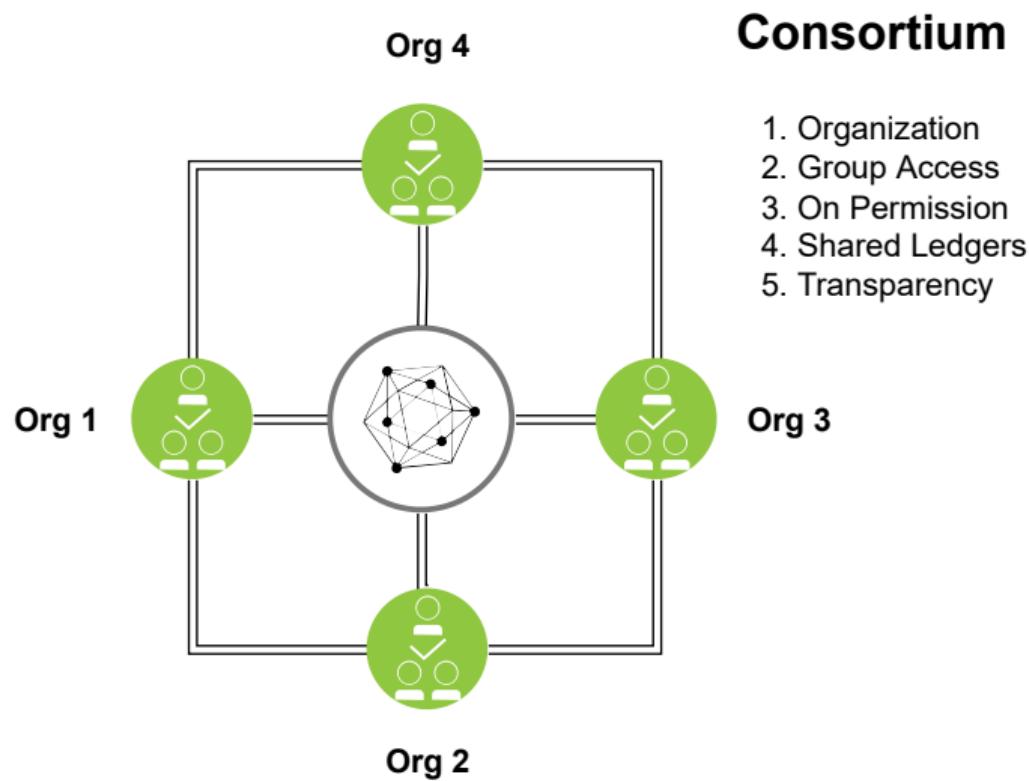
<https://example.org/ark:/12345/x54xz321/s3/f8.05v.tiff>

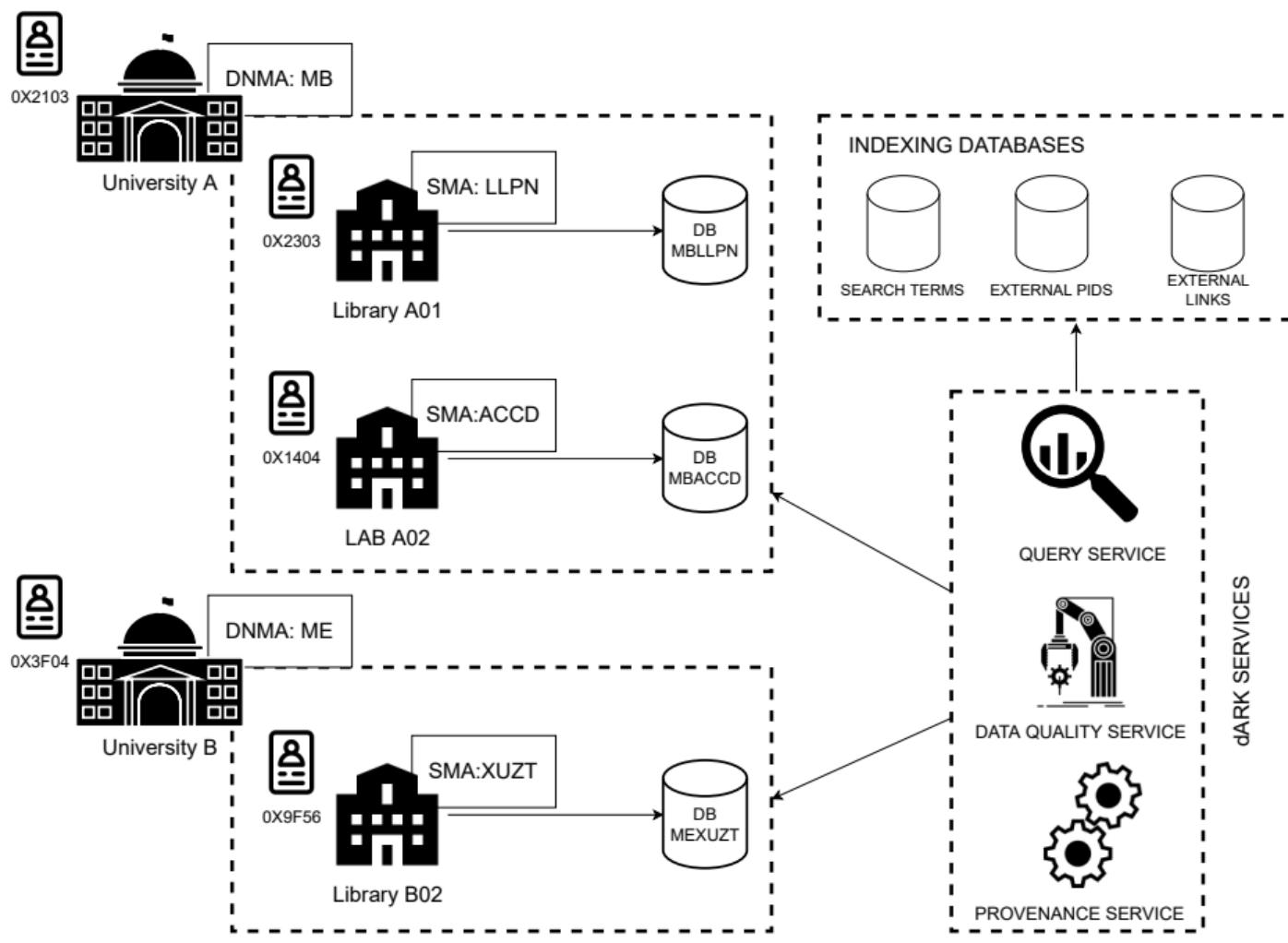


Huynh, Kyle, et al. "Simplifying ARK ID management for persistent access to digital objects." *Code4Lib Journal* 54 (2022).

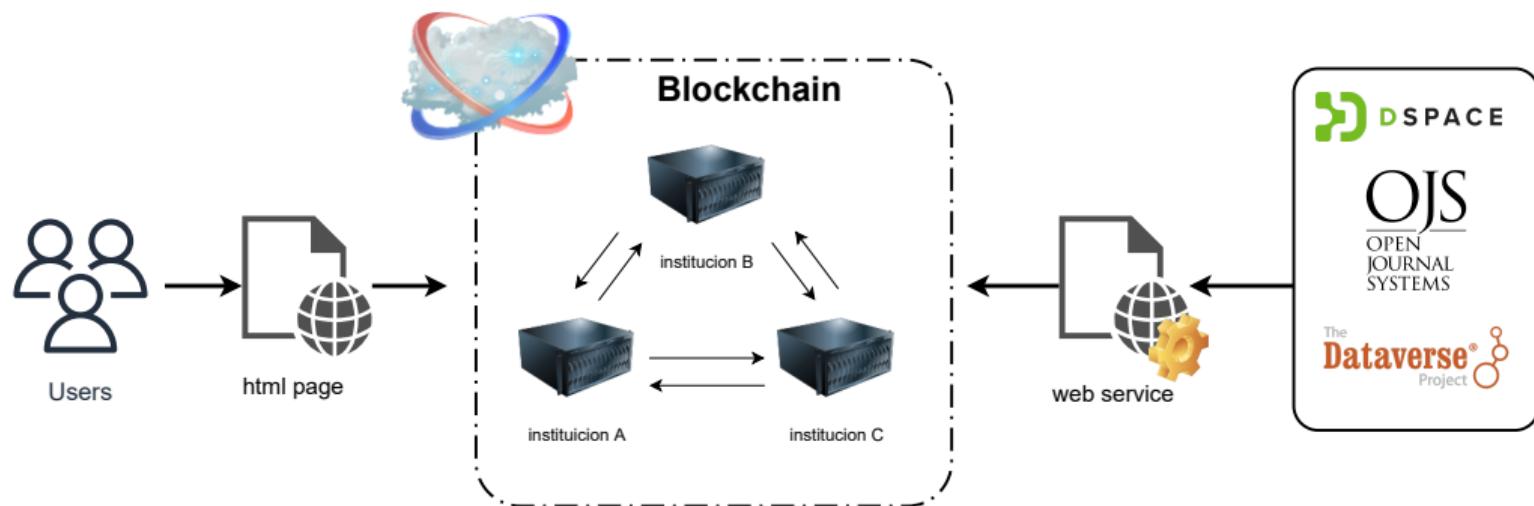


The blockchain consortium model





The dARK PoC



Response example

```
{  
  "noid": "8003/fkwff300001v",  
  "external_pids": [  
    {  
      "id": "0x6cd32058785840306ddaa126d57999315739722ce57131dde41aa3690a8afdd",  
      "schema": "DOI",  
      "value": "10.1016/J.IS.2021.101826",  
      "owner": "0xf17f52151EbEF6C7334FAD080c5704D77216b732"  
    }  
  ],  
  "payload": "{ 'title': 'Blockchain-based Privacy-Preserving Record Linkage: enhancing data privacy in an un  
' Thiago Nóbrega', 'type': 'Article'}",  
  "external_links": [  
    "https://doi.org/10.1016/j.is.2021.101826"  
  ],  
  "owner": "0xf17f52151EbEF6C7334FAD080c5704D77216b732"  
}
```

Conclusion

- Cooperation (in Brazil, Latin America and outside) between **institutions** and formation of **collaborative networks**
- Development of **models**, **standards** and **aggregation tools / treatment / retrieval / data visualization** and **information**
- Solutions, with **distributed management**, that are **simple** to use and have **low computational cost**

Acknowledgements

	Oasisbr	DSpace 7 updating	Dataverse Network	BrCris	Laguna	dARK
Lautaro Matas	X				X	X
Márcio Gurgel		X				X
Renê Júnior			X		X	
Thiago Nóbrega						X
Edilson Júnior						X
Jesús M.Chalco				X	X	X
Thiago Dias				X	X	
Vivian Silva	X			X		
Jesieli Viana	X			X		
Adilson Pinto				X	X	
.	.	.				
.	.	.				
.	.	.				

Thank You



washingtonsegundo@ibict.br