The Digital Object Identifier: From Ad Hoc to National to International

By Norman Paskin

Historical context

The Digital Object Identifier (DOI®) system[1] originated as a consortium standard from publishing trade bodies in the late 1990s. It has since moved through formal standardization both nationally and internationally and so demonstrates each form of consensus building, and how these relate to each other in the process of developing a standard. Two qualifications about the title of this study need to be noted:

- "Ad hoc" generally signifies a non-generalizable solution designed for a specific problem or task. Although DOI *originated* in a specific problem, it was a fundamental tenet of its development that digital convergence and multimedia availability required a generic framework for managing identification of content over digital networks. Although DOI originated from publishing trade bodies, their view was that it should be extensible to other areas as a *generic solution* to a perceived problem.
- The process of moving from ad hoc standards forum to national to international standard proceeded in parallel with the development of the DOI system; and so each stage was not a linear progression simply formalizing the preceding step for a larger audience, but provided further depth or endorsement of the development to date. The key formal standards activities provided firm waypoints, each establishing a secure foundation for going further by the International DOI Foundation.

It is useful to consider the "problem situation" both at the time and in hindsight; as a problem is tackled, it is further refined, and looking at how we would now describe the problem can shed light on why a particular path was taken[2]. We need to beware "the Whig interpretation of history"— studying the past only with reference to the present, while ignoring failed solutions and dead ends.[3] A practical example of this may be helpful. The DOI system was seen in 1996 as a response to the shortcomings of the URL specification; yet one cannot fully judge the DOI effort by comparing it to the current specifications of URI, which have changed in that period[4] and are likely to change further.[5]

A good summary of the problem as it was seen at the outset of the DOI effort is given in an early article by Davidson and Douglas,[6] specifically *how can scholarly publishers avoid the long-term chaos resulting from unstable web links giving "404 not found?"* The outlined solution was to marry technologies for managed identifier registries that have been proven in print publishing, like ISBN and ISSN, with technologies for managing identifiers on digital networks.

The problem from the perspective of hindsight is somewhat more complicated. We now see it as "*how can we design and use identifiers within the digital network in the digital supply (and rights) chain of all content, supporting the highest level of automation, trust, and accuracy?"*[7] This perspective highlights a number of issues that were only vaguely perceived, or not perceived at all, in the original problem view: specifying appropriate metadata, interoperability

with other identifiers, and technical and social infrastructures to support registries and persistence commitments. All of these have been part of the evolution of the DOI system, and its standardization.

The initial problem had been recognized in several places in the early 1990s, which led to convergence through the mediation of publishing trade bodies:

- Some publishers had already recognized the need for new means of identification of digitized copyright content, resulting in collaboration to develop a Publisher Item Identifier.[8]
- Douglas Armati, following work on digital copyright issues at Murdoch University in Western Australia in 1990-91, promoted the concept of a Unique Digital Identifier (UDID) for copyrighted content in several venues, including the annual meeting of the International Association of Scientific, Technical and Medical Publishers (STM) held at the 1994 Frankfurt Book Fair.[9] STM formed a task group on the subject in collaboration with the International Publishers Association (IPA), and commissioned a further report from Armati,[10] whose work had drawn attention to the wider nature of the problem.
- The Association of American Publishers (AAP) tasked its Enabling Technologies Committee with specifying a system that would protect copyright while facilitating commercial transactions. Armati produced for the AAP a similar study paper[11] to the one for STM, and the most significant step taken by AAP (with the leadership of one of its staff members, Carol Risher) was to enlist a consultant, Chris Burns, who conceived a pilot technical answer involving a collaboration between R. R. Bowker (the US ISBN agency, representing a system proven and trusted by publishers) and the Corporation for National Research Initiatives® (CNRI [12], the developer of the Handle System®[13] for managing identifiers on digital networks).

The late Charles Ellis[14] was a senior figure in AAP, IPA, and STM and played a key role in the unification of their efforts, leading to a joint statement of support for the DOI system in April 1997 and a commitment by all those parties at the Frankfurt Book Fair that year. [15] The International DOI® Foundation (IDF)[16] was created to develop and manage the system, with Ellis as its first chair (and Norman Paskin as Director). The IDF has been at the heart of the standardization of the system. The Handle system was devised by one of the pioneers in internet development, Dr. Robert E. Kahn;[17] this helped to add credibility to the DOI system on its launch, and Bob Kahn has been a firm and valued supporter of DOI.

The "ad hoc": post Frankfurt 1997

The announcement of the DOI system in 1997 was a statement of intent, and required the launch of an organization (the IDF) to obtain funding for activities to meet the commitment of the founders. A first step was deepening understanding of what would need to be agreed as a consensus.[18] Some of this work was in collaboration with NISO; for clarity, an account of that work is deferred to the next section. Initially, there was no thought of moving DOI to a formal standard, but it was recognized that there would be merit in opening up the discussion and participation with standards bodies. Activities prior to the launch of the DOI included some "content-led" activities that were contributions to a potential solution, among them the *Publisher Item Identifier* (PII)[19] (conceived by

an informal collaboration of Elsevier Science and several US scientific society publishers), and the *Serial Item and Contribution Identifier* (SICI) .[20]   None of these activities had however come to terms with the burgeoning uptake of web technologies.   Key early topics for the new Foundation were:

- *What does a DOI identify*? It soon became clear that DOI was to be interpreted as a "digital identifier of an object" rather than an "identifier of a digital object;" and that the object (the *referent* of a DOI), often thought of as a "physical thing" or, in the case of internet, a "digital thing," usually turns out to be an abstract thing—a class, e.g., an ISBN identifies the class comprising all copies of an edition.

- *A clear distinction between what an identifier refers to and what it may be resolved to* on a digital network. It was seen that there should be a lot of freedom in what the DOI resolves to (multiple things, be it the object itself, a homepage, or order form, etc.) and this could be allowed to evolve as best practice in a community, but that it was essential for a given DOI to unambiguously *refer* to one and only one object.   The distinction was not clear initially to many people, since URL practice resolved an identifier to one single location, and thereby often there was conflation of the identified entity and the resolution result, whereas the Handle system has the capability of direct *multiple resolution* (from one identifier to several results) and so makes this distinction clearer.[21]

Completion of consensus on these issues, plus agreed standardization of the DOI syntax (see next section) led to the publication on the DOI website of the first version of the *DOI® Handbook*[22] in July 2000.

The aim of the International DOI Foundation was to avoid re-invention of technologies or practices, where these already existed and had been proven, so as to harness best practice to meet the needs of the DOI system. Given the importance of the migration to digital technologies in this period, it is not surprising that the Foundation saw a need to coordinate extensively with a wide range of other initiatives. Among the most productive of these were discussions with the ISBN and EAN bar code[23] communities on models for how identifier registries could be funded; active discussions with communities involved with identification of rights in content[24] (which led to the indecs framework[25]); and understanding more about the potential of the Handle system's multiple resolution capability for resolving identifiers on networks.[26]

The most important development of this "ad hoc" period was an initiative launched early in 1999 by the AAP's subcommittee on DOI (in collaboration with IDF)[27] to build a prototype DOI-cross-citation (DOI-X) project[28]. The DOI system itself is optimized to providing speedy and accurate resolution of a DOI to data; the reverse of this process (look-up from data to a DOI) was a service that should not be built and run by the IDF (as this risked the creation of a vast "universal content" database and duplication of existing effort) and would be best federated to individual registration agencies to run as services for their communities, where appropriate existing content databases might also be harnessed. The DOI-X pilot demonstrated the feasibility of this concept for prospective users and provided an answer to consistent cross citation of digital content for STM publishers with an urgent need for such a tool. Emerging from this successful prototype was a commitment by several STM publishers to create a third-party consolidation service to take on this role, which became the first use of DOIs: CrossRef.[29]

The "national": ANSI/NISO Z39.84

It was recognized that there would be merit in opening discussions with standards bodies with an interest in the content sector. NISO, the National Information Standards Organization in the US, had close connections with the community from which the problem statement had originated, and represented the most important market for STM information at the time. Joel H. Baron,[30] NISO Chair for the two years from June 1997 and a member of the new IDF Board, suggested that the IDF should work with NISO as a way of quickly establishing a community of interest pending the development of a full membership participation structure in the IDF.   A number of IDF working groups (open to any interested participants) were established and "*Discuss-DOI; the general DOI discussion forum* mailing list was started (in January 1998 as a NISO list and transferred to the control of the IDF in June 1998), open to anyone who wished to make a contribution.[31] The list continued for six years, finally closing in February 2004 after several hundred postings, by which time much had been achieved and most development activities had migrated to within the IDF. Among the successes of the list was the publicizing and organization of some early and well-attended workshops, and using the list members in testing proposed specifications.

This formed a natural link to the next step, which was a suggestion that once some basic specification for DOI had been established, it should be standardized through NISO (even if not yet a complete system), as a clear waypoint. It became clear that a significant part of the initial work could be separated out and usefully standardized: the syntax of the DOI name. A NISO committee was formed, chaired by Ed Pentz,[32] tasked with creating a *"Standard Syntax for the Digital Object Identifier, to clearly delineate the alpha/numeric string which identifies the object being pointed to within the DOI system and the rights holder of the object."*   After successful ballot in 1999, the standard was published as ANSI/NISO Z39.84-2000. The standard was revised with a minor modification in 2005 and reconfirmed in 2010.[33]

Note what was included in the syntax and why was it needed, as well as what was excluded from the NISO standard. The syntax was a straightforward application of the Handle system chosen as an underlying mechanism for the AAP pilot. At the time the Handle system syntax was well established and available on the Handle site,[34] but not yet a published standard;[35] and the DOI system used a restricted implementation (for example, it restricted the Handle's Unicode possibility to just alphanumeric characters and specified encoding rules for some common special key characters). Publishing the syntax as a NISO standard also made potential users in the library and publishing community more comfortable with using this hitherto unfamiliar specification, especially at a time when CrossRef was about to launch the DOI-based citation linking service.

The syntax standard omitted the committee's initial mandate to identify "the rights holder of the object," as it quickly became clear from discussions with rights communities that this was not a simple matter.[36] The standard did not make any mention of accompanying metadata (necessary in order to define and distinguish what was being identified). Nor did it cover the mechanism and procedures by which DOI names would be allocated. With hindsight, one such recommendation would already have been useful in the standard: that of encouraging many prefixes to be allocated. We now see that there is no penalty to using as many prefixes as useful, and current DOI users are

encouraged to work in this direction, allocating at least one separate prefix for each customer, and where appropriate more than one (e.g., a single company that has three easily identifiable divisions, imprints, product lines etc. may best be considered three "customers" for the purposes of prefix allocation on the assumption that at some point in the future the company could split along those lines). Such recommendations now form part of the DOI System's policies.

The "international": ISO 26324

The ISBN agencies had been involved in the conception of the DOI pilot and in 2004 Brian Green, then Managing Agent of the International ISBN Agency and chair of the International Organization for Standardization (ISO) Technical Committee 46 (Information and documentation)/SubCommittee 9 (Identification and description)[37] responsible for information identifier standards, invited IDF to make a proposal to SC9 to standardize the full DOI system as an International Standard. DOI development had by then reached a stage sufficient to make that step meaningful, adding value by codifying the DOI-agreed specifications for metadata and registries, indicating international acceptance, and complementing the existing ANSI/NISO Z39.84 syntax through expansion to cover a detailed extensible metadata schema and discussion of the guarantees provided regarding persistence. The IDF felt that it was possible that, as had been the case with the NISO standardization, such a step would instill further confidence in the user community, acting as a guarantee of good governance, and provide a template for future uses by new applications beyond those by then well-established CrossRef. It was however a requirement that, as with the NISO process, standardization did not remove the momentum (some 12 million DOIs had been assigned by mid-2004) that the DOI system had achieved.

The IDF had two reservations about this proposal: (1) that the ISO process could be notoriously time-consuming; and (2) that the DOI system could allocate identifiers to material already in the scope of other SC9 identifiers (e.g., a book might have both an ISBN and a DOI) and so might be wrongly perceived as a problem, rather than an opportunity[38]. Both concerns proved justified. In October 2004 a presentation was made to SC9, which accepted the idea in principle. Jane Thacker, then secretary of SC9, provided a significant amount of effort in drafting an initial New Work Item Proposal, issued in May 2006, based on the then current DOI Handbook. Following approval by SC9 of the proposal, a working group (ISO TC46/SC9/WG7) was initiated in September 2006 and a Committee Draft was completed in late 2007. During the working group process, two significant improvements were made: (1) removal of detailed resolution and metadata procedures, elevating the standard to an abstract specification to be met by any appropriate technology matching the specifications; and (2) clarification of how other identifier systems can work with the DOI System. The draft standard included recommendations on incorporation of other schemes into the syntax or metadata of the DOI System, and "the scope of the DOI system is not defined by reference to the type of content (format, etc.) of the referent, but by reference to the functionalities it provides and the context of use."[39] In April 2008 the Committee Draft was approved by the national standards body members of TC46/SC9 as a Draft International Standard (DIS); and the Final Draft International Standard was unanimously approved in a ballot closing in Nov 2010.[40] ISO 26324 was finally published on May 1, 2012,[41] after a further delay caused by ISO revising its generic Registration

Authority agreement (an issue which was not specific to the DOI system, but which affected all the SC9 standards that require a Registration Authority).[42]

As noted earlier, the ANSI/NISO Z39.84 standard did not cover DOI metadata or registration procedures. By 2004 significant progress had taken place with both, allowing their inclusion in the ISO draft standard. This was significant, since the DOI should be extensible to all media types (in recognition of digital convergence), and so its metadata should be similarly extensible. Whereas in the case of DOI identifier management we had been able to adopt a highly efficient and proven technology (Handle), no such existing solution for metadata was immediately available. We were fortunate that we were not the only group searching for such a solution: the early years of the International DOI Foundation coincided with the indecs project (1998-2000), devised to provide such a scheme and since widely adopted[43] and continued in activities such as the Linked Content Coalition[44]. IDF became an early participant and subsequent supporter of the indecs approach. The metadata specifications of ISO 26324, a DOI "Kernel" of controlled metadata elements plus extensible options based on a data model, is a development of indecs.

At the time when the DOI community was looking for the metadata methodology to specify interoperability rules and applications, one approach considered and rejected due to its limited scope was the Dublin Core (DC), which was devised as a metadata set for searching for bibliographic resources on the internet. The IDF participated in some early DC meetings and the debate about whether the "fifteen elements in search of a data model" could be truly reverse engineered into an extensible framework—a challenge which was not met. DC has since been formally standardized in various ways[45] and it put certain elements (like *dc:creator* and *dc:title*) into quite widespread use, which promotes a certain amount of interoperability. Some basic tags can be used for common terms in multiple schemas for lightweight search interoperability. But its limitations in terms of vagueness and ambiguity cause problems (e.g., the arbitrary distinction of *dc:creator* and *dc:contributor* that can be interpreted quite differently by different users, or the extreme vagueness of *dc:date*). For use in depth, as with DOI, it must be extended and there is no common model to ensure that extensions are compatible. Few in-depth content metadata standards developed since Dublin Core have built on it, either in the content creator/publisher world (e.g., ONIX, DDEX, PRISM, PLUS, etc.) or recent major bibliographic developments (e.g., FRBR and RDA). By contrast, the indecs principles[26] and approach have been adopted and developed in, among others, the Vocabulary Mapping Framework, RDA/ONIX Framework for Resource Categorization, DDEX (Digital Data Exchange) Music industry messaging and data dictionary application, ONIX (Online Information Exchange) standards for the use of publishers in distributing digital metadata about their products, and is at the heart of the Linked Content Coalition framework for a fully interoperable and fully connected standards-based communications infrastructure for rights management.

Reflections

Some observations from this use case:

- The activities in early years of DOI "ad hoc" development overlapped substantially with both AAP and NISO workshops and with many common players; it becomes hard to distinguish the formal standards process from the "ad hoc" development that strongly influenced it. I

believe that is a good sign: it shows ongoing engagement with the market need for a standard.

- The choice of a national standards forum (NISO) for the initial syntax standardization was coincidental. What was important was NISO's role in the information community and its overlap with the community developing the DOI, not its national role in the US.

- The step from syntax to full system standard merited international standardization, rather than another national NISO standard, as a peer of similar standing to other ISO identifiers used in the community and likely to be of interest for interoperability with DOI.

- The choice of the ISO route was a logical one given the early role of the International ISBN Agency, NISO, and others in the international ISO community. But it is also possible that we could have gone down a route of IETF RFC or OASIS[46] standardization. The ISO choice was largely due to connections both organizational and personal, and the focus on the content communities as the key rather than the technology used.

- Unlike simple specifications (e.g., for socket dimensions), information identifiers are not simple specifications that anyone can adopt. They require a controlled registry mechanism (with close attention to what *precisely* is being identified, which in the case of abstractions can be unclear) and thus ultimately a registration authority to manage and implement the standard. This brings the potential problem of possible vested interests in commercial implementation of the standards (and thus by participants in the development process). It was thus necessary during the standards development process to pay attention to intellectual property policies and declarations of interest (and, perhaps especially, undeclared interests).

- Information identification in a connected world necessitates semantic interoperability, which is knowledge representation and thus requires a relatively heavyweight data model.[47] This is specialist work, not ideally suited to committee work; in the case of DOI, this work was done by experts commissioned by the IDF and submitted as a near-final form in the standard proposal, not developed as part of the standards process. The same may be true of detailed network technologies; we are approaching the point where the guideline "as simple as possible, but not simpler" necessitates informed specialist effort in the standards process (and therefore may give difficulties in the review process).

- In the move from ad hoc to national to international, the DOI opened up the audience for feedback to wider groups at each stage, benefiting from testing by constructive criticism. The NISO route, since it was a simple standard and close to the originating community, produced relatively little disagreement and change to the proposed specification, but good endorsement through informed participation. By contrast, the ISO route (with the more complex standard) produced substantive discussion and improvements (the move to an abstract specification, interoperability with other identifier scheme, etc.); but the downside to this were:
    - Fears concerning the length of time needed for ISO standardization were justified. The full ISO process for DOI took seven and a half years (October 2004 to May 2012) from inception to completion.
    - The ISO process requires voting by national standards bodies. Not all national bodies have the interest or expertise to produce well-informed review. This can produce something of a herd mentality where a substantial number may vote for what seems the least troublesome course of action; as H.L. Mencken put it, *"There is*

> *always a well-known solution to every human problem—neat, plausible, and wrong."*
> In the case of DOI, a few national bodies were very helpful, despite occasional dangers of straying into these waters. In my view this diversity in understanding and participation remains a weakness of the ISO process.

- The decision taken by IDF to look beyond the short-term problem statement (consulting with other industries beyond text publishing for example) and focus on the long term (avoiding pinning the DOI to a specific technical solution and building on enduring principles) was proven so far to be correct. Today we can look back at development of a standard that necessitated no significant U-turns; is a proven reliable system, which in 2012 reached 1 billion resolutions of DOIs per year; and has an increasing coverage across scholarly publishing, scientific data, and the entertainment industry among others.

- Globalization will increase the importance of international standards. The IDF now has four Registration Agencies operating in non-Roman text environments (in China, Japan, and Taiwan). It is possible that future standards will need to pay more attention to international issues such as Unicode, internationalized resource identifiers (IRIs), and internationalized domain names.[48]

- During the development process, the standards community and the publishing community moved closer. Consider for example the preceding PII effort, where there was neither perceived need from the publishers nor any interest from the standards community[49] to develop this as a formal standard, and compare that to the current wave of activities on new identification efforts (ISSN-L, ISTC, ISNI, ORCID, etc.). This is certainly not due only to the DOI; the increasing use and depth of messaging (e.g., ONIX standards) and web technologies (e.g., linked data) has brought standards to the attention of the publishing sector.

- Reviewing the history of the DOI, one sees the key roles played by individuals in making connections; such continuity is a great advantage. Some of the key people have already been mentioned; it is significant that the current IDF retains some Board members[50] and several operational participants[51] from the early launch period.

About the Author: Dr. Norman Paskin (n.paskin@tertius.ltd.uk) is the founding Director and Managing Agent of the International DOI Foundation. He was part of the STM Task Force on Information Identifiers, which was one of the originators of the DOI system, and convenor of the ISO working group that standardized the DOI System as ISO 26324.

Notes

---

[1] The unqualified term "DOI" alone (which was used in the early years of the system's development) is now deprecated, as a potential source of confusion. The preferred usage is with a qualifier to refer to either specific components of the

DOI System (e.g., "DOI name", the string that specifies a unique referent within the DOI System); or the system as a whole ("DOI System"), which is the sense implied here if not otherwise stated.

2   The analysis of problems as the basis of historiography was championed by Karl Popper, See for example: Karl R. Popper, *The Myth of the Framework* London/New York: Routledge, 1994, Chapter7, "A pluralist approach to the philosophy of history."

3   Herbert Butterfield coined the term "Whig history" in his influential book, *The Whig Interpretation of History* (London: G. Bell and Sons, 1931). See especially Chapter 2, "The Underlying Assumption" available in the Electronic Library of Historiography (ELIOHS) at: http://www.eliohs.unifi.it/testi/900/butterfield/chap_2.html

4   The current specification is IETF RFC 3986, *Uniform Resource Identifier (URI): Generic Syntax* (International Engineering Task Force, January 2005), available at: http://www.ietf.org/rfc/rfc3986.txt. For a discussion on how the URI specifications have evolved and the historical ambiguity and confusion in the use of the terms URI, URL, and URN, see" *Factsheet: DOI® System and Internet Identifier Specifications (International DOI Foundation) available at:* http://www.doi.org/factsheets/DOIIdentifierSpecs.html

5   Larry Masinter's November 2, 2012 posting "obsoleting 3986 -- what would it look like?**"** to the W3C URI list (available at: http://lists.w3.org/Archives/Public/uri/2012Nov/0000.html) proposes the creation of a new RFC "to obsolete 3986 (URI) with a document that combined it with 3987 (IRI) [Internationalized Resource Identifier, a generalization of URI allowing the use of Unicode], reverts to the "URL" name, and gave updated parsing advice… removing any basis for support of using http URLs to "mean" abstractions or people."

6   Davidson, Lloyd A. and Kimberly Douglas. "Digital Object Identifiers: Promise and Problems for Scholarly Publishing." *Journal of Electronic Publishing,* 4 (2), December, 1998. http://dx.doi.org/10.3998/3336451.0004.203

7   A good statement of the current perspective is that of the Linked Content Coalition (www.linkedcontentcoalition.org/), *Linked Content Coalition: Principles of Identification, Version 1.1, April 2014, including Appendices.* Editors Norman Paskin & Godfrey Rust. *DOI: 10.1000/283* *http://doi.org/10.1000/283*

8   Publisher Item Identifier as a means of document identification http://web.archive.org/web/20031013073003/http://www.elsevier.nl/inca/homepage/about/pii/

9   Armati, Douglas. *A Uniform Approach to Identification of Digitized Copyright Content*? *STM Newsletter* 95. November 1994. [Transcript of paper delivered at 26th STM General Assembly, Frankfurt, October 1994.] http://www.doi.org/topics/Armati_STMNews.pdf

10 Armati, Douglas. Information Identification: Report for the STM International Group of Scientific Technical and Medical Publishers Task Force on Information Identifiers and Metering Systems in the Electronic Environment. Amersfoort, The Netherlands, June 1995. http://www.doi.org/topics/Armati_Info_Identification.pdf

11 Armati, Douglas. *Uniform File Identifiers: Report to the Association of American Publishers*. Commissioned by the AAP Enabling Technologies Committee, Washington, October 1995.

12 Corporation for National Research Initiatives [website] http://www.cnri.reston.va.us

13 Handle System [website]. http://www.handle.net

14 Weber, Bruce. "Charles R. Ellis, President of John Wiley & Sons, Dies at 72 [Obituary]".*The New York Times*, May 10, 2008. http://www.nytimes.com/2008/05/10/business/10ellis.html

15 *Ensuring the Future of Electronic Publishing: Ground-breaking Digital Object Identifier System Leads Publishing Industry into 21st Century*. Press Release. Washington, DC: Association of American Publishers, September 22, 1997.http://www.doi.org/press-9-22.html

16 International DOI Foundation [website]. http://www.doi.org. The IDF is a not-for-profit membership body, the DOI system registration authority and maintenance agency, and the central body that governs the DOI system. It is the common management and coordination body for DOI Registration Agencies and also manages those aspects of the DOI system that are put through external standardization procedures.

17 *Robert E. Kahn* [bio]. http://www.cnri.reston.va.us/bios/kahn.html

18 Detailed historical summaries of the DOI System development can be found in the IDF's archive of 79 Reports to IDF Members issued by the IDF between June 1998 and January 2009, together with key presentations, reports of members meetings, and other key documents (currently available only to IDF members at: http://www.doi.org/idf-members/rpts-old.html).

19 http://en.wikipedia.org/wiki/Publisher_Item_Identifier . The PII had been developed because of perceived problems limitations with the SICI for digital use; it was never formally standardised, as it was superseded by the use of DOI in the CrossRef application to solve the same problem, but continues in informal use in several publishers internal registries.

20 *Serial Item and Contribution Identifier (SICI).* ANSI/NISO Z39.56-1996 (R2002). Version 2. Bethesda, MD: National Information Standards Organization, August 14, 1996, reaffirmed August 30, 2002; withdrawn August 27, 2012. http://www.niso.org/standards/z39-56-1996r2002/

21 This is necessarily an approximation: to add to the complexity, a given physical server can easily be the end point for routing to multiple IP addresses. This reinforces the point that identifiers require a specific dedicated mechanism beyond DNS.

22 The DOI® Handbook (http://www.doi.org/hb.html) has been revised many times since then and remains a key resource for the system.

23 EAN bar codes are now managed by GS1. http://www.gs1.org/

24 See for example: Rust, Godfrey. "Metadata: The Right Approach, An Integrated Model for Descriptive and Rights Metadata in E-commerce." *D Lib Magazine,* 4 (7/8), July/August 1998. http://www.dlib.org/dlib/july98/rust/07rust.html

25 Rust, Godfrey and Mark Bide. *The <indecs> metadata framework: Principles, model and data dictionary*. WP1a-006-2.0. Indecs Framework Ltd., February 26, 2000. http://www.doi.org/topics/indecs/indecs_framework_2000.pdf

26 DOI Discussion paper: *From One to Many*. International DOI Foundation, August 2000. http://www.doi.org/one2many.pdf

27 The AAP's Enabling Technologies Committee that had launched the DOI pilot was now (1995-1998) chaired by Craig Van Dyck, who subsequently has been a long term Board member of both IDF and CrossRef.

28 Atkins, Helen, et al. *"Reference Linking with DOIs: A Case Study."* D-Lib Magazine, 6 (2), February 2000. http://www.dlib.org/dlib/february00/02risher.html

29 *The Formation of CrossRef: A Short History.* CrossRef, 2009. http://www.doi.org/topics/CrossRef10Years.pdf

30 At the time, publisher of *The New England Journal of Medicine.*

31 The list is archived a: http://www.doi.org/mail-archive/discuss-doi/maillist.html.

32 Then at Academic Press and subsequently Executive Director of CrossRef.

33 *Syntax for the Digital Object Identifier*, ANSI/NISO Z39.84-2005 (R2010), Bethesda, MD: National Information Standards Organization, September 30, 2005, reaffirmed May 13, 2010, http://www.niso.org/standards/z39-84-2005/. The 2005 revision removed the option of case sensitivity, i.e., 10.123/ABC as different from 10.123/AbC. Although the Handle System is configurable by service so as to be either case sensitive or case insensitive, practical DOI use was entirely case-insensitive (http://www.doi.org/doi_handbook/2_Numbering.html#2.4). The 2010 reaffirmation was made without further change, on the assumption that the inclusion of exactly the same syntax in the more detailed ISO standard then underway would render any changes superfluous.

34 *Handle System: Handle Syntax*, CNRI, April 10, 2000. http://web.archive.org/web/20000510172643/http:/www.handle.net/overviews/handle-syntax.html

35 It was subsequently standardized as: *Handle System Overview.* IRFC 3650. Internet Engineering Task Force, November 2003. http://www.ietf.org/rfc/rfc3650.txt

36 This issue is only now being robustly tackled through the Linked Content Coalition (http://www.linkedcontentcoalition.org/) and its implementations, notably The Copyright Hub (http://www.copyrighthub.co.uk/)

37 ISO TC 46/SC 9 Identification and description [webpages]. http://www.iso.org/iso/home/standards_development/list_of_iso_technical_committees/iso_technical_committee.htm?commid=48836

38 The reason for this is that, despite being the same object, the object is used in different contexts by different parts of the supply chain of information. The position can be analogized in brief as: "My passport does not replace, substitute, or compete with my Amex card, though they both have the same referent (me) ."

39 From the *Introduction* of ISO 26324.

40 A significant delay was caused by a failed appeal against the standard in 2009 by AFNOR, the French national standards body, due to concerns and misunderstandings about how the DOI relates to other existing identifiers.

41 *Information and documentation — Digital object identifier system*. ISO 26324:2012. Geneva: International Organization for Standardization, April 23, 2012. http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=43506

42 There are only 67 Registration Authority Agreements among ISO's 19,000 standards, but these include all of the information identifiers in the TC46/SC9 portfolio, such as ISBN, ISSN, ISRC, etc.

43 *indecs Content Model*, Wikipedia, the free encyclopedia. http://en.wikipedia.org/wiki/Indecs_Content_Model

44 Linked Content Coalition http://www.linkedcontentcoalition.org/

45 *The Dublin Core Metadata Element Set* has been published as the following standards: ANSI/NISO Z39.85 (2001, 2007, and 2012); IETF RFC 5013 (2007); and ISO 15836 (2003 and 2009).

46   The Organization for the Advancement of Structured Information Standards (OASIS) was at the time an independent standards consortium. In 2013, it was accredited by the American National Standards Institute (ANSI) to develop and maintain American National Standards (ANS). NISO has been ANSI accredited since 1939.

47   Sowa, John F. *Knowledge Representation: Logical, Philosophical, and Computational Foundations.* Pacific Grove, CA: Brooks Cole Publishing Co., 2000. ISBN: 978-0534949655 (Summary at: http://www.jfsowa.com/krbook/)

48   See: *Unicode* (http://en.wikipedia.org/wiki/Unicode), *Internationalized resource identifier* (http://en.wikipedia.org/wiki/Internationalized_Resource_Identifier), and *Internationalized domain name* http://en.wikipedia.org/wiki/Internationalized_domain_name

49   The PII was developed by Tim Ingoldsby (AIP), Lorrin Garson (ACS), and Norman Paskin (Elsevier Science); the sponsoring group of publishers saw no need for making it openly available as a specification. The ISSN was contacted (as the PII specification built on ISSN), but they had no interest in an article-level identifier. The need for abstract-level identification and grouping was not tackled by ISSN until recently, with the ISSN-L in the 2007 revision of ISO 3297.

50   Craig Van Dyck and Ed Pentz are current Board members who have been involved since the outset.

51   Norman Paskin as Managing Agent, CNRI staff as technical services provider, and Godfrey Rust as metadata advisor.