



TM

The International DOI Foundation

Washington & Oxford

DOI® System Applications

Version 1.1

The DOI® System has the flexibility to deliver identification and resolution services that fulfil the requirements of any application domain. However, these don't come "in a box". Someone needs to build the specific social and technical structures to support the particular requirements of a community. The rules about what is identified, and whether two things being identified are (or are not) "the same thing", are made at a lower level: in a specific application of the DOI System. This is a role of Registration Agencies. This provides an identification system of enormous flexibility and power – while hugely increasing the importance of an explicit structured metadata layer, since without this the identifier essentially can have no meaning at all outside a specific application.

In designing a DOI System application several questions need to be addressed, including:

- *What are we identifying with this DOI® name ?*
- *What are we resolving to from this DOI name?*
- *What, if any, explicit metadata are we making available through this DOI name?*
- *How will the cost of providing the application using the DOI System be met?*

What are we identifying with this identifier?

This deceptively easy question is one of the most difficult encountered in all discussions about identifiers (but the one most commonly overlooked) and an answer is often much more difficult than it might at first appear. The problem is compounded by the closely related question of when two things are "the same thing" (the answer to which is entirely contextual).

- The DOI System has a core strength: a DOI name can be used to identify anything within its very broad "creation" scope (which may also sometimes be perceived as a weakness, since it sometimes makes it very difficult to explain!). This means it can be used widely, and enable interoperability. The rules about what is identified, and whether two things being identified are (or are not) "the same thing", are made at a lower level: in a specific application of the DOI System. Attempting to extend the use of the identifier without explicit metadata risks a breakdown, since the implicit metadata needs to be made explicit to allow such interoperability away from the original context. If that breakdown happens, the "identifier system" will appear to have failed; when what has failed is the design of the appropriate implementation.
- For example: imagine an Application A which enables a user to purchase an e book in a specific format, pdf; that book, in that format, needs to be identified (separately from say the paperback physical format or the

The DOI® System

Website: www.doi.org

DOI® and DOI.ORG® are registered trademarks and the DOI> logo is a trademark of The International DOI Foundation. Other trademarks may be the property of their respective owners.

audiobook format) for the purposes of a purchase using Application A. Further up the rights chain, the rights in that book will have probably been managed at the abstraction level (identifying the work, irrespective of all purchasable formats which are derived from that abstract work). An Application B elsewhere (which A may not be cognisant of, and vice versa) for the purposes of rights clearance could consider all formats to be the same. Application A and B will be able to interact only if each is precise about what it is identifying; if they are not, then mixing the two (in order to generate e.g. royalty payments in a third application C) may result in chaos through misunderstanding of what is being identified in each case.

- The means by which the assigner defines what entity is being identified is to specify the properties of the entity in terms of an underlying ontology. Assigning DOI System kernel metadata is a convenient starting point for doing so and ensures conformance without the need to be an expert in the underlying ontology. Ultimately, to ensure that a DOI name can be applied with any necessary degree of precision to any entity we use a model of making (from which is derived the Index Data Dictionary: see “DOI[®] System and Data Dictionaries”) which deals with all possible entities through their context: the events which result in their formation (the interaction of agent, resource, time and place), and which therefore can relate two entities through the events of derivation of one from another.

What are we resolving to from this identifier?

A DOI name can resolve to anything. At minimum it will resolve to a URL, but there may be multiple URLs or multiple other data types returned on resolution. The issue is therefore not “what can a DOI name resolve to” in a generic way, but rather what it does resolve to in any specific context. It depends on what the assigner wants it to resolve to in a particular application. The solution lies in a specific application domain rather than in the DOI system.

- Note specifically that what a DOI name identifies may not be what the DOI name resolves to. This initially may seem counter-intuitive but is obvious when one considers that (1) resolution may return multiple things; or (2) a DOI name may identify e.g. an abstract work, which by definition can only be perceived through a derived manifestation; or (3) a DOI name may identify a person, who is more than some set of data returned on resolution.
- To return to the previous example: an identifier of a work may well resolve to a specific manifestation (CrossRef DOI names are an example), yet this does not mean that another different manifestation of the same work would need a different DOI name.

What metadata are we making available through this DOI name?

Most DOI names are not yet used for widespread interoperability, but are used within specific applications. They do not need to reveal explicit structured metadata and have no defined (externally readable) Application Profile (they are formally called “zero AP” applications).

- Such existing applications include:
 - Using a DOI name which resolves to one (base) URL. The assigner ensures that what is resolved to is a single URL, and that this is appropriate for that application; but makes no guarantee that it is suitable for other applications. No explicit metadata is necessary as this is carried inside the application. (Example: CrossRef).

- Using DOI names which resolve to multiple resolution data, but have no defined Application Profile. The assigner ensures that what is resolved to is a multiple set of data; that this data is appropriate for that application; but makes no guarantee that it is suitable for other applications. No explicit metadata is necessary as this is carried inside the application. [examples: many of the CDI applications using Multi-link™]
- Interoperability is enabling information that originates in one context to be used in another in ways that are as highly automated as possible. Using DOI names for full interoperability implies that DOI names assigned in one context may be encountered, and may be re-used, in another context (or application) – without consulting the assigner. Interoperability of this sort is the case in the illustrative examples of applications A, B and C given earlier. Hence if the DOI name is designed to support interoperability, the metadata describing what the DOI name refers to must be explicit: the information necessary must be discoverable from the DOI name. The DOI System allows this in the form of an Application Profile (see below).
- This does not imply that assigning DOI names which can be used interoperably requires the explicit declaration of all metadata the assigner collects; but it does imply that some (at minimum a kernel) is declared explicitly. This will prevent the confusion of e.g. work and manifestation cited in the earlier example (A,B,C).

Interoperability of DOI names

Any DOI name which is intended for interoperability – that is, which has the possibility of use in services outside of the direct control of the issuing Registration Agency (RA)- is subject to DOI System metadata policy, to achieve two objectives:

1. To ensure that metadata held by different RAs is not fundamentally inconsistent, and

2. To ensure that an efficient and extensible means of interchange exists for transporting metadata between RAs (and in future other service providers).

- DOI names are grouped through an Application Profile (AP) - a group of DOI names with some shared characteristic(s). This grouping will be used to connect DOI names to Services which are common to the group. A Service is a specific transaction that can be performed on or with the DOI name. Each DOI name is associated with one or more Application Profiles, and each AP is associated with one or more defined Services.
- The DOI® Data Model defines the way in which DOI names, Application Profiles, and Services relate. APIs have been built which allow application programmers to use the data model, using DOI® APs and Services within their applications.
- One of the services, possibly the only one for some DOI names, is the provision of kernel metadata for each DOI name. This is associated with AP 1. Other sets of metadata may also be available for some DOI names and this, as with other services, would be known through the inclusion of a given DOI name in an AP and the association of that AP with the given service.
- Following DOI System metadata policy and making available a kernel of declared metadata allows added value additional information (such as non-public metadata) to be built on top of the DOI name with confidence that it can be used interoperably as and when required.
- The provisions above do not apply to DOI names registered under the "Zero AP" which are not designed for open interoperability.

How will DOI names assigned in an application be paid for?

A cost is associated with managing persistence and with assigning identifiers and data to the standards needed to ensure long-term stability, because of the need for human intervention and support of an infrastructure. The DOI System operates on the basis that such costs are borne by the assigner of the DOI name. The way in which these costs are recouped depends on the application.

- DOI names are assigned on behalf of registrants (content owners and other bodies) by approved Registration Agencies (RAs). These RAs contribute to the IDF to support the infrastructure (technical and social) which underpins the DOI System.
- The three principle areas of cost are incurred in
 - Number and metadata registration (maintenance of resolution destinations; declaration of metadata; validation of number syntax and of metadata; liaison with the IDF registry)
 - Infrastructure (resolution service maintenance, scaling and further development; customer guidance and outreach; marketing; administration)
 - Governance (common "rules of the road"; further development and support of the system)
- RAs are free to establish their own business model for the allocation of DOI names. The services offered by a DOI RA will include more than simple provision of a DOI name: these value added services may include data, content or rights management. DOI RAs may also choose to collaborate with others in business agreements and services, as well as partake of the generic facilities offered by interoperability features of DOI names.
- There is no single business model applicable to all DOI RAs, and consequently no single answer to the question of how a DOI name is paid for and what it costs.

This is one of a series of DOI System factsheets. To see the latest version of this factsheet online, and to see the other factsheets go to:
<http://www.doi.org/factsheets.html>