

# DOI

**Paul Walk**

**p.walk@ukoln.ac.uk**

**UKOLN is supported by:**



**JISC**

Supported by



Museums, Libraries and  
Archives Council

**[www.ukoln.ac.uk](http://www.ukoln.ac.uk)**

**A centre of expertise in digital information management**



DOL is a **business**  
proposition

# overview (I)

- *Digital Object Identifier*
- “a managed system for persistent identification of content-related entities on digital networks”
  - **not** people, places, organisations....
- based on the *Handles* system & informed by the *indecs* content model
- governed by the *International DOI Foundation*
  - a non-profit organisation created in 1998
  - I sit on the board as JISC’s representative
- implemented through a federation of *Registration Agencies (RA)*

# overview (2)

- *Handles* = a system for minting, managing & resolving persistent identifiers
- *indecs* = interoperability of data in e-commerce systems
  - (it's trendily lowercase, so you realise immediately that it won't be at all boring)
- RAs: supporting DOI 'registrants' - allocating prefixes, registering DOI names and supporting the management of associated metadata

# overview (3)

- DOI is a syntax
- DOI = a name (identifier)
- DOI is a system

# syntax (I)

**prefix**

(assigned by an RA  
to a Registrant)

**suffix**

(assigned by registrant  
& unique within their prefix)



# syntax (2)

- cited as:
  - `doi:10.314/56789`
- (hyper)linked as:
  - `http://dx.doi.org/10.314/56789`
- structured suffixes:
  - `/whole_work.next_level.next_level`
- suffixes embedding other schemes, e.g. ISBN:
  - `10.1002/0471758132`

# URN?

- DOI is **not** registered as a URN namespace
- fulfils all the functional requirements
- however, URN registration appears to offer no advantage to the DOI System....
  - ~~urn:doi:10.314/56789~~
- if at some point in the future there exists a widely deployed Resolution Discovery Service (RDS) then DOI *could* be registered with this

# metadata & services

- DOI names belong to one or more *application profiles* which define extended metadata schemes to arbitrary levels of complexity (can be nested)
- most DOIs currently associated with single URL only, but must have kernel metadata to support the management of the identifier
- application profiles, in turn are associated with one or more services
- the minimum, default service is the return of kernel metadata
- there exist (Handle) APIs to access and manipulate all of this

# example metadata

## DOI Record

**DOI:** [10.1045/june2001.iannella](https://doi.org/10.1045/june2001.iannella)

**URL:** <http://www.dlib.org/dlib/june01/iannella/06iannella.html>

**Registration Agency:** 10.ra/1

**Profile DOI:** [10.profile/0](https://doi.org/10.profile/0)

- **Encoding:** ref.url
- **Data:** <http://www.dlib.org/dlib/june01/iannella/06iannella.html>+

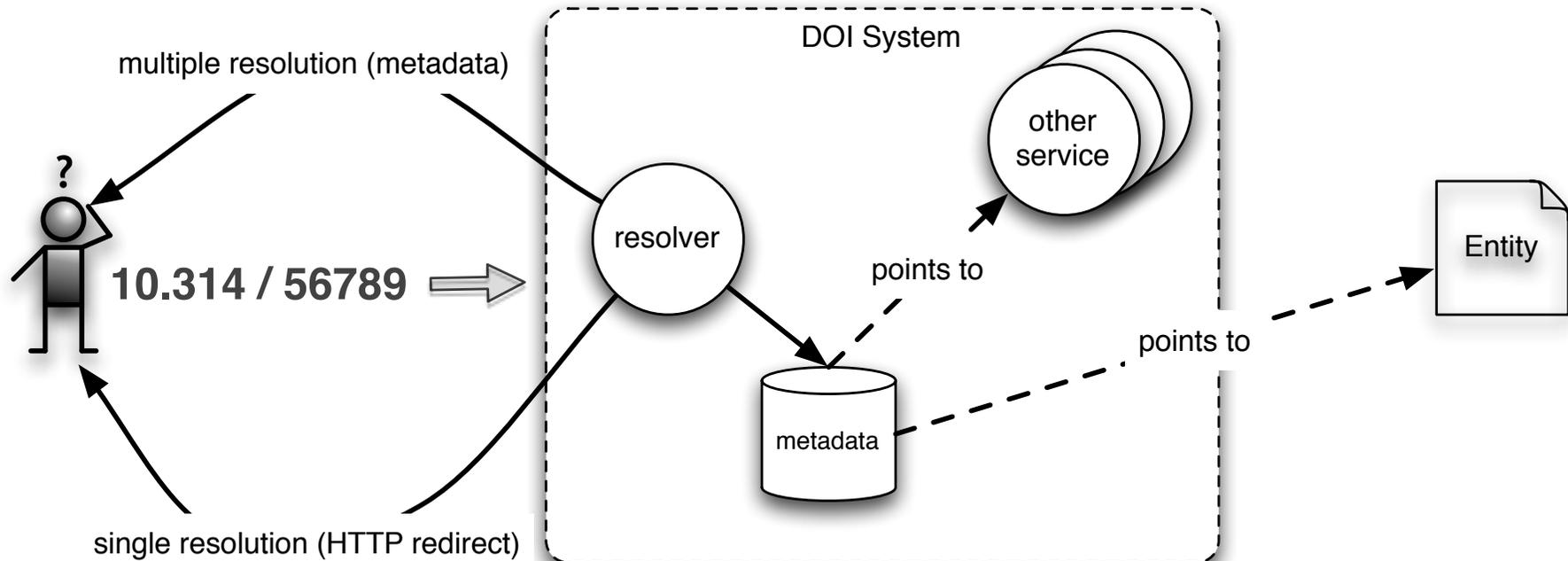
**Profile DOI:** [10.profile/1](https://doi.org/10.profile/1)

- **Encoding:** val.xml.xsd.doikernelmetadata
- **Data:** <10\_profile\_1><http://www.dlib.org/dlib/june01/iannella/06iannella.html>10.ra/1  
10.1045/june2001.iannella ISSN 1082-9873 DigitalResource Abstract Journal Article Digital  
Rights Management (DRM) Architectures Main Title Person Renato Iannella  
renato@iprsystems.com email 10.ra/1 Author +

**Profile DOI:** [10.profile/dlib0](https://doi.org/10.profile/dlib0)

**Profile DOI:** [10.profile/dlib1](https://doi.org/10.profile/dlib1)

# system



- indirection mitigates 'brittleness'
- multiple resolution allows rich metadata + wider service provision

# business model

- generally operated on a not-for-profit, cost recovery model
- federated through RAs
- RAs charge a fee for maintenance of a registrant's code, and a fee for each name registered
- RAs pay a fee to IDF, partly based on usage

# current coverage

- scholarly papers (articles, conf proceedings etc.)
  - *CrossRef, Wanfang Data (in China)*
- research datasets
  - *DataCite, Wanfang Data (in China)*
- multi-lingual, European content
  - *mEDRA*
- EU publications
  - *OPOCE (Office des publications EU)*
- books & serials
  - Bowker

# requirements?

- What are the requirements? In general, for UKHE in particular
  - persistent access, persistent citability, confidence in curation
- What are the problems and opportunities that face the sector?
  - funding, merging/loss of HEIs, rising prominence of data as a cited resource
- What needs do teachers, learners and researchers have that PIDs can help with?
  - resource discovery, resource relationship traversal, referencing, auxiliary metadata

# value proposition

- technology enabling an organisational/social/business commitment to persistence
  - “Cool URIs don’t change”
  - Not all URIs are ‘cool’ - ‘link rot’ is a reality - what do we do about the ‘unhip’ identifiers?
- richer description for management, discovery and re-discovery
  - metadata closely associated with identifier
  - services loosely coupled with identifier
  - resource relationships expressed in curated metadata
- a level of confidence in PI curation (confidence in federation of IDF and RAs)

# conclusions

- despite reservations about divergence from Web Architecture, DOI cannot and should not be ignored
- DOI satisfies concerns of particular domains - discussions about PIs need to be situated in business contexts
- While the Web continues to grow rapidly, DOI is growing also: we should accept that there will be a 'mixed economy' for some time to come

the persistent identifier  
system which refused to  
DOI....

(said in a Brummie accent!)