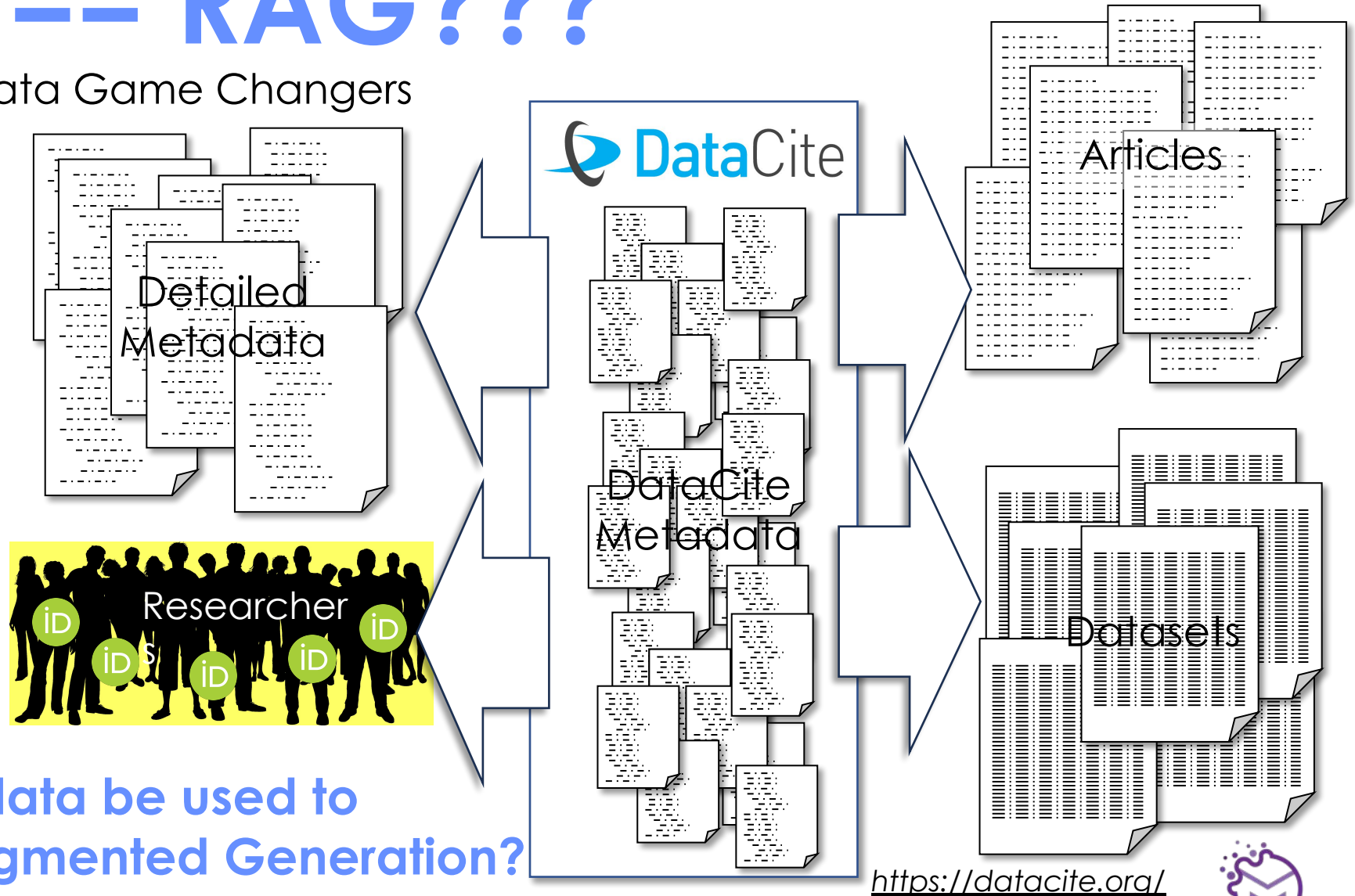


DataCite == RAG???

Ted Habermann, Metadata Game Changers

A DOI provider and identifier metadata repository with ~1700+ members.

DataCite includes over 110,000,000 resources with 28 different types and 36 different connection types.



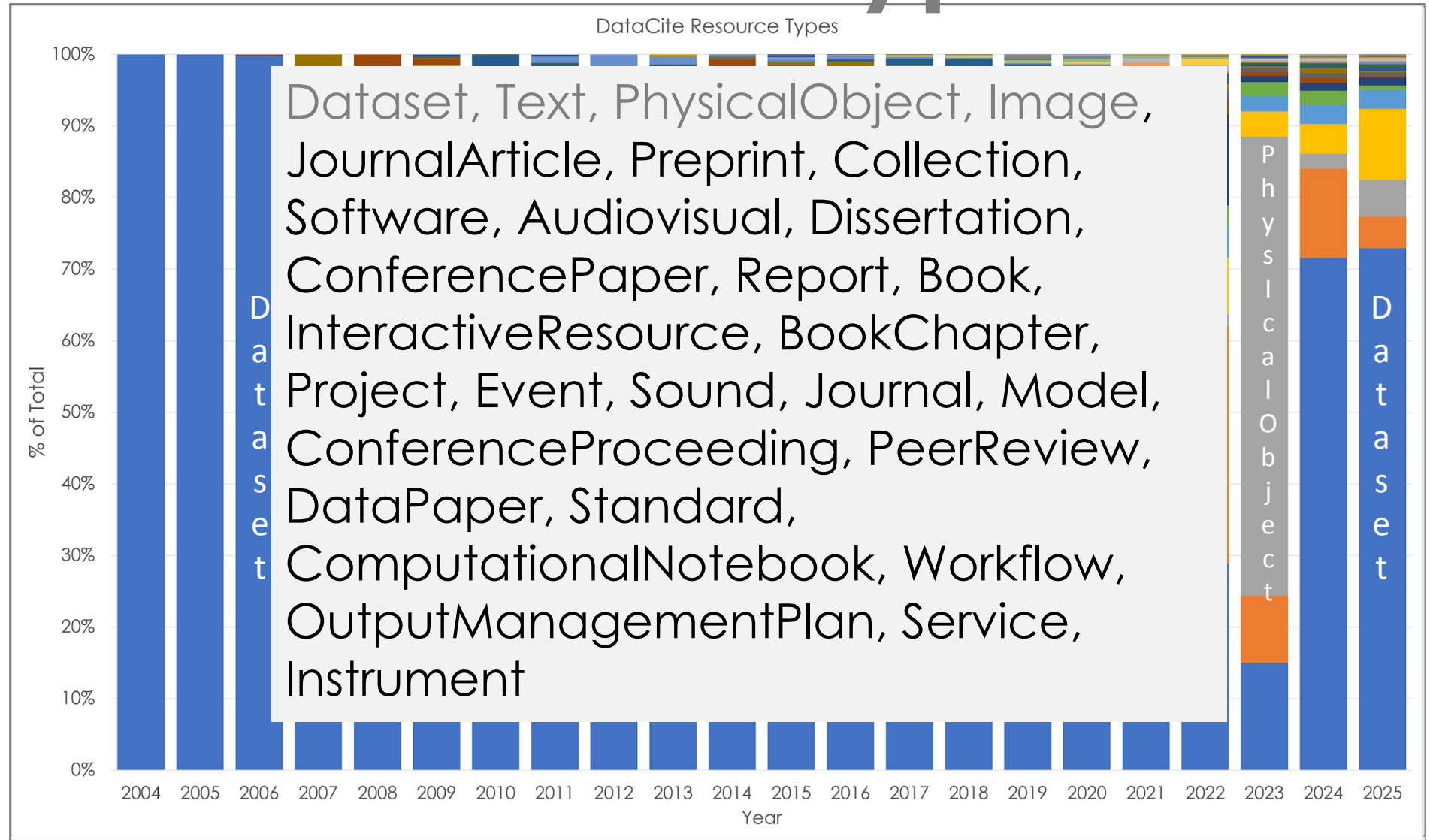
Can DataCite metadata be used to support Retrieval Augmented Generation?

Records and Resource Types

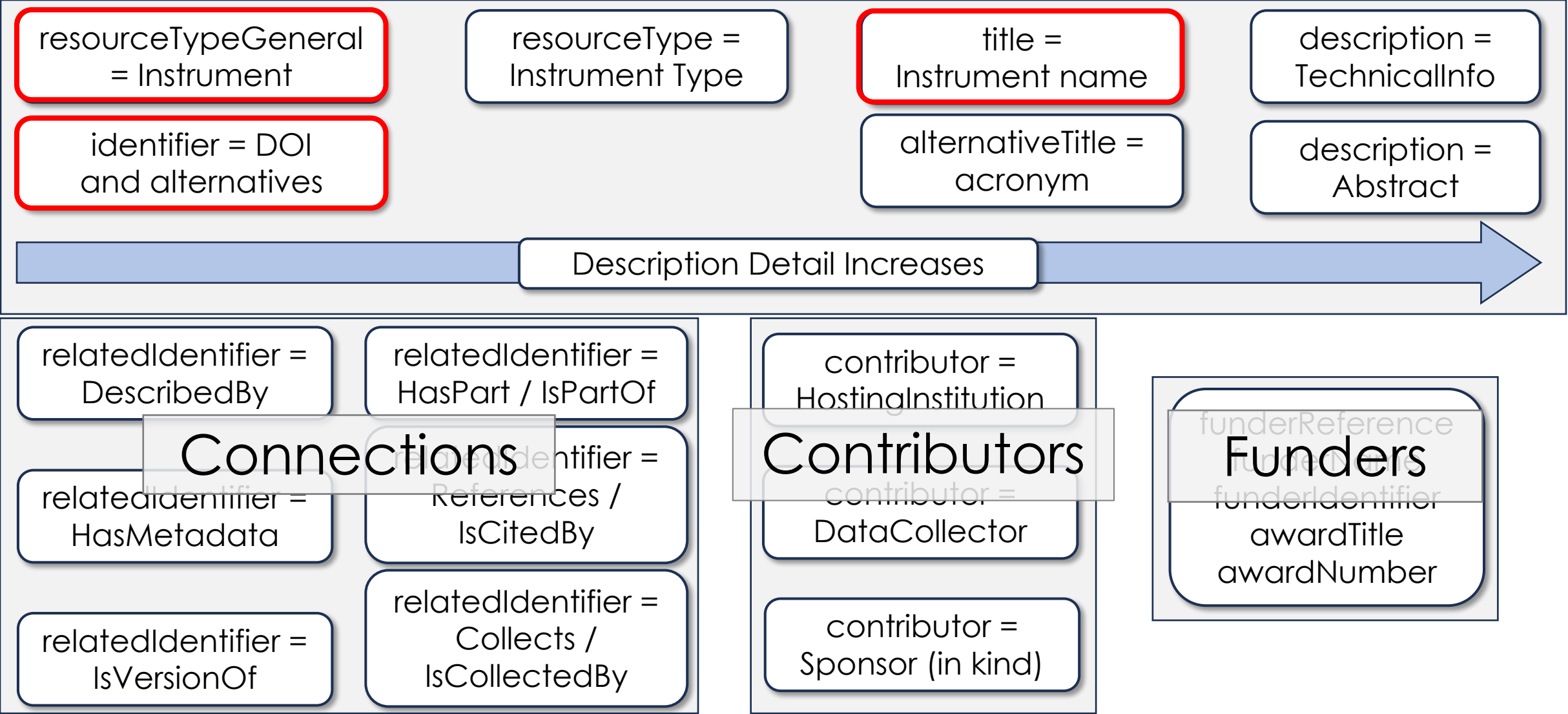
DataCite has grown steadily since 2004, passing 100 million records last year.

How Many When (2025 Update): Dataset Bright Spot in the Driver's Seat

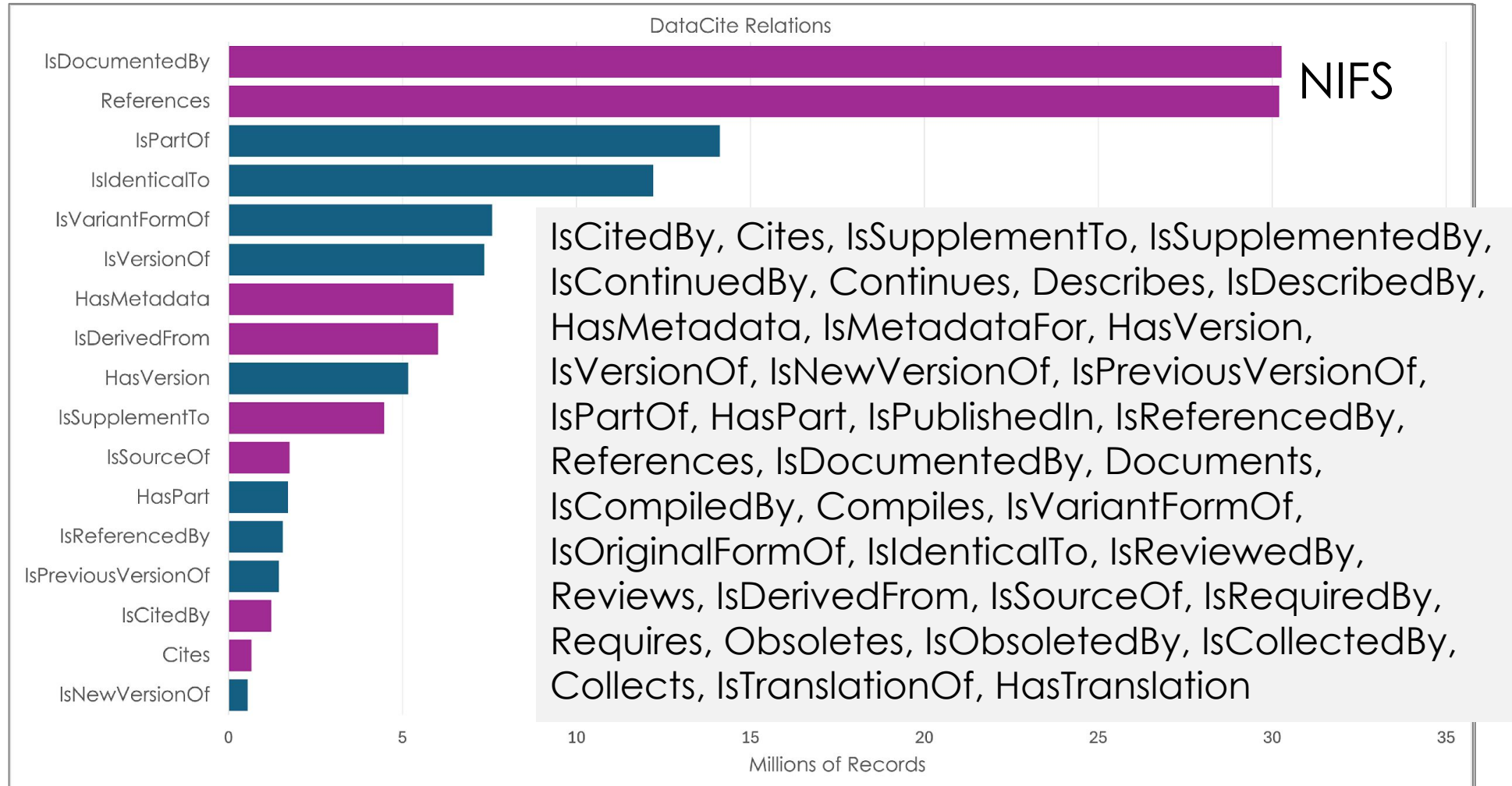
Scaling for Impact: How NIFS and DataCite Are Advancing Global Discoverability of Japanese Research



Metadata Detail @ DataCite



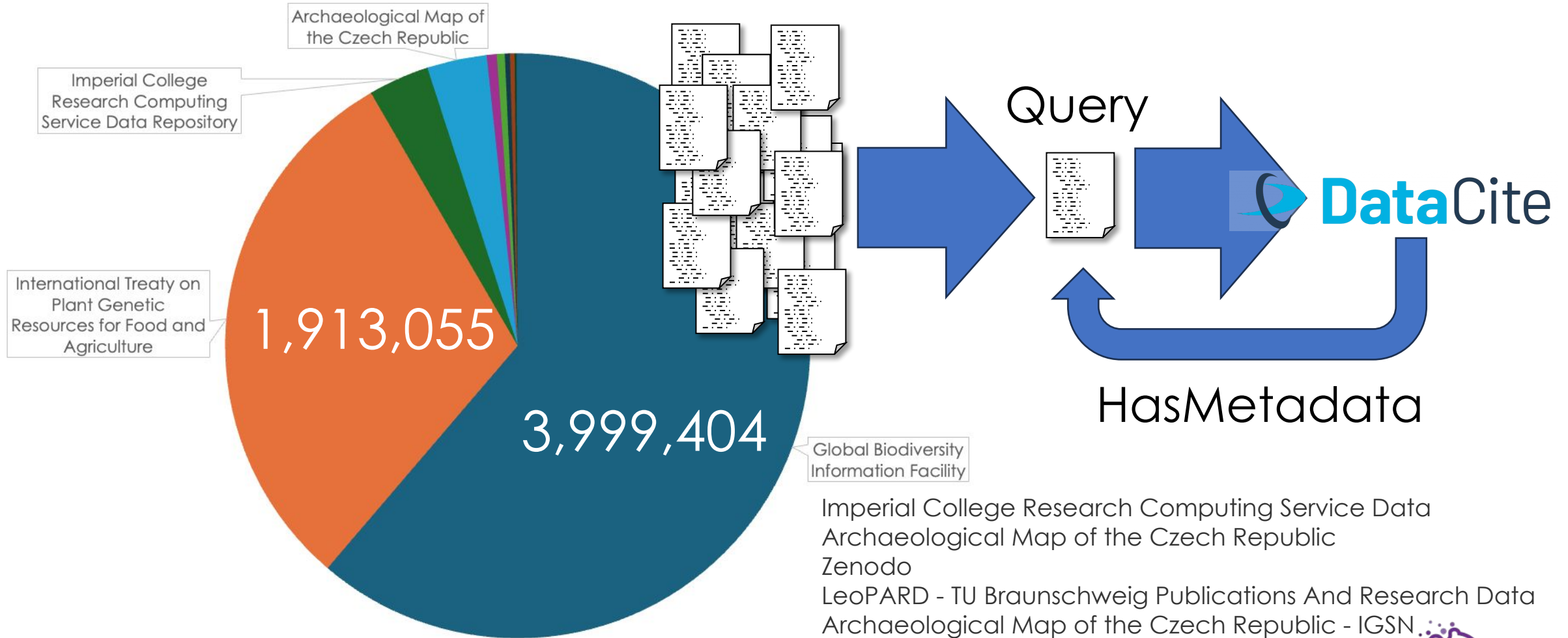
Connection Types



DataCite supports 38 connection types. These occur over 500,000 Times.

These provide more metadata or documentation.

HasMetadata = 6,460,569



International Treaty on Plant Genetic Resources for Food and Agriculture fao.itpgrfa

"schemeUri": "http://www.fao.org/plant-treaty/areas-of-work/global-information-system/descriptors",
"schemeType": "XML",

"relationType": "HasMetadata",

"relatedIdentifier": "https://glis.fao.org/glisapi/v1/pgrfas?doi=10.18730/1N4C34",

"relatedIdentifier": "http://rs.tdwg.org/dwc/terms/guides/text/index.htm",

"relatedIdentifier": "schemeType": "DwC-A",

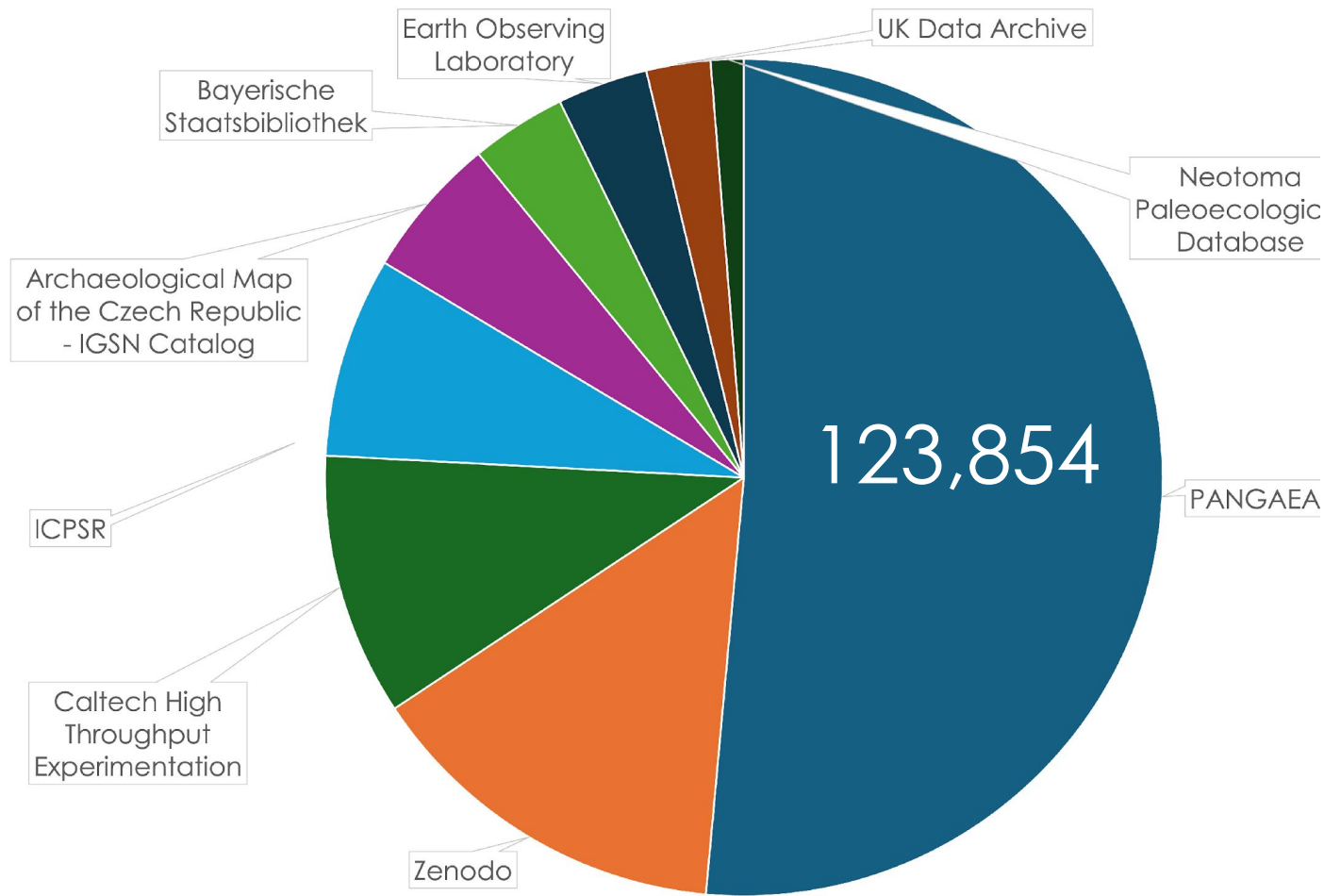
relationType : HasMetadata ,

"relatedIdentifier": "https://glis.fao.org/glisapi/v1/pgrfas?_format=dwc&doi=10.18730/1N4C34",

"relatedIdentifierType": "URL",

"relatedMetadataScheme": "Darwin Core Archive"

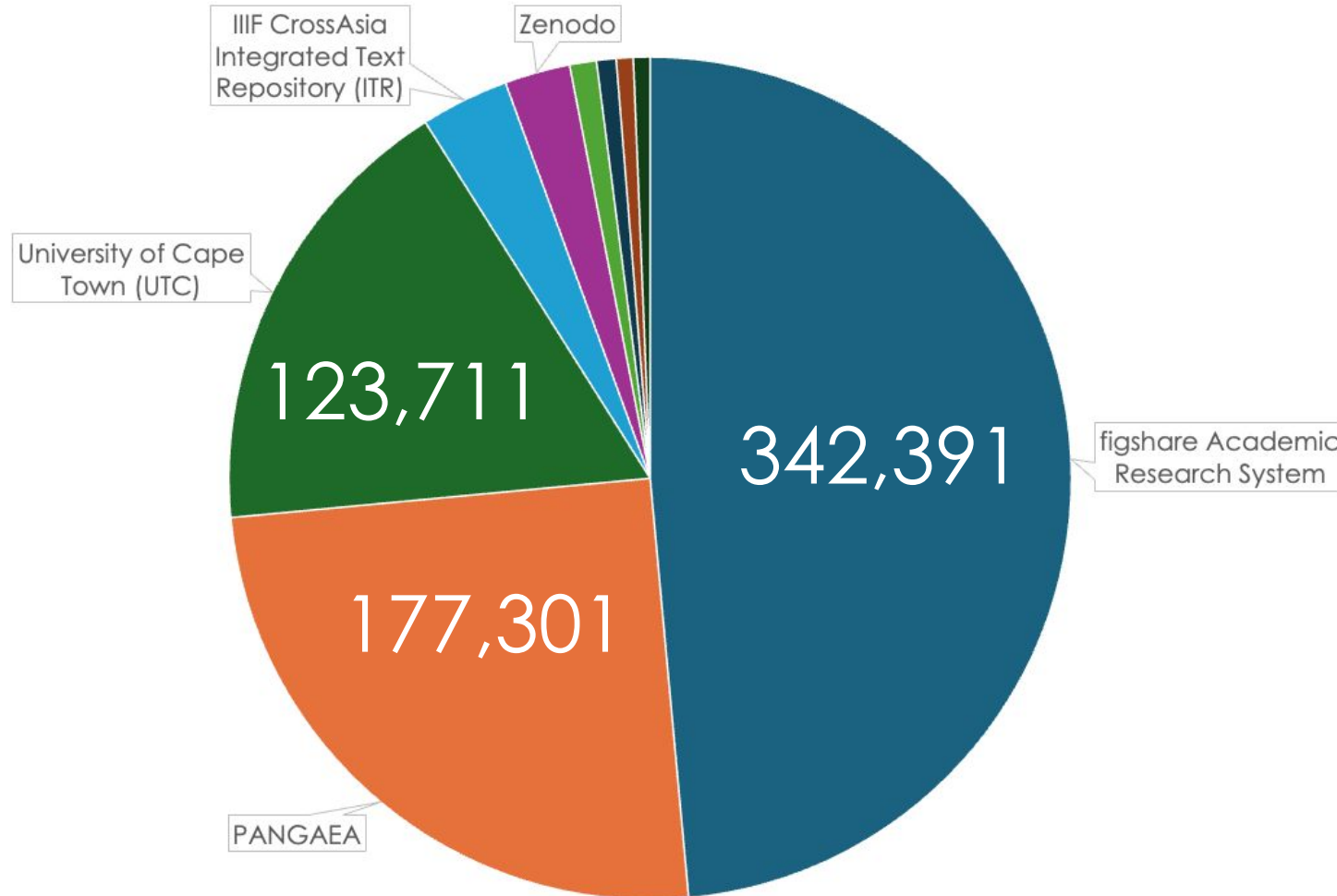
IsDocumentedBy = 30,262,853



IsDocumentedBy
– connecting
metadata to
documentation

NIFS – 30,000,000

References = 30,199,213



References –
connecting
dataset
metadata to
articles.

NIFS – 29,422,133

New Identifiers in Open Metadata



New Identifiers in Metadata

Occurrences

Publisher (RORs)

117,525

Affiliated Organizations
(RORs)

116,512

Authors (ORCIDs and RORs)

116,172

Rights Holders (RORs)

122,269

Total

472,483



Nearly 500,000 ORCIDs and RORs were added throughout the National Geophysical Facility DataCite metadata records during 2025.

ted@metadatagamechangers.com
erin@metadatagamechangers.com

<https://orcid.org/0000-0003-3585-6733>
<https://orcid.org/0000-0001-9998-0114>





Questions?

Recent Work:

- [Connecting PIDs in DataCite Metadata](#)
- [University and College Connectivity @ DataCite](#)
- [BrightSpots Get Brighter](#)
- [The Global Research Infrastructure and the NSF Public Access Repository](#)
- [FAIR DataCite Metadata - University and College Bright Spots](#)
- [Funder Acronyms Are Still Not Enough](#)

Work with us on:

- Repository Re-Curation
- Repository and Journal Connectivity
- FAIR Metadata evaluation and improvement
- Community building strategy
- International Metadata Standards (ISO, DataCite, schema.org)
- Workshop design and facilitation
- Community conventions / profiles
- Leadership coaching



Your
Repository
Insights!

ted@metadatagamechangers.com

erin@metadatagamechangers.com