The Copyright Hub and DOI

‘What is the Copyright Hub and how does it fit with the DOI ecosystem’

Mark Hewis | Technical Architect
@MarkHewis

21st Nov 2014
Our Goal.....

https://vimeo.com/111096150

http://copyrighthub.co.uk/video
Our History.....
Our Architecture…

- Data Source
- Source Repository
- Onboarding
- Hub Model Repository
- Hub Registry
- Conflict Reporting
- Hub Index
- Discovery
- Transforming
- Hub Application
- Mapping
- Hub Ecosystem

Connected Digital Economy
Key Terms.....

- **HubKey** – index key within Hub Ecosystem maps to specific Entity in a Repository
- **Source ID** – any identifier used by a Source which may also be “public” (e.g. DOI, ISBN)
- **Service Resolution URL** – the API Endpoint at which a Service is found
- **Index Service** – routes requests using these IDs to Services – “hub of the Hub”
- **Registry Service** – maintains mappings and repository configuration used by above
How to Join:

Now comes the process which is the key to the way the Hub is “bound together” - the **Hub Index**.
the source Repository automatically updates the Index with the IDs of all the Creations for which it has information which it wants to make accessible.

How to Join:

Hub Ecosystem

Hub Model Repository

Hub Index

Load data
Because the aggregator may offer different kinds of Services, the Index is also told what those services are, and a Web Address (URL) at which the service can be found...
...so the Index stores “what” and “where”. For example:

ID = ISBN 9781234567890
Service = License Offer
Service 1
At = www.abcd.co.uk
Hub Model
Repository

Load
data

Hub Index

If a Repository has more than one ID for the same Creation (which is common) these are also stored at the Index. This allows the Index to build up a network of “Same As” links, and to detect errors in these.
How to Join:

The Services have already been registered with the Hub Registry, so the Index is ready to take queries.
Joining workflow in Summary

- Repository registers with Registry Service – includes information below
  - ID Types they support
  - Method to update Index Service (push, pull, live)
  - URLs to Repository API endpoints and types
  - Authentication and Security Information
- Index service
  - Updated by ‘push’ from repository or ‘pull’ based on schedule in registry
  - Matches source IDs between different repositories
  - Alerts on conflicts in particular incorrect ‘same_as’ assertions
  - Optimised for multiple ID lookups
  - Can Query repositories in real time if desired (live)
Sample Record in Index service

source_id:doi:10.1038/issn.1476-4687

same_as:issn.1476-4687

hub_key:chub:repo1:doi:10.1038/issn.1476-4687

hub_key:chub:repo2:issn.1476-4687

service:api:repo1.org/rights_api

service:api:repo2.org/rights_api

service:website:repo1.org
Querying workflow in Summary

- User or service queries index service with a ‘source ID’ or ‘Hub Key’
- Index returns with following depending on type of query
  - Web page to present rights information based on ID if via a web browser
  - API response containing related IDs, Hub_Keys and service URL endpoints which can provide rights information or services within the eco-system
What we are investigating

• Underlying technology behind DOI (Handle) – can it be used for Index Service?
• Resolution ‘types’ model supported by Handle/DOI
• More complex assertions of ID relationships beyond ‘sameAs’

Find out from yourselves - DOI Operators and Registries - about how it works in practice
Summary

What the Hub is:

• A core set of Open Source Services
• An eco-system of Companies and Services connected to above – the Hub(s)
• Route Rights/License Information between services in eco-system
• Mapping of Rights/License between different formats using the LCC model

Current Status: Completed Alpha phase working on Beta now using Agile/Continous Pipeline methodology
Any questions?
More Information

• The Digital Catapult: www.cde.catapult.org.uk
• The Copyright Hub: http://www.copyrighthub.co.uk/
• Mark Hewis: mark.hewis@cde.catapult.org.uk @Mhewis

Contact myself to join the Copyright Hub Technical Working Group
• Further information from Dominic Young Dominic.Young@copyrighthub.co.uk