

Search-ability Issue Report

This report is focused specifically on the areas where improving searchability may be helpful.

Table of Contents

Improving Search-ability of WHOIS Databases	2
Current Policy Environment	3
What are Some Reasons for Searching WHOIS data?.....	3
Who Benefits from Enhanced Search-ability?.....	4
Concerns Expressed to the TF About Enhancing Search-ability:.....	4
What Could Enhance Search-ability?	5
3.3.1.	6
The use of elements other than domain names as query keys.	6
3.3.2	7
Prompt updates of data elements.	7
3.3.3	7
Subcontracting Public Access to WHOIS databases.	7
3.3.4	7
Distributed query-based search functionality across all registrars.	7
3.3.5	8
Terms and conditions for query-based access.	8
Appendix A: Relevant Provisions from the RAA	9
Authors of this Report (Working Group Three)	15

Search-ability Issue Report

Improving Search-ability of WHOIS Databases

Our Task Force examined three kinds of improved search-ability:

- (A) Centralized public access to WHOIS databases across each gTLD,
- (B) The use of data elements different from the domain name as query keys, and,
- (C) The provision of still more advanced database query capabilities and centralized search services across Top Level Domains, including Country Code TLDs.

Tactics included collecting comments and generating statistics by first developing a survey, working with the data collected from it and comments about it from a variety of sources. We advanced our work agenda in accordance with survey statistics and responses to both the survey and comments sent to the General Assembly, specific comment areas for our Whois Task Force and teleconferences with invited guests, like the Registry Constituency, CCTld Constituency, and the CRISP working group among others.

Our [Survey responses](#) indicate that, among respondents, there is demand and support for all these services. In addition, the survey respondents expressed concern about privacy issues related to searching WHOIS data. Additional communications resulted in a nearly identical list of wants, needs, and concerns.

The current policy environment supports our findings; but is not being enforced. We have explored short and long term solutions to foster development of an open standards based mechanism for “centralized access” which allows several queryable elements, restricts output to desired information only, and searches on a per TLD level for now with more robust capabilities in the nearest future. This solution is sensitive to a variety of privacy concernsⁱ and supports work in progress at the IETF and after being further advanced has been included in this report.

This document contains summaries of what we learned over the last year about searching WHOIS databases, the variety of search returns, RFC's, RAA provisions, opinions of our constituency members, responses and opinions of our survey respondents as they relate to WHOIS search-ability within top level domains and across them. We also recommend further work efforts in order to answer key questions related to legitimate use, differentiated access, privacy, and new protocols that support our policy environment.

Search-ability Issue Report

Current Policy Environment

The current policy environment is documented in Louis Touton's [letter to the ICANN WHOIS Committee](#) from January 1, 2000:

An overall goal of the Whois provisions of the Registrar Accreditation Agreements was to help restore the InterNIC Whois service that existed in .com, .net, and .org prior to the introduction of multiple registrars. This service is described in Section 6.4 of [RFC 1580](#). With the advent of multiple, competitive registrars in .com, .net, and .org, contact data on domain-name holders was broken up into separate databases maintained by each sponsoring registrar. As a result, searches that were previously possible (e.g., a search for all .com/.net/.org entries that reference a particular person) were no longer possible on a TLD-wide basis. The approach of the agreements was to require, as an immediate measure, the provision of Whois service from each registrar's database ([subsection II.F.1](#)), and to provide a pathway toward restoration of a TLD-wide Whois capability.

The relevant WHOIS provisions from the May 2001 RAA are included in [Appendix A](#) of this document.

What are Some Reasons for Searching WHOIS data?

The WHOIS database was not ever intended for use by marketers: it was intended for use by all those who participate in activities online; other than spamming and UCE. The WHOIS database is the place where identifiable information should be found.

- Consumers need WHOIS to discover who they are dealing with online and where to seek redress for problems.
- Parents need WHOIS to find out who is responsible for sites their children are visiting and to restrict children's access to inappropriate material.
- Intellectual Property owners need WHOIS to determine the identity of those conducting piracy operations over the Internet and to identify cyber-squatters.
- Law enforcement authorities need WHOIS to investigate illegal activities taking place online, from child pornography to consumer fraud to spreading viruses.

Search-ability Issue Report

- Technical Contacts, Administrative Contacts need to be able to communicate with other network operators should problems arise.
- Registrants, Billing Contacts may review their domain status and prepare for payment, be sure contact data is correct, etc.
- All Internet users need WHOIS to provide transparency and accountability on the Internet.

The above is not a comprehensive list; instead provided as an example. There are many reasons people need or want to search WHOIS data, we found identifying specific reasons to be helpful when making decisions related to searching WHOIS data.

Who Benefits from Enhanced Search-ability?

- Technical and Administrative Contacts
- Registrants and Billing Contacts
- Registrars and Registries
- Internet Users in general

Concerns Expressed to the TF About Enhancing Search-ability:

- A great amount of people are significantly affected by use of WHOIS data.
- Enhancing search-ability could enable inappropriate data mining.
- The way WHOIS data is searched may impact willingness of Registrants to maintain accurate contact information and Registrars to enforce or in some instances begin to support enforcement of accurate WHOIS data.
- Cost of implementation and maintenance is a concern; although there are those who believe this is simply a cost of doing business for the Registry and their corresponding Registrars. (who should pay was a question with many responses from our survey, responses were inline with the idea that whomever pays will inevitably be the end user. We have heard the suggestion that ICANN create a process and templates for use by those who implement the solution to help alleviate cost.)

Search-ability Issue Report

- We raise questions about what may have been meant by the phrase: “centralized WHOIS database”.
 - Some members of the TF questioned either the feasibility of such an approach or the necessity.
 - The TF suggests a portal approach which gives the output of a “centralized search” across all registrations in a single TLD, and which is largely the approach taken today by available services, is a more useful approach than consideration of recentralizing WHOIS data.
 - Further developments in technology present options which may be more feasible than an actual centralized repository for all WHOIS data pertaining to one TLD.
 - Recentralization of the data also raises other questions to the TF, including how to ensure accuracy, costs for updates and corrections, given the distribution of responsibilities between registrars and registries.
- Output varies greatly per TLD. Consideration should be given to developing a means of meeting the stated desire of the survey respondents for centralized access to WHOIS data on a per TLD basis, including standardization of format and data output.

What Could Enhance Search-ability?

This portion of the document contains our recommendations for future work items concerning enhanced search-ability of WHOIS databases.

Search-ability Issue Report

3.3.1.

The use of elements other than domain names as query keys.

The first provision includes a mandate for registrars provision of “an interactive web page and a port 43 WHOIS service providing free public query-based access to up-to-date (i.e. updated at least daily) data concerning all active SLD registrations sponsored by Registrar in the registry for the .com, .net, and .org TLDs.” In addition this provision includes information about accessibility of many data elements.

- We do not suggest enforcing all elements be made into query keys.
- We do suggest Registrants and those with NIC handles should be permitted to search their own information thoroughly.
- Consideration should be given to those without NIC handles or who are not Registrants. Much research is needed in this area.
- Our survey respondents indicated they would like to be able to use the following elements as query keys for searches:
 - Registrant Name
 - Technical Contact Name or Handle
 - Administrative Contact Name or Handle
 - Primary Name Server or IP Address
 - Secondary Name Server or IP Address
- We suggest reducing the query requests and responses to pertinent information only.
- We suggest an investigation of role-based access to WHOIS information.
 - For example, the Technical Contact level might permit access to information relevant to networking, while the Billing Contact level might permit access to information relevant to the needs of finance, and so on.
 - We recommend an investigation exploring how security could be leveraged against privacy concerns.

Search-ability Issue Report

3.3.2

Prompt updates of data elements.

Quite a bit of work needs to be done in this area. Please consider the following:

- Work with registrars to find ways to facilitate registrant updates of WHOIS data.
- Discover ways to enforce this provision of the RAA.
- Establish a separate complaint mechanism for this problem.

3.3.3

Subcontracting Public Access to WHOIS databases.

A third party, “may subcontract its obligation to provide the public access...”. We support the market’s ability to fulfill this obligation and suggest this service be provided for clients who wish to search their own information.

- Thought should be given to legitimizing custom queries via a third party service using bulk access instead of query-based.
- Registrants and those with NIC handles should be permitted to search their own information thoroughly.

3.3.4

Distributed query-based search functionality across all registrars.

The existing gTLD registry agreements provide for access to each registrar's database via a WHOIS service and specifically contemplate the possibility of a distributed cross-registrar WHOIS service.

As a first step, further discussion should be undertaken regarding 3.3.4. Clearly, it is in the interests of the Registrars to provide the WHOIS service themselves. IF, after reasonable exploration of this approach, it appears that no progress will be forthcoming, THEN,

Search-ability Issue Report

Consideration should be given to moving toward a consensus policy as foreseen by the first sentence of this provision (i.e., requiring registrars to cooperatively implement a portal approach to offering centralized access to WHOIS data across multiple TLDs). Such consideration would require a further work effort and should be based on non-proprietary open standards based solutions.

There are several services who currently offer a form of centralized search service across some gTLDs and some ccTLDs. Most of these are able to return searches for the larger gTLDs and some ccTLDs. Before undertaking further recommendations, the Task Force suggests a brief examination of any barriers to further additions to these services be undertaken.

- Consideration should be given to revising the wording of this provision so the definition of "centralized" is in reference to report display and request pages rather than the database.
- The Task Force recommends further examination of the role of standards (as applicable to searchability) especially to continue to address the issues of more advanced database query capabilities.

3.3.5

Terms and conditions for query-based access.

- Consideration should be given to the removal of the word "mass".
- Consideration should be given to add postal mail to the list under (a).

Search-ability Issue Report

Appendix A: Relevant Provisions from the RAA

3.2 Submission of Registered Name Holder Data to Registry. During the Term of this Agreement:

3.2.1 As part of its registration of Registered Names in a TLD as to which it is accredited, Registrar shall submit to, or shall place in the Registry Database operated by, the Registry Operator for the TLD the following data elements:

3.2.1.1 The name of the Registered Name being registered;

3.2.1.2 The IP addresses of the primary nameserver and secondary nameserver(s) for the Registered Name;

3.2.1.3 The corresponding names of those nameservers;

3.2.1.4 Unless automatically generated by the registry system, the identity of the Registrar;

3.2.1.5 Unless automatically generated by the registry system, the expiration date of the registration; and

3.2.1.6 Any other data the Registry Operator requires be submitted to it.

3.3 Public Access to Data on Registered Names. During the Term of this Agreement:

3.3.1 At its expense, Registrar shall provide an interactive web page and a port 43 Whois service providing free public query-based access to up-to-date (i.e., updated at least daily) data concerning all active Registered Names sponsored by Registrar for each TLD in which it is accredited. The data accessible shall consist of elements that are designated from time to time according to an ICANN adopted specification or policy. Until ICANN otherwise specifies by means of an ICANN adopted specification or policy, this data shall consist of the following elements as contained in

Search-ability Issue Report

Registrar's database:

3.3.1.1 The name of the Registered Name;

3.3.1.2 The names of the primary nameserver and secondary nameserver(s) for the Registered Name;

3.3.1.3 The identity of Registrar (which may be provided through Registrar's website);

3.3.1.4 The original creation date of the registration;

3.3.1.5 The expiration date of the registration;

3.3.1.6 The name and postal address of the Registered Name Holder;

3.3.1.7 The name, postal address, e-mail address, voice telephone number, and (where available) fax number of the technical contact for the Registered Name; and

3.3.1.8 The name, postal address, e-mail address, voice telephone number, and (where available) fax number of the administrative contact for the Registered Name.

3.3.2 Upon receiving any updates to the data elements listed in Subsections

3.3.1.2, 3.3.1.3, and 3.3.1.5 through 3.3.1.8 from the Registered Name Holder, Registrar shall promptly update its database used to provide the public access described in Subsection 3.3.1.

3.3.3 Registrar may subcontract its obligation to provide the public access described in Subsection 3.3.1 and the updating described in Subsection

3.3.2, provided that Registrar shall remain fully responsible for the proper provision of the access and updating.

3.3.4 Registrar shall abide by any ICANN specification or policy established as a Consensus Policy according to Section 4 that requires

Search-ability Issue Report

registrars to cooperatively implement a distributed capability that provides query-based Whois search functionality across all registrars. If the Whois service implemented by registrars does not in a reasonable time provide reasonably robust, reliable, and convenient access to accurate and up-to-date data, the Registrar shall abide by any ICANN specification or policy established as a Consensus Policy according to Section 4 requiring Registrar, if reasonably determined by ICANN to be necessary (considering such possibilities as remedial action by specific registrars), to supply data from Registrar's database to facilitate the development of a centralized Whois database for the purpose of providing comprehensive Registrar Whois search capability.

3.3.5 In providing query-based public access to registration data as required by Subsections 3.3.1 and 3.3.4, Registrar shall not impose terms and conditions on use of the data provided, except as permitted by policy established by ICANN. Unless and until ICANN establishes a different policy according to Section 4, Registrar shall permit use of data it provides in response to queries for any lawful purposes except to: (a) allow, enable, or otherwise support the transmission by e-mail, telephone, or facsimile of mass, unsolicited, commercial advertising or solicitations to entities other than the data recipient's own existing customers; or (b) enable high volume, automated, electronic processes that send queries or data to the systems of any Registry Operator or ICANN-Accredited registrar, except as reasonably necessary to register domain names or modify existing registrations.

3.3.6 In addition, Registrar shall provide third-party bulk access to the data subject to public access under Subsection 3.3.1 under the following terms and conditions:

3.3.6.1 Registrar shall make a complete electronic copy of the data available at least one time per week for download by third parties who have entered into a bulk access agreement with Registrar.

3.3.6.2 Registrar may charge an annual fee, not to exceed US\$10,000, for such bulk access to the data.

Search-ability Issue Report

3.3.6.3 Registrar's access agreement shall require the third party to agree not to use the data to allow, enable, or otherwise support the transmission by e-mail, telephone, or facsimile of mass, unsolicited, commercial advertising or solicitations to entities other than such third party's own existing customers.

3.3.6.4 Registrar's access agreement shall require the third party to agree not to use the data to enable high-volume, automated, electronic processes that send queries or data to the systems of any Registry Operator or ICANN-Accredited registrar, except as reasonably necessary to register domain names or modify existing registrations.

3.3.6.5 Registrar's access agreement may require the third party to agree not to sell or redistribute the data except insofar as it has been incorporated by the third party into a value-added product or service that does not permit the extraction of a substantial portion of the bulk data from the value-added product or service for use by other parties.

3.3.6.6 Registrar may enable Registered Name Holders who are individuals to elect not to have Personal Data concerning their registrations available for bulk access for marketing purposes based on Registrar's "Opt-Out" policy, and if Registrar has such a policy, Registrar shall require the third party to abide by the terms of that Opt-Out policy; provided, however, that Registrar may not use such data subject to opt-out for marketing purposes in its own value-added product or service.

3.3.7 Registrar's obligations under Subsection 3.3.6 shall remain in effect until the earlier of (a) replacement of this policy with a different ICANN policy, established according to Section 4, governing bulk access to the data subject to public access under Subsection 3.3.1, or (b) demonstration, to the satisfaction of the United States Department of Commerce, that no individual or entity is able to exercise market power with respect to registrations or with respect to registration data used for development of value-added products and services by third parties.

3.3.8 To comply with applicable statutes and regulations and for other

Search-ability Issue Report

reasons, ICANN may from time to time adopt policies and specifications establishing limits (a) on the Personal Data concerning Registered Names that Registrar may make available to the public through a public-access service described in this Subsection 3.3 and (b) on the manner in which Registrar may make such data available. In the event ICANN adopts any such policy, Registrar shall abide by it.

3.7.7 Registrar shall require all Registered Name Holders to enter into an electronic or paper registration agreement with Registrar including at least the following provisions:

3.7.7.1 The Registered Name Holder shall provide to Registrar accurate and reliable contact details and promptly correct and update them during the term of the Registered Name registration, including: the full name, postal address, e-mail address, voice telephone number, and fax number if available of the Registered Name Holder; name of authorized person for contact purposes in the case of an Registered Name Holder that is an organization, association, or corporation; and the data elements listed in Subsections 3.3.1.2, 3.3.1.7 and 3.3.1.8.

3.7.7.2 A Registered Name Holder's willful provision of inaccurate or unreliable information, its willful failure promptly to update information provided to Registrar, or its failure to respond for over fifteen calendar days to inquiries by Registrar concerning the accuracy of contact details associated with the Registered Name Holder's registration shall constitute a material breach of the Registered Name Holder-registrar contract and be a basis for cancellation of the Registered Name registration.

3.7.7.3 Any Registered Name Holder that intends to license use of a domain name to a third party is nonetheless the Registered Name Holder of record and is responsible for providing its own full contact information and for providing and updating accurate technical and administrative contact information adequate to facilitate timely resolution of any problems that arise in connection with the Registered Name. A Registered Name Holder licensing use of a Registered Name according to this provision shall accept liability for harm caused by wrongful use of the

Search-ability Issue Report

Registered Name, unless it promptly discloses the identity of the licensee to a party providing the Registered Name Holder reasonable evidence of actionable harm.

3.7.7.4 Registrar shall provide notice to each new or renewed Registered Name Holder stating:

3.7.7.4.1 The purposes for which any Personal Data collected from the applicant are intended;

3.7.7.4.2 The intended recipients or categories of recipients of the data (including the Registry Operator and others who will receive the data from Registry Operator);

3.7.7.4.3 Which data are obligatory and which data, if any, are voluntary; and

3.7.7.4.4 How the Registered Name Holder or data subject can access and, if necessary, rectify the data held about them.

3.7.7.5 The Registered Name Holder shall consent to the data processing referred to in Subsection 3.7.7.4.

3.7.7.6 The Registered Name Holder shall represent that notice has been provided equivalent to that described in Subsection 3.7.7.4 to any third-party individuals whose Personal Data are supplied to Registrar by the Registered Name Holder, and that the Registered Name Holder has obtained consent equivalent to that referred to in Subsection 3.7.7.5 of any such third-party individuals.

3.7.7.7 Registrar shall agree that it will not process the Personal Data collected from the Registered Name Holder in a way incompatible with the purposes and other limitations about which it has provided notice to the Registered Name Holder in accordance with Subsection

3.7.8 Registrar shall abide by any specifications or policies established according to Section 4 requiring reasonable and commercially practicable (a) verification, at the time of registration, of contact information

Search-ability Issue Report

associated with a Registered Name sponsored by Registrar or (b) periodic re-verification of such information. Registrar shall, upon notification by any person of an inaccuracy in the contact information associated with a Registered Name sponsored by Registrar, take reasonable steps to investigate that claimed inaccuracy. In the event Registrar learns of inaccurate contact information associated with a Registered Name it sponsors, it shall take reasonable steps to correct that inaccuracy.

Authors of this Report (Working Group Three)

Abel Wisman, GA

Kristy McKee, GA

Tim Denton, Registrar

Troy Dow, BC

Special thanks to Antonio Harris for Appendix A's content.

ⁱ Identifying privacy issues is an item we recommend should be an ongoing process. Our solutions support the development of policy that enables country specific privacy implementation so as to avoid any unnecessary skirmishes and keep things as simple as possible; but maintain the opinion gTLDs will need a privacy policy of their own.