

Testimony of Benjamin Edelman

before the

United States House of Representatives
Committee on the Judiciary

Subcommittee on Courts, the Internet, and Intellectual Property

September 4, 2003

Benjamin Edelman
Berkman Center for Internet & Society
at Harvard Law School
<<http://cyber.law.harvard.edu/edelman>>

Baker House – Harvard Law School
1587 Massachusetts Avenue
Cambridge, MA 02138

Chairman Smith, Ranking Member Berman, Members of the Subcommittee:

My name is Benjamin Edelman, and I am a fellow at the Berkman Center for Internet & Society at Harvard Law School, where I write software to study the Internet. Among my research interests is the Internet's domain name system, and I have written a series of articles about flaws in the Whois system, about domain name registrants who exploit these flaws, and about possible means of detecting and preventing such exploits. My full biography and publication list are available at <http://cyber.law.harvard.edu/edelman> .

In the interests of full disclosure, let me pause to note that I previously worked for ICANN. I designed and operated webcasts of fully a dozen ICANN's meetings – so that anyone interested could watch, read, and even ask a question from home or office, without traveling to a far-flung meeting site. In 2000-2001, I also briefly served as a consultant to ICANN as to technical issues associated with the introduction of new top-level domains as well as with certain security and competition concerns .

Today the subcommittee considers the accuracy of the Whois database, and the role of the Department of Commerce, ICANN, registries, and registrars in assuring the accuracy of Whois data.

My bottom line:

As the DNS is currently structured, registrants are under only an honor system to provide accurate Whois data. Meanwhile, it makes no economic sense for registrars to enforce Whois accuracy. The result is that in terms of accuracy, when compared with other compilations of public data (such as driver's licenses and trademark registrations), the Whois database is substantially fiction.

Despite years of inquiry by this subcommittee, in addition to numerous ICANN working groups and other discussions, intentionally invalid Whois data remains widespread. But failure to solve this problem so far doesn't mean the subcommittee must give up. Instead, new efforts at detection could better find invalid domain names, while new incentive systems assure that registrants provide accurate data and that registrars confirm that they do so.

My specific suggested incentives include 1) a reduction in the lenience of opportunity to "cure" intentionally invalid data, 2) for registrants with multiple domain names with intentionally invalid data, forfeiture of all domains when any are to be cancelled, 3) statistically valid surveys of registrars' Whois accuracy, with public reporting of each registrar's accuracy, 4) public reporting of Whois accuracy complaints and their dispositions, and 5) financial and other penalties to registrars with poor Whois accuracy records.

Scope of the Problem of Invalid Whois Data

The Internet's domain name system (DNS) currently includes approximately thirty million domain names within the top-level domains of .COM, .NET, and .ORG. Under ICANN policy, passed on to domain registrants through contracts via registry and registrar, each of these domains must report the name, address, telephone number, and email address of its technical and administrative contacts, as well as the name and address of the its registrant. This information must be published via the so-called "Whois" database operated by domain name registrars.

It has long been known that a large number of domain names offer invalid Whois contact information. In some instances, the invalidity may be unintentional; registrars' data systems occasionally corrupt registrant contact information, and registrants (especially non-native English speakers) might misunderstand registration forms. In general, though, invalidity is thought to be intentional, reflecting registrants' desire to keep their identities confidential. This inference is particularly strong when Whois data is obviously intentionally invalid ("123 Main Street" or "0 Does Not Exist Lane"), when invalid Whois data is combined with controversial content (pornography, cybersquatting, etc.), or when the invalid information and associated web sites are clearly the work of sophisticated registrants.

In the past, some have attributed Whois accuracy shortfalls to difficulty in determining whether specified addresses are valid. After all, if a registrar cannot determine if a given address is accurate, the registrar cannot enforce accuracy requirements. However, automated systems are increasingly well able to cross-check registrant name, address, and postal code, all with minimal delay and low cost, at least as to registrations in industrialized countries. A new service called Fraudit (from a DNS service provider called Alice's Registry¹) performs precisely these functions, using only publicly-available databases. Credit card verification software typically uses similar methods, and registrars have been using card verification software for some time in order to reduce "chargeback" penalties and confirm validity of customer credit cards. However, I know of no registrar currently using these methods to prevent invalid Whois data.

Using a variety of methods of locating suspicious registrations, my prior research identifies thousands of domains with intentionally invalid Whois data. For example, in my May 2002 *Large-Scale Intentional Invalid Whois Data: A Case Study of "NicGod Productions" / "Domains For Sale"*,² I identified a total of 2,754 domains registered by a single registrant – but using addresses in at least ten countries, registered via at least eleven registrars. Similarly, my January 2003 *Large-Scale Registration of Domains with Typographical Errors*³ reports more

¹ <http://www.ar.com>

² <http://cyber.law.harvard.edu/people/edelman/invalid-whois>

³ <http://cyber.law.harvard.edu/people/edelman/typo-domains>

than 8,800 domains registered by a single registrant using at least six pseudonyms, using addresses in at least six countries, and using at least four registrars.

Intentionally invalid Whois data is often associated with other controversial registration practices. This is perhaps not surprising – after all, registrants with something to hide are particularly likely to conceal their true contact information. My *NicGod* research found clear evidence of bulk registration of domains previously used by other registrants, then allowed to lapse (typically mistakenly, e.g. by administrative error), subsequently captured by NicGod, which then attempts to resell them to their original registrants after markups on the order of 5000%. My *Typographical Errors* research found registrations of strings that are small variations on well-known marks (e.g. cartoonnetwork.com [sic]), and the resulting domains were typically redirected to sites offering pornography, online gambling, filesharing, or other controversial materials. These are troubling practices – practices which force small business owners to pay thousands of dollars to retain the domains they previously used, and practices which expose Internet users to pornography as a penalty for small mistakes in typing URLs.

Incentives for Registrants to Provide Accurate Data

That registrants provide invalid Whois data should perhaps come as no surprise. After all, domain name registrants have only limited incentives to provide accurate Whois data.

1. Accurate Whois data is not necessary in order to pay for a domain name. Even when contact information is cross-checked with credit card records at the time of domain registration, it is typically possible to modify contact information subsequent to registration.
2. Registration agreements, typically accepted by registrants by pressing an “I agree” or similar button during the domain name registration process, oblige a registrant to provide accurate Whois information. But few registrants typically read these agreements, and the format of these agreements rarely places special emphasis on Whois accuracy.
3. Even when registrars send periodic reminders that Whois data must be kept up to date, as is required under ICANN’s Whois Data Reminder Policy,⁴ registrants are likely to ignore the reminders. This too is no surprise – particularly since Whois reminders are widely thought not to be supported by active investigation or enforcement.

⁴ “At least annually, a registrar must present to the registrant the current Whois information, and remind the registrant that provision of false Whois information can be grounds for cancellation of their domain name registration. Registrants must review their Whois data, and make any corrections.” <http://www.icann.org/registrars/wdrp.htm>

4. When a registrar receives a complaint as to the accuracy of a registrant's Whois data, the registrar typically grants the registrant an opportunity to cure the problem by correcting the invalid entry. Anticipating this opportunity, a registrant need not offer accurate information in the first instance. Instead, the registrant can provide invalid Whois data, to be corrected only upon complaint. In addition, some registrants provide a series of invalid contact names and addresses, a problem recently faced by staff of the OECD.⁵

In short, the current registration scheme fails to set incentives for registrants to provide accurate Whois data. The system provides no incentives for registrants to provide accurate data in the first instance – for registrants always receive an opportunity to cure invalid entries, without penalty. Furthermore, the system allows bulk registrants to sacrifice a disputed domain rather than share their true identities – for domain cancellations are limited to the specific disputed domains and do not extend to other domains registered by the same registrant using the same invalid Whois data.

The following modifications would correct these incentive problems

1. When a registrant's Whois data is found to be intentionally inaccurate, penalize the registrant in some way before (or instead of) offering an opportunity to correct the error. The penalty could consist of charging a fee for investigation, or forfeiting some period of prepaid registration service.
2. When a given domain name is to be cancelled for offering invalid Whois data, also cancel all other domain names registered with identical invalid Whois data.

Incentives for Registrars to Assure Accurate Data

Registrars' failure to enforce Whois accuracy is also predictable, for registrars face equally limited incentives to provide accurate Whois data.

1. Registrar contracts with ICANN oblige registrars to include certain language in their contracts with registrants, asking registrants to provide accurate Whois data.⁶ But this requirement extends only to language in registration agreements – not to actual efforts at enforcement. Neither do other sections of registrar contracts with ICANN require specific enforcement procedures as against registrants who provide invalid Whois

⁵ "Cybersquatting: The OECD's Experience and the Problems It Illustrates with Registrar Practices and the 'Whois' System" <http://www.oecd.org/dataoecd/46/53/2074621.pdf>

⁶ "Registrar shall require all Registered Name Holders to enter into an electronic or paper registration agreement with Registrar including at least the following provisions ..."
<http://www.icann.org/registrars/ra-agreement-17may01.htm#3.7.7>

data.⁷ In fact, ICANN specifically allows registrars to maintain domains even in the face of intentionally invalid Whois data constituting a material breach of the domain registration agreement.⁸

2. For failure to assure Whois accuracy, registrars face only a toothless sanction from ICANN, and ICANN isn't even making meaningful use of this approach. Pursuant to ICANN's April 3, 2003 advisory to registrars,⁹ and as took place in September 2002,¹⁰ ICANN may present a registrar with a formal notice of breach if, in ICANN's view, the registrar "appears to routinely ignore reports of inaccurate and incomplete contact data in its Whois database." However, only one such notice has been issued to date; it reported inaccuracies in only seventeen domains; its recipient was a registrar not typically thought to harbor particularly egregious cases of invalid Whois data; the only resulting sanction was brief public embarrassment for the registrar, without financial penalty. Registrars are unlikely to respond to such sporadic enforcement by ICANN.

In contrast, registrars face clear incentives to allow inaccurate Whois data.

1. The costs of inaccurate Whois data fall not on registrars but on others – on law enforcement officials, on consumers, and on those wishing to pursue copyright, trademark, and other claims against domain name registrants.
2. A registrar that enforces Whois accuracy requirements faces increased costs relative to a registrar that ignores Whois accuracy. Increased costs include staff time to seek out errors and respond to customer complaints, as well as software to automate these processes.
3. A registrar that enforces Whois accuracy requirements may face lost revenue by driving certain customers to other registrars. In particular, large registrants with systematic intentionally invalid contact information (such as the registrants described in my *NicGod* and *Typographical Errors* research) are likely to select registrars that allow or tolerate invalid contact information.

The following policy changes would correct these incentive problems:

1. ICANN could commission audits of Whois accuracy, using statistically valid methods to examine a significant sample of domains. Results would be tabulated and published on a per-registrar basis, allowing comparisons of Whois accuracy among registrars.

⁷ e.g. "take reasonable steps to investigate" – <http://www.icann.org/registrars/ra-agreement-17may01.htm#3.7.8>

⁸ <http://www.icann.org/announcements/advisory-03apr03.htm>

⁹ <http://www.icann.org/announcements/advisory-03apr03.htm>

¹⁰ <http://www.icann.org/correspondence/touton-letter-to-beckwith-03sep02.htm>

2. ICANN could use the results of Whois accuracy audits to present registrars with formal notices of breach of their contracts with ICANN.
3. If formal notices of breach fail to encourage registrars to improve their performance on Whois accuracy audits, ICANN could implement a system of graduated financial sanctions, consistent with ICANN's practice for registry service level agreements.¹¹
4. ICANN could post periodic statistics as to Whois Data Problem Reports, Registrar Problem Reports, and registrars' actions taken in response to these complaints.

All registrars would face these policy changes simultaneously and equally. Across-the-board enforcement of Whois accuracy would prevent registrants from switching registrars to avoid Whois enforcement efforts.

Privacy Concerns Reflect Misguided Overemphasis

In response to calls for increased Whois accuracy and enforcement, some have raised privacy concerns.¹² Their typical worry is that an emphasis on Whois accuracy would purportedly prevent individuals from registering domains for purposes that are in some way controversial yet simultaneously commendable (e.g. political dissent, whistle-blowers, or other anonymous speech).

Policymakers rightly encourage the use of the Internet for activities legitimately requiring anonymity. However, such activities are in no way incompatible with accurate Whois data. Domain registrants who wish to keep their name and address confidential can register names through one of several third-party services specializing in privacy protection¹³ or can register names through an attorney or other representative. It is not necessary to sacrifice Whois accuracy in order to preserve the possibility of anonymous publication on the web.

Distinct from the privacy concerns typically offered in response to calls for Whois accuracy, are concerns as to publication of truthful email addresses, for fear of receiving unsolicited email. In the past, such emails have included offers from registrars and web hosting companies. More recently, email worms and viruses have proven particularly disruptive. I am sympathetic to these concerns, but the proper response is not to discard all calls for Whois accuracy. Indeed, email concerns in no way lessen the need for accurate registrant name, address, and telephone information. Instead, those who find bulk email problematic can route their email through any of various mail filtering services, or can rely on temporary "alias" email addresses. Certain registrars already offer this email alias feature,

¹¹ e.g. <http://www.icann.org/tlds/agreements/name/registry-agmt-apppe-02jul01.htm>

¹² e.g. Electronic Privacy Information Center – Whois. <http://www.epic.org/privacy/whois/>

¹³ e.g. GoDaddy's Private Registration service, https://registrar.godaddy.com/dbp.asp?isc=&se=%2B&pl_id=1&prog_id=GoDaddy

typically at no additional charge.¹⁴ In any case, recent research indicates that Whois records are not a significant source of spam.¹⁵

Trends in Registrar Compliance with ICANN Policies

I understand that the subcommittee is also concerned about the possibility that certain registrars systematically tend not to comply with relevant ICANN policies. In particular, despite obligations under the Registrar Accreditation Agreement, certain registrars apparently ignore selected UDRP judgments calling for transfer of domains away from the registrars and their registrant customers. New York attorney Martin Schwimmer publicly raised this issue in a blog entry of June 4, 2003,¹⁶ and I have subsequently attempted to quantify the scope of the problem in my *Compliance with UDRP Decisions: A Case Study of Joker.com*.¹⁷ I have found significant evidence that registrar Joker.com, perhaps among others, systematically fails to abide by its contractual obligation to transfer domains subsequent to orders received from UDRP panels.

To assure that registrars comply with their contractual obligations to ICANN, ICANN could establish a procedure for formally receiving, processing, and acting on complaints against registrars, ultimately upon threat of termination of an offending registrar's Accreditation Agreement. At present, ICANN's investigative procedures are ad hoc, and many complaints therefore fall through the cracks – with extended delays before ICANN takes action, if it does so at all. A more formal method of passing complaints to ICANN – complete with web-based publication of complaints, status, and disposition – would assure that ICANN acts promptly and transparently in resolving these situations.

The Special Problems of .US

The Department of Commerce has a special ability to shape policy in the United States' country-code top-level domain, .US. In particular, the DoC has a direct contractual relationship with .US registry Neustar, allowing DoC to directly specify .US policies. (In contrast, DoC's influence over policies in .COM, .NET, etc. require passing through DoC's Memorandum of Understanding with ICANN and subsequently through ICANN's policy-making process.) In this context, it is particularly desirable to assure that .US Whois rules and associated registration policies are fully in order.

¹⁴ e.g. <http://www.namescout.com/master/privacyfeatures.asp>

¹⁵ "Addresses posted in ... Whois domain name registries ... did not receive any spam during the six weeks of the investigation." <http://www.ftc.gov/bcp/conline/pubs/alerts/spamalrt.htm> "Domain name registration does not seem to be a major source of spam."

<http://www.cdt.org/speech/spam/030319spamreport.pdf> .

¹⁶ <http://trademark.blog.us/blog/2003/06/04.html#a646>

¹⁷ <http://cyber.law.harvard.edu/people/edelman/udrp-compliance>

.US Whois Policy

Neustar's .US Policies page¹⁸ makes no mention of a .US policy as to Whois accuracy or registry procedures for assuring Whois accuracy. NeuStar's Registration Review Policy¹⁹ references "Accuracy of information," but places this section at heading six on page three of a PDF file (easy for registrants to miss) and fails to use the word "Whois" to make clear to registrants what specific information is at issue. Improvements in these areas are necessary to assure .US's position as a leader in Whois accuracy.

.US Nexus Requirements

Closely related to .US Whois rules are .US nexus requirements for registration. Under the .US Nexus Requirements,²⁰ .US domains may be registered only by 1) US citizens or residents, 2) US entities or organizations, and 3) foreign entities or organizations with a bona fide presence in the US. In practice, however, .US domains are registered by a variety of entities meeting none of these criteria. Furthermore, these entities often register a large number of domains – as many as 800 per registrant, in my research – and their domains often infringe on the marks of others. These practices are documented in my *Survey of Usage of the .US TLD*.²¹ However, despite discussion list coverage of this research, as well as numerous personal emails from concerned citizens to staff at the Department of Commerce and at Neustar, I gather these registrations remain in effect, in many instances with new invalid Whois information replacing the old.

If existing procedures fail to separate US from non-US registrants – on the basis of what could initially have been presumed to have been truthful registrant contact information – their ability to perform the more subtle task of separating valid Whois contact data from invalid entries ought to be very much in question. Here too, improvement likely requires setting appropriate incentives – requiring Neustar to face a penalty when it allows the registration of scores of domains with invalid Whois data or with invalid nexus qualifications.

The Unavailability of the .US Zone File

The Department of Commerce's agreement with Neustar apparently fails to provide the public with access to the .US zone file (the list of all registered .US domain names). Zone files are essential for conducting research as to trends in domain registrations, and public access to zone files is therefore a cornerstone of all ICANN contracts with gTLDs. However, Neustar reports that the DoC has failed to provide for such access under its .US contract with Neustar, and Neustar staff refuse to distribute the file to the public until the DoC and Neustar agree on terms for doing so. As a result, research and public critique of .US registrations and policies are rendered considerably more difficult, and it was

¹⁸ <http://www.nic.us/policies>

¹⁹ http://www.nic.us/policies/docs/registration_review_policy.pdf

²⁰ http://www.nic.us/policies/docs/ustld_nexus_requirements.pdf

²¹ <http://cyber.law.harvard.edu/people/edelman/dotus>

only with considerable additional effort that I was able to conduct the *Survey* referenced above.

Lack of Related Efforts by the .US Policy Council

.US policy is to be set in consultation with a .US Policy Council, formed by Neustar in 2002. However, the status of this Council is unclear, with no meeting minutes posted since January 2003.²² My sense is that this period has brought a similar lack of forward progress on .US Whois accuracy, nexus requirements, zone file availability, and other .US policy issues.

²² <http://www.neustar.us/policycouncil>