

GAO

Report to the Subcommittee on Courts,
the Internet, and Intellectual Property,
House of Representatives

November 2005

INTERNET MANAGEMENT

Prevalence of False Contact Information for Registered Domain Names





INTERNET MANAGEMENT

Prevalence of False Contact Information for Registered Domain Names

Highlights of [GAO-06-165](#), a report to the Subcommittee on Courts, the Internet, and Intellectual Property, Committee on the Judiciary, House of Representatives

Why GAO Did This Study

Individuals or organizations seeking to register the names of their Web sites may provide inaccurate contact information to registrars in order to hide their identities or to prevent members of the public from contacting them. Contact information is made publicly available on the Internet through a service known as Whois. Data accuracy in the Whois service can help law enforcement officials to investigate intellectual property misuse and online fraud, or identify the source of spam e-mail, and can help Internet operators to resolve technical network issues.

GAO was asked, among other things, to (1) determine the prevalence of patently false or incomplete contact data in the Whois service for the .com, .org, and .net domains; (2) determine the extent to which patently false data are corrected within 1 month of being reported to ICANN; and (3) describe steps the Department of Commerce (Commerce) and ICANN have taken to ensure the accuracy of contact data in the Whois database.

What GAO Found

Based on test results, GAO estimates that 2.31 million domain names (5.14 percent) have been registered with patently false data—data that appeared obviously and intentionally false without verification against any reference data—in one or more of the required contact information fields. GAO also found that 1.64 million (3.65 percent) have been registered with incomplete data in one or more of the required fields. In total, GAO estimates that 3.89 million domain names (8.65 percent) had at least one instance of patently false or incomplete data in the required Whois contact information fields. The table below shows the estimated number of instances of patently false data for each of the three types of contact information within each generic top-level domain.

Of the 45 error reports that GAO submitted to the Internet Corporation for Assigned Names and Numbers (ICANN) for further investigation—one for each domain name with patently false contact data that GAO found in a random sample of 900—11 domain name holders provided updated contact information that was not patently false within 30 days after GAO submitted the error reports to ICANN. One domain name, which had been pending deletion before submission to ICANN, was terminated after GAO submitted the error report. The remaining 33 were not corrected.

Commerce and ICANN have taken steps to ensure the accuracy of contact data in the Whois database. In addition to implementing a Registrar Accreditation Agreement that requires registrars to investigate and correct any reported inaccuracies in the contact information, they have amended their memorandum of understanding to require ICANN to continue assessing the operation of the Whois service and to implement measures to secure improved accuracy of data.

Commerce and ICANN officials generally agreed with a draft of this report.

Prevalence of Patently False Contact Information (in millions; percentages in parentheses)

Data	Registrant			Administrative contact			Technical contact		
	.COM	.ORG	.NET	.COM	.ORG	.NET	.COM	.ORG	.NET
Not patently false	33.13 (92.65)	3.29 (93.69)	5.34 (94.26)	31.90 (89.20)	3.15 (89.77)	5.21 (91.88)	32.18 (89.98)	3.18 (90.63)	5.29 (93.37)
Patently false	1.18 (3.30)	0.10 (2.97)	0.05 (0.89)	1.86 (5.20)	0.22 (6.25)	0.18 (3.13)	1.50 (4.18)	0.19 (5.51)	0.16 (2.76)
Incomplete	0.27 (0.76)	0.07 (2.09)	0.17 (2.98)	0.83 (2.31)	0.11 (3.09)	0.18 (3.13)	0.91 (2.54)	0.10 (2.97)	0.11 (2.01)
Unable to access Whois data	1.18 (3.30)	0.04 (1.25)	0.11 (1.86)	1.18 (3.30)	0.04 (1.25)	0.13 (2.24)	1.18 (3.30)	0.04 (1.25)	0.13 (2.24)

Source: GAO analysis of test results.

Note: Margin of error is ±5 percent or less at the 95 percent confidence level. Some domain names contained both patently false and incomplete information and so percentages do not add up to 100.

www.gao.gov/cgi-bin/getrpt?GAO-06-165.

To view the full product, including the scope and methodology, click on the link above. For more information, contact Linda Koontz, 202-512-6250, koontzl@gao.gov, or Keith Rhodes, 202-512-6412, rhodesk@gao.gov.

Contents

Letter

1

Appendix

Appendix I: Prevalence of False Contact Information for Registered Domain Names

5

Abbreviations

DNS	domain name system
ICANN	Internet Corporation for Assigned Names and Numbers
IP	Internet Protocol
IRIS	Internet Registry Information Service
MOU	memorandum of understanding
RAA	Registrar Accreditation Agreement

This is a work of the U.S. government and is not subject to copyright protection in the United States. It may be reproduced and distributed in its entirety without further permission from GAO. However, because this work may contain copyrighted images or other material, permission from the copyright holder may be necessary if you wish to reproduce this material separately.



United States Government Accountability Office
Washington, D.C. 20548

November 4, 2005

The Honorable Lamar Smith
Chairman
Subcommittee on Courts, the Internet,
and Intellectual Property
Committee on the Judiciary
House of Representatives

The Honorable Howard Berman
Ranking Minority Member, Subcommittee on Courts
the Internet, and Intellectual Property
Committee on the Judiciary
House of Representatives

Individuals or organizations seeking to establish sites on the World Wide Web are required to register the names of the sites with authorized domain name registrars. These registrars, who operate under agreement with the Internet Corporation for Assigned Names and Numbers (ICANN), also collect contact information from the registrants and make the information publicly available on the Internet through a service known as Whois. Although registrants are required to provide accurate contact information during the domain name registration process, they may supply false or incomplete information in order to hide their identities or to shield themselves from being contacted by members of the public.

This report responds to your request that we (1) determine the prevalence of patently false¹ or incomplete contact data in the Whois service for the three “legacy” generic top-level domains (.com, .org, and .net); (2) determine the extent to which patently false data identified through our analysis were corrected within 1 month of being reported to ICANN and the types of businesses associated with the domain names containing patently false data; (3) describe the steps the Department of Commerce (Commerce) and ICANN have taken to ensure the accuracy of contact data in the Whois database; and (4) describe the tools and techniques intended to reduce the amount of false information in the Whois service.

¹For the purpose of this report, we define “patently false data” as data that appeared obviously and intentionally false without verification against any reference data.

To address the first objective, we obtained “zone files” maintained by Verisign, Inc., and the Public Interest Registry.² These files listed all registered Internet domain names for the three legacy generic top-level domains as of February 2005. After selecting random samples of 300 domain names from each of the three zone files for .com, .net, and .org, we performed online Whois searches to obtain contact information for each domain name. Finally, we assessed the contact information for each domain name in our random samples to identify data that are incomplete or patently false. To address the second objective, we submitted error reports to ICANN for Whois data entries we identified as patently false and re-examined the same entries after 30 days to determine whether actions had been taken to correct the false data. For the third objective, we interviewed officials from federal agencies and ICANN to identify actions taken to improve the accuracy of contact data in the Whois database, and reviewed the memorandum of understanding between Commerce and ICANN and other contractual agreements. For the final objective, we obtained and documented information from federal agency officials and selected registrars regarding the availability of tools and technologies that could aid in reducing the false contact data in the Whois service. We completed our work in Washington, D.C. between December 2004 and August 2005 in accordance with generally accepted government auditing standards.

In summary, we estimate that 2.31 million domain names (5.14 percent) have been registered with patently false data in at least one of the required contact information fields. In addition, we estimate that 1.64 million domain names (3.65 percent) have incomplete information in one or more of the required fields. In total, we estimate that 3.89 million domain names (8.65 percent) had at least one instance of patently false or incomplete data in the required Whois contact information fields.

Of the 45 error reports that we submitted to ICANN for further investigation—one for each domain name with patently false contact data that we found in our random sample of 900—11 domain name holders provided updated contact information that was not patently false within 30 days after we submitted the error reports to ICANN. One domain name, which had been pending deletion before our submission to ICANN, was terminated after we submitted the error report. The remaining 33 were not

²Verisign, Inc. is the designated administrator (called a registry) that is responsible for managing domain names and setting policy for the .net and .com top-level domains. The Public Interest Registry is responsible for managing the .org domain.

corrected. Of the 45 domain names, 19 were Web sites that were unavailable, under construction, or had no significant content, while 6 had unknown foreign-language content. The remaining 20 were associated with a wide variety of businesses, including Web search portals, adult content and merchandise, IT consulting services and information, general information, retail merchandise, and other online services.

Commerce and ICANN have taken steps to ensure the accuracy of contact data in the Whois database, including implementing a Registrar Accreditation Agreement that requires registrars to investigate and correct any reported inaccuracies in Whois contact information for the domain names they register, and an amendment to their memorandum of understanding that required ICANN to implement measures to improve the accuracy of Whois data. ICANN has also published additional information and guidance for registrars regarding their obligations to investigate and correct data inaccuracies, and implemented a system to receive and track complaints about inaccurate and incomplete data. ICANN recognizes that more can be done and is planning to take further steps, including enhancing the system, hiring additional staff to conduct follow-up to ensure that reported inaccuracies are addressed, and seeking recommendations from a task force formed to address data accuracy issues.

We identified two technologies and tools intended to help reduce false contact information in the Whois database. They are (1) the Internet Registry Information Service protocol, which provides tiered access to sensitive contact information and, thus, would encourage the submission of more accurate information; and (2) Support Intelligence's *Trust Factor* product, which could be used to assess the validity of contact information against public information stored in commercial databases. While both tools have the potential to help reduce false contact information, neither is widely implemented by registrars and registries. We did not determine the effectiveness of such technologies and tools in reducing inaccuracies in the Whois service.

On August 30, 2005, we provided your staff with briefing slides on the results of our study. This report provides you with the published briefing slides, included as appendix I to this report. We received comments, via E-mail, on a draft of this report from the Deputy Chief Counsel of Commerce's National Telecommunications and Information Administration, and ICANN's Deputy General Counsel. Both Commerce

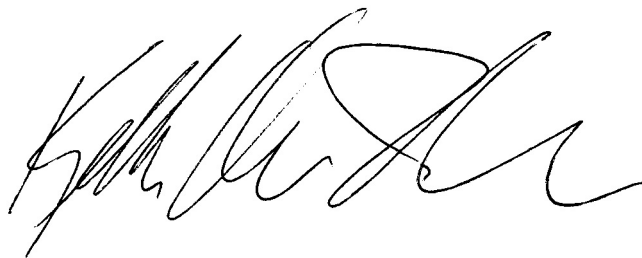
and ICANN generally agreed with the information presented in the draft report. A technical comment provided by Commerce has been addressed as appropriate.

As we agreed with your office, unless you publicly announce the contents of this report earlier, we plan no further distribution until 30 days from the date of this letter. At that time, we will send copies of this report to the Secretary, Department of Commerce; Chairman and Ranking Minority Members, House Committee on the Judiciary; and other interested congressional committees. Copies of this report will also be made available to others upon request. In addition, this report will be available at no charge on the GAO Web site at <http://www.gao.gov>.

If you or your staff have any questions concerning this report, please contact Linda Koontz at (202) 512-6240 or koontzl@gao.gov; or Keith Rhodes at (202) 512-6412, or rhodesk@gao.gov. Contact points for our Offices of Congressional Relations and Public Affairs may be found on the last page of this report. Other major contributors to this report included James Ashley, Barbara Collier, John de Ferrari, Mark Fostek, Wilfred Holloway, Steven Law, and Amos Tevelow.



Linda D. Koontz
Director, Information Management Issues



Keith A. Rhodes
Chief Technologist
Director, Center for Technology and Engineering

Prevalence of False Contact Information for Registered Domain Names



Internet Management

Prevalence of False Contact Information for Registered Domain Names

Subcommittee on Courts, the Internet, and Intellectual Property
House Committee on the Judiciary

August 30, 2005



Table of Contents

Introduction

Objectives

Scope and Methodology

Results in Brief

Background

Prevalence of Patently False Contact Information

Correction of Reported Patently False Contact Information

Steps Taken to Ensure Accuracy

Tools and Technologies That Could Reduce False Contact Information

Summary

Agency Comments



An individual or organization seeking to establish a site on the World Wide Web is required to register the name of the site with an authorized domain name *registrar*.

These registrars operate under agreement with the Internet Corporation for Assigned Names and Numbers (ICANN), which is charged with administering the Internet's name and address system, known as the *domain name system* (DNS).¹ ICANN's authority is based on a memorandum of understanding it has with the Department of Commerce.²

Based on their accreditation agreements with ICANN, registrars require all prospective Web site *registrants* to provide contact information, which is then made publicly available on the Internet through a service known as *Whois*.

¹The DNS is an Internet directory service that controls the delivery of electronic mail and translates domain names into numerical Internet Protocol (IP) addresses, which computers use to communicate with each other over the Internet.

²For more information on Commerce's relationship with ICANN, see GAO, *Internet Management: Limited Progress on Privatization Project Makes Outcome Uncertain*, GAO-02-805T (Washington, D.C.: June 12, 2002).



The Whois service was originally intended as a source of contact information that technicians could use to reach each other when necessary to troubleshoot problems with Internet connectivity or functionality. However, users of the Whois service have broadened over time to include law enforcement officials, owners of intellectual property, and others seeking contact information about Web site owners for a variety of reasons.

Although registrants are required to provide accurate Whois contact information, they may supply false or incomplete information in order to hide their identities or to shield themselves from being contacted by members of the public.



Objectives

Our objectives were to

- determine the prevalence of “patently false” or incomplete contact data in the Whois service for the three “legacy” generic top-level domains;³
- determine the extent to which patently false data identified through our analysis were corrected within 1 month of being reported to ICANN and the types of businesses associated with the domain names containing patently false data;
- describe steps the Department of Commerce and ICANN have taken to ensure the accuracy of contact data in the Whois database; and
- describe tools and techniques intended to aid in reducing the amount of false data in the Whois service.

³“Legacy” generic top-level domains consist of all Internet addresses that end in .com, .org, and .net.



To address these objectives, we took the following actions:

- We obtained “zone files” maintained by Verisign, Inc., and the Public Interest Registry, which list all registered Internet domain names for the three legacy generic top-level domains (.com, .org, and .net) as of February 2005. Verisign, Inc., and the Public Interest Registry reported that there were 44.93 million registered domain names for these three domains in February 2005.⁴
- We selected random samples of 300 domain names from each of the three zone files and performed online Whois look-ups to obtain contact information for each domain name.
- We assessed the contact information for each domain name in our random samples to identify data that are incomplete or patently false—data that appeared obviously and intentionally false without verification against any reference data, such as “(999) 999-9999” for a telephone number, “asdasd” for a street address, or “XXXXX” for a postal code.

⁴Verisign Inc., *Registry Operator’s Monthly Report* (Dulles, Va.: February 2005) and Public Interest Registry, *Registry Operator’s Report* (Reston, Va.: February 2005).



Scope and Methodology

- We submitted error reports to ICANN for Whois data entries we identified as patently false and re-examined the same entries after 30 days to determine whether actions had been taken to correct the false data. We did not assess the reasons why reported inaccuracies were not all corrected.
- We interviewed federal officials from Commerce, the Department of Justice, the Federal Trade Commission, the U.S. Securities and Exchange Commission, and ICANN to identify actions taken to improve the accuracy of contact data in the Whois database.
- We reviewed the memorandum of understanding between Commerce and ICANN and other contractual agreements.
- We obtained and documented information from federal agency officials and selected registrars regarding the availability of tools and technologies that could aid in reducing the amount of false contact data in the Whois service.



Scope and Methodology

Based on results from our random samples, we estimated the number of domain names with patently false and incomplete data. All estimates in this report have a margin of error of plus or minus 5 percent or less at the 95 percent confidence level.

Our work was completed in accordance with generally accepted government auditing standards in Washington, D.C., between December 2004 and August 2005.



We estimate that 2.31 million (5.14 percent) domain names have been registered with “patently false” Whois contact data—data that appeared obviously and intentionally false without verification against any reference data—in at least one of the required contact information fields. In addition, we estimate that 1.64 million domain names (3.65 percent) have incomplete information in one or more of the required fields.

Of the 45 error reports that we submitted to ICANN for further investigation—one for each domain name with patently false contact information that we found in our random sample of 900—11 domain name holders provided updated contact information that was not patently false within 30 days after we submitted the error reports to ICANN. One domain name, which had been pending deletion before our submission to ICANN, was terminated after we submitted the error report. The remaining 33 were not corrected. Of the 45 domain names we submitted reports on, 19 were for Web sites that were unavailable, under construction, or had no significant content, while 6 had unknown foreign-language content. The remaining 20 were associated with a wide variety of businesses, including Web search portals, adult content and merchandise, IT consulting services and information, general information, retail merchandise, and other online services.



Commerce and ICANN have taken the following steps to ensure the accuracy of contact data in the Whois database:

- ICANN implemented a Registrar Accreditation Agreement that requires registrars to investigate and correct any reported inaccuracies in Whois contact information for the domain names they register. It also published additional information and guidance for registrars regarding their obligations to investigate and correct data inaccuracies, and implemented a system to receive and track complaints about inaccurate and incomplete data.
- In a September 2003 amendment to its memorandum of understanding (MOU) with ICANN, Commerce required ICANN to continue assessing the operation of the Whois service and to implement measures to secure improved accuracy of Whois data.

ICANN recognizes that more can be done to ensure the accuracy of Whois data and is planning to take further steps, including having its staff follow up to ensure that reported inaccuracies are addressed.



Available technologies and tools intended to reduce the amount of false contact information entered into the Whois service include the following:

- The Internet Registry Information Services protocol, which provides for tiered access to sensitive contact information, could be used to restrict public access to sensitive personal information within the Whois data. According to proponents, this restriction would encourage the submission of more accurate data.
- Commercial screening tools, such as Support Intelligence's *Trust Factor* product, could be used to assess the validity of contact information as it is entered by registrants by verifying the registrant information against public information stored in commercial databases.

Neither the protocol nor the screening tools are widely implemented by registrars and registries. We did not determine the effectiveness of such technologies and tools in reducing inaccuracies in the Whois service.

Both Commerce and ICANN generally agreed with the information presented in a draft of this briefing and provided technical comments which have been addressed as appropriate.



Background
The Internet Domain Name System

The *domain name system* (DNS) is a vital component of the Internet that works like an automated telephone directory, allowing users to reach Web sites using easy-to-understand domain names (e.g., www.gao.gov) instead of the numeric Internet Protocol (IP) addresses (e.g., 161.203.16.2) that computers use when communicating with each other.⁵ The DNS consists of a series of *name servers* that store data linking numeric IP addresses with their associated domain names.

The letters at the far right of a domain name (e.g., “gov” in www.gao.gov) represent *top-level domains* and include well-known generic domains such as .com, .net, and .org. The next string of text to the left (“gao” in www.gao.gov) in an address is called a second-level domain and is a subset of the top-level domain.

⁵For more information on the domain name system, see GAO, *Internet Management: Limited Progress on Privatization Project Makes Outcome Uncertain*, GAO-02-805T (Washington, D.C.: Jun. 12, 2002).



Background
The Internet Domain Name System

Each top-level domain has a designated administrator (called a *registry*) that is responsible for managing domain names and setting policy for the domain. For example, the Public Interest Registry—a not-for-profit corporation—manages the .org top-level domain, and Verisign, Inc. manages the .net and .com registries.

Registrars are organizations (usually private companies) that support registries by selling domain name registration services to *registrants* (the “owners” of specific domain names). During the registration process for .org, .net, and .com names, registrars collect information from registrants that includes three types of contacts:

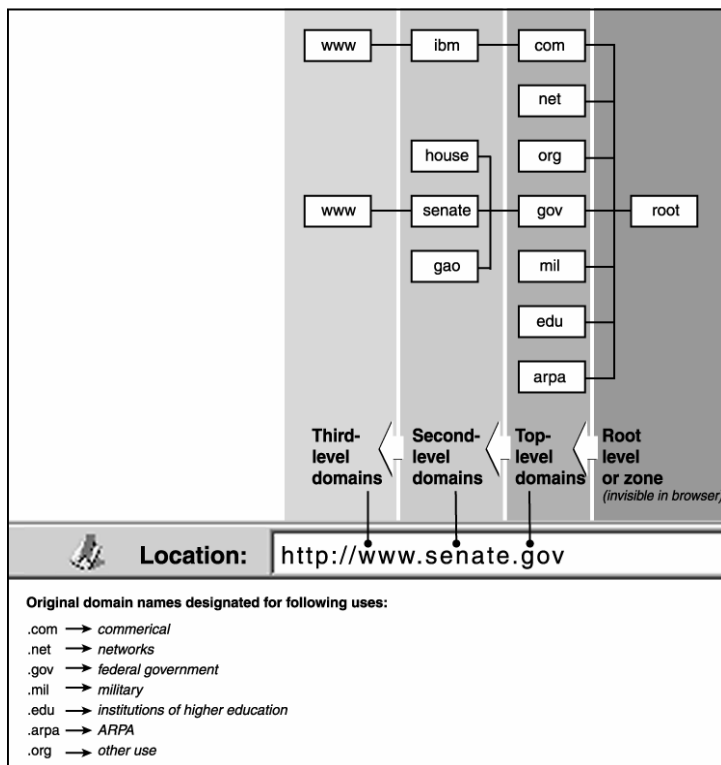
- the domain name registrant,
- an administrative contact, and
- a technical contact.

Registrars maintain this information in their individual databases and make it available to the public through their own Whois service. There is no unified Whois service containing all registrant data for the .org, .net, or .com registries.



Background The Internet Domain Name System

The figure depicts the hierarchical organization of Internet domain names for the original seven generic top-level domains, including the three “legacy” domains, .com, .org, and .net.⁶



⁶In November 2000, seven additional top-level domains (.info, .biz, .name, .aero, .museum, .coop, and .pro) were introduced.

Source: GAO.



Background
Relationship between Commerce and ICANN

The U.S. government supported the development of the Internet and the DNS and has the authority to make key decisions affecting the DNS.

In 1997, the President tasked the Department of Commerce with transitioning the DNS to private management. Commerce selected ICANN, a not-for-profit private corporation, to carry out the transition and to demonstrate that it had the resources and capability to manage the DNS.⁷ Accordingly, in November 1998, Commerce entered into an agreement with ICANN, in the form of an MOU,⁸ to jointly develop “mechanisms, methods, and procedures” necessary to transfer DNS management to the private sector.

The MOU states that before making a transition to private sector management, Commerce requires assurance that the private sector has the capability and resources to manage the DNS. To gain this assurance, Commerce and ICANN agreed in the MOU to complete a set of transition tasks.

⁷For more information on the relationship between Commerce and ICANN, see GAO, *Department of Commerce: Relationship with the Internet Corporation for Assigned Names and Numbers*, GAO/OGC-00-33R (Washington, D.C.: Jul. 7, 2000).

⁸The original MOU was set to expire in September 2000 and has been amended six times. The latest amendment, in September 2003, will expire on September 30, 2006.



Background
Relationship between Commerce and ICANN

As established in the MOU, Commerce and ICANN agreed to perform the following activities in support of the joint DNS project:

- Commerce is to provide advice, coordinate with foreign governments, and generally oversee activities conducted by ICANN as part of the MOU.
- ICANN is to design, develop, and test procedures for managing the DNS.

In June 2005, the Assistant Secretary of Commerce's National Telecommunications and Information Administration stated that Commerce would **continue to** provide oversight indefinitely to ensure that ICANN continued to focus on meeting its core technical mission.



Background
Relationship between ICANN and Registrars

In 1999, ICANN developed a Registrar Accreditation Agreement (RAA) that sets the terms by which accredited registrars are authorized to register domain names within the generic top-level domains.

As of May 2005, about 400 accredited registrars from the United States and foreign countries offered domain name registration services for the generic top-level domains.

As part of the terms of the RAA, each registrar is to provide a Web-based Whois service that offers free access to contact information on all active registered domain names sponsored by the registrar.



Background
Whois Provisions in the RAA

The ICANN RAA specifies that registrars are required to collect the following information, as a minimum, from each registrant:

- the domain name being registered;
- the names of the primary name server and any secondary name servers;
- the identity of the registrar;
- original creation and expiration dates of the registration;
- the registrant's name and postal address; and
- the name, postal address, electronic mail address, telephone number, and fax number (optional) for both the technical contact and the administrative contact for the domain name.

This information is then made publicly available through the Whois service.



Background
The Whois database Service

Created in the 1970s, Whois began as a service that Internet operators could use to identify and contact individuals or entities responsible for the operation of a computer on the Internet when an operational problem arose.

Since then, the Whois service has evolved into a tool used for many purposes, such as determining whether a domain name is available for registration, identifying the source of spam⁹ e-mail, enforcing intellectual property rights, and identifying and verifying online merchants.

The Whois service is not a single centrally managed database but consists of linked information that is collectively maintained in distributed databases by domain name registrars and registries.

⁹Spam is unsolicited “junk” e-mail that usually includes advertising for some product.



Background
Whois Data Accuracy

Data accuracy is important to the effectiveness of the Whois service in helping Internet operators to resolve technical network issues, as well as helping law enforcement officers to investigate such things as intellectual property misuse or online fraud. According to federal agency officials, accurate Whois data have the potential to allow law enforcement officials to identify individuals involved in criminal activities on the Internet more quickly than if such information were not available.

The ICANN RAA directs registrars to

- require registrants to agree that willfully submitting inaccurate contact details (or failing to respond within 15 days to an inquiry regarding accuracy) shall be a basis for cancellation of the registration and
- take reasonable steps to investigate and correct contact information in response to any reported inaccuracy.



Prevalence of Patently False Contact Information

Based on our test results, all three legacy generic top-level domains contained names with patently false Whois information. We estimate that patently false registrant data were entered for 1.18 million (3.30 percent) .com names, 100,000 (2.97 percent) .org names, and 50,000 (0.89 percent) .net names.

The table on the following page shows the estimated number of instances of patently false data (and the corresponding percentage of the entire sample) for each of the three types of contact information (“registrant,” “administrative contact,” and “technical contact”) within each generic top-level domain.

As previously described, “patently false” contact data in the Whois service are data that appear obviously and intentionally false without verification against any reference data. For example, “(999) 999-9999” for a telephone number, “asdadasd” for a street address, and “XXXXX” for a postal code represent patently false data.

**Appendix I
Prevalence of False Contact Information for
Registered Domain Names**



Prevalence of Patently False Contact Information

Data classification	Registrant			Administrative contact			Technical contact		
	.COM	.ORG	.NET	.COM	.ORG	.NET	.COM	.ORG	.NET
Not patently false	33.13 92.65%	3.29 93.69%	5.34 94.26%	31.90 89.20%	3.15 89.77%	5.21 91.88%	32.18 89.98%	3.18 90.63%	5.29 93.37%
Patently false	1.18 3.30%	0.10 2.97%	0.05 0.89%	1.86 5.20%	0.22 6.25%	0.18 3.13%	1.50 4.18%	0.19 5.51%	0.16 2.76%
Incomplete	0.27 0.76%	0.07 2.09%	0.17 2.98%	0.83 2.31%	0.11 3.09%	0.18 3.13%	0.91 2.54%	0.10 2.97%	0.11 2.01%
Unable to access Whois data	1.18 3.30%	0.04 1.25%	0.11 1.86%	1.18 3.30%	0.04 1.25%	0.13 2.24%	1.18 3.30%	0.04 1.25%	0.13 2.24%

Source: GAO analysis of test results.

Note: The estimated numbers in the table are in millions with a margin of error of ± 5 percent or less at the 95 percent confidence level. Also, some domain names contained both patently false and incomplete information. As a result, the percentages do not add to one hundred.



Prevalence of Patently False Contact Information

Looking across the domains, we estimate that 2.31 million (5.14 percent) domain names had at least one data field populated with patently false data for the registrant, administrative contact, and/or technical contact. The estimates by domain are as follows:¹⁰

- For the .com domain, 1.86 million domain names (5.20 percent) contained patently false data in at least one contact information field.
- For the .org domain, 0.25 million domain names (7.17 percent) contained patently false data in at least one contact information field.
- For the .net domain, 0.20 million domain names (3.50 percent) contained patently false data in at least one contact information field.

¹⁰In some cases, patently false data had been entered for more than one of the three types of contacts. For this reason, the total estimates for each domain are less than the sum of the corresponding numbers in the previous table.



Prevalence of Patently False Contact Information

In addition to patently false Whois data, we estimate that 1.64 million domain names (3.65 percent) contained incomplete information in one or more of the required contact fields.

In total, we estimate that 3.89 million domain names (8.65 percent) had at least one instance of patently false or incomplete data in the required Whois contact information fields.

In addition to the domain names involving patently false or incomplete contact information, we estimate that 1.35 million (3.00 percent) domain names contained inaccessible Whois data for various reasons, such as registrars restricting access to the information and domain names that had expired or were deleted.

The remaining 39.70 million domain names (88.35 percent) contained no patently false or incomplete contact information in required data fields.¹¹

¹¹We did not verify that the “not patently false” information was accurate. It is possible that contact information that appeared to be valid for many domain names was not, in fact, accurate.



Correction of Reported Patently False Contact Information

The following actions occurred in response to the error reports that we submitted to ICANN¹² for the 45 domain names in our random sample of 900 that had patently false contact information:

- Within 30 days, 11 domain name holders (24 percent) had submitted updated contact information that was not patently false.
- One domain name in the .net domain was terminated (the domain name had been marked for deletion before our submission).

Contact information for the remaining 33 domain names (73 percent) remained unresolved 30 days after we submitted the error reports.

According to ICANN officials, registrars have the discretion to deactivate domain names by placing them on “registrar hold” as part of their investigation/response to reported inaccuracies. Of the 45 domain names with patently false contact information, ICANN stated that 11 were inactive.

¹²As discussed below, ICANN operates a Whois Data Problem Report System to accept reports of inaccuracies from the general public and forward them to registrars for resolution.



Correction of Reported Patently False Contact Information

The Web sites associated with the 45 registered domain names containing patently false contact information fell into the following categories:

- **Miscellaneous online services:** The domain name was associated with Web sites offering specific services, including movie downloads, electronic mail, online messages, real estate information, online gaming, and access to an organization's Intranet (6 Web sites or 13 percent).
- **Foreign language site (content unknown):** The domain name was associated with a Web site developed in a foreign language, and we were not able to determine the content of the site (6 Web sites or 13 percent).
- **IT consulting services and information:** The domain name was associated with a Web site offering IT consulting services or providing IT-related information (4 Web sites or 8.9 percent).
- **Web portals with online search capability:** The domain name was associated with a Web portal offering search capability for various services (3 Web sites or 6.7 percent).



Correction of Reported Patently False Contact Information

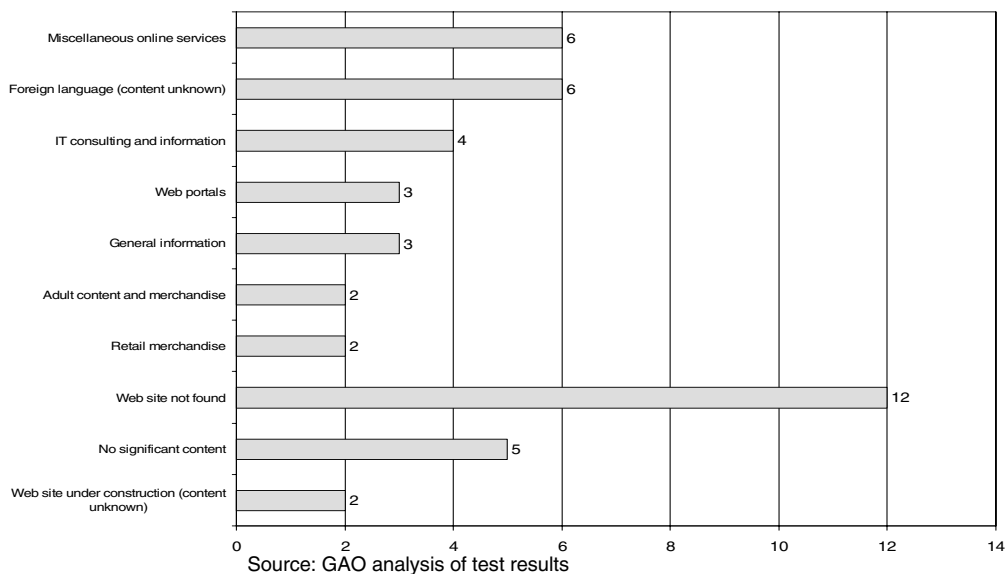
- General information: The domain name was associated with a Web site providing general information on miscellaneous topics, such as music or scouting (3 Web sites or 6.7 percent).
- Adult content or merchandise: The domain name was associated with a Web site offering adult content or merchandise for sale (2 Web sites or 4.4 percent).
- Retail merchandise: The domain name was associated with a Web site offering retail products for sale to the public (2 Web sites or 4.4 percent).
- Web site not found: The domain name was not associated with an active Web site; an error message was displayed when we attempted to access it (12 Web sites or 26.7 percent).
- No significant content: The domain name was associated with a Web site that was blank or had only limited content, such as a symbol or phrase (5 Web sites or 11 percent).
- Web site under construction (content unknown): The domain name was associated with a Web site designated as being under construction (2 Web sites or 4.4 percent).

Appendix I
Prevalence of False Contact Information for
Registered Domain Names



Correction of Reported Patently False
Contact Information

The results of our analysis are shown in the following chart.



Of the 45 domain names, 20 (44 percent) were associated with Web sites containing readily identifiable content, such as IT consulting services and information, Web search portals, adult content and merchandise, general information, retail merchandise, and other online services.



Steps Taken to Ensure Accuracy
Department of Commerce and ICANN

The Department of Commerce and ICANN have taken the following steps intended to ensure the accuracy of contact data in the Whois service:

- Commerce amended its MOU with ICANN to include a provision that ICANN continue assessing the operation of the Whois service and implement measures to improve the accuracy of Whois data.
- According to the amended MOU, ICANN was to publish (1) a report providing statistical and narrative information on its experience with the operation of the Whois Data Problem Report System and (2) a report providing statistical and narrative information on the implementation of the Whois Data Reminder Policy.
 - To meet its obligations under the MOU, ICANN published two reports on its experience with the report system (in March 2004 and March 2005) and one report on implementation of the reminder policy (in November 2004).
 - According to the March 2004 and 2005 reports, at least a quarter of the complaints submitted through the report system resulted in the correction of data or removal of a domain name.



Steps Taken to Ensure Accuracy
Department of Commerce and ICANN

- The November 2004 report, which was based on survey responses received from 254 out of 364 registrars (about 70 percent), concluded that it was difficult to determine the impact of the reminder policy on improving the accuracy of Whois data.

According to Commerce officials, one way the agency judges the success of ICANN is by the number of milestones it meets in fulfilling its obligations under the MOU. Thus far, according to the officials, ICANN has met its obligations regarding the Whois data accuracy provisions of the MOU by implementing a Whois Data Problem Report System and publishing annual reports on its experience with the operation of the system and the implementation of the Whois Data Reminder Policy.



Steps Taken to Ensure Accuracy
ICANN

ICANN's Registrar Accreditation Agreement requires each of its accredited registrars to investigate and correct any reported inaccuracies in Whois contact information for the domain names that they register.

After establishing the agreement, ICANN published the following four notices to provide additional information or guidance to registrars regarding their obligation to investigate and correct data inaccuracies:

- *Registrar Advisory Concerning Whois Data Accuracy*, May 10, 2002,
- *Steps to Improve Whois Data Accuracy*, September 3, 2002,
- *Registrar Advisory Concerning the "15-day Period" in Whois Accuracy Requirements*, April 3, 2003, and
- *Whois Data Reminder Policy Posted*, June 16, 2003.



Steps Taken to Ensure Accuracy
ICANN

In May 2002, ICANN published an advisory intended to aid registrars in understanding their obligations under the RAA regarding the accuracy of Whois data. Specifically, the advisory required all ICANN-accredited registrars to

- provide public access—through a Web-based Whois service—to contact information for all top-level domains covered under the RAA;
- require each registrant to submit (and keep updated) accurate contact information;
- notify registrants that willfully submitting inaccurate contact information (or failing to respond within 15 days to an inquiry regarding accuracy) would be a basis for cancellation of the registration; and
- take reasonable steps to investigate and correct Whois data in response to reported inaccuracies.



Steps Taken to Ensure Accuracy
ICANN

Subsequently, in September 2002, ICANN announced that it had implemented the Whois Data Problem Report System, designed to process complaints about inaccurate and incomplete Whois data. According to the announcement, many registrars have responded to reports of inaccurate and incomplete data submitted through the system.

The system is a Web-based tool accessible through the Web site <http://wdprs.internic.net>.

The tool is intended to help registrars meet their responsibilities to correct inaccurate or incomplete data.

Users submit complaints about false or inaccurate Whois data for a particular domain name to ICANN through the report system, and ICANN in turn notifies the appropriate registrar.



Steps Taken to Ensure Accuracy
ICANN

In April 2003, ICANN published an advisory aimed at providing clarification to registrars concerning the 15-day time limit for registrants to respond to an inquiry regarding accuracy before facing possible cancellation of a domain name registration.

- According to ICANN, the RAA does not require a registrar to cancel a registration in the event that a registrant fails to respond within 15 days.
- According to the RAA, registrars shall, upon notification by any person of an inaccuracy in the contact information associated with a registered name sponsored by the registrar, take reasonable steps to investigate the alleged inaccuracy.
- The requirement that registrars “take reasonable steps” is intended to give registrars the flexibility to determine what action should be taken when a registrant fails to respond to a notification of inaccuracy.



Steps Taken to Ensure Accuracy
ICANN

In March 2003, ICANN adopted the Whois Data Reminder Policy, which is also intended to contribute to improved Whois data accuracy.

- The policy requires registrars to present current Whois information to registrants for verification and remind them that submitting false Whois information can be grounds for cancellation of a domain name's registration.
- Because ICANN has few agreements with operators of country code top-level domains (2-letter top-level domains reserved mainly for disposition by national governments, such as .us for the United States), these operators are not required to comply with the policy. (ICANN's RAA currently applies only to registrations within generic top-level domains, not country code domains.)



Steps Taken to Ensure Accuracy
ICANN

In addition to the steps already taken, ICANN recognizes that more can be done to ensure the accuracy of Whois data and plans to implement additional measures:

- According to ICANN's proposed budget for fiscal year 2005–2006,¹³ it plans to improve the report system by having staff conduct follow-up to ensure that reported data inaccuracies are corrected or the domain names deleted. (As noted above, 33 of the 45 reports we submitted to the system remained uncorrected after 30 days.)
- According to ICANN officials, ICANN is recruiting three new employees specifically to promote registrar and registry compliance with ICANN policies and agreements.
- In addition, ICANN staff are to conduct studies to determine the overall data accuracy in the Whois service and develop a plan for improvement.
- ICANN has chartered a task force to recommend additional ways to improve the processes for notifying a registrar of inaccurate data, and for investigating and correcting inaccurate data in the Whois service.

¹³Internet Corporation for Assigned Names and Numbers, *Proposed Budget: Fiscal Year 2005-2006* (May 17, 2005).



Tools and Technologies That Could Reduce False Contact Information

According to registrar officials, false data are easily entered into the Whois service because most domain name registration systems are automated and do not check the accuracy of data submitted by registrants. Based on information collected from federal officials and one registrar, we identified the following two technologies and tools intended to help reduce false contact information in the Whois database:

(1) The *Internet Registry Information Service (IRIS)* protocol¹⁴ was originally designed to address privacy concerns associated with making Whois contact information publicly available.

- IRIS includes authentication mechanisms that allow for “tiered” access to Whois data. Under IRIS, only limited Whois information—such as the name and contact information for the registrar—would be available to the general public (the “bottom” tier of users). A second “top” tier implementing stricter access controls could provide more specific registrant information to a limited number of users, such as law enforcement officials conducting online fraud investigations.

¹⁴The Internet Registry Information Service is a protocol developed by the Internet Engineering Task Force’s Cross-Registry Internet Service Protocol Working Group.



Tools and Technologies That Could Reduce False Contact Information

- According to agency officials, implementing the IRIS protocol could encourage more registrants to submit accurate data into the Whois database.
- ICANN is investigating proposals to implement a tiered access system. However, according to the American Intellectual Property Law Association,¹⁵ a number of issues would need to be addressed before implementing the system, including
 - defining criteria for users to access data in the top tier,
 - establishing an authority to determine whether a potential user meets those criteria, and
 - determining the cost to implement such a system and who will bear it.

¹⁵Michael K. Kirk, Executive Director of the American Intellectual Property Law Association, *Comments of the American Intellectual Property Law Association On the Request for Comments Regarding The Preliminary Reports of the Whois Task Forces*, (Arlington, Va.), Available: <http://gnso.icann.org/mailing-lists/archives/whois-tf2-report-comments/doc00007.doc>, downloaded on April 28, 2005.



Tools and Technologies That Could Reduce False Contact Information

(2) Support Intelligence¹⁶ developed a commercial product called *Trust Factor*, (formerly known as *Fraudit*), which is intended to assess Whois data accuracy at the time of registration.

- The tool conducts a number of automated checks on a registrant's contact data during the registration process to identify a range of potential problems, such as invalid or undeliverable postal and e-mail addresses, and telephone numbers that are inoperable or not consistent with the addresses provided.
- To accomplish this, *Trust Factor* attempts to verify a registrant's information against a number of commercial databases of public information. The tool can correlate postal and e-mail addresses and can match telephone numbers with network addresses to verify that they are all valid.
- Some federal officials and registrars questioned the general approach used by *Trust Factor* because it could involve potentially incorrect assumptions about a registrant's location and how it correlates with addresses, ZIP codes, and phone numbers.

¹⁶Support Intelligence is a private firm headed by the chief executive officer of Alice's Registry, an accredited registrar.



Tools and Technologies That Could Reduce False Contact Information

- In addition, according to one registrar, performing checks of the kind done by *Trust Factor* could produce incorrect results in such cases as when an individual submits contact information for a third party registrant.
- According to registrar officials, an automated system that implements technologies similar to those in *Trust Factor* could incorrectly flag accurate information as false and thus prevent customers from legitimately registering domain names. Their registration systems generally include no more than a simple check to ensure that all required data fields are filled in (no blank data fields). However, as our test results showed, even this check is not always implemented.
- Furthermore, registrars stated that they did not have the resources to manually check Whois contact information for inaccurate data because of the volume of registrations they process on a continuing basis.

At the time of our review, neither the IRIS protocol nor the *Trust Factor* tool had been widely implemented by registrars and registries. We did not determine the effectiveness of these tools in reducing inaccuracies in the Whois service.



Based on our test results, we estimate that 5.14 percent of domain names had at least one instance of patently false contact information in the required Whois contact data fields. As identified in our random sample, 45 domain names had at least one data field populated with patently false data. In addition, for the 45 corresponding error reports we submitted to ICANN, most false Whois contact information (73 percent) was unresolved after 30 days.

ICANN recognizes that more can be done to ensure the accuracy of Whois data and has taken steps to address these concerns by establishing registrar policies and issuing advisories on how to address false or inaccurate Whois data. In addition, ICANN implemented an automated system for users to report false or inaccurate Whois data for further investigation and correction. ICANN is planning to take further steps to improve data accuracy through enhancement of the system and by seeking recommendations from a task force formed to address accuracy issues.

Given that inaccurate and incomplete data persist in the Whois service, ICANN's continued attention to this issue will be critical to achieving improvements in data accuracy.



Although we identified two kinds of technological tools intended to help reduce the amount of false contact information, neither of these tools has been widely implemented by registrars, and we did not determine the effectiveness of the tools in reducing data inaccuracies.



We received oral comments on a draft of this briefing from the Deputy Chief Counsel of Commerce's National Telecommunications and Information Administration and comments via e-mail from ICANN's Deputy General Counsel. Both Commerce and ICANN generally agreed with the information presented in the draft briefing and provided a number of technical comments which have been incorporated where appropriate.

GAO's Mission

The Government Accountability Office, the audit, evaluation and investigative arm of Congress, exists to support Congress in meeting its constitutional responsibilities and to help improve the performance and accountability of the federal government for the American people. GAO examines the use of public funds; evaluates federal programs and policies; and provides analyses, recommendations, and other assistance to help Congress make informed oversight, policy, and funding decisions. GAO's commitment to good government is reflected in its core values of accountability, integrity, and reliability.

Obtaining Copies of GAO Reports and Testimony

The fastest and easiest way to obtain copies of GAO documents at no cost is through GAO's Web site (www.gao.gov). Each weekday, GAO posts newly released reports, testimony, and correspondence on its Web site. To have GAO e-mail you a list of newly posted products every afternoon, go to www.gao.gov and select "Subscribe to Updates."

Order by Mail or Phone

The first copy of each printed report is free. Additional copies are \$2 each. A check or money order should be made out to the Superintendent of Documents. GAO also accepts VISA and Mastercard. Orders for 100 or more copies mailed to a single address are discounted 25 percent. Orders should be sent to:

U.S. Government Accountability Office
441 G Street NW, Room LM
Washington, D.C. 20548

To order by Phone: Voice: (202) 512-6000
TDD: (202) 512-2537
Fax: (202) 512-6061

To Report Fraud, Waste, and Abuse in Federal Programs

Contact:

Web site: www.gao.gov/fraudnet/fraudnet.htm

E-mail: fraudnet@gao.gov

Automated answering system: (800) 424-5454 or (202) 512-7470

Congressional Relations

Gloria Jarmon, Managing Director, JarmonG@gao.gov (202) 512-4400
U.S. Government Accountability Office, 441 G Street NW, Room 7125
Washington, D.C. 20548

Public Affairs

Paul Anderson, Managing Director, AndersonP1@gao.gov (202) 512-4800
U.S. Government Accountability Office, 441 G Street NW, Room 7149
Washington, D.C. 20548