

## Key facts on *Digital Object Identifier System*



### Key concepts and abbreviations

- DOI = Digital Object Identifier
- IDF = International DOI Foundation (operating and governing organisation): [www.doi.org](http://www.doi.org)
- RAs = DOI Registration Agencies (= members of IDF offering the system to customers who wish to assign DOI names)

### Status: operational system

- Foundation launched to develop system in 1998. First applications launched 2000
- Currently used by c. 4,000 naming authorities (assigners) e.g., 3,300 publishers, EU documents, science data sets, etc.
- Over 50 Million DOI Names assigned to date
- Over 210,000 DOI name prefixes (naming authorities within the DOI system)
- Around 100 million DOI resolutions per month
- DOI names have been assigned by 12 RAs (past and current)
- Well established in professional information sector; best known applications are CrossRef ([www.crossref.org](http://www.crossref.org)) and DataCite ([www.datacite.org](http://www.datacite.org))
- Initial applications are simple redirection
- More sophisticated functionality available e.g. multiple resolution, data typing, "Application Profiles"
- International Standard (ISO: in publication)

### Scope

- *Digital Identifier of an Object* (not "Identifier of a Digital Object")
- Object = any entity (thing: physical, digital, or abstract)
  - Resources, parties, licences, etc.
- Digital Identifier = network actionable identifier ("click on it and do something")
- Initial focus on entities was documents/media e.g., articles, data sets
  - Now also moving into parties and licences
  - Extending to other sectors, e.g., Movie industry? Financial sector? Music? Newspapers?
- Extensible by design (as e.g., URI): not intended as a publishing-only solution (digital convergence)
- International – e.g., in 2007 appointed China RA

### What it does

- provides an *actionable, interoperable, persistent* link:
- *actionable* - through use of identifier syntax and network resolution mechanism (Handle System<sup>®</sup>)
- *persistent* - through combination of supporting improved handle infrastructure (registry database, proxy support, etc) and social infrastructure (obligations by Registration Agencies)
- *interoperable* - through use of a semantically interoperable data model and grouping mechanisms.

### Governance

- IDF = operating and governing organisation
- Provides the social infrastructure
  - e.g., obligations for persistence, back-up, in event of failure, etc
- Proven model: successfully transitioned the management of persistent identifiers between different registrants and between different RAs
- US "Not for profit" open membership (with membership fee)
- Federation of Registration Agencies makes up significant part of the IDF (possibly 100% eventually)
- Elected Board, working groups (including RA Working Group)
- No full time staff (contracted managing agent, outsourced functions)

### Business model

- IDF receives membership fees from RAs, contracts technical operator
- RAs are members of IDF and pay a fixed fee per year
- Costs of operating the system are divided across RAs so that IDF is cost-neutral
- Assigners are customers of RAs
- RAs might have their own existing numbering scheme, existing communities etc. – which can be integrated with a DOI Application, not replaced by it (e.g. ISBN)
- RAs are *autonomous independent bodies*. They offer services to assigners using DOI names
  - RAs' business model with their customers is entirely autonomous
  - RAs only obligation to IDF is a licence/operating agreement

- o RAs may choose to put DOI names "under the hood"
- Inspired by bar code model, ISBN etc.: assigner pays
- Some RAs are commercial; others are themselves member communities (e.g., CrossRef)

#### Technical infrastructure

- *Handle System*: persistent identification in digital networks (devised by TCP/IP co-inventor)
- *Indecs*: principles of contextual ontology data model for associated metadata("interoperability of data in e-commerce systems"), implemented in Vocabulary Mapping Framework (<http://www.doi.org/VMF/index.html>)
- Both used elsewhere: aim was not to re-invent the wheel

#### Standardisation

- ISO 26324, Information and Documentation – Digital Object Identifier System: currently in publication.
- URI (within info-URI scheme)
- Mechanism for, and emphasis on, enabling re-use of other existing identifier schemes e.g. ISBN: see <http://www.doi.org/factsheets/DOIIdentifiers.html>

#### Documentation

- Website <http://www.doi.org/>
- Factsheets – recommended for coverage of topics in detail: <http://www.doi.org/factsheets.html>
- Summary articles: [http://www.doi.org/about\\_the\\_doi.html](http://www.doi.org/about_the_doi.html)
- DOI® Handbook at <http://www.doi.org/handbook2000/> – awaiting updating (when ISO DIS is approved we intend to reconfigure the Handbook around it)

#### Origin

- 1996 proposal from the three major international publishing trade associations to develop infrastructure for digital publishing; they brought together expertise in numbering content (the ISO standard ISBN) and expertise in digital network technology (CNRI)
- Need in the digital supply chain for an equivalent of the analogue bar code: migration from analogue to digital networked content cannot rely on URLs as identifiers (e.g. due to "linkrot": "404 not found")

#### Relation to other schemes

- Strong focus on interoperability and on working with existing and new schemes
  - o Technical, syntactic and semantic interoperability
  - o Use of DOI names by multiple RAs (through APs and Services)
- Involvement with key activities in the content sectors: ACAP, ONIX, ARK, PURL, info URI, URN, Open URL, GS1, MPEG-21, IETF, RDA, DCMI, FRBR, ITU-T Idm, ICSTI, CENDI, HSAC, indecs, CONTECS-DD (etc!)
- Adopt existing proven components
- Note potential confusion: the term "identifier" can mean several different things – not always clear like-for-like comparison (e.g. URI v ISBN...)

#### Intellectual property considerations

- IDF owns DOI®, a registered trademark for the system
- IDF does not have any patents (or patent applications) on DOI System
- IDF collectively licences appropriate technology from suppliers on behalf of members (CNRI Handle System global license, data dictionary tools, etc.)
- Patent policy in place: <http://www.doi.org/news/050914PatentPolicynews.html>
- All RAs must sign RA agreement re use of DOI System
- Optional legal "Community of Interest" agreement to enable sharing of information on relevant patent issues – findings confidential to signatories
- IDF is a participant in related semantic vocabulary work e.g. CONTECS:DD, Vocabulary Mapping Framework