NAMING AND MEANING: KEY TO THE MANAGEMENT OF INTELLECTUAL PROPERTY IN DIGITAL MEDIA

Norman Paskin, Tertius Ltd
n.paskin@tertius.ltd.uk

Abstract

Digital information needs to be a first class citizen in the networked environment. The fundamental characteristic of digital information is that it is processable data, enabling re-use and hence new forms of electronic commerce, creativity and social benefit. Managing these units of digital information, the "citizens" in the network, requires that they have unique *names* denoting a specific referent. Equally, these names have to have some agreed *meaning* so that one computer system knows what the names and attributes from another computer system denote. As applications become more sophisticated, objects may be representations of people, resources, licences, avatars, sensors, etc., which requires the ability to identify them by name and to have these names specify identity (what is named).

Naming is a prerequisite for management of digital information entities: as a means of storing, accessing, disseminating and exchanging them. Meaning is a prerequisite for enabling them to interact: as a means of interoperability and digital policy management[1]. In this paper I discuss a successful approach for each in the context of intellectual property in digital media:

- Several existing and emerging applications have successfully named and managed information in the form of digital objects, which are stored, accessed, disseminated and managed. A digital object is a data structure whose principal components are digital material, or data, plus a unique identifier (lexical token or name) for this material. A digital object architecture provides naming conventions for identifying and locating digital objects, a service for using object names to locate and disseminate objects, and access protocols, forming an infrastructure that is open, and which supports a large and extensible class of distributed digital information services. Such a naming system is agnostic as to technology (web, mobile, P2P, etc) and assumes only the existence of the internet protocol.

- Meaning has been tackled through semantic interoperability in a series of developments building on the indecs (interoperability of data in e-commerce systems) project, resulting in successful deployment of a context-based ontology mapping.

1.   NAMES

The name assigned to an item of digital media we wish to manage should be a first class name: one that has an identity independent of any other item. The identity allows the item to persist when its attributes change, and other items to claim relationships with the item. As a general rule, first class items represent things rather than relationships.

- A URL is not a first class name, but is an attribute: a location of a file on the WWW, currently based on the DNS (Domain Name System) – although the URL specification allows addressing by full path to host (IP address), this is rarely used. If the content, but not location, of the file is changed, a user may not know this; if the content of the file is moved, the URL link won't find it ("404 not found", or manual redirection, or automated redirection which may not persist). All URLs at one location have to be ultimately managed by the same domain name owner: the owner of the domain name has final control over all the URLs beginning with that name, which makes URLs especially brittle for any piece of content which could possibly change owners.

- a URN is a naming convention for the content of files on the internet. Although designed so that it is independent of any underlying technologies such as DNS, the only present technique of resolving URNs on the internet is based on DNS. There are no widely standardised ways of using this: e.g. you can't type URNs into browsers except in certain special circumstances.

- URI is the collective name for URN and URL schemes.

- A Handle (discussed below) is a name for entities, designed for use on IP (Internet Protocol) networks (i.e. the internet) which (a) can be used with the DNS, but is not DNS-based; (b) can redirect to a URL and is managed to be persistent even if the URL moves; (c) can have additional features of granularity of

management, structured metadata, scalability, reliability, etc

- A Digital Object Identifier (discussed below) is a Handle implementation with additional features designed for the management of intellectual property entities in digital networks.

Resolution is the process in which an identifier is the input - a request - to a network service to receive in return a specific output of one or more pieces of current information (state data) related to the identified entity: e.g. a location. The most common mechanism for resolution of names on the Internet is the Domain Name System (DNS). The DNS administrative model has disadvantages as a general-purpose name system: DNS administration typically requires a network administrator, and has no provision for administration per name by anyone other than a network administrator.). URLs are grouped by domain name and then by some sort of hierarchical structure, originally based on file trees, now possibly unconnected from that but still a hierarchy. Handles offer a more finely grained approach to naming where each name stands on its own, unconnected to any DNS or other hierarchy. This offers beneficial flexibility, especially over time, as the document origins reflected in that hierarchy lose meaning, such as a change in ownership reflected in DNS. DNS also has well-recognised problems of security and updating, and potentially of scalability in the face of new technologies[2], which suggest that it will not be sufficient to assume that existing DNS technology should be adapted to deal with new requirements, rather than inventing something new[3]: peer-to-peer networks already presage this. It is interesting to note that DNS resolutions don't take up as high a percentage of internet traffic as they used to (BitTorrent is now a substantial load).

1.1 Names and Digital Object Architecture

A Digital Object Architecture provides a means of managing digital information in a network environment, by viewing a digital object as a machine-independent and platform-independent structure that allows it to be identified, accessed and protected, as appropriate. A digital object may incorporate not only informational elements, i.e., a digitized version of a paper, movie or sound recording, but also the unique identifier of the digital object and other metadata about the digital object. The metadata may include restrictions on access to digital objects, notices of ownership, and identifiers for licensing agreements, if appropriate.

The most widely cited Digital Object Architecture Project, at the Corporation for National Research Initiatives (CNRI)[4], is based on the fundamental work of Kahn and Wilensky[5]. It describes an infrastructure of services that provide access to distributed and secure digital objects: networked objects that are instantiated by an infrastructure service, which includes

- the concept of digital objects;
- the Handle System for providing persistent names for Internet resources. It is a highly reliable, high performance, distributed system.;
- the Repository for storing them and making them available over the Internet. The Repository provides network based storage and access to digital objects: all access to digital objects passes uses a simple repository access protocol and is subject to access controls established by the manager of the repository. The repository provides an information service that removes host, application, and storage dependencies while providing consistent access to digital resources and makes them virtually accessible and manageable anywhere over the Internet as if they were located on one's local desktop.
- the Registry: a specialized repository that provides secure registration and authentication of digital objects. The registry also provides a way to search for registered objects, depending on what sort of metadata is provided to the registry. This would be community specific.

Although designed as a complementary set of components, individual elements from the architecture have been used alone in applications: the Handle system is one such, described in more detail below.

In the integrated architecture concept, digital objects provide access to their content using an extensible and secure dissemination mechanism (disseminations can be thought of as high level types that are uniquely distinguished by a combination of operations, and types of data the latter are performed on). Current ongoing research includes the development of dissemination registries, infrastructure searching, security and scalability.

1.2 The Handle System

The Handle System[®][6] is a general purpose distributed information system used to assign, manage, and resolve persistent identifiers, known as "handles," for digital objects and other resources on the Internet. A handle may be a long-lasting reference to digital material that can be used to locate the material even when it changes location or owner. It can also be used to return metadata related to the material. Over 600 million digital objects are managed today through a globally distributed set of handle service sites. It offers general purpose efficient, extensible, and secure identifier and resolution services for use on networks such as the Internet. It includes an open set of protocols, a namespace, and a reference implementation of the protocols. The protocols enable a distributed computer system to store identifiers, known as handles, of arbitrary resources and resolve those handles into the information necessary to locate, access, contact, authenticate, or otherwise make use of the resources. This information can be changed as needed to reflect the current state of the identified resource without changing its identifier, thus allowing the name of the item to persist over changes of location and other related state information.

The Handle System was developed by CNRI under the overall direction of Dr. Robert Kahn, a pioneer in open-architecture networking, the co-inventor of the TCP/IP protocols, and the originator of the DARPA program which developed the Internet. It is described in a series of informational RFCs [7] [8] [9]. The Handle System has been widely used for educational, research, and other experimental purposes; CNRI 's Handle System Public License allows commercial and non-commercial use of both its patented technology and reference implementation of the Handle.net software, and will allow the software to be freely embedded in other systems and products. The license comes with a service agreement requiring those who use the system to provide identifier and/or resolution services to obtain a unique prefix from the Handle System administrator and to follow certain system-wide guidelines to insure the overall integrity of the system. Users of the system will not encounter charges for its use, however, service providers will be charged an annual fee which will be used to defray costs for the global root of the system, known as the Global Handle Registry[®]. CNRI has established an interim Handle System Advisory Committee to move the Handle System to an independent governance model.

The Handle System includes an open set of protocols, a namespace, and an implementation of the protocols. The protocols enable a distributed computer system to store handles of digital resources and resolve those handles into the information necessary to locate and access the resources. This associated information can be changed as needed to reflect the current state of the identified resource without changing the handle, allowing the name of the item to persist over changes of location and other state information. Each handle may have its own administrator(s), and administration can be done in a distributed environment. The name-to-value bindings may also be secured, allowing handles to be used in trust management applications. The Handle System is an infrastructure on which applications serving many different purposes are being built. Users of the Handle System include the content industries and related sectors through the Digital Object Identifier System (DOI[®]) system; the Library of Congress; the U.S. Defense Technical Information Center; the US Department of Defense distributed learning systems [10];MIT's DSpace digital library system; the Globus Alliance (which produces a leading open source toolkit for building computational grids); the Advanced Distributed Learning (ADL) CORDRA project for federating learning object repositories; the Australian Research Repositories Online to the World (ARROW) project; CNNIC (China Internet Network Information Center: see below), and various other efforts. Handle applications can e.g. federate a group of other identifier schemes[11].

1.3 Internationalisation

"The internet is a global revolution in communication - as long as you use letters from the western alphabet. [There is] growing pressure for a net that recognises Asian, Arabic and Hindi characters, too" [12]. The DNS only recognizes ASCII characters A-Z, 0-9 and the hyphen, the characters used in primarily Latin-based languages; it does not recognize other character sets. Around 33% of the current online population are native speakers in non-Roman character language zones (Arabic, Chinese, Farsi, Hebrew, Japanese, Korean, Malay, Thai, Vietnamese) - an estimated 240 million people[13], a figure likely to grow especially as online transactions keep moving into traditional areas, e.g. finance and

consumerism. Handles may consist of any printable characters from the Universal Character Set (UCS-2) of ISO/IEC 10646, which is the character set defined by Unicode v3.0. The UCS-2 character set encompasses most characters used in every major language written today. The Handle System has been documented in Chinese[14][15][16]. To allow compatibility with most of the existing systems, and to prevent ambiguity among different encodings, the Handle System protocol mandates UTF-8 to be the only encoding used for handles. The UTF-8 encoding preserves any ASCII encoded names so as to allow maximum compatibility with existing systems without causing naming conflict.

CNRI have been working with China Internet Network Information Center (CNNIC), the state network information center of China. CNNIC is China's domain name registry to operator and administrator of the ".CN" country code top level domain (ccTLD) and Chinese Domain Name (CDN) system. A Handle-DNS integration system has been developed (described below) which will result in a deployment of integrated Handle-DNS through the .cn domain. The International DOI Foundation, a Handle implementation, is discussing the establishment of a DOI Registration Agency in China.

1.4 Handle-DNS integration

The CNNIC/CNRI collaboration[17] takes advantage of the Handle System to provide a security service for the DNS namespace, including secured DNS resolution (whenever needed), discretionary administration & dynamic update, access control & privacy protection, delegation & real-time credential validation. This service may co-exist with the existing DNS operation: there is no need to change the DNS client.

The abstract Handle System is specified in RFC3650,3651,3652. CNRI have developed, and distribute, a reference implementation of the specification, available as Java through open source distribution. A perfectly compliant handle service built without using any of the reference implementation code would not by definition be distinguishable from the standard version from the outside: the global Handle records give IP address, port numbers, public keys, etc., but nothing about the internals of the machinery with those handle service attributes. (However, the onus would be on the developer to ensure that this

assertion of compliance was true). CNNIC developed a Handle Server in a new implementation in C/C++ (server/client) integrated with BIND 9.3.0 standard distribution, and additional modules offering improved performance. A prototype application offers secured DNS resolution via a Handle protocol interface. Further work will package the Handle-DNS software for public release; deploy the Handle-DNS server in ".cn" TLD registry and its subsidiaries; and establish an ENUM service and client software based on the Handle-DNS interface.

The DNS/Handle integration enables an identifier service for any digital resource over the Internet, with a distributed, scalable service infrastructure similar to DNS with additional features:

- Efficient name-resolution and administration, supporting both TCP and UDP.
- Built-in security options for both name resolution and administration.
- Secure handle resolution, including data confidentiality and service integrity checking
- Discretionary namespace and identifier attribute administration, independent from host-admin, which allows creation, deletion, and modification of identifier and/or identifier attributes (this level of granularity is a requirement for any truly sophisticated extensible management of digital media objects)
- Standard access control model per individual identifier attribute (essential for privacy protection applications).
- A mechanism for credential validation per individual handle attribute.

1.5 Digital Object Identifiers

The International DOI Foundation has developed DOI names (Digital Object Identifier names)[18] as actionable persistent identifiers for content-related entities. Note that "DOI" is construed as *digital identifier of an object* (not "identifier of a digital object"); and the term *"object"* here is used in the accepted ontology sense of an entity which may be abstract, physical or digital[19], since all these forms of entity are of relevance in content management (e.g. people, resources, agreements)[20] and may be manifested in, or compounded within, a particular object (see (2) below). The DOI system is an implementation of the Handle System, but that is only one of

the four components of the system which comprises:

- *Numbering syntax*: rules for assigning an alphanumeric string (a number or name) to the intellectual property entity that the DOI name string identifies. The naming mechanism follows a syntax standardised as ANSI/NISO Z39.84-2006. The number may incorporate any existing identifier scheme (thereby retaining its construction, check digits, etc.) though for the purpose of the system the string is "opaque" or meaningless. DOI names are not case-sensitive and have no fixed field length.
- *Description* of the entity that has been identified with a DOI name, through associated metadata. The DOI data model is based on the <indecs> framework, described later in this paper, and provides a data dictionary to precisely define referents, and a grouping mechanism (Application Profiles) to relate sets of DOI referents with common properties.
- *Resolution*: the Internet technologies that make the identifier "actionable" on digital networks, by providing resolution services, using the Handle System.
- *Policies*: the rules that govern the operation of the system, in a social infrastructure. The social infrastructure defines the funding and ongoing operational requirements of the system as well as its day-to-day support and management. One of the key features of the DOI system is a co-funded social infrastructure to ensure consistency and quality. This also ensures a fair distribution of funding for the required technical and social infrastructure needed for the system. As with other identifiers such as ISBN, ISSN, etc, the only persons permitted to register DOI® identifiers are Registration Agencies that have been authorized by IDF, or persons acting under the authority of Registration Agencies

The DOI system is unique in bringing together all four components in a fully implemented and managed system[21]. It is currently undergoing standardisation through ISO TC46/SC9.[22]

A DOI name persistently identifies an entity of relevance in an intellectual property transaction and associates the entity with relevant data and services. An entity can be identified at any arbitrary level of granularity. DOI names can be used to identify, for example, text, audio, images, data[23], software, etc., and in future could be used to identify the agreements and parties involved, though initial implementations have focussed on "creations". While the scope of intellectual property transactions is quite broad, it is unlikely that DOI names would be appropriate for identifying entities such as people or natural objects or trucks unless they are involved in such a transaction. DOI names can be used to identify free materials and transactions as well as entities of commercial value[24].

The DOI system provides:

- *Persistence* – DOI names resolve to information (metadata) about the referent (identified object) in a manner that persists over changes in location, ownership, description methods, and other changeable attributes. If the object ceases to be available, the DOI name at minimum indicates a valid but now defunct identifier.
- *Interoperability* – providing tools to enable the use of the identified referent in services beyond the assigner's direct control, which enables rich interlinking with related content, so as to increase the content's usefulness and visibility;
- *Extensibility* - the ability to later add new features and services
- *Efficiency* - Through single management of data for multiple output formats (platform independence) and class management of applications and services, efficiency is gained;
- *Dynamic updating* - metadata, applications and services need to be quickly and easily updated.

The benefits of this functionality, because it is essentially generic and so rather abstract, needs to be translated into specific illustrations that make sense for a particular community. For example, DOI names in enterprise content management convey the benefits of knowing what you have and being able to find and use it efficiently[25][26]. DOI names for publishers provide improved discoverability, longer shelf life for access, and linking to related offerings[27]. DOI names for citations improve the ability to create cross-links in the publishing production process[28][29].

The initial simple implementation of DOI names as persistent names linked to redirection continues to grow, with over 20 million assigned by the end of 2005 from several hundred organisations through a number of Registration Agencies in USA, Europe, and Australasia, supporting large scale business uses. Implementations of more sophisticated applications (offering associated services) have been developing well but on a smaller scale. A number of issues remain to be solved: these are no longer technical in nature, but more concerned with perception and outreach to other communities. Persistent, actionable identifiers with a fully managed sustainable infrastructure are not appropriate for every activity; but they are suitable for many, and where they are used, the key to providing a successful and widely adopted system is encouraging economy of scale (and so, where possible, convergence with other related efforts), flexibility of use, and a low barrier to use.

A common mistake is to compare DOI and/or Handle to 'other web identifiers' such as URI: it is a false comparison. The World Wide Web is a communication medium and a highly successful one; it is not an information management system (for example, it hasn't made databases obsolete). The DOI System, especially as it has evolved, has much more in common with an information management system or inventory system or distributed database than it does with web publishing. It is easy to misunderstand this because what is primarily managed at the moment is the web publishing aspect (if you want to get that article on the web, which will naturally involve using web protocols, go here). But the goal is to provide a management framework for the identified entities. Making them available on the web will naturally involve using web tools and protocols, which is unsurprisingly what is now happening.

1.6 Deployment of naming mechanisms

In any new functionality offering, software distribution is an issue. Any new technology has to face the issue of the "calcification" of the current deployed base technology; the fact that existing technologies are in use and will not be easily displaced, and the difficulty of deployment to potential users of new software necessary to take advantage of better technologies. As noted above, URI (URL+URN) specifications are agnostic as to resolution method (e.g. a URN implementation

could be based on the handle system); yet in practice on the Web, the deployment is of a DNS-based implementation of the specifications.

Handle clients can be embedded in end user software (e.g. a web browser) or in server software (e.g. a web server). The choice is one of embedding functionality in individual clients (which puts it closer to the end user, and simplifies the architecture, but means that you have to deploy and maintain the software using plug-ins etc.) versus simpler maintenance of a centralized piece of middleware (which means that the users must all then talk to that middleware). Handle client software libraries in both C and Java are freely available.

CNRI runs a proxy server system, a collection of web servers that understands the handle protocol and knows how to talk to the Handle System[30]. Many implementations of the Handle System intended to help manage web content use handles embedded in URLs on web pages, and for the convenience of their customers, use the proxy server (or a similar implementation) for resolution. A growing ecology of other tools for handles is developing, both from CNRI and from outside parties: examples include integration into next generation technologies such as GRIDs[31]; "Sente" a Mac OS X application that incorporates the handle resolver for any handles and DOI names that it finds[32]; discussions to get a built-in handle client for Acrobat 9 (requirements gathering is about to start); a demonstration plug-in is already available for Adobe Acrobat which embeds native handle functionality into links within PDF documents; several DOI-specific tools are available (see below); etc.

The DOI system deals with the problem of software distribution by making DOI names usable in both native protocols or by a common proxy: many implementations of the DOI system intended to help manage web content use DOI names embedded in URLs on web pages, and for the convenience of their customers, use a proxy server implementation (dx.doi.org) which has the functionality of the general Handle System proxy but may have additional functionality added in the future. A DOI name takes the form of a URL when the proxy is involved (e.g. *doi: 10.1234/abcd* becomes *http://dx.doi.org/10.1234/abcd*) but this resulting URL will never change even if the actual content location changes. A growing number of specific DOI tools are becoming

available[33]: some deploy native handle resolution, whilst others make use of proxies easier - e.g. Connotea, a free online reference management and social bookmarking service[34], recognises and stores DOI names, enabling bookmarking a DOI name directly in web browsers.

1.7   Names and internet governance issues

The Handle RFCs contain an IESG Note: "Several groups within the IETF and IRTF have discussed the Handle System and its relationship to existing systems of identifiers. The IESG wishes to point out that these discussions have not resulted in IETF consensus on the described Handle System, nor on how it might fit into the IETF architecture for identifiers. Though there has been discussion of handles as a form of URI, specifically as a URN, these documents describe an alternate view of how namespaces and identifiers might work on the Internet and include characterizations of existing systems which may not match the IETF consensus view".

Internet naming standards do not yet specify a satisfactory approach for naming objects consistently[35]. Handles, for example, are capable of being used in any specification that is finally be endorsed. Until a clear consensus is reached in the internet communities on which approach is to be preferred, handle applications remain agnostic as to formal registration as a generic scheme such as URI or URN, but useable and widely implemented for millions of objects. Ongoing debates about the nature of URIs, URNs, and URLs (which sometimes approach the character of religious wars and have been ongoing for over ten years) and the references to an undefined "IETF architecture for identifiers" suggest that improved standards of clarity and process (e.g., what is the consensus?) would be beneficial to any development which, like the DOI system, attempts to build constructively on existing infrastructure.

There is a danger that the current dominance in internet governance and, perhaps more importantly, in internet funding, of organisations reliant on one naming mechanism, domain naming (a mechanism which makes it particularly difficult to identify digital content independent of location and at appropriate levels of administrative granularity) may be problematic in introducing complementary alternative naming

mechanisms[36]. That some internet applications (e.g. peer-to-peer, or multiplayer games) do not rely on DNS demonstrates that DNS cannot be a necessary required component of any future development (for example, P2P, at its core, does not use DNS. There are probably entry point web sites for most services, some of which may be obvious and some of which may not, but e.g. your Skype identity is not based on a domain name, and that's not how Skype finds you; and in fact such systems are designed so that users never even see a recipients IP address). The internet is not DNS but the global information system that is logically linked by a globally unique address space and communications using the Transmission Control Protocol/Internet Protocol (TCP/IP) suites and provides high level services layered on these (or successors)[37]. The Domain Name System and its disputes as to governance through recent WSIS summits have overshadowed the real issues of efficient naming here; DNS is receding in real importance at the same time that governance issues increasingly look at DNS as the thing to govern. There is a danger that the attempt to control uses of internet protocols too closely may impede the evolution of new technologies and even endanger the nature of the internet[38][39].

2.   MEANING: SEMANTIC INTEROPERABILITY

How does one computer system know what the terms from another computer system mean? If A says "owner" and B says "owner", are they referring to the same thing? If A says "released" and B says "disseminated", do they mean different things? Are two identifiers from different schemes actually denoting the same referent[40]? To answer such questions, terms (and relationships between terms) must be logically defined; if two terms are from different sources, a logical agreed definition for mapping must be bilaterally agreed, so that A or B (or anyone else) can make use of one another's metadata with confidence and in a highly automated way[41]. Clear declaration and definition of identifiers associated with an object is required even if "what it is" appears to be apparent (e.g. obtained by resolution from a found identifier, or from a physical object in hand). A given instance of an object will encapsulate several related identifiers of different entities inherent in the intellectual property it represents, any of which might be exemplified in the object. For example, a pdf text file may embody an abstract "work"; a particular publication edition of that work; and

a format of that edition. Incorrect assumptions about the referent of the identifier will lead to error (fig 1). An identifier does not necessarily resolve to its referent, but may often resolve to something we understand to represent it as part of the compound object (akin to the literary figures of speech metonymy: the use of a word referring to an element or attribute of something to mean the thing itself, as in "the kettle is boiling", and synecdoche: allusion to the part used to imply the whole).
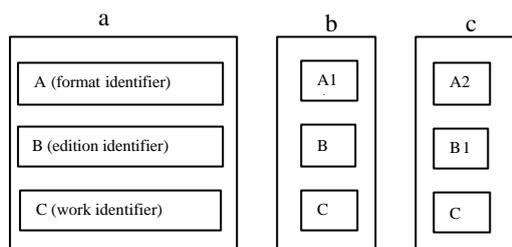
a   b   c

| A (format identifier) |
| B (edition identifier) |
| C (work identifier) |

| A1 |
| B |
| C |

| A2 |
| B1 |
| C |

*Fig 1 The pdf file **a** is obtained by resolution of found identifier **A**. Assuming (incorrectly) that A denotes the work will not distinguish it from **b** (a different format) or **c** (a different published edition).*

The meaning of A (in this case, an identifier of the pdf format) must be declared precisely

Most objects of interest in transactions have this compound form, simultaneously embodying several referents: a digital object may be seen as both a simple "bag of bits" (a string of 1s and 0s with a unique identifier), and also as a higher order structure such as a pdf representation of a document; just as a physical object may be seen as both a set of atoms and as a discrete physical object - a wedge for example. The most useful description of the behaviours of the object may be in terms of the higher order structure (the physics of a wedge, versus the interaction of many atoms), without implying that the lower level description is incorrect ("downward causation"). In intellectual property transactions, the higher order structure may be an abstraction, manifested in some digital or physical form, which makes the interplay of identified entities even more complex than in description of straightforward physical systems.

The only way of unambiguously deciding if one term means the same as another, irrespective of what it is called, is by sharing a single frame of reference: a structured ontology (an explicit formal specification of how to represent the entities that are assumed to exist in some area of interest and the

relationships that hold among them) with an underlying model that allows the generation of consistent new relationships, and a method of recording the agreement between the parties whose terms are included in it.

The indecs project[42] considered logical definition for intellectual property entities through a *Model of Making*[43], relating the various types of creations which are the intellectual content of digital media: performances, fixations and abstractions. One phylum of development resulting from the indecs project, a contextual-based ontology approach for creating data dictionaries, is now well established and in practical use in several major applications. Context has a specific meaning in this analysis: "An intersection of *time* and *place*, in which *entities* may play roles". The most highly developed form of this analysis, the Contextual Ontologyx Architecture (COA)[44], is a generic ontology-based metadata framework comprised of a set of defined types of *Entity* and *Attribute*, and the *Relators* which link them within a contextual model structure. In this analysis every entity belongs to at least one of five primary classes: context, time, verb, place or resource. The underlying central ontology that COA builds is called Ontology_X. It is a proprietary data model, with origins in the development of the indecs metadata framework. It may be expressed in e.g. OWL (web ontology language) for use in Semantic Web applications.

2.1   The origins: indecs

The indecs project developed an analysis of the requirements for metadata for e-commerce in intellectual property in the network environment, and received widespread support. At its heart, indecs proposed a simple generic model of commerce (the "model of making"), paraphrased as: "people make stuff; people use stuff; and (for commerce to take place) people make deals about the stuff". If secure machine-to-machine management of commerce is to be possible, the stuff, the people and the deals must all be securely identified and described in standardised ways that machines can interpret and use. This metadata is crucial to all e-commerce, but is particularly relevant to commerce in intellectual property where the goods being traded are intangible rights rather than tangible goods. With the increasing granularity of the intellectual property being traded, metadata is never likely to come from a single source or to follow a single standard for identification and

description. If metadata from different sources is to interoperate successfully, it must, though, be developed within a coherent and consistent view of the things that are being described so that such views can be successfully mapped to others.

Central to the analysis is the assumption that it is possible to produce a generic mechanism to handle complex metadata for all different types of intellectual property. So, for example, instead of treating sound carriers, books, videos and photographs as fundamentally different things with different (if similar) characteristics, they are all recognised as creations with different values of the same higher-level attributes, whose metadata can be supported in a common environment. Metadata in the network environment needs to support interoperability of at least five different types: across

- media (such as books, serials, audio, audiovisual, software, abstract works, visual material).
- functions (such as cataloguing, discovery, workflow and rights management).
- levels of metadata (from simple to complex).
- semantic barriers.
- linguistic barriers.

The indecs framework showed how such interoperability could be achieved. Several principles were stated that have proved to be key to the management of identification:

- *Unique Identification*: every entity should be uniquely identified within an identified namespace.
- *Functional Granularity*: it should be possible to identify an entity whenever it needs to be distinguished.
- *Designated Authority*: the author of an item of metadata should be securely identified.
- *Appropriate Access*: everyone requires access to the metadata on which they depend, and privacy and confidentiality for their own metadata from those who are not dependent on it.
- *A definition of metadata*: An item of metadata is a relationship that someone claims to exist between two entities (a recent suggested nuance improvement is "to exist between two referents", where a referent is the thing that is identified by an identifier). [This is simply the common definition of metadata as "data about data" made more precise:

"data A is about data B" = "I claim that there is a relationship between A and B"; in computerised systems A and B must be denoted by a referent].

This set of principles provides a concise paraphrase of much of the indecs framework. It stresses the significance of relationships, which lie at the heart of the indecs analysis. It underlines the importance of unique identification of all entities (since otherwise expressing relationships between them is of little practical utility). Finally, it raises the question of authority: the identification of the person making the claim is as significant as the identification of any other entity.

2.2 Methodology for contextual data dictionaries

Logical definition of terms in an ontology-based dictionary allows relationships to be denoted: terms can be "mapped" to other terms[45]. The indecs framework was developed further, with the specific aim of responding to one such mapping requirement, the MPEG-21 Rights Data Dictionary[46], which used an early version of the COA. Whilst "crosswalks" can be constructed to compare terms in any two schemes, the total number of such crosswalks grows much faster as the number of schemes grows linearly (N schemes require (N/2)(N-1) mappings). The existence of one dictionary "hub" reduces this to N mappings, one for each scheme. Bilateral agreement between dictionary and scheme ensure that the existence of agreed mapped terms enables extensibility - mapping to another scheme - without reference to the originators of each scheme. Such mappings will increasingly be computable and thus automated. Mapping through a hub only works if the hub is sufficiently rich (one to one mappings are preferable to mapping through an inappropriate hub), as in the COA. Ontologies are the key to semantic automation[47].

2.3 Applications

The subsequent development of COA has been advanced by, and has in turn influenced, a number of other metadata and identifier schemes and projects, including the bibliographic initiative FRBR[48], ABC-Harmony[49], and others detailed below. The COA also has a good deal in common with the museum/archive community's CIDOC Content Reference Model (CRM)[50], although the development of the two has been entirely

independent. Using one technology initiative applied to a range of problems generates both (economies of scale for implementation and critical mass of support[51]. Some initiatives adopting a standards-based ontology approach are:

- The MPEG-21 Data Dictionary (ISO/IEC 21000-6). The Moving Picture Experts Group (MPEG)[52] is best known for compression standards for audio; MPEG now includes the MPEG-21 "Multimedia Framework", which includes several components of digital rights management technology standardisation. One component is a Rights Data Dictionary[53] to support activities such as the MPEG Rights Expression Language. This is supported by an ontology based on the COA.
- ISO TC46/SC9 identifier interoperability. Work is underway on the practical implications of interoperability across the family of ISO TC46/SC9 identifiers (better known as the ISBN and related identifiers)[54]; this will need a registry for metadata semantics for all its content identifiers, which would add considerable value. The COA approach, such as the ISO/IEC 21000-6 dictionary, could support typed links between identifiers (relators), a schema to facilitate interoperability between reference descriptive metadata sets (hub and spoke mapping), and the development of a taxonomically structured glossary including reference descriptive metadata to support the development of all TC46 SC9 standards.
- Digital Object Identifier (see 1.5 above) metadata elements are mapped through a Data Dictionary that is supported by an ontology based on the COA and includes as a subset the ISO MPEG 21 Rights Data Dictionary ISO/IEC 21000-6
- The ONIX family of standards is an ontology-based and well-accepted tool for electronic commerce in the book and serials sectors. ONIX. EDItEUR[55], the body co-ordinating the development, promotion and implementation of ONIX, is now developing further standards for licensing and multimedia, both of which require a rich semantic interoperability, including ONIX for Licensing Terms, building on earlier

joint EDItEUR/NISO work and the work of the Digital Libraries Federation's Electronic Resource Management Initiative (ERMI)[56].
- DDEX, the Digital Data Exchange (building on the earlier Music Industry Integrated Identifiers Project (MI3P) is developing and maintaining a robust framework of data exchange Standards for the exchange of information relating to digital media content with the initial focus on music and music-related assets. DDEX is expected to release some standard XML message formats by the end of 2006. At foundation of these messages will be a data dictionary which defines elements to be used in the message standards and other framework components and is supported by an ontology based on the COA. In addition two other key standards have been developed in related activities. IFPI (a trade organisation for the international recording industry has developed the Global Release Identifier Standard (GRid) to identify electronic releases that might embody sound recordings, music videos and other digital content and CISAC (a trade organisation for collective rights management organisations worldwide) has developed the Musical Work Licence Identifier Standard (MWLI) designed to identify licences issued in respect of the musical works contained within those Releases.

2.4   Metadata as the "Lifeblood of e-commerce"

The key to potential use of identified entities in meaningful management of intellectual property in digital media is not the assignment of an identifier per se, which is relatively simple, but the definition and assignment of corresponding metadata which defines what that identifier references: without this secure binding, any use of the identifier itself will be fragile. Metadata is the lifeblood of e-Commerce; especially in "Web 2.0", the evolving ecosystem of networked enterprises that use the web as a platform, such as Blogger, Flickr, YouTube, GoogleMaps, del.icio.us, etc. eBay works in part because participants are willing to invest the effort to virtually "wrap" their items in metadata. But we can't find and use entities in such services in any sense that is reliable and useful enough

for digital re-use: if e.g. I want to transact the rights to use a photo on Flickr, I can't do it. Every posting there could /potentially/ be "wired" into a service (or services) whereby I could (a) obtain descriptive information, (b) obtain basic rights information, and (c) initiate a rights transaction (including an open-ended or "free text" request). But they aren't: it awaits throwing the requester and owner into a transactional context, based on associated metadata, that facilitates a dialogue and ultimately a transaction. A potential alternative has been described[57]: when a photo is posted, it might have a tag associated with it that indicates that it is transactable. There would be a third-party service (or more likely several, from amateur grade to professional) driven by the metadata persisted in Flickr, and additional metadata that enables the transaction. The unique identifier is the crucial link: that "point in space" that aggregates the critical metadata for this object. When the object is first "put out there" that point is established or born; as services are added, that point is embellished or "decorated."

2.5 Semantic interoperability and language

English is the operational language of the dictionaries built on the basis of indecs to date. It would however be possible to map a schema in another language into the dictionary. A set of terms in another language would be mapped just as with any other namespace: through bilaterally, mutually agreed mapping of terms into the underlying ontology - so that "what I call *book* is what you call *livre*" is not a simple convention but an agreed analytic ontology definition. Mapping into Chinese would be possible (though this would raise the usual character encoding issues, and the social interaction difficulty of agreeing a mapping might be significant). Mapped operational definitions are just that: they do not imply more than the pragmatic operational use of the term in the ontology context, and so they avoid, for that particular purpose, the philosophical problem of "indeterminacy of translation"[58] (pithily put by Wittgenstein: "If a lion could talk, we would not understand him"[59]). Nor, for the same reason, do terms necessarily correspond with their use in a legal namespace, though they may do in a namespace that considers this purpose.

2.6 Meaning, digital policy enforcement and governance issues

There is a need for digital rights management infrastructure, as a tool for content management (both commercial and non-commercial). But digital rights management, even in the limited context of the management of "content" on the network, has at least four different components[60]:

- A "policy metadata" layer, which allows for the structured description of policies – what permissions relate to this item of content, under what conditions of use (for example, attribution, period of use, payment), and what is not permitted (for example, adaptation);
- An "authentication, authorisation and access" layer – which allows for the structured identification and authorisation of different users (or classes of users) and the matching of their privileges with the permissions relating to content;
- An "enforcement" layer, which is the technology most commonly associated with the acronym "DRM" – the technology which allows policies relating to content to be enforced even after content has been released from a controlled local network into the (uncontrolled) global network;
- An "audit" layer, which allows activities to be recorded and compliance with policies to be monitored.

Mechanisms which would allow these layers to be created – such as "structured identification and authorisation of different users (or classes of users)" – have application far beyond content protection, and identification of users and licences raises issues of privacy and governance. The "rights" that we should manage in the network are not simply therefore those of traditional content management (such as copyright enforcement, as seen in the recent music and motion picture industry concerns over piracy). The same layers apply also to rights of civil society: personal and collective rights to privacy and protection from fraud and other crime. In the absence of a common trusted infrastructure, the future potential benefits of the global network will be increasingly curtailed[61] [62]. One of the practical elements of a trusted infrastructure is the structured description of entities, allowing the analysis of meaning. The governance issues

around the concepts of these technical means of interoperable metadata as a vocabulary for intellectual property rights are significant, since any formal analysis of meaning is underpinned by the question of "who says": who has the right to authorise semantic mappings and to undertake analyses; who is *allowed* to say.

3. CONCLUSIONS

1. We should facilitate diversity through evolution of the internet while retaining continuity with what has gone before (interoperability), and beware of calcification (introducing measures which limit change and "generativity"[63] e.g. through firewall design). "The Internet has now been around long enough that it is easy to take it as given, rather than question why it is the way it is. In fact, the Internet was invented only about 30 years ago by some designers exploring a very different way of conceiving what a communication network should be. It was very much a clean-slate design, and while the Internet has been wildly successful, there is no reason to think we got it exactly right on the first try. In fact, it is pretty clear that we did not get it right in all respects" (D. Clark, op.cit).

2. The dominance in internet governance and funding of organisations reliant on one naming mechanism, domain naming, may inhibit complementary alternative naming mechanisms. In 2006, domain names passed the 100 million mark[64]. The monetisation of this system (with typical registration fees of several tens of dollars per year per domain name) is substantial. That can be an incentive to investment, but can also become an inhibitor of change if support of it becomes the status quo. In addition to the large for-profit registrar entities (notably Verisign with 49.7m .com and 7.3m .net registrations[65]) it is notable that ISOC (a body whose "principal purpose is to maintain and extend the development and availability of the Internet and its associated technologies and applications"[66]) receives the bulk of its income from .org registrations (in 2006, there were 4.4m such domains). Although direct comparisons are difficult, alternative mechanisms offering naming advantages are already known to cost one or two orders of magnitude less.

3. There is a need for a trusted infrastructure for content management, personal and collective rights to privacy, protection from fraud etc and other crime. One of the practical elements of a trusted infrastructure is the structured description of entities, allowing the analysis of meaning. In developing applications such as content licensing it has become apparent that there is a need to identify entities of all forms. These include abstract and physical as well as digital entities; they include content, parties and agreements; and they need to consider relative identity (contextual resolution). As applications become more sophisticated, objects may be representations of people, resources, licences, avatars, sensors, etc., and machine to machine interaction is likely to predominate (where domain names may not be the most useful mechanism).

4. That identifiers can be placed in a web context through expression as DNS-based URIs (or as some would say, URNs) is excellent. To claim that identifiers *must* be expressed in these forms is short-sighted. The web is not the net: examples of identity in a non-web context include Skype identities; examples of identity in a non-net context include resolution of GS1 barcodes, RFIDs, etc – these may use the DNS, but should not have to, if they may also run compatibly on the base level TCP/IP.

5. Organisations promoting new evolutionary conjectures often face the problem of "herding cats" in developing an appropriate social infrastructure. Nor is there any guarantee that success will be to the best technical design. Any new technical solution may need to go to great lengths to ensure continuity and interoperability (http proxies, DNS clients, etc) yet still faces the problem of technical deployment and more importantly mindshare; it is salutary to note that IP (internet protocol) took some 15 years to become established over proposed (at the time) competitors (D. Clark op cit).

6. There are unmet needs in the current internet infrastructure - notably, international language use in naming and appropriate granularity of naming; and precise definition of meaning - which may have more elegant solutions than by building solely on workarounds of the current limitations.

Note

Handle System, Handle.net and Global Handle Registry are CNRI trademarks registered in the U.S. Patent and Trademark Office. DOI® and DOI.ORG® are registered trademarks and the doi> logo is a trademark of The International DOI Foundation.

3. REFERENCES

[1] N. Paskin, "Components of DRM Systems: Identification and Metadata": Norman Paskin, in E. Becker et al (eds), "Digital Rights Management: Technological, Economic, Legal and Political Aspects in the European Union" in the series Lecture Notes in Computer Science (Springer-Verlag, 2003), pp 26-61. http://www.doi.org/topics/drm_paskin_200301 13_b1.pdf

[2] J. Evers, "Will Vista stall Net traffic?" CNET News.com (2006) http://www.zdnetasia.com/news/software/0,39 044164,39433774,00.htm

[3] E. Dyson, "Online registries: The DNS and Beyond…" Release 1.0, Vol 21 No 8 (2003) EDventure Holdings, New York www.edventure.com

[4] Corporation for National Research Initiatives web site: http://www.cnri.reston.va.us/

[5] R. Kahn & R. Wilensky, "A framework for distributed digital object services": Technical Report tn95-01, Corporation for National Research Initiatives, 1995 http://www.cnri.reston.va.us/k-w.html (re-published, with an additional introduction by the authors, as the keynote paper in a special issue on "Complex Digital Objects": International Journal on Digital Libraries (2006) 6(2): 115-123. doi:10.1007/s00799-005-0128-x

[6] Handle System web site: www.handle.net

[7] S. Sun, L. Lannom, B. Boesch, "Handle System Overview". Internet Engineering Task Force (IETF) Request for Comments (RFC), RFC 3650, November 2003 ftp://ftp.rfc-editor.org/in-notes/pdfrfc/rfc3650.txt.pdf

[8] S. Sam; S. Reilly, L. Lannom, "Handle System Namespace and Service Definition". Internet Engineering Task Force (IETF) Request for Comments (RFC), RFC 3651, November 2003 ftp://ftp.rfc-editor.org/in-notes/pdfrfc/rfc3651.txt.pdf

[9] S. Sun, S. Reilly, L. Lannom, J. Petrone, "Handle System Protocol (ver 2.1) Specification". Internet Engineering Task Force (IETF) Request for Comments (RFC), RFC 3652, November 2003. ftp://ftp.rfc-editor.org/in-notes/pdfrfc/rfc3652.txt.pdf

[10] Dept of Defense Instruction 1322.26, "Development, Management, and Delivery of Distributed Learning", section 6.7 (2006) http://www.dtic.mil/whs/directives/corres/html /132226.htm

[11] Australian Research Repositories Online to the World (ARROW) project, http://www.arrow.edu.au/about/

[12] McCarthy, Kieren: "Divided by a common language" http://technology.guardian.co.uk/weekly/story/ 0,,1830481,00.html

[13] Global Internet Statistics (by Language) http://global-reach.biz/globstats/index.php3 (2006)

[14] M. Wei, S. Sun, F. Wang (2003): Introduction to the Handle System. http://www.cnnic.net.cn/download/2003/11/27/ 173300.pdf

[15] Online course on the Handle System, Shanghai Science and Technology Center (2001): http://www.stcsm.gov.cn/learning/lesson/xinxi/ 20010613-1/20010613.asp

[16] Handle System Specification (based on Handle System RFC3650, RFC5651, RFC3652) (from the Chinese Digital Library project) http://cdls.nstl.gov.cn/mt/blogs/2nd/archives/d ocs/%ca%fd%d7%d6%d7%ca%d4%b4%ce%a 8%d2%bb%b1%ea%ca%b6%b7%fb%bd%e2 %ce%f6%cf%b5%cd%b3%d3%a6%d3%c3% b9%e6%b7%b6.pdf

[17] S. Sun, Experiences Securing DNS through the Handle System http://middleware.internet2.edu/pki06/proceedi ngs/ . (2006)

[18] Digital Object Identifier System: http://www.doi.org.

[19] J. F. Sowa, "Knowledge Representation: Logical, Philosophical and Computational Foundations. " Brooks/Cole, Pacific Grove, CA, 2000. http://www.jfsowa.com/krbook/index.htm

[20] G. Rust and M. Bide, The <indecs> Metadata Framework: Principles, model and data dictionary (2000): http://www.indecs.org/pdf/framework.pdf

[21] N. Paskin, "DOI: a 2003 Progress Report" D-Lib Magazine, Vol. 9, No. 6, June 2003. doi:10.1045/june2003-paskin

[22] ISO/WD 26324: Information and documentation — Digital Object Identifier (DOI) system. http://www.collectionscanada.ca/iso/tc46sc9/

[23] N. Paskin, "Digital Object Identifiers for scientific data" Data Science Journal, Vol. 4, pp12-20, 28 April 2005. http://www.doi.org/topics/050428CODATAart icleDSJ.pdf (also available at www.datasciencejournal.org)

[24] N. Paskin, "Digital Object Identifiers and Digital Preservation of the Record of Science": ICSTI Symposium on Digital Preservation of the Record of Science, February 2002; IOS Press, 2002 http://www.doi.org/topics/020210_CSTI.pdf

[25] B. Rosenblatt, "Enterprise Content Integration with the Digital Object Identifier: A Business Case for Information Publishers", June 2002. http://doi.contentdirections.com/mr/cdi.jsp?doi=10.1220/whitepaper5

[26] S. Sieck, "Economic Benefits of Digital Object Identifier Applications in Content Marketing. Part I. Using the DOI to Improve Profitability In Publishers' E-Commerce Operations", EPS White Paper, 16 December 2003. http://dx.doi.org/10.1220/eps1

[27] S. Sieck, "Economic Benefits of Digital Object Identifier Applications in Content Marketing: II. Using the DOI to Improve Profitability in Content Distribution", EPS White Paper, 20 July 2004. http://dx.doi.org/10.1220/eps2

[28] N. Paskin, "E-citations: actionable identifiers and scholarly referencing". Learned Publishing, Volume 13, Number 3, July 2000. doi:10.1087/09531510050145308

[29] Crossref: "Info for publishers": http://www.crossref.org/02publishers/index.html

[30] Handle System Proxy Server System: http://www.handle.net/proxy.html

[31] The Globus Alliance: Globus Toolkit-Handle System project (2006): http://dev.globus.org/wiki/Incubator/gt-hs

[32] Third Street Software: "Sente", http://www.thirdstreetsoftware.com

[33] International DOI Foundation, "DOI System Tools", http://www.doi.org/tools.html

[34] B. Lund et al., "Social Bookmarking Tools (II) A Case Study – Connotea" D-Lib magazine (2005) doi:10.1045/april2005-lund

[35] International DOI Foundation Factsheet: "DOI and Internet Identifier Specifications" http://www.doi.org/factsheets/DOIIdentifierSpecs.html

[36] N. Paskin, "Naming and meaning: requirements for internet progress". Internet Governance Forum, Oxford Internet Institute, 2006.

[37] R. Kahn, V. G. Cerf, "What is the Internet (And What makes it Work)?". Internet Policy Institute, December 1999. http://www.cnri.reston.va.us/what_is_internet.html

[38] K.N. Cukier, "The Internet™: A solution for openness through closedness". Internet Governance Forum, Oxford Internet Institute, 2006.

[39] D. Clark, "Why the Internet is the way it is (and why it will be very different in ten years)" http://webcast.oii.ox.ac.uk/?view=Webcast&ID=20060428_143

[40] N. Paskin, "On making and identifying a "copy"". D-Lib Magazine, Jan 2003 doi: 10.1045/january2003-paskin

[41] N. Paskin, "Towards a Rights Data Dictionary - Identifiers and Semantics at work on the net." imi insights, June 2002 http://www.epsltd.com/

[42] <indecs> project home page: http://www.indecs.org.

[43] G. Rust "The Model of Making in indecs and RDD" ISO/IEC JTC1/SC29/WG11, document M12159

[44] The detailed documentation of the COA is not yet published. For further information see www.rightscom.com

[45] International DOI Foundation Factsheet: "DOI and Data Dictionaries" http://www.doi.org/factsheets/DOIDataDictionaries.html

[46] CONTECS:DD Press Release, 3 June 2004: "ISO Publication of the MPEG Rights Data Dictionary Standard, ISO/IEC 21000-6": http://www.doi.org/news/contecsdd_standard_pr.html

[47] I. Horrocks, "Ontologies and the Semantic Web" http://www.epsg.org.uk/pub/needham2005/.

[48] IFLA (1998): "Functional Requirements for Bibliographic Records", IFLA Study Group on the Functional Requirements for Bibliographic Records, 1998 http://www.ifla.org/VII/s13/frbr/frbr.htm

[49] ABC Harmony Data Model Version 2 http://www.metadata.net/harmony/ABCV2.htm

[50] The CIDOC Conceptual Reference Model http://cidoc.ics.forth.gr/

[51] M. Bide, "Rights metadata: Standards at last?" Imi insights, May 2005, pp.1-4 http://www.epsltd.com/

[52] MPEG Home page: http://www.chiariglione.org/mpeg/.

[53] ISO/IEC Information technology - Multimedia framework (MPEG-21) - Part 6: Rights Data Dictionary, ISO/IEC 21000-6. http://iso21000-6.net/

[54] N. Paskin, "Identifier Interoperability: A Report on Two Recent ISO Activities" D-Lib Magazine, April 2006 doi:10.1045/april2006-paskin

[55] Editeur Home Page: http://www.editeur.org.

[56] DLF Electronic Resource Management Initiative, http://www.diglib.org/standards/dlf-erm02.htm.

[57] J. Erickson: "Handle Records, Rights and Long Tail Economies" D Lib magazine, September 2006. doi:10.1045/september2006-erickson

[58] W. V.O. Quine, "Word and Object" (1960); "Ontological Relativism" (1977).

[59] L. Wittgenstein, **"**Philosophical Investigations", II, xi, p. 223 (1953)

[60] M. Bide, "Digital Rights Management: preventing or enabling access?" Serials 17:2 (2004)
http://uksg.metapress.com/(f14cboawwef5mpuwk4qkxbbq)/app/home/contribution.asp?referrer=parent&backto=issue,7,17;journal,6,34;linkingpublicationresults,1:107730,1

[61] J. Hoff & F. Bjerke, "Fences and gates in cyberspace: Is the Internet becoming a threat to democracy?" Information Policy 10 (2005) 141-151 IOS Press (2005)

[62] P. Woodward, "Challenges of the Internet Dream in Asia: Freedom compromised?". imi, Sept 2006, pp1-4
http://www.epsltd.com/accessArticles.asp?articleType=2&articleID=393&imiID=84

[63] J. Zittrain, "The Generative Internet"
http://papers.ssrn.com/sol3/papers.cfm?abstract_id=847124

[64] Verisign Inc, "Verisign Domain Name Industry Brief", August 2006:
http://www.verisign.com/static/039111.pdf

[65] M. Zook, "History of gTLD domain name growth"
http://www.zooknic.com/Domains/counts.html

[66] Internet Society: Purposes and Goals of the Organization.
http://www.isoc.org/isoc/mission/principles/