ESTHER DYSON'S MONTHLY REPORT

Release 1.0®

VOLUME 21, NO. 8 | 16 SEPTEMBER 2003 | www.edventure.com

Online Registries: The DNS and Beyond...

INSIDE

BY ESTHER DYSON

As the world grows more connected and more complicated, we all need ways of defining, identifying and keeping track of things and cross-referencing them with their owners. The simplest way to do that is with registries – everything from the Domesday Book, a medieval registry of land, property and people; to current-day auto registries on the one hand and the worldwide Domain Name System on the other. In today's computer world, registries are gaining visibility, even though in some form they have been around for a long time (cf. Novell's Directory, Microsoft's Active Directory, and even IBM's old mainframe catalogue).

But now, companies and organizations have to keep track of ever more things and people, not just inside their walls but across extended organizational boundaries. Call this new wrinkle an "external registry." Finally, they may want to interact with things and people, rather than just look them up, via an "active registry." The DNS is the best-known example of an active, external registry. Its technical function, automatically translating the logical address of a URL into a (semi-) physical IP address, supports the Net's rapid growth while the mnemonic character of URLs facilitates the e-mail and Web connections among businesses and people that have driven that stunning growth.

Several new online registries are positioning themselves to play analogous roles in the growth of specific new capabilities on the Net. The implementation of ENUM, a new standard for turning phone numbers into domain names, could vastly expand the use of new ways to connect people such as instant messaging, voice over IP (VoIP) or video conferencing services, and promises to enable a more real-time Internet. Other registries such as the Handle System

{ *continued on page 2* }

(PAGE 22) and AutoID's Object Name System (ONS) (PAGE 12) will help enable new product and service offerings such as content purchases or product-tracking updates through indirect addressing that references data and functions rather than servers and locations. The kind of growth the DNS supported for the Internet as a whole may be fostered in other marketplaces through registries that enable interconnection among people and applications with functions specific to particular industries.

But these capabilities raise hoary questions about information ownership, access and modification control, version management. . .and the neutrality of the registries in the markets they serve. In short, registries can't simply be inserted into the infrastructure and work (even though the software itself is simple enough). Registry systems need corresponding infrastructure: standards, federated identity management (SEE **RELEASE 1.0**, JUNE 2002), business "protocols" and policies, and sometimes even political buy-in or active regulation. Lack of appropriate governance can foment pointless and costly disputes that benefit almost no one other than lawyers, along with meretricious business models based on artificial economics.

Yet most of the governance questions about active, external registries are still unresolved. The "experiment" of the Internet Corporation for Assigned Names and Numbers' (ICANN) oversight of the DNS is still not conclusive, but ENUM and to some extent ONS rely on the DNS. Meanwhile, the Handle System and its "digital object architecture," developed by Bob Kahn (who also has much of the Internet's architecture to his credit), offers both better technology and "lessons-learned" governance. . .but unfortunately it lacks the visibility of the DNS, which is an order of magnitude larger in the number of things registered (approaching 200 million vs. less than one tenth of that). In this issue we start with a look at ICANN and the Domain Name System, and use it as a lens through which to assess the promise and challenges of other registries.

Registering the concept

Registries come in many flavors, and the word loosely covers many kinds of things that don't qualify for the term as we use it here. For

Release 1.0® (ISSN 1047-935X) is published monthly except for a combined July/August issue by EDventure Holdings Inc., 104 Fifth Avenue, New York, NY 10011-6987; 1 (212) 924-8800; fax, 1 (212) 924-0240; www.edventure.com. It covers the worlds of information technology and the Internet, including wireless communications, security, business models, online services, tracking systems, identity management and other unpredictable topics. . .and the policy issues they raise.

EDITOR: Esther Dyson (edyson@edventure.com)

PUBLISHER: Daphne Kis (daphne@edventure.com)

MANAGING EDITOR: Christina Koukkos (christina@edventure.com)

CONTRIBUTING WRITERS: Eric Dean (ecdean99@hotmail.com), Dan Gillmor (dan@gillmor.com), John Hagel (john@johnhagel.com), JC Herz (jc@joysticknation.com), Clay Shirky (clay@shirky.com), Jeff Ubois (jeff@ubois.com)

CIRCULATION MANAGER: Beckie Jankiewicz (beckie@edventure.com)

SYSTEMS MANAGER: Geoff Clarke (geoff@edventure.com)

CONSULTING EDITOR: Bill Kutik (bill@kutik.com)

our purpose, registries generally provide a cross-reference between an identifier for something – a domain name, an RFID tag number in the case of manufactured goods (SEE **RELEASE 1.0**, JUNE 2003), a digital object identifier (DOI) in the case of books or music – and the thing itself. A registry can point directly to the item registered via a URL or an IP address, or it can point to a proxy for the item, usually controlled by the owner, that can contain a variety of information including the object itself (if it's digital), with access controls and other functionality around it. Likewise, the identifier can be a unique, arbitrary string, or it can be a name, such as a second-level domain name (edventure in edventure.com), or even a full URL (HTTP://WWW.EDVENTURE.COM/RELEASE1/ABSTRACTS.CFM?COUNTER=6805767) – which de facto depends on a named location (a website) rather than an identity. In the case of the Handle System, the identifier is actually part of the digital object (which could in turn refer to a physical item). All these seemingly minor distinctions end up making a large difference.

These registries are an important factor in realizing the dream of standard, open protocols and the Semantic Web or Kahn's digital object architecture, for which communication protocols or even APIs alone are not enough. Interconnections among digital objects are defined more and more by the concrete specifics of the things they model and their business roles, rather than by technical considerations. Registries give Juan and Alice a way to know they are talking about the same thing, and to the right person (with proper identity management).

External registries can support new market dynamics for existing markets or services. For example, until there was a registry outside individual phone companies' internal systems that could cross-reference telephone numbers with telephone-circuit switching information, there was no way for people to keep their phone numbers if they changed service providers, and hence no real open market for local phone service. Likewise, the continued growth of online content delivery, including game-playing services, would benefit from reliable and comprehensive registries of copyrighted materials, combined with digital-rights management software (SEE PAGE 26).

Further, the existence of external registries may be a key enabler for new services and new markets. For example, broad use of RFID tags coupled with externalized registries pointing to product information and practical (human) identity management could enable everything from valuable information about patient-specific drug interactions and less valuable but newly affordable information about body-type-specific shirt and trousers interactions, to better maintenance histories for a range of products. (SEE **RELEASE 1.0**, JULY 2002 AND JUNE 2003.)

Empires and umpires

In addition to technology and standards, registries need viable business models to recoup the costs of creating and maintaining the registry. Such models need not be for-profit, or even paid for by the registrants who own or control the objects registered. Often, registries are managed by for-profit organizations - what you could call "subset" registries, who manage a portion of an overall registry database, or registrars, who interact directly with the registrants of the objects. (For example, VeriSign's Network Solutions runs the .com subset registry of the overall DNS; it resells domain name registrations through 100-odd ICANN-certified registrars who interact directly with the registrants.)

There also needs to be some source of the registry's authority. Sometimes the authority is clear: US state governments are the authoritative source for automobile licenses by law. Sometimes, the authority is less clear and arises from market success or consensus among interested parties. And sometimes the authority is unclear and there is no consensus, with various parties contending for control, as with the DNS.

The DNS and the phone-number-portability registry comprised their entire universes from the start, while most other registries comprise some set of things that already exist; they need to penetrate their potential installed bases. Nonetheless, aside from the number-portability registry, registries have a way of starting out small. By the time things get to full scale, considerations that seemed like minutiae at the outset, such as the use of real words in DNS identifiers, or whether registrants' contact information is held by a registry or by its registrars, turn out to be major sources of dispute. That's why it behooves the new registries to pay attention to history.

ICANN and the DNS: Unfinished History

The DNS is a model of a working registry, necessary but not sufficient for everything the Internet delivers today and still promises. It comprises around 180 million domain names, with billions of references to URLs from the more than 500 million Internet users brokered daily.

But in aspiring to the DNS's operational success, other registries no doubt want to avoid the governance problems and technical limitations that have beset it.

MEA CULPA

It's hard to write about ICANN objectively, and perhaps I should not even try. However, even as I bear responsibility for many of its deficiencies, the experience taught me things that I hope can be useful for it and other registry oversight bodies. As founding chairman of ICANN, I saw some of the problems discussed below and missed others. I thought I could fix it best from inside, but I wish I had known then what we all have learned since. While ICANN's struggles are specific to ICANN's purpose and history, they also illuminate the challenges other registries can encounter. Here's hoping we can all learn from the mistakes I helped make.

— Esther Dyson

The DNS's design, with meaningful names instead of mere numerical (or arbitrary) identifiers, has led its overseer ICANN into the maws of intellectual property (IP) disputes that in an ideal world wouldn't concern it at all. In addition, users searching for information often use generic terms – such as "books" or "travel" or "chicagoapartments" – as a de facto search term by adding .com to the end, creating marketing value beyond what the DNS strings have as mere identifiers. And finally, the use of names and (website) locations in identifiers leads to problems because they tend to change over time.

The landscape is further complicated by the fact that policies concerning domain names vary across the different TLDs (Top-Level Domains), which partition the DNS and are administered by different subset registry operators. De facto, anyone can register a domain name in most of the generic TLDs (gTLDs) as long as the name is not someone else's trademark, but the administrators of various country-code TLDs, or ccTLDs, and some gTLDs impose additional restrictions on who may hold a domain name and what names they can use.

There is further contention around transparency and privacy of personal data (who owns a domain name, who is responsible for the relevant DNS servers, who can get at that information, etc.) and political quarrels about who's in control, regardless of what they are in control of. Since the DNS is a global function, these issues have implications and complications at both national and international levels. (By contrast, ICANN's technical standards work is managed by the Internet Engineering Task Force (IETF) and the European Telecom Standards Institute and gets little press outside the technical community.)

Learning by example

The disputes are of interest to the few; their outcomes may affect the many. If ICANN were more accommodating, would the DNS make a good technical underpinning for the variety of registries described below? The current spec for ENUM (PAGE 14), for example, explicitly assumes use of the DNS, as do the plans for the

AutoID ONS registry of products (PAGE 12). Before moving forward, let's consider some of the lessons ICANN can offer.

- **IP and the desirability of new generic top-level domains**
  The biggest substantive conflict within the ICANN community – registry operators, registrars, technology providers and users – concerns the desirability of new TLDs, beyond the seven new ones launched in 2000: .aero, .biz, .coop, info, .museum, .name and .pro. Those with an existing stake – both current registrants with IP to protect and current registries and registrars who benefit from a closed market – would mostly prefer to see the number of new TLDs limited, although some registries and registrars would be happy to participate in a growing market. On the other side, newcomers and people left out thus far would generally like to see more names so that they can participate.

  The challenge is that the real shortage is not so much of namespace as of space in people's heads. Putting names into the DNS doesn't create brands; it works the other way around. There are arguments that too many TLDs would strain the DNS and place too great a burden on the root, but there should be some way to expand the system at a moderate pace without the extensive rigmarole currently required. That might result in a fractionated world, but the market would probably do its work: The more reliable registries would grow and the rest would wither. It would be nice for the Internet to be perfect, but the fact is that sections of it go down every day without much overall damage. At worst, the new TLDs proved that they could be added without much fuss other than policy disputes. Technically, they mostly work fine. (SEE PAGE 10.)

  In short, the conflict over new gTLDs and broader more granular disputes over individual domain names, are mostly conflicts over IP and freedom of speech issues – conflicts that should be resolved elsewhere. The challenge is that the obvious "elsewhere" is the World Intellectual Property Organization, a treaty organization that is perceived to represent just one side of the debate. The ICANN community comprises both trademark holders and those who would challenge them – free-speech advocates, "little guys" and those who are simply opposed to the IP interests – so it can't reach a consensus on where to delegate those questions.

  Lesson for other registries: Anything other than first-come/first-served in a registry leads to problems. Registries should provide the tools to enforce policies about intel-

lectual property and other political/commercial disputes, but they should not themselves be involved in deciding the policies. For the DNS, that's IP issues about names. For music registries, that's copyright issues. And for ENUM, this means avoiding sticky issues around personal identity or access to ENUM numbers.

NOTE: Google is facing these same questions of exclusion and access as it becomes more dominant. Precisely because it is not a registry, but rather a service that explicitly resolves names and concepts (rather than identifiers) to content, it will not be able to avoid these questions.

- **Accommodation with the managers of the country-code TLDs**
  One of ICANN's biggest problems has been to get the country-code registry managers to join. . . partly because they were not invited to join and participate in ICANN's decision-making, but rather to [paraphrase] "sign here and agree to follow our rules." With new management and staff, ICANN is now more welcoming. Meanwhile, the cc registries are coming under increasing scrutiny from their own governments, many of whom had previously paid little attention to them. ICANN is an alternative to direct government control, and looks more appealing by comparison. So far, 15 ccTLDs (out of about 250) have come into the fold, and discussions with more are underway.

  Lesson for other registries: Welcome your entire potential market into the governance structure; don't create enemies or second-class citizens.

- **Policies concerning the nature of and access to data in the Whois**
  Briefly, the Whois contains the contact details for individuals associated with domain names, whether individual registrants or employees of registrant organizations. It is a valuable resource for vigilantes and spammers (though perhaps not as much as commonly complained of; there are plenty of other lists around), but also for law-enforcers, consumers checking out websites they're not familiar with, journalists doing research, and others with legitimate and illegitimate interests in such information.

  Right now, much of the information in the Whois is missing, out of date or inaccurate on its face. Currently, registrants who want their privacy are hard to find. . .whether by strangers, or by registrars seeking to renew them. But there's no real policy, and almost everyone is unhappy. It's hard to predict how this will

be resolved short-run, but long-run it will probably turn into something that allows "tiered access": i.e., law enforcement gets to see what it wants, while other kinds of people have limited access on request. That will require some kind of corresponding authentication of the entities making the queries.

Lesson for other registries: Any successful registry will manage valuable or sensitive information. It needs clear policies about access to the information and mechanisms to enforce them, including identity/authentication/authorization systems.

- **Marketing and business practices of registries and registrars**
  In many ways, the privatization of the DNS has led to the realization of critics' worst fears. Yes, second-level domain name prices have dropped (from $35 per year for .com names to $10 or less). Rather than compete on quality or service, most registrars compete on price – and with sleazy marketing tactics. It's easy to register a domain name but very hard to switch registrars. Meanwhile, registrars are constantly trying to steal one another's customers with switching pitches disguised as renewal notices. ICANN is not a consumer-protection agency, and few people think it should be one, but its limited set of rules are generally ignored and registrants have little recourse for mistreatment. ICANN's primary enforcement mechanism is to cancel a registrar's accreditation, but so far, few if any ICANN registrars have lost their accreditation for any reason other than going out of business or failing to pay their fees. Most of ICANN's registry/registrar members would like to see stronger rules enforced, but meanwhile they feel forced to descend to industry practices to compete.

  Lesson for other registries: It's in the registry's and registry operators' own interest to set and visibly enforce reasonable rules of conduct, with a range of sanctions that allow responses proportionate to violations (so they have a chance of actually being imposed when there are violations), or the customers will ultimately lose faith or choose an alternative if an alternative exists.

- **Who are the constituents and how can they have their say?**
  Rather than making its way into a competitive market and proving itself by being the best at what it does, or growing out of a regulated industry such as the telephone business where authority was clear, the DNS started out without much legal identity. When ICANN was created to oversee the DNS, it was supposed to represent the public interest through its board, with a membership

composed primarily of industry participants (including users) who would develop policies by consensus for ratification by the board.

However, ICANN and its board were perceived as a US-run commercial monopoly. Among those three words most participants found at least one to deplore.

One of ICANN's first missions was to break up Network Solutions' monopoly (although it was a legal monopoly, under a US government contract). But ICANN itself also controls a de facto monopoly. Although it could not totally eliminate its competition, ICANN too behaved in the heavy-handed way common to monopolies.

Certainly, ICANN primarily represents commercial interests, especially those of vendors, but also of large IP-holders. Big business and especially trademark-owners have their own representation within ICANN – and the time and funds to send people to its meetings to represent their interests. By contrast, the little guys, whether small businesses, noncommercial organizations or individuals, have little time to spend on ICANN business, so their cause was often pressed by people who have little better to do and lacked credibility with ICANN's board and management. ICANN has now created an At-Large Advisory Committee (ALAC) that is setting up a structure whereby individuals can participate in ICANN's decision-making and board selection. (DISCLOSURE: ESTHER DYSON SITS ON THE ALAC.) Now it is up to individuals themselves and organizations representing them to participate. That's a challenge, because individuals are busy. It's hard to claim early signs are wildly encouraging, but at least the opportunity is now open. Moreover, ICANN's new board itself represents a broader range of interests and includes people who come from outside the traditional ICANN community.

And finally, ICANN still betrays its US heritage. Although only four of its 13 board members are American and its new CEO is Australian, ICANN is a US corporation, based in California. Only 7 of its 20 employees are from outside the US. ICANN's Governmental Advisory Committee (GAC), created to mollify governments who objected to ICANN's independent, non-treaty status (and its US flavor), has a delicately vague role in ICANN's affairs: It can't formally control what ICANN does, but ICANN has to pay attention to it because after all, ICANN's constituents all live or operate somewhere, under some government's authority. ICANN's openness about how this works in practice will have a lot to

do with the intensity of the perception that ICANN is US-run and US-centric (though the GAC itself is even less transparent than ICANN).

Under new management and with a new, restructured board, ICANN is starting to fix some of these problems, but it bears a heavy burden of history.

Lessons for other registries: Avoid looking or behaving like a monopoly. Give your registrars' customers an explicit role in policy-making. And if your market is global, be sure you think and are constituted globally.

New TLDs and New.Net: Is the club worth joining?

Since the establishment of ICANN, only a few new players have been allowed to enter the game of operating a TLD registry, a fact critics cite in suggesting that ICANN is nothing but an old boys' club. While a server or even a website can exist without the DNS, a domain name, for practical purposes, cannot. Thus it's hard for any real competition to exist – either to constrain ICANN or to spur it to improve.

But ICANN is at least creating a little internal competition. As noted, ICANN led a selection process and then accredited operators for seven new generic TLD registries in 2000. The organization still has not delivered the full evaluation of the new TLDs that it promised, but it is moving ahead towards creating a few more new TLDs, picked from the list of unsuccessful bids in the previous round.

Some of these new TLDs are overseen by industry associations such as SITA, the airline industry's network consortium, for .aero, and the Museum Domain Management Association (MuseDoma) for .museum (SEE **RELEASE 1.0**, JANUARY 2003). Afilias, a consortium of registrars, manages .info and also runs the back-end for (but does not manage) .org. Venture-funded Global Name Registry (SEE **RELEASE 1.0**, JULY 2002) manages .name, and publicly traded Register.com manages .pro. Several private companies also hold registry-management contracts for ccTLDs, such as NeuStar (SEE PAGE 20) for the external gateway to China's .cn, and Afilias for .vc (St. Vincent and the Grenadines).

One company in this business that has taken a different path is New.net – but it now may be coming into the tent. New.net was created by Bill Gross's Idealab!, in a story akin to that of the protester who makes money selling protest T-shirts while his pals go to jail. Many people protested against ICANN; few had the initiative to start a new business in opposition to it.

New.net's basic business model is to create its own set of non-ICANN TLDs that can be resolved at its own domain name servers, outside the DNS root. It offers a greater variety of meaningful domain names than ICANN's seven new TLDs, along with a set of ancillary services. All that was needed was for each user's browser, his ISP's local DNS server, or each enterprise's DNS server to redirect the call to New.net's domain name server. About 90 million potential users have installed plug-ins that automatically append "new.net" to the relevant URLs or e-mail addresses. About 90 million more are able to reach New.net TLDs – whether they know it or not – because their ISPs support New.net. Those ISPs include Earthlink, Tiscali, NetZero, Juno, Prodigy and Tutopia, the largest Latin American ISP.

They can communicate with a universe of about 30,000 active SLDs out of about 100,000 registered, in 87 TLDs such as .travel, .shop and .voyage (French for travel). Most of the registrants are small businesspeople who take advantage of ancillary services offered by New.net that raise the visibility of their online businesses and help drive traffic via New.net's Quick! search service and portal. Other consumer-oriented initiatives include .movie, which allows users to type anyname.movie and get directly to the official movie page. Partners who sell New.net TLD registrations include some ICANN-accredited registrars such as BulkRegister, EasySpace and Dodora.

NEW.NET INFO

Headquarters: Pasadena, CA
Founded: May 2000
Employees: 25
Key metric: estimated $10 million in revenues in 2003; profitable
Funding: undisclosed amount from IdeaLab!
URL: www.new.net

It's fair to say that New.net's impact has been more political than economic, yet the company has grown to about $10 million in annual revenues, has been profitable since the middle of 2002 and is now debt-free, says (new) ceo Dan Sheehy.

Meanwhile, New.net has evolved – under new management – from an outsider trying to shout down ICANN, into an outsider trying to get in. Originally, says Sheehy, "We spent a lot of time defining ourselves politically. It was not the smartest way to establish ourselves."

While it has successfully made the economic argument that someone wants what it supplies, it could offer a more attractive product if its TLDs were recognized within the standard root. New.net has had to build and seed the world with redundant infrastructure in order to make its service work and to solve a problem that shouldn't exist (just like an alternate telephone system, though in this case it's limited to software). The whole exercise looks pointless from an economic point of view – unless you believe that the presence of New.net may have moved ICANN to behave a little better in some way.

The irony is that in the process of opening up and awarding a second round of new TLDs, ICANN could damage New.net. Among the new TLDs ICANN may award is .travel, a key offering for New.net. The only potential bidder is Tralliance, a commercial start-up with support from the International Air Transport Association (which has withdrawn its own earlier bid for .travel to work with Tralliance). New.net is talking with Tralliance, says Sheehy: "We'd be happy to work something out [with Tralliance] where we'd stop new registrations in .travel and hand over our existing customers [for some consideration]. Instead of making life tough for our customers, it would be a really nice precedent." Is that extortion? Or is it merely a reasonable, market-based resolution of a problem that shouldn't have existed in the first place?

Object Name System: The story thus far

One of the major new registry efforts, the Object Name System (ONS) for manufactured consumer goods, is moving forward – with implementation details that illustrate the technical and commercial realities of such initiatives. What started as a set of research projects is becoming a reality, even as the parent organization, the Auto-ID Center at MIT, undergoes its own transformation: The academic research center, a major promoter of the use of RFID in consumer goods and retail supply chains (and sponsored by major corporations such as Wal-Mart), is spawning an industry standards and coordination body under the aegis of the Uniform Code Council and the European Article Numbering Association, while the academic research will continue at MIT and other research institutions. The new nonprofit company, to be born as AutoID Inc. on November 1, will coordinate the ONS, the registry of electronic product codes (EPCs) to be embedded in RFID tags, and will license the enabling intellectual property. (SEE **RELEASE 1.0**, JUNE 2003.)

Intended as a registry for manufactured items, the ONS supports a cross-reference between the individual EPC of a particular manufactured item (this specific shirt, that package of razor blades) and the URL for a website or ideally a particular Web page that has information about that item. (This is subject to all the DNS issues around broken links, supplier name changes and the like.)

A trial, small-scale version of ONS is currently running at MIT. It is performing as designed, including automated registration of EPCs from participating product owners as well as online resolution of EPCs to URLs.

One example of the compromises practicality has demanded, and of the complications large-scale implementations always face, is a change from the fixed-size EPC of

the initial design. It turns out that the number of bits that have to be transmitted from an RFID tag matters: Most scanning technology currently operates at low bandwidth, so real-life performance of the scanners is gated, in part, by the size of the EPC. The result is EPCs of varying sizes: The portion of the code that indicates the original manufacturer varies in length, which means the translation is no longer a simple lookup, but first a context-sensitive parsing and then a simple lookup.

These near-term pressures of practical implementation will result in variable-length EPCs going into production, which means that even if in the next few years improvements in the scanning technology make it unnecessary to restrict the number of bits, there will be a legacy variable-length installed base to deal with.

Similar practical considerations led the ONS design team to implement a simple cross-reference between EPCs and URLs rather than something more generalized and persistent – and therefore more complex – such as the Handle System underlying the DOI environment (PAGE 22). Auto-ID Center chairman Sanjay Sarma notes that the larger players, notably AutoID sponsor Wal-Mart, already have capabilities in place for communicating with their suppliers about goods moving through the supply chain. The ONS's role initially is not to displace those established channels for real-time and proprietary tracking data, but rather to open the door to easy lookups of product information, whether by someone in the middle of the supply chain or a consumer at the end it.

That more modest short-term ambition provides the opportunity to build up the registry of products and the infrastructure for product information services without tackling the infrastructure and institutional issues that an implementation aimed more directly at the supply chain would face. Such a system would require shared identity management services to confirm the identity/credentials of the person making a query. Today such information is generally provided indirectly, i.e. by the companies involved connecting specific internal systems to one another and then trusting each other's internal log-on controls.

It's not surprising that the AutoID initiative wouldn't start there. After all, volume is needed to sustain the start-up economics for everything from the tags to the readers to the systems that process the resulting data. That volume is going to come only from the large manufacturers and retailers.

Nonetheless, the basics are being put in place that could someday enable a new generation of applications, both in the supply chain and for consumers.

From Names to Numbers

ENUM: Balkan voices?

Potentially the most exciting new registry is ENUM – or whatever will fulfill its role of enabling the real-time, executable Internet: real-time because it supports real-time sessions between Internet endpoints, and executable because it enables those sessions between applications as well as people. ENUM is the proposed standard (RFC 2916, for the cognoscenti) for representing telephone numbers anywhere in the world as domain names so that they can be reached through the Internet rather than through the phone system, thus supporting end-to-end voice over IP (VoIP) among other capabilities.

ENUM is worth getting right, because it could do for real-time communication what the DNS has done for e-mail and the Web: enable both people and applications to connect in real time. You could call it a real-time, external active registry. Though part of the magic is in the Session Initiation Protocol (SIP, see box), ENUM is the registry that supports SIP interactions with people and applications outside private network boundaries. Even in its first, voice-only instantiation, loosely known as Voice over IP (VoIP), the concept shifts the phone system from what David Isenberg famously calls the "stupid network," with all the power held by the operators, towards a smart, edge network where the intelligence lies in the hands of users, and of entrepreneurs who can offer independent, third-party services from which users can pick and choose.

ENUM number/domain names combine an existing phone number (in reverse order) and an allocated SLD of e164.arpa. Each ENUM URL contains as its right-most digits an international dialing code (e.g. 1 for the US, Canada or Caribbean, 7 for Russia, 3.3 for France) that indicates which country it "belongs" to. There are a couple of non-country ENUM subdomains, but the presumption is loud and clear that ENUMs belong to the telephone world.

Meanwhile, the e164.arpa SLD is assigned to the US military, but in practice it will be run like a set of separate DNS subset registries (called "Tier 1"). RIPE NCC, the not-for-profit organization that manages IP addresses for Europe and Africa, is in charge of delegating each country's registry, with approval from the International Telecommunication Union (ITU), which regulates telephone numbering systems. (The ITU, for what it's worth, has offered to take over ICANN's regulatory duties, but has been firmly rejected.)

SIP ALL THE WAY DOWN

The SIP protocol is to ENUM what http is to the DNS – the workhorse communication standard. It stands for Session Initiation Protocol; more complex than http, which supports only one kind of communication, SIP enables negotiation of the parameters of a real-time session between endpoints over the Internet, whether voice or various kinds of data communication. Negotiating the nature of a session before it is established allows the entity receiving a "call" to decide whether to accept it, and also allows a calling system with more advanced features (e.g. video and voice) to step down to the capabilities of the receiving system (e.g. voice only).

Such capabilities are at the core of two urgent issues for these sorts of new services: user control/privacy and coexistence with the installed base.

Most of the things you can do with ENUM are possible now, but the SIP protocol makes them easy to set up within an organization, and ENUM will make them easier to set up with (relative) strangers. Moreover, in an open market, users will be able to pick the functions they want from competing third-party providers, rather than relying on a particular phone-system operator to supply them.

Some examples:

- A proxy server could give differential access to Juan depending on whether Alice or someone else is calling and at what time, along with other presence and access management services.
- A third party could offer various communication services, including translation from one format (text, say) to voice or from SMS to e-mail, (video) conference call set-up and recording, or automated call distribution for organizations.
- If you have phone capability on your PC, you can make and receive phone calls from your PC at your regular "number" as long as you have Internet access, wherever you are in the world (in a hotel room in Delhi, at Starbucks using a WiFi connection, or waiting in the lobby of the client you are calling on).
- Dell's support technicians (or HP's) could have preferential access to its hotlines to hand over a customer case to a new department, instead of being queued up just like ordinary callers.
- We could even fix the perennial problem of sending a cell-phone photo to a recipient on a different network!

Currently there's a lot of activity around the ENUM standard: There's an ENUM Forum with about 30 members including AT&T, and 20 countries plus two organizations (NeuStar, below, and VisionNG in Austria) are running ENUM testbeds.

There's political momentum as well: Long negotiations among the Internet Architecture Board (loosely linked to ICANN), the ITU and a variety of governments have resulted in a standard based on the DNS, but with implementations controlled country-by-country by whoever allocates phone numbers in each country – most likely phone incumbents or ccTLDs working with what the ITU calls "National [telecommunications] Regulatory Authorities."

Coordination and control: I-CANN/TU

ENUM highlights the delicate tensions between ICANN and the ITU. Although the overall ENUM registry is technically a part of the DNS falling under ICANN's oversight, its numbers will be assigned under the ITU's authority. The worldwide telephone system is a different proposition from the Internet from a governance point of view. Historically, it has been organized by geography – politically, economically and operationally. De facto, rather than bringing the phone system onto the Internet, use

of the ENUM standard would put part of the Internet (and what could amount to a new identity system) under the control of the worldwide telephone community.

Thus, the use of the DNS to store and manage the ENUMs would initially make things easier (and integration with the Internet simpler) by using an established, global, high-performance address translation system (the DNS) to establish linkages between the calling party and the recipient.

Meanwhile, telephone numbers are less complicated than names, but they are still too tied to something "real," rather than being mere identifiers with no other role. For the ENUM standard to have the positive impact some predict, its identifiers should not be tied to the world of telephony. (What happens when someone changes a phone number – especially to a phone number in a different country?) And for ENUMs that are not tied to actual telephones and their implied geography, then who will manage the allocation of such ENUM numbers? Coordinating this effectively will be essential to the reliability that is one of the fundamental charters of the government regulators and both a public and commercial imperative.

"The telco baggage is inherent in what ENUM was created to do: put real phone numbers in ENUM to ensure interoperability across PSTN [public switched telephone networks] and wireless as well as VoIP," says Mark Foster, CTO of NeuStar (below), which is running an ENUM testbed with ITU approval and also some internal ENUM-like registries for various unidentified carriers and manufacturers. "This in turn mandates that ENUM be delegated on a national basis, as administrative rules for numbers vary radically country by country. The solution is to have other ENUM-like services that resolve other forms of subscriber addresses (e-mail, IM addresses, etc.) that are free of the international and national regulatory baggage that phone numbers have."

But in fact ENUM could be useful for much more than phone calls, and long-term it would be an interference with the market – and an unjustified benefit for phone companies – to leave them the privilege of allocating and charging for most ENUMs.

There are other implementation questions, too: Should ENUM point directly to a user's phone number and other data – to be held in the registry – or should it merely point to a proxy server where the data can be held and access granted according to rules set up by the individual ENUM holder? In theory, the user should decide, but it's not clear whether that will be the case in each country's implementation.

All these questions have answers of sorts, but the current answers may not be the best ones in the long run. There's a distinct possibility that ENUM will not be the market success that its proponents envision, and that instead we'll have a long interregnum during which a number of ENUM-like registries will proliferate, each slightly different and controlled by a small group of vendors or users. Think of the world of e-mail before MCI, CompuServe, AOL and others adopted DNS-based e-mail addresses – or the current situation of competing, non-interoperable instant-messaging registries. Reuters (using an IM platform from Microsoft) and AOL have just announced plans to interoperate their IM systems, creating a bigger island – but an island nonetheless.

Show me yours first

ENUM's connection with the DNS raises another interesting set of questions. The DNS so far has primarily been a means of identifying Internet hosts – and indirectly, through the Whois, exposing some information about related individuals who are the holders or designated contacts for each domain name. A website is in one way or another a publishing mechanism, whether for an individual or an institution; a domain name identifies a server, whereas an ENUM ID identifies (from the user's perspective) a potential peer for a peer-to-peer connection. A "phone call" – even if it ends up being an IM session or a video conference or a money transfer – is inherently something real-time and interactive: a conversation rather than an interface.

So the question for ENUM implementations is how much information to expose in the publicly accessible Whois registry – and how that interacts with other identity information an ENUM registry could contain or point to. From being an institutional holding or personal "real estate," the object that an ENUM domain name points to becomes something close to a personal proxy, and the ENUM registry could become in essence its own (more accurate and far-reaching) Whois.

Questions about what to include involve not just engineering issues about how to establish a session, but public policy and consumer questions about privacy and control. (While initially ENUM is for setting up communication, there's no reason its identity information couldn't be used for confirming financial transfers.)

Under one ENUM DNS entry you could maintain pointers to a variety of possible forms of connection – one or more phone numbers, fax numbers, e-mail addresses, etc. – along with presence information and also access controls. . .but that's probably not a good idea. Such services will meet individuals' needs best if they are supplied

by independent providers, with a variety of capabilities, that ENUM could point to. Individuals could then choose among them (or use several, federated ones), rather than be forced to rely on a one-size-fits-all registry that includes applications.

There are identity questions on the "incoming" side, too. First of all, as with the DNS, there needs to be some way to authenticate the parties using the ENUM registry or the services it points to, whether to reach someone or to change their own information – either short-term presence information or more persistent information such as a corporate e-mail address, a credit-card number, or a personal white list. And more: Assuming an individual's work number is included in the ENUM entry, can the company change the entry when that individual is fired? (How do ENUMs interact with corporate provisioning systems?)

As NeuStar's Mark Foster puts it: "The challenge posed by ENUM is to keep it focused on what it's good for, service end-point resolution, not a generalized all-in-one directory in the sky." The less the ENUM registry contains, the easier it will be for it to evolve. For example, one could imagine ENUM using the Handle System (PAGE 22) to point to a variety of third-party proxy servers. Another ENUM registry, independent of the phone-based registries run by countries, could potentially federate current instant-messaging registries.

In short, tempting as it may be to see ENUM as a universal rolodex and a universal person locator system, the administrators of ENUM registries will foster wide adoption by limiting its functionality. Such a minimalist position would encourage entrepreneurs to build useful applications such as rolodexes and sophisticated presence management tools for those who want them. Then separate corporate and personal identities can be federated but not combined (or confused), and control can remain with the appropriate party(ies). This will foster competition and innovation as vendors respond to, support and take advantage of the evolution of pragmatic network identity management.

The market

How all this plays out and when provokes intense speculation in some quarters. The biggest players currently in the telephony market – the incumbent carriers – generally have little interest in a more open market, even though they increasingly send their own traffic over internal IP backbones. (They also raise the issue of 911 calls, which have to remain geographically bound.) As David Passmore of the Burton Group writes: "It's not currently in the best interests of the Incumbent Local

Exchange Carriers to share the contents of their [phone] network directories with anyone else, nor today would they be interested in populating the DNS with phone number information. So how might VoIP islands coalesce into a critical mass of VoIP users that would require widespread adoption and support for ENUM?"

By contrast, the cable companies now offering broadband service see Internet telephony and other communications services as a way to absorb their unused capacity. Cable operators such as Comcast and Cox are eagerly looking for ways to add value and redeem their reputations. In Japan, 2.7 million of Yahoo! Broadband's 3 million subscribers use its VoIP service, the company says. "Voice is the killer application for DSL!" says one wag.

Third-party services, such as Critical Path (SEE **RELEASE 1.0**, JUNE 2002 AND MARCH 2003) and various other hosted messaging services, see opportunities for the kinds of applications listed above, as do start-ups such as SoloMio (SEE **RELEASE 1.0**, MARCH 2002), who can sell to and through the hosted services and perhaps reach customers on their own. Many of the pioneers in these services are likely to be either acquired or crushed by the big guys, who will be provoked into providing such services themselves.

And many organizations are using Internet telephony systems internally; the US government's own Department of Commerce, JetBlue Airways and bank holding company SouthTrust of Birmingham, AL, were listed in a recent New York Times story. It also cited market research from In-Stat/MDR that about 2 percent of US businesses are implementing some form of Internet telephony internally. That is, they're not just purchasing telephone service, but implementing their own over the Internet.

Microsoft will accompany its new Office 11, due in October, with a new SIP-based server, Live Communications Server 2003. The idea behind it is to be able to link up with others through a variety of applications rather than through dedicated "communication tools": For example, soon (though not in this year's release), if you click on the balloon linked from a nitpicking revision of cell 56g of your stunning business plan spreadsheet, you'll be able to tell not only who did it, but whether they are online right now. Depending on your relative rank, you can reply to that individual with a scathing rebuttal or a constructive compromise – and hope to get a reply in real time. "Up to now, communication tools have been vertical," says general manager Gurdeep Singh Pall. "This makes communication horizontal." But though the new Office is what Pall calls "SIP-centric," it does not support ENUM; instead, it uses Active Directory, which is the basis of Microsoft's own instant messaging directory.

And finally, the combination of entrepreneurial innovation enabled by the end-point control model of the Internet and the price leverage it enables (leverage which may dissipate if current downward pricing trends for phone service continue) will continue to propel even the established phone carriers towards a more Internet-like model.

NeuStar: Have registries, will travel

One of ENUM's biggest proponents is NeuStar, based in Sterling, VA. In the still-small world of external registries, NeuStar wants to be everyone's plumber – expert at solving problems and rarely visible to the millions who depend on it. Formed seven years ago as a business unit of Lockheed/Martin, the company spun off in 1999 as a registry supporting phone-number portability – the technology that allows US phone subscribers to switch vendors without switching phone numbers. The management buyout was funded by Warburg Pincus, which sees the company as a key to the Net's emerging intellectual infrastructure.

With about $100 million in annual revenues, and with measurable but undisclosed profits, NeuStar sees its work with the telephone industry as a solid, recurring-revenue foundation from which the company can branch out into other similar businesses. Over the past couple of years, NeuStar has made its way into the domain- name registry business. It runs the registry for the .us and .biz domains, provides the external gateway to China's .cn domain, and is running a variety of ENUM and ENUM-like registry services. And finally, it operates one of the first Liberty Alliance-based identity management services.

NEUSTAR INFO

Headquarters: Sterling, VA
Founded: spun off from Lockheed/Martin in 1999
Employees: 324
Key metric: $100 million in annual revenues; profitable
Funding: undisclosed amount from Warburg Pincus
URL: www.neustar.biz

On the telecom side, NeuStar's number portability registry contains all 483 million US phone numbers, and links them to their corresponding Local Routing Numbers (LRNs, another 10-digit number referring to a physical line termination). "Two billion phone calls hit our database every day," says NeuStar's Mark Foster.

As things worked before, every phone number was essentially a physical address, managed internally by the phone company that "owned" it and allowed a customer to use it. The first three digits of a 10-digit US telephone number, the "area code," stood for a particular segment of the network of the phone company that serviced a geographic area.

But for customers to switch phone carriers without switching numbers, the numbers had to be disconnected from their physical meaning. To enable the physical connection of a phone call (which still mostly requires an end-to-end circuit between caller and called), there still had to be something that represented the actual physical routing. Enter the LRN and the need for a cross-reference between telephone numbers and LRNs.

By now, a large proportion of US phone numbers are different from their LRNs, and virtually all exchanges contain at least some ported numbers. NeuStar's primary role was to set up the new registry, provide copies to its thousands of phone-company customers, and keep all the copies updated thereafter. Using data supplied by those same customers, NeuStar updates its own registry and broadcasts the changes to each copy; that process happens many thousands of times a day. The daily number of changes is likely to increase later this year when mandated cell-phone number portability goes into effect. It is the volume of changes that is NeuStar's core revenue stream, from the phone companies, who in turn charge their customers.

Number portability was mandated by Congress as part of the Telecommunications Act of 1996. A telecommunications industry cooperative called the North American Portability Management Company (NAPM) acts as the oversight body for this registry, with authority delegated from the Federal Communications Commission. Membership in it is de facto mandatory for any licensed facilities-based carrier in North America.

NeuStar has a master contract (renewable yearly) with NAPM covering pricing, service levels and penalties for non-performance, and each US carrier has a corresponding contract with NeuStar. Prices, set by members of the NAPM consortium, are designed to support the service plus a small profit margin to induce NeuStar to stay in the game.

Regulations prevent NeuStar from using the phone-number information for any purpose other than number portability, but it can use its relationships with the carriers to sell other services, which it does. For example, it also provides a clearing house for operational support services (basically ordering, billing, and other administrative and support services) between interested carriers. For that, NeuStar can charge what the market will bear. Although regulations keep NeuStar from trading on its access to data, it can certainly leverage its relationships.

NeuStar and the digital closing: Digital Domesday

Elsewhere, NeuStar provides circle-of-trust administration support for online real estate closings – illustrating the interaction of identity management and (in this case) real-estate registries. Working with Affiliated Computer Services (ACS) of Dallas, TX, a $3-billion IT service provider with a specialty in government outsourcing, NeuStar is implementing the Liberty Alliance federated identity standards in order to make the time-consuming, paper-intensive process quick. . .and secure. It's one thing to send documents and to check local real-estate registries electronically. It's another thing entirely to update them online, and to ensure that the electronic signatures belong to people with the authority to commit to what they are signing. For example, is Sue still working for Ma&Pa's Title Search Inc., and is she still authorized to sign title-search reports? (This problem exists in the physical world, too; one benefit of the electronic closing, beyond cost-savings, is that it can settle such questions more reliably.)

ACS and NeuStar are applying federated identity technologies to this problem. Each of the participants in the closings in a given county – the bank, the title insurer, the county, etc. – will federate their internal identity systems. Digital identity meets the descendent of the medieval Domesday book.

Overall, NeuStar sees itself as a model for a new kind of business – a disinterested third-party operator of registries and other common infrastructure. Its technology skills are part of its edge; it knows how to define requirements and meet them. But its people also have great skills in managing political and consensus processes. Certainly the challenges of number portability are not as great as those facing ICANN and the ENUM registries, but NeuStar has a knack for engendering compromise and working effectively with a broad range of constituencies.

Digital Object Architecture and the Handle System

Both the DNS and the phone system offer ways to reach specific Internet locations or phone numbers, at a given time. Telephone numbers change, as do the devices they're attached to; website content changes and moves, from, say "Today's hot tip" to "Last week's highlights" to "Archive," or from "Team member" to "CEO" to "emeritus" to "who he?" Owners of websites change over time, whether through corporate reorganizations or the sale of a catchy domain name from one speculator

to another (or perhaps to an actual website operator); people switch (or accumulate) e-mail addresses.

But the people and content you can reach these ways persist through time, even as site names and numbers change. How can we register persistent things in a more persistent way? One man with an answer is Bob Kahn, a veteran of the Internet. He and Vint Cerf, now chairman of ICANN and SVP of Internet Architecture and Technology at MCI, jointly developed TCP/IP for the US Defense Department's Advanced Research Projects Agency (DARPA) in the early 1970s.

In 1986, Kahn founded the nonprofit Corporation for National Research Initiatives (CNRI) in order to pursue the notion of a national information infrastructure. (Cerf had left DARPA to join MCI in 1982; he rejoined Kahn as vice president of CNRI from 1986 to 1994 before returning to MCI.) Over the last two decades CNRI has worked with government, educational and private organizations on a variety of R&D initiatives, and provides the secretariat for the IETF. It patents its R&D where it can, says Kahn, but primarily for defensive purposes – to ensure that its work is available to anyone, retains its integrity and remains unrestricted by other patents.

One of CNRI's leading initiatives is the "digital object architecture" and its implementation through the system of unique, persistent identifiers it calls handles. The Handle System is somewhat like the DNS, in that it provides a registry to find things online. But it has important differences, says Kahn: "While the DNS registers machines [with domain names], the Handle System registers digital objects. The DNS served its purpose well, but now we have something better and different, with persistence and location independence. The digital object architecture is a reconceptualization of the Internet to deal with specific information objects, instead of just flows of packets between servers. It takes transport out of the picture [although you still need it underneath]."

History that hasn't happened. . .yet?

The Handle System has been operational since 1994, and growing steadily, with somewhere upwards of 10 million handles now registered. Says Kahn, "In the late 90s, Jon Postel [who oversaw the DNS until ICANN took over] and I had come to an agreement to consider using the Handle System to replace or at least augment the DNS, but he couldn't make the decision by himself. . . ." And then came ICANN, and Postel's death. The existing system persisted.

"Understanding the digital object architecture is a bit like understanding the Internet," Kahn continues. "When they finally got it, a lot of people started to do interesting things with the Internet. We hope the same thing will happen with handles. The idea is that you can invest in content independent of platforms and of where you store or it or make it available: Anyone can find it, with certainty, not with the ambiguity of a search engine or dependency on some website."

Although the implementation is complex, its use is not. The result is as simple as, say, a Pobox.com or alum.XXX.edu e-mail account/ID that you can keep for life (updating the information behind your unchanging e-mail ID to point to the new address). In the same way, for example, you might use a handle for your blog, so that you could keep the same handle and update its location information (rather than mailing all your friends) when you switch from one blog-hosting service to another.

Handles work seamlessly with the DNS: For human use, they can be embedded in URLs and resolved through a proxy server (such as CNRI's hdl.handle.net), so that they look just like a "regular" URL: e.g. HTTP://HDL.HANDLE.NET/4263537/1001, which currently resolves to the URL, HTTP://WWW.CNRI.RESTON.VA.US/FCC/REPLYTOCOMMENTS-FCC.PDF.

What's different from the DNS is that the Handle System's location independence, including identifiers that have no meaning (and therefore do not need to be changed, when, say, an owner changes its name or an object changes its location). That provides for persistence as well. For example, if CNRI decided to rename the document above, "RejectedPetitions," for example, or to put it elsewhere on the Net, the handle wouldn't change – and it would still resolve to the same object at its new URL (assuming proper updating of the handle record).

Thus, the Handle System obviates the DNS problem of broken URL links. That problem is a side-effect of the fact that the URL is used as an identifier and typically specifies the local file name at a location where you once were able to find that object.

Resolved but not seen

Handles generally are not meant to be seen by humans; they usually hide behind clickable buttons in applications. When a browser enabled with a handle plug-in encounters a handle, it resolves the handle via the registry to an appropriate URL or IP address for that object. Other, non-browser applications can interact directly with the Handle System using software routines that CNRI makes available for download from its site. When a handle server resolves an identifier, it returns a handle record

which contains owner-specified information for the object, which could be location information, a PKI key, or a variety of different kinds of information or functions about the object such as links to digital rights management or purchase information. (It's very flexible!) It's up to the user or the user's software to extract the appropriate information from the handle record. In short, the Handle System supports application interactions rather than just queries.

The organization side

Currently, CNRI runs the handle servers, but the system is designed so that any entity can set up its own registry and run its own handle servers without continued involvement from CNRI.

Like the DNS, the Handle System is logically a single registry, but unlike the DNS it is not hierarchical. It can scale and still operate "flat." While both the DNS and the Handle System can be partitioned into subset registries that are managed separately, TLDs and second-level domains are hierarchical subsets, while the Handle System has local handle servers that interconnect horizontally.

Like a second-level domain name, a handle has two parts, before and after a slash (vs. a dot in a domain name). But again, there's a difference. If the handle is in the global registry, the entire handle is resolved. If the handle is stored only in a local registry, set up independently, the part of the handle before the slash may be used to determine which local registry can resolve the entire handle.

Because the handles are simply arbitrary numbers, about as exciting as IP addresses, there's none of the vying for "space" that afflicts the DNS. "It's a green-eyeshade job to allocate them," says Kahn. "You do have to pick a registry, or run your own," but that just means using the global handle registry or selecting from a list of registrars certified by the International DOI Foundation (IDF, see below) or another organization. . .or taking a risk on whatever registry you choose. Whether or not it's as simple as that, the IDF stands ready to provide oversight of one part of the Handle System – and if users don't like the IDF's approach, they are free to start their own registries and their own governance system.

One of the Handle System's most valuable uses could be for managing a public key infrastructure or for identity management more generally. Or it could provide a valuable degree of persistence and indirection for ENUM. "I've talked to people connected with ENUM," says Kahn, "but the implementation decisions involved political

considerations, and they are reluctant to open them again at this point." The Handle System could also provide an effective underpinning for AutoID's ONS.

However, the Handle System lacks the visibility of the DNS; after all, it is so open that no one has seen much edge in marketing it until now. In addition to the IDF, a number of organizations including the Department of Defense Technical Information Center (DTIC) and the Library of Congress are running their own handle servers as internal registries. And groups in several countries, including CNNIC (which also handles China's .cn TLD), have been experimenting with it on a national basis.

But there is now a broader, outward-facing implementation, overseen by the IDF. Created by major players in the publishing industry, the IDF oversees the most visible, external registry in the Handle System, for what it calls "Digital Object Identifiers," or DOIs, a branded, trademarked version of handles. The IDF board includes the Association of American Publishers, The International Publishers Association, and various publishers including John Wiley & Sons and Springer Verlag as well as technology companies including Microsoft and Hewlett-Packard. IDF performs the sorts of functions ICANN does for the DNS, including certifying registrars for its local registry (but unlike ICANN, it does not control the global registry). It also licenses the DOI trademark and establishes policies for use of DOIs. The DOI registrars include CrossRef, a consortium for journal publishers; The [UK] Stationery Office, a recently privatized government bureau; the Copyright Agency Ltd. in Australia; Enpia Systems, a South Korean scholarly publishing group; MEDRA, the EU-funded Multingual European DOI Registration Agency, comprising publishers' associations in France, Germany, Italy and Spain; and the leading for-profit, commercial entry, Content Directions, Inc.

Content Directions, Inc.: Objects get respect

Of the seven registrars for DOIs certified by the IDF, Content Directions, Inc. (CDI) is the most active commercially. The story begins in 1996, when CDI's founder and CEO, David Sidman, was director of New Publishing Technologies at John Wiley & Sons and an active participant in the Association of American Publishers (AAP). The AAP was looking for a persistent registry system for content – more specifically, for a way to publish and sell high-value content electronically without losing it to piracy. The scientific community in particular was pushing science journal publishers to find a way to put their very high-value content online in a controllable way. They wanted it online to facilitate sharing. They wanted it controllable both to help

ensure proper credit in the academic world and also for the usual commercial reasons of protecting intellectual property.

The new digital numbering scheme had to be more granular than, say, the ISBN (International Standard Book Number) codes used to identify books, for references to scientific articles and other such content sections rather than entire books or journals. Also, it had to be permanent. Scientific journals still refer to Einstein's 1905 paper on Special Relativity, for example. The system also had to handle very high volumes, of both registrations and references. Millions of books are published in the US each year alone; if you count chapters, articles and the like, the Lexis-Nexis database lists over 4 billion [sic] publications. Finally, it had to handle any language and alphabet in which things might be published.

The Handle System was the only system that met all the requirements. The AAP – and Sidman (then at Wiley) – worked closely with CNRI on an implementation. The first registrar was CrossRef, a non-profit consortium of about 200 scientific journal publishers. "They happily used it for cross-referencing all their articles," says Sidman, "but I saw the potential for a broader range of uses. The killer application for DOI is not online footnote resolution, but rather full e-commerce, eventually including digital rights management." So in 2000 he left Wiley to start CDI, which is now a leading commercial registrar of handles.

Digital objects and digital rights

CDI has used the Handle System's extensibility to develop what it calls "MultiLinks": basically, a specific use of the DOI system to manage and present a variety of functions and contents that a single DOI can link to. . . including the content itself, related content (author information, for example), as well as links to a copyright administration server or to retailers' online shopping carts. (Note that the digital rights administration server may need an identifier for the person who wants access to the content. . .another example of the ubiquitous need for identity management.)

Publishers and other owners of content pay an initial setup fee for registration of the objects, including the MultiLinks, and then they pay for maintenance of the DOIs and their MultiLinks in the global directory, including, for example, constantly changing links to retailers or to related content. The pricing is negotiated; publishers typically pay in the tens or hundreds of thousands of dollars per year, says Sidman.

Some big names in the publishing business have signed up with CDI thus far: McGraw-Hill broke the ice last year by embracing DOIs for all its published content. Since then, Thomson Learning, Corbis (Bill Gates's stock photo agency), RAND, Penguin Putnam and others have followed suit. Even the AAP itself has embraced DOI and uses CDI as the registrar for its association journal and other publications.

But Sidman has eyes on more: "The flexibility of the MultiLinks [and of the underlying Handle System] just opens whole worlds of applications." CDI and a large, unnamed music company have initiated a pilot under which CDI will register the label's offerings for several major artists, with three kinds of objects – for the artists, for their albums, and for individual songs. Their goal is to improve the online content experience, to make it easier and more fun for customers to hear their favorite artists and look things up about them online at the same time, find related performances. . .and to support all this, of course, without losing control over their copyrighted material. "We're getting a good hearing at the other music companies too," says Sidman. "We didn't want to venture into that world without a solid success behind us, but now we're ready." And perhaps the music companies are ready, too. But in the meantime, there's Gracenote (PAGE 29), a do-it-yourself registry of music published on CD.

CONTENT DIRECTIONS, INC. INFO

Headquarters: Brooklyn, NY
Founded: August 2000
Employees: 6 full time, 10 contractors "plus a halo of 40 to 50 consultants who implement systems"
Key metric: customers include McGraw Hill, Thomson Learning, Corbis, RAND and Penguin Putnam
Funding: "$several million" from angels including former Dun & Bradstreet chairman Bob Weissmann, hedge fund wizard D.E. Shaw and Marty Kahn
URL: www.contentdirections.com

What the content providers really want is to make another sale when friends pass each other a song or a book or a magazine article. The vendors' dream of user-oriented, flexible digital rights management, of the sort that will require a registry like the DOI registry, is precisely to make it easy for that secondary sale to happen. The question, of course, is whether friends are really willing to be sales agents. When Juan sends Alice a copy of something, he is showing off what he already bought, just as he might drive by her house in his new red Infiniti. How likely would he be to drive by her house in his new car if the manufacturer tried to collect a royalty from Alice when he did?

Finally, CDI is also talking to a maker of networking equipment, which is looking to use DOIs to track individual pieces of equipment, allowing various parties to see or update information about the individual item's history – manufacture, journey through a variety of sales channels, configuration, repair and maintenance. The idea is to manage after-sales support more effectively – much the same idea as motivates many of the RFID applications covered in our June 2003 issue – and above in the section on ONS.

Gracenote: Do-it-yourself registry

As noted above, the music industry may be about to do something about creating a registry for itself, starting with the unnamed company working with CDI. But in the meantime, the industry's users have taken matters into their own hands and have registered CDs without help (or permission) from their owners.

The company holding the database, Gracenote, started out in 1995 as something called the CD database, or CDDB. A compilation of information about music on commercially produced CDs, CDDB was the brainchild of Ti Kan and co-founder Steve Scherf, UNIX kernel programmers at a now defunct Silicon Valley computer vendor. The database itself was born in 1993, when Kan and Scherf began building it to keep track of the contents of CDs they played on their computers.

Interestingly, CDs for the most part do not contain any information about the music on them, other than a table of contents (TOC) that has only the start and end times of each track – no text, artist names or titles. But it turns out that the exact sequence of start and end times for each track on a CD (measured in 1/75ths of a second) constitutes a unique identifier: As a practical matter, no two CDs have exactly the same sequence of track lengths (although some with only four or five tracks occasionally get pretty close and can be distinguished only by complex algorithms or human intervention, says Scherf).

GRACENOTE INFO

Headquarters: Emeryville, CA
Founded: 1998
Employees: 55
Key metric: approx. 1 million daily users and 2.1 million CDs registered
Funding: undisclosed amount from Sequoia, Simon Investments, Scott A. Jones and Silicon Valley Bank
URL: www.gracenote.com

By manually entering the descriptive information for the CD and indexing it to the CD's TOC timing signature, Kan and Scherf were able to start building a database that could be used to automatically display the information about a CD when it was inserted into a computer's CD drive. They put the database online in 1995, and soon thousands of people were not only looking up their CD information from the database, but contributing more information for CDs not already in the database. (Note that the database does not provide access to the music itself, though it can be used by, say, a retailer pointing to purchasable content.)

Today, Gracenote's database contains detailed information about every track on essentially every CD produced just about anywhere in the world. About a million users per day (often several times a day) use it to find out about their tracks. They also contribute about 7000 submissions per day to the database by clicking on the "submit" button usually included with music playback applications that reference CDDB. Of those 7000 submissions, about 1500 make it into the database each day.

Those 1500, a mix of new CDs and updates to already-cataloged CDs, have survived several thousand filters that weed out spurious submissions, automated voting logic to select the most likely accurate version from among near duplicates, plus human screening (done in the US, Japan and China) where needed. Director of marketing Ross Blanchard says that within hours of a popular new CD's release, Gracenote receives between 10 and 100 submissions of information about it.

Annotation is the sincerest form of flattery

"The labels don't seem to be threatened by what we do, and why should they?" says Scherf. "In fact, they have been quite friendly towards us. We offer a great service that enhances a person's music listening experience and ultimately even drives sales."

He adds: "There is a lot more to 'publishing' a registry than just making a giant spreadsheet of discographies. Even if the labels did manage to put out a unified set of album information, it takes a huge expenditure of money and manpower to invent the necessary technology to perform media recognition functions, and to design, build and maintain an Internet service. In doing what we do, we ultimately save the labels the effort of doing it themselves. I don't think it would ever be possible for the labels to successfully organize such an effort. Besides the big five (or four these days?), there are hundreds, if not thousands of smaller labels across the world. Only a community effort like the one we've built would be capable of collecting information from such a massive set of sources."

Meanwhile, he adds, Gracenote makes its money not off the data, but from software that it sells to music-player manufacturers and to online services. The functionality includes media management, file retagging, playlist generation, data-linking and the like. It and its partners use what Gracenote calls TUIs (title-unique identifiers) to resolve to everything from liner notes and album covers, to AOL's Media Browser purchasing gallery or reviews. "The TUIs are our key to the universe," says Blanchard.

For commercial developers, Gracenote charges per unit, on a sliding scale, ranging from tens of cents to tens of dollars. It gives away database access to end-users, and licenses the software free to non-commercial developers – notably college students who often develop media players for class assignments. Giving away database access to consumers helps keep up the interest that stimulates the 7000 contributions a day that keep Gracenote's database perhaps the most comprehensive and current database of published CDs in the world. That, in turn, makes Gracenote's software and database so valuable to the commercial developers it sells to.

Gracenote's software is now preinstalled with commercial music players that mine the Gracenote database to display information about the CD they're playing. Apple's iTunes, RealNetworks' RealOne player and AOL's WinAmp, as well as a number of home-grown and noncommercial applications, use the Gracenote database.

Most of Gracenote's users are online, but that's changing. Gracenote also counts among its customers hardware manufacturers such as Sony, Philips, Yamaha, Samsung, Simple Devices of San Mateo, CA, and others. In Japan, Pioneer has incorporated music playback capabilities in the on-board disk-based navigation systems prevalent in Japanese cars, and uses Gracenote's software and a copy of the database to display which tunes are playing. It keeps the database up to date either through a wireless phone application or through what amounts to a flash card that can be removed from the car, plugged into the owner's PC and updated online.

The un-registry

Gracenote is not really a registry in the sense we've been using the term. For one thing, though Gracenote allows labels to directly update the database and to lock it to ensure it doesn't get changed by users, fewer than 10 percent actually do.

For another, Gracenote's database is used to display information to humans, rather than as an indirect addressing mechanism for retrieving the objects themselves. As Blanchard says, "We're not pursuing digital rights management applications. If our customers were to come to us looking for something, that might be a different matter, but for now we're not pursuing it."

COMING SOON

- Nonstructured data analysis.
- Maps and geolocation.
- The future of e-mail.
- State of the Net: Is our infrastructure sound?

- And much more. . . (If you know of any good examples of the categories listed above, please let us know.)

Moreover, Gracenote's database records more information about the CDs than registries generally hold. Nonetheless, Gracenote's business illustrates the sorts of creative applications that registries, and databases linked to them, can enable. Moving beyond the simple cross-reference of CD tracks with their descriptive information, the other characteristics in the database (including recording dates, genre information, instrumentation, etc.) can be used to develop sophisticated playlist generators (e.g. play all the blues I've got that was recorded between 1963 and 1968 in Chicago that includes a harmonica). It can also be used to develop multimedia applications that search the Web for related information such as album cover art, music videos, artist bios, latest news – all of which could be sold if the music industry took advantage of the opportunity.

Registries: The Sequel

Our experience with ICANN and the DNS has given us an abiding interest in registries – what they can do for the world, and how they can be better organized and run. In particular, we were intrigued by ENUM and what a worldwide registry could do for real-time Internet communications, the promise of the Handle System, and the use of registries in the context of digital rights management and identity management. While most of the new registries we cover here are attending to the governance and broad technical questions they face, they will need to sustain that attention and respond effectively to evolving circumstances over time. Registries typically begin from a point where they are a curiosity of interest to the few, and the main issue is to persuade object-owners to register their objects and object-users to look to the registry as a source for them. But as the registries approach success, playing a key role and helping to hold a market together, they will need to pay special attention to scalability, persistence and accountability to an entire marketplace of users as well as providers. ■ R 1.0

Acknowledgment:
Thanks to Eric Dean for his research and resilience. He helped immeasurably to ensure the accuracy and accountability of this issue of Release 1.0.