

WHOIS TASK FORCE 1

RESTRICTING ACCESS OF WHOIS FOR MARKETING PURPOSES

PRELIMINARY REPORT

I. INTRODUCTION

To become an accredited domain name registrar for any of the existing top-level domains (“TLDs”), all registrars are required to enter into a Registrar Accreditation Agreement (Agreement) with the Internet Corporation for Assigned Names and Numbers (ICANN). Under that Agreement, registrars are required to provide an on-line, interactive Whois database. This database contains the names and contact information - - postal address, telephone number, electronic mail address and in some cases facsimile number - - for registrants who register domain names through the registrar, as well as the domain names’ administrative, technical and, in some cases, the billing contacts. The Agreement also requires registrars to make the database freely accessible to the public via its web page and through an independent access port called port 43. These query-based channels of access to the Whois database allow any person or entity to collect registrant contact information for one domain name at a time by entering the domain name into the provided search engine.¹

In addition, many of the new unsponsored TLD registries approved in November 2000, including .biz, .info, .name and .pro as well as the recently transitioned .org registry are based on “fat” or “thick” registry models, meaning that the standard Whois service provides a central location for all authoritative data for its respective TLD. As such, the ICANN Registry Agreements for each of these TLDs require that the registries provide RFC 954 conformant service, including making such service accessible through their own front-end web interface as well as over Port 43.

Although initially developed for purely technical purposes, to contact the owner of a domain name or other network resource to aid in the resolution of technical issues with respect to the domain name,² the Whois database over the years has become an important resource to Internet users, ISPs, governmental users, intellectual property holders and to registrars. Despite its utility to such users, the reality is that Whois data consists of personal and business contact information, including their telephone numbers, physical and e-mail addresses, most of which registrants provide not knowing that such information can be accessed by any individual or entity.

¹ See <http://www.icann.org/registrars/register.com-verio/order-08dec00.htm#V3>.

² The basis for establishing a WHOIS service was set forth as early as 1985 in RFC 954 (<http://www.ietf.org/rfc/rfc0954.txt?number=954>) which dictates that for each domain name, certain information about the registrant should be able to be retrieved in “human-readable form”. Although the RFC itself does not set forth the reasons why such data should be collected, it has been argued by many in the technical community, that the use of WHOIS information for purposes other than to aid in the resolution of technical issues, was not its original intention.

A. Background

On February 8, 2001, the Domain Name Supporting Organization (now called the Generic Names Supporting Organization) commissioned a task force to “consult with the community with regard to establishing whether a review of any questions related to ICANN’s Whois policy is due and if so to recommend a mechanism for such a review.”³ This process took over two years to produce any consensus-based recommendations by the GNSO Council to the ICANN Board. One such recommendation that was promulgated by the GNSO Council was to emphasize a condition already contained within Registrar Accreditation Agreement involving the registrars providing Whois information in bulk to any entity that so requests for a maximum price of \$10,000. That condition was that the use of Whois information for marketing should not be permitted.⁴ This Consensus Policy was later adopted by the ICANN Board at its meeting in Rio de Janeiro, on March 27, 2003.⁵

The recommendation, however, did not address a number of key issues, including what exactly was meant by “marketing purposes.” In fact, the committee designed to implement the recommendations of the Whois Task Force specifically concluded that “there is a need to clarify the definition of ‘marketing purposes.’ This may require a small working group to define, possibly just in the form of examples (but not limited to) of marketing activities covered.”⁶

More importantly, from our perspective, is that neither the recommendations by the Whois Task Force nor the GNSO Council addressed the use of Whois information acquired through other contractually required means of access for marketing purposes. This includes how to restrict the access of information acquired through front-end web interfaces or over Port 43 from being used for marketing purposes.

B. The Problem: Data Mining.

Many believe that bulk access under license may be only a minor contributor to the perceived problem of use of Whois data for marketing purposes. A subset of a registrar’s Whois database that is sufficiently large for data mining purposes may be obtained through other means, such as a combination of using free zonefile access (via signing a registry zonefile access agreement to obtain a list of domains, and then using anonymous (public) access to either port-43 or interactive web pages to retrieve large volumes of contact information. Once the information is initially obtained it can be kept up-to-date by detecting changes in the zonefile, and only retrieving information related to the changed records. This process is often described as “data mining”. The net effect is that large numbers of Whois records are easily available for marketing purposes, and generally on an anonymous basis (the holders of this information are unknown).

The above scenario of Whois data mining is not a new phenomena. In fact, in the case of *Register.com v. Verio, Inc.* filed in 2000 in the United States District Court for the Southern District of New York, Register.com sued Verio for allegedly mining Register.com’s Whois data

³ <http://www.dnso.org/dnso/notes/20010208.NCtelecon-minutes.html>

⁴ <http://www.icann.org/gns0/whois-tf/report-19feb03.htm#III>.

⁵ <http://www.icann.org/minutes/prelim-report-27mar03.htm>

⁶ <http://www.icann.org/gns0/whois-tf/report-19feb03.htm#III>

base for the purpose of soliciting Register.com's customers through e-mail, facsimile and direct mail. On December 8, 2000, the court granted an injunction against Verio from engaging in such practices.⁷ The Court described the process used by Verio as follows:

In general, the process worked as follows: First, each day Verio downloaded, in compressed format, a list of all currently registered domain names, of all registrars, ending in .com, .net, and .org. That list or database is maintained by Network Solutions, Inc. ("NSI") and is published on 13 different "root zone" servers. The registry list is updated twice daily and provides the domain name, the sponsoring registrar, and the nameservers for all registered names. Using a computer program, Verio then compared the newly downloaded NSI registry with the NSI registry in downloaded a day earlier in order to isolate the domain names that had been registered in the last day and the names that had been removed. After downloading the list of new domain names, only then was a search robot used to query the NSI database to extract the name of the accredited registrar of each new name.⁵ That search robot then automatically made successive queries to the various registrars' Whois databases, via the port 43 access channels, to harvest the relevant contact information for each new domain name registered. Once retrieved, the Whois data was deposited into an information database maintained by Verio. The resulting database of sales leads was then provided to Verio's telemarketing staff.⁸ [citations omitted]

The Purpose of Whois Task Force 1 is to determine what contractual changes (if any) are required to allow registrars and registries to protect domain name holder data from data mining for the purposes of marketing.

C. Summary of Findings:

1. Although there are mechanisms currently employed by Whois Providers that have had limited success on the amount of data mining, such mechanisms alone, are insufficient to prevent data mining of the Whois database for marketing purposes.
2. The output of this Whois Task Force depends heavily on the output of Whois TF 2 (which data elements are included in the publicly available Whois).
3. Subject to the exceptions set forth in the remainder of this report, To the extent that data deemed to be sensitive by the Internet community is recommended to be publicly disclosed by Whois TF 2, then at a minimum, the requestor of Whois information should be required to identify itself to the Whois Provider (i.e., the Registrar or the Registry [in the case of thick registries]) along with the reasons for which it seeks the data. Representatives from the Noncommercial, ALAC, Registrar and Registries Constituencies believe that such information should be made available to the registrant whose Whois information is sought, whereas representatives of the from the Intellectual Property, Commercial and Business Users Constituency and Internet Service Providers disagree with the requirement

⁷ <http://www.icann.org/registrars/register.com-verio/order-08dec00.htm#V3>

⁸ <http://www.icann.org/registrars/register.com-verio/order-08dec00.htm#V3>

that notice be provided to the registrant. They believe that an acceptable alternative to the notice requirement could be to require the preservation of some form of audit trail so that in the rare case in which Whois access were abused, it could be established who had made the request.

4. It is not possible to create technical restrictions under the current port 43 specifications that will limit port 43 access to a specific type of purpose such as “non-marketing uses.”
5. To the extent sensitive data is required to be displayed, Port 43 should only be used by registrars to facilitate transfers if no other mechanism is available to registrars for this purpose.
6. Some members of the Task Force stated that they may not be fundamentally opposed to having an automated mechanism to retrieve sensitive data for identified requestors with approved purposes provided that certain terms and conditions (set forth below) apply.
7. A Cost benefit analysis and a feasibility study should be done when considering any significant changes in Whois requirements.

II. PROCESS:

Initially convened on 2 December, 2003, this Task Force engaged its work in a serious and diligent manner. The Task Force held weekly meetings and established a schedule for addressing the milestones outlined in the Description of Work⁹.

The Purpose of the Task Force is to determine what contractual changes (if any) are required to allow registrars and registries to protect domain name holder data from data mining for the purposes of marketing. The focus is on the technological means that may be applied to achieve these objectives and whether any contractual changes are needed to accommodate them. The Task Force was given three milestones, namely: (1) collect the “needs and Justifications” for Whois information for “nonmarketing purposes; (2) review general approaches to prevent automated data mining; and (3) determine whether any changes are required in the contracts to implement an approach to prevent automated data mining.

The Task Force prepared a survey seeking to identify non-marketing uses of Whois data, and the methods of accessing that data. The survey was distributed to a group of companies identified by the Task Force participants as likely to have non-marketing uses of Whois-type data, to the gTLD Registries and Registrars through their constituencies, and to the other GNSO constituencies. Additionally, the survey was posted to the GNSO website for public use.

A total of ten unique replies were received to the survey, four from Registrars, six from general respondents. The response to the survey was insufficient to be an effective tool for evaluating all non-marketing uses and needs for Whois data, but did provide interesting information regarding specific uses by those who replied.

⁹ <http://gns0.icann.org/issues/whois-privacy/tor.shtml>.

The Task Force reviewed prior work that had been done on the issue of Whois privacy and access, particularly reviewing the materials from the Whois Workshops given during the Montreal ICANN meeting.

Constituency statements were received from all GNSO constituencies, and from the At-Large Advisory Committee. Using the statements and other materials, the Task Force members worked cooperatively through discussion and debate to prepare the Preliminary Report.

III. ANALYSIS

Needs and Justification

Principles for the use of Whois

Whois TF 1's goal was, consistent with its mandate, to balance the concerns and needs of domain name registrants, legitimate Whois data users, registrars and registries.¹⁰ We recognize the need to take into account issues of privacy and data protection, data accuracy, continued flows of data, registrant accountability, and system burdens. We also recognized the need to ensure that whatever process we developed must not prevent exchanges of information needed to make the DNS as a technical system operate smoothly and efficiently. In addition, the task force considered the effects of proposed changes to the Whois service on the ability of groups such as law enforcement, intellectual property owners, Internet service providers, and consumers to continue to retrieve information necessary to perform their functions.¹¹

In statements collected by the Task Force from the previous Whois task forces as well as ICANN workshops and our recent survey, some groups have indicated that access to accurate, up-to-date, and reliable Whois data has become an important tool for a variety of Internet users. Consumers as well as consumer protection authorities frequently use Whois to discover with whom they are dealing online. The United States Federal Trade Commission, for example, often uses Whois to investigate online fraud, identity theft, and "phishing"¹² scams, particularly in the cross-border context. Law enforcement officials likewise access Whois information to combat online crimes. Intellectual property owners, both copyright and trademark owners, use Whois as a tool to fight online piracy and cybersquatting. In addition, trademark owners and other

¹⁰ The representative from the Intellectual Property Constituency believes that this paragraph significantly misstates TF1's sole goal, which is defined in the Description of Work as "to determine what contractual changes (if any) are required to allow registrars and registries to protect domain name holder data from data mining for the purpose of marketing. The focus is on the technological means that may be applied to achieve these objectives and whether any contractual changes are needed to accommodate them."

¹¹ There was disagreement within the group as to whether the new policies needed to ensure that these latter groups could continue to access such information. The Commercial and Business Users, Internet Service Providers and Intellectual Property Constituencies strenuously argued in favor of continuing to make such information available to the aforementioned groups, while the ALAC and Noncommercial Constituency User Constituency argued that such information was not necessary for performing their functions.

¹² "Phishing" involves setting up a phony but very realistic website in order to deceive people into providing identity information, such as social security numbers, credit card account numbers, and passwords. Major online businesses have had considerable problems with this type of scam. Often, phishing uses many exploits of browsers to hide the true name of the domain so it's not easy for naive users to tell that they are not at the official site.

business users use Whois as a way of managing trademark portfolios, conducting due diligence for the purpose of corporate acquisitions, and identifying company assets in bankruptcies or insolvencies.

Individuals responsible for network security have stated that access to Whois data is needed to prevent denial of service attacks¹³ and identify other threats to networks stability. ICANN's Security and Stability Advisory Committee recently noted the importance of Whois data and recommended that "[t]he accuracy of Whois data used to provide contact information for the party responsible for an Internet resource must be improved, both at the time of its initial registration and at regular intervals."¹⁴

Registrars currently utilize Whois information in the transferring of domain names from one registrar to another. Registrars are required to obtain confirmation from the domain name holder (or one that has apparent authority over the domain) in order to complete a transfer request. For registries that do not employ authorization codes, the gaining registrar must access the Whois information from the losing registrar so that it can send a confirmation message to the registrant confirming transfer. For registries that do employ authorization codes, the gaining registrar must still have access to the Whois information because, in compliance with the registrar's ICANN contracts, the gaining registrar must store (presumably in the gaining registrar's Whois database) the losing registrar's pre-transfer Whois information for any transferred-in domain. For thick registries it can obtain this information from either the registry's or the losing registrar's Whois database.¹⁵

B. General Approaches to Prevent Data Mining

Today, registrars use a combination of techniques in an effort to thwart data mining. One of these techniques, "CAPTCHA" (completely automated public Turing test to tell computers and humans apart), is where the registrar displays, for example, a gif image of a series of letters, and the user must decipher the image and manually enter the series of letters displayed in order to gain access to the registrar's web-based interface to its Whois service. Unfortunately, this technique is unable to be used for accessing Whois information over port-43.

In addition, to CAPTCHA, registrars typically monitor the number of requests made from IP addresses and limit the amount of queries from particular IP address. This technique has often

¹³ Although these statements were made in that materials we were provided, none of the materials set forth exactly how Whois information is used to prevent these attacks. Task Force 1 seeks comment from those individuals responsible for network security on how Whois information prevents denial of service attacks.

¹⁴ See Whois Recommendation of the Security and Stability Advisory Committee, at <http://www.icann.org/committees/security/sac003.htm>.

¹⁵ In accordance with the Description of Work, Whois Task Force 1 undertook an effort to review the World Wide Web Consortium's (W3C) Working Draft of 11 March 2004 entitled "Web Content Accessibility Guidelines 2.0" to ascertain that "requirements for access are met (including accessibility requirements for those that may for example be visually impaired)" The W3C guidelines are a technical exposition of the standards that may be applied to all forms of information disseminated by the World Wide Web. The guidelines do not specifically refer to Whois, but are applicable to Whois as well as to any other data available on the web. We believe that any web-based Whois service should comply with the W3C recommendations, as should all information disseminated on the World Wide Web.

been referred to as "speed bumping". Unlike CAPTCHA, speed bumping can be used to limit the amount of queries on the web-based Whois service as well as Port-43.

Neither of these techniques is foolproof when used to prevent data mining. For example, speed bumping can be defeated because those more experienced data miners can easily gain access to multiple IP addresses (in some case numbering in the thousands) and perform automated Whois lookups from each IP address. Although in the aggregate the number of queries are above the speed bump limit, because the number of queries from each individual IP address is below the threshold, such data miners pass through the system. With regards to CAPTCHA, data miners are still able to work around the system because miners are able to use sophisticated OCR (optical character recognition) software to decipher the image. Despite these flaws, registrars believe that these mechanisms do have value and that they do prevent data mining in most scenarios, especially with respect to web-based public Whois interfaces at a very minimum increase the costs to the data miner. In summary, these approaches work in most cases but are not enough to entirely solve the data mining issues.

C. Policy Recommendations

1. Dependence on Whois Task Force 2

The output of this Whois Task Force depends heavily on the output of Whois TF 2 (which data elements are included in the publicly available Whois). The more sensitive the data: (a) the more value there is to the data; (b) the more likely such data is to be mined, (c) the more this impacts the privacy rights of individuals and (d) creates an incentive for the registrant to make the data inaccurate. In such cases, there may be a need to restrict access to that data.¹⁶

¹⁶ Although arguably outside the scope of the mandate of our task force, our Task Force had a number of productive discussions on which data we believed would be classified as sensitive, and which we would consider to be non-sensitive. Although the Task Force was unable to form a consensus on each element of the current Whois database, we were able to generally agree on which data we believed, at a minimum, could be considered "non-sensitive." We engaged in this exercise not to duplicate the work of Whois Task Force 2, but rather to understand the different classifications for data in order to develop the policies contained herein. We will coordinate with Whois Task Force 2 after the Preliminary Report is completed.

Non-Sensitive Data

- Domain Name
- Domain ID
- Registrar Name
- Registrar ID
- Name Server Name
- Technical Contact Name
- Technical Contact Address (Full information)
- Technical Contact Phone Number
- Technical Contact Fax Number
- Technical Contact e-mail address

The Task Force noted that in many cases registrars ask for one set of contact information from its registrants and populate all of the required Whois contact fields (administrative, technical and billing contacts) with such information. Often times, the registrant information provided for the one contact, may not actually be the technical

2. *Value of Whois Data*

It is believed that if only data deemed to be non-sensitive by the Internet community ("Non-Sensitive Data") were to be publicly displayed (whether on the Web, Port 43 or other automated process), the data itself has little value, is less likely to be data mined, and has little effect on privacy rights. Therefore, imposing restrictions on access to Non-Sensitive data may not be necessary.

Note, that we have assumed that the less sensitive the data is, the less valuable the data will be and the less data mining will occur. However, there is still some value to the information and therefore, there may be a need for query limits to prevent denial of service attacks.

3. *National Law Applies*

To the extent that restrictions are imposed on access to Whois information, this should not be taken to mean that we are addressing all of the privacy implications nor the entire problem of data mining. In addition, as in all cases, National law, as applicable, should be taken into consideration.

All Registries and Registrars are currently required to provide access to WHOIS information via web-based access and Port 43 regardless of their applicable national laws on privacy. In fact, some have argued that complying with their ICANN Agreements placed them in a position of choosing whether to violate their ICANN Agreements or violating national law. On the other hand, the Task Force did note that allowing each registrar or registry to rely on its own "national law" could have significant impacts on competition among registrars and even within the registries. Comment is sought by the Task Force on how to balance the requirements of national law with the ICANN mission of promoting competition.¹⁷

4. *Identification of Requestor and Notification to Registrant*

To the extent that data deemed to be sensitive by the Internet community ("Sensitive Data") is recommended to be publicly disclosed by Whois TF 2, then at a minimum, the requestor of Whois information ("Requestor") should be required to identify itself to the Whois Provider (i.e., the Registrar or the Registry [in the case of thick registries]) along with the reasons for which it seeks the data. Representatives from the Noncommercial, ALAC, Registrar and Registries Constituencies believe that such information should be made available to the Registrant whose Whois information is sought,¹⁸ whereas representatives of the from the Intellectual Property, Commercial and Business Users Constituency and Internet Service

contact for the domain name. For purposes of the above classification, we assumed that the person or entity listed as the "Technical Contact" was indeed the true and correct technical contact not merely the domain name holder.

¹⁷ The representatives from the Intellectual Property and Business Constituencies believe that contents of these paragraphs are outside of the scope of TFI believes they are addressed in the TF2 report.

¹⁸ The Registrar representative to the Task Force stated that if registrants are to be notified that their WHOIS information was sought, that this communication be done only by the sponsoring registrar of that particular registrant.

Providers disagree with the requirement that notice be provided to the registrant. They believe that an acceptable alternative to the notice requirement could be to require the preservation of some form of audit trail so that in the rare case in which Whois access were abused, it could be established who had made the request.. The group recognizes, however, that an exception may need to be granted for certain law enforcement investigations (including civil investigations), who may need the information without having to provide the reasons to the Registrant.¹⁹

If this method is to be employed, the members of the Task Force believe that there should be some sort of authentication mechanism to prove the identity of the Requestor to minimize the chances for fraud. Otherwise, we can envision parties abusing the system in order to obtain the Sensitive Data of registrants. In addition, several members of the Task Force suggested that there should only be a limited number of “purposes” for which a Requestor could seek the Sensitive Data and that such purposes should be provided in the form of a multiple choice list. The Task Force seeks comment on this proposal.

The representatives from the Intellectual Property, Commercial and Business Users Constituency and Internet Service Providers disagree with the requirement that notice be provided to the registrant. They believe that an acceptable alternative to the notice requirement could be to require the preservation of some form of audit trail so that in the rare case in which Whois access were abused, it could be established who had made the request. According to the Intellectual Property representative, a notice requirement would substantially undermine the value of Whois data for a host of legitimate purposes; would be likely to add considerable cost and delay in obtaining access to Whois data and would do little if anything to discourage data mining. Finally, according to this representative, a notice requirement would entirely abolish anonymous access to Whois data, in direct contravention of the Task Force’s terms of reference, which state that “the task force should not study the amount of data available for public (anonymous) access for single queries.”

5. *Changes should apply to all forms of access*

To the extent that we are recommending any changes to access of Whois information, such changes need to be applied to all forms of access to Whois, whether Web-based, Port 43-based, or through any other mechanism.

6. *Future of Port 43 Access*

Based on input from the community, TF 1 has come to the conclusion that it is not possible to create technical restrictions under the current Port-43 specifications, that will limit port 43 access to a specific type of purpose; e.g., "nonmarketing uses." We have concluded that any access restrictions imposed on Port 43 by TF1 will apply to any Whois user, regardless of their purpose. In order to prevent abusive data mining by some on Port 43, we are required to develop access restrictions on Port 43 that affect all users and all purposes.

- a) Currently, Port 43 does not provide a way for a requestor to identify him or herself or the reasons for which it is seeking the data.

¹⁹ The Task Force seeks comment on the breadth of this exception and who would qualify for this purpose.

- b) If only Non-sensitive Data is displayed, there is little reason to change anything with respect to Port 43²⁰.
- c) If Sensitive Data will be displayed, then Port 43 would not be able to provide the functionality described in Section 4 above.
- d) Port 43 should, however, not be shut down completely. The Task Force believes that unless other mechanisms were available to the Registrars to retrieve sensitive data, Port 43 should be available to Registrars solely for the purpose carrying out its obligations with respect to transfers of domain names between registrars.²¹

7. *Automated Access to Whois*

Some members of the Task Force stated that they may not be fundamentally opposed to having an automated mechanism to retrieve Sensitive Data for approved Requestors with approved purposes provided that:

- The Requestor is asked to sign (or “click”) an electronic license agreement for the Sensitive data promising:
 - To use the data for only the purpose(s) indicated;
 - That the Whois data will not be used for marketing purposes; and
 - That the Requestor shall be prohibited from compiling, leasing, sublicensing, reselling or otherwise transferring the data to any third party (except to comply with law).²²
- The Requestor is identified to the Whois Provider;
- The Requestors identity and purposes for such information is disclosed to the Registrant.
 - The group recognizes, however, that an exception may need to be granted for certain law enforcement investigations (including civil investigations), only when notification of the registrant will defeat the purpose of the investigation.²³
- The Sensitive Data is provided to the Requestor in human-readable format only²⁴ (and not computer readable).

²⁰ The Task Force notes that nothing in this report shall limit the right of Whois Providers to institute mechanisms, as they deem appropriate, to prevent denial of service attacks (i.e., by imposing rate and query limits).

²¹ The representative from the Intellectual Property Constituency noted that the inability of Port 43 to provide the functionality described in Section 4 is not dispositive since there is a split of views with regard to Section 4. They also note that the proposal to shut down Port 43 except for registrar access is outside the scope of the Task Force and unworkable for the reasons stated above.

²² This does not prohibit third parties from acting as agents from obtaining the information and transferring such information to their principle, so long as the principal is identified as the Requestor and the agent complies with the limitations in this section, passes the Whois data to the principal in full and without modification, and does not store or use the information for its own purposes.”

²³ See Footnote 11.

²⁴ "Human readable format" means an output format that is not easily parsed by computer, such as a graphical (example ".gif") or audio format file, but that is easily interpreted by a human.

8. *Approval Process for Automated Searches to prevent data mining*

If there were to be an automated process available to retrieve sensitive data, like that currently provided under Port 43, with the functionality described in Section 7 above, the group discussed two alternative methods of regulating access to sensitive data.

White List. One would have a central authority (not a registry or registrar) approve entities that could use this automated process. This option became known as a "White List" of IP addresses. In this scenario, a White List would be created of Requestors that are believed to be nonmarketing users of Whois information (i.e., Law Enforcement, Consumer organization, Intellectual Property Organizations, etc.) This list would be provided to the registries and registrars and only those Requestors sending requests through the automated process would be allowed to access the sensitive Whois information. Questions arose concerning (a) who would operate this White List, (b) what would be the criteria for being on this White List, (c) whether it was actually feasible to implement; (d) secondary use of access, and (e) a process for dealing with abuses.

Individual Use List. The other alternative would approve specific individual uses of sensitive Whois data rather than giving blanket approvals to user entities. Each time a requestor wanted to gain access to Whois information it would submit an automated request to the Whois Provider. The Requestor would identify itself to the Whois Provider and also identify the specific purpose for which the data was requested (i.e., suspected trademark infringement, a desire to contact the domain name holder for sale of the name, suspected consumer fraud, etc.). This option would give all Internet users the same rights to access sensitive Whois data, but would require them to authenticate their identification. It would also require the creation of a "list of approved purposes" as described above.

A minority of the Task Force constituencies, including those representing the Noncommercial Constituency and the At-Large Advisory Council believe that the creation of a White List would be impractical and would place a large burden on the entity handling requests to be on the White List. In addition, they do not believe that any Requestor should be entitled to the Sensitive Data unless retrieval of such information was pursuant to a formal request by law enforcement (i.e., subpoena).

A majority of the Task Force constituencies, including those from the Commercial and Business users, ISPs, gTLD Registries and Intellectual Property Owners do not fundamentally oppose the "White List", but believe that it is essential for those legitimate Whois users to obtain the Sensitive Whois information in a timely and reliable manner. Moreover, these representatives questioned whether the cost of implementing such a system would be one which could be borne by the current funding models, and encourage that a cost-benefit analysis be undertaken before any such system is approved and implemented.

Finally, if there is a "White List" or "Individual Use List," the Task Force emphasized the need that a mechanism be employed to authenticate the identity of the Requestor to the entity administering either alternative.

With respect to the alternatives presented above, the Task Force seeks comment on this entire section, including the following questions:

- If there were a White List or Individual Use List, who would serve as the central authority (“Authority”) that determines the eligibility for entities to be on these lists?
- Does this same Authority maintain the centralized white-list or Individual Use List database/system?
- What are the criteria that the Authority uses to determine who is eligible to be on either list?
- Is there a limit of the number of entities that can be on the White or Individual Use Lists?
- Who pays for the implementation of either system? Would there be a contribution paid by the members of the either list?
- If entities on the White or Individual Use List must give the reasons for their queries, how does (or can) that information be delivered to the registrants?

Other Considerations

10. A technical means of providing this tiered access (i.e., allowing these parties to access the information, while preventing others from getting the information) could be through the IRIS protocol developed by the CRISP working group of the IETF. When finalized, we believe that a comprehensive review of this technical solution be undertaken. We believe a more detailed effort is needed to identify any specific parties that need access to selected elements and what information should be obtained about such access.

11. A Cost benefit analysis should be done when considering any significant changes in Whois requirements. Such analysis should include how the costs are distributed and who bears such costs.

12. Finally, careful consideration should be given to the feasibility of registrars and registries to implement any proposed changes in Whois requirements including but not limited to enforcing such requirements. And sufficient time should be allowed for any associated migration.

IV. IMPACT ON CONSTITUENCIES - TBD

CONSTITUENCY STATEMENTS

AT-LARGE ADVISORY COMMITTEE STATEMENT ON TF 1

Policy proposal

We recommend a simple two-tiered system.

Tier 1 -- public access. Users who access a future Whois-like system anonymously get access to non-sensitive information concerning a domain name registration, to be defined in detail by task force 2.

Tier 2 -- authenticated access. Users who want to access a more complete data set (to be defined in detail by task force 2) need to reliably identify themselves, and indicate the purpose for which they want to access the data. The identity of the data user and their purpose is recorded by registrars and registries, and made available to registrants when requested. This information could be withheld for a certain amount of time if the data user is (1) a law enforcement authority that is (2) accessing the data for law enforcement purposes.

Implementation remarks

We do not recommend any particular implementation of this proposal, but note that "reliable identification" could be provided by commercially available SSL certificates. In general, we would favor implementation of our proposal in a dedicated protocol (such as IRIS) over implementation through Web forms.

Rationale

The key aspect for deciding whether access to data gathered by registrars can be given to a third party is the purpose for which this data is going to be used. Obviously, registrars have no way to verify the purpose for which Whois data is being accessed.

The best heuristic we know of is to hold data users accountable for their activities, and to put enforcement of purpose limitations into the hands of registrants. This can be achieved by reliably identifying data uses and putting their identity, contact information, and purpose indication in the hands of registrants.

At the same time, a tiered system -- if implemented reasonably -- could preserve the ability of data users to automatically access Whois data in reasonable quantities. Registrars, on the other hand, would be enabled to limit the amount of data any particular party can access in a given interval of time.

Identifying data users and their purposes would also enable registrars to comply with legal obligations to make this kind of information available to data subjects.

Discussion of other proposals

There have been suggestions that "automated access" could be used as a heuristic to determine illegitimate access. In this scheme, automated access is blocked by attempting to require human attention with all queries. One set of implementations of these kinds of tests is known as CAPTCHA.

There is evidence that automated access is also being used for legitimate purposes; on the other hand, there is publicly available information on how CAPTCHA-like tests are being circumvented in other contexts. The circumvention here is based on a fundamental design problem of CAPTCHAs.

<http://boingboing.net/2004_01_01_archive.html#107525288693964966>

One particularly popular CAPTCHA has been broken in academic more than a year ago, but is still being used by registrars. <<http://www.cs.berkeley.edu/~mori/gimpy/gimpy.html>>

Accessibility problems posed by CAPTCHA-like tests are not fully understood by now; we note, though, that purely visual tests are insufficient from an accessibility point of view.

<<http://www.w3.org/TR/turingtest/>>

In conclusion, CAPTCHA tests address the wrong problem, and they address it badly. We strongly recommend against going down this path.

Task Force 2: Data elements displayed and collected

Policy proposal

We recommend that the mandatory collection and display of personal information about registrants be reduced as far as possible. What information is actually required for placing a domain name registration should be a matter of registrars' business models, and of applicable law, not of ICANN policy.

We consider the removal of the following data elements from registrars' and registries Whois services (in a tiered model, from *all* tiers) a priority:

- registrant name, address, e-mail address, and phone number, unless registrant has requested that this information be made available.

- administrative contact name, address, e-mail address, and phone number, unless registrant (or admin-c) has requested that this information be made available.

- Billing contact. These data are traditionally not published by registrars, but are included in many thick registries' public Whois services.

For the purposes of a tiered access system (see recommendations for task force 1), we would recommend that the following information be included in a public tier:

- Registrar of record.
- Name servers.
- Status of domain name.

- Contact data, if the data subject specifically requests that these data be included in the public tier.

Implementation remarks

None.

Rationale

For personal registrations, the registrant, administrative contact, and billing contact data sets are most likely to concern sensitive information, such as the registrant's home address and phone number.

We recognize that domain name registrations by online merchants often imply less privacy concerns; it has been argued that online merchants must make privacy information public in many jurisdictions. We are confident that businesses will also follow these duties by requesting registrars to make contact information about them available publicly. Conversely, if bad actors decide not to make contact information publicly available, that could actually make bad actors more easily recognizable, and provide consumers with a "red flag."

Discussion of other proposals

At the Whois workshop in Rome, we have heard several lawyers praise the usefulness of registrant and other telephone numbers in Whois services. That way, we were told, many cases could be settled by a single phone call. The easier the contact, we were told, the merrier.

This argument is troubling: What we were hearing there is a request to ICANN to enable lawyers to make off the record contact with other parties to a dispute that may not have a lawyer readily available, and to make this contact in a way which makes it hard for the registrant to get legal counsel involved in early negotiations arising out of the dispute.

Telephone numbers of registrant and administrative contacts should be *removed* from Whois services for precisely this reason: Forcing the non-registrant party to a dispute to open up that dispute by on-the-record means (e-mail, fax [not universally available], postal mail) ensures that registrants have an opportunity to retain legal counsel in these disputes, and to fully understand any claims made by the non-registrant party. It also helps to avoid legal bluff and plain bullying.

To summarize, it may be true that availability of phone numbers enables quick settlement. But availability of phone numbers also favors situations in which these settlements are achieved by dubious means, to the detriment of the registrant.

COMMERCIAL AND BUSINESS USERS

In order to provide input to all three Task Forces (TF) and provide a broader statement from the Commercial and Business User Constituency (hereafter Business Constituency or BC), we have consolidated our input into a single document.

Members of the Business Constituency use the Internet to conduct business. The Business Constituency is a constituency representing customers of providers of connectivity, domain names, IP addresses, protocols and other services related to electronic commerce in its broad sense. The BC membership includes corporations, entrepreneurs, and associations.

The BC recognizes that the Internet is changing and evolving into a more commercial and widely used communication mechanism, and that the characteristics of the Internet users are also changing, over time. It is generally agreed that more and more users are registering domain names for a wider and wider variety of purposes. As the user characteristics are changing and the Internet is growing, it is important to keep in mind the key issues of Internet stability. The BC believes that accurate Whois data is an essential element to that core value. In examining the possibility of changes in the Whois, the BC believes that better mechanisms are needed to ensure accurate Whois data, while balancing the needs of the full set of stakeholders and affected parties.

Principles for the use of Whois

Striking a balance among concerns and needs of the different stakeholders related to accuracy, reliability, access and privacy issues is the goal. This is consistent with the OECD Guidelines on the Protection of Privacy and Trans-border Data Flows of Personal Data, the international consensus, that works to strike a balance between effective privacy protection and the free flow of information.

Purposes of Business User access to Whois:

Business users access the Whois database to obtain registrant contact information for the following reasons:

1. to verify the availability of a name they might wish to register
2. to thwart security attacks of their networks and servers
3. to validate the legitimacy of a website for transactions
4. to identify consumer fraud and cyber-scam incidents
5. to undertake routine reviews to protect their brands
6. to support UDRP and other infringement proceedings
7. to combat spam.

The BC's guiding principles related to Whois are:

1. **Accuracy and access.** Accuracy and access to accurate data are the top priorities. Enforcement of accuracy requirements is essential.

2. **Use of data.** It is key to find a balance between data use for legitimate purposes and avoiding unwelcome or illegal use.
3. **Balance of Stakeholder needs.** Any changes in access to Whois must be balanced across the needs of all stakeholders and take into account the costs to the registries/registrars to maintain more complex systems, as well as the burden on the legitimate users of Whois.
4. **Marketing.** Whois data should never be used for marketing purposes. This includes precluding the use of Whois data for marketing by the registry or registrar other than for services that are **directly** applicable to registration or other purposes that are not inconsistent with the original purpose [*see* OECD Guidelines] or for which the registrant has explicitly opted-in.
5. **Scope.** The focus for now should be ensuring a consistent system of Whois across generic top-level domain names. Any discussion of Whois policies that might affect Whois within country-code domain names should be addressed later and through the new Country Code Names Supporting Organisation.

Task Force One: What contractual changes, if any, are needed to protect domain name holders from data mining for the purpose of marketing?

The BC notes:

Concerns arise from marketing use. The BC has previously stated that marketing uses of Whois data should be prohibited. The basis of much data protection law is that data should only be used for the purpose directly applicable to registration or other purposes that are not inconsistent with the original purpose [*see* OECD Guidelines] or for which the registrant has explicitly opted-in.

- **Spam.** Confusion exists today regarding whether and to what extent Whois data is used for the development of Spam. Data indicates that the involvement is small, but in any case, it is important to not allow contamination of the issues relating to Whois by the issue of spam prevention. Regardless of the limited degree of impact, mechanisms to limit any use should be supported.

The BC therefore proposes:

- **Eliminate marketing.** The BC believes that Whois data should never be used for marketing purposes. This includes precluding the use of Whois data for marketing by the registry or registrar, other than for services which are directly applicable to registration or for which the registrant has explicitly opted-in.
- **Limit access to Port 43 access.** Although it does not appear that Whois is a significant contributor to Spam, the BC supports the limitation on port 43 access (an Internet-based access used by registrars and others) to discourage any use for that purpose. Also, this will limit uses of port 43 for other marketing purposes.

- **Creation of a White list approach for “legitimate use”.** There are legitimate uses of Whois, which should be supported, including uses facilitated by bulk access. Such uses include research, creation of third party value-added services, etc. The BC therefore supports the creation of a list of legitimate uses, and recommends that such uses be limited via registry/registrar/third party contract when bulk access is provided to such third parties. Specific conditions as to use should be specified in the contractual terms.
- The BC therefore proposes that the examination of such a white list process should be referred to Council for consideration as a policy development process.

Task Force Two: data collection and display of data elements

The BC notes:

- **Privacy concerns:** The question of whether and how Whois data should be made public has been raised. It is unclear whether this question pertains to a broadly held governmental concern with all Whois data or whether the question relates to the narrow class of registrations by individuals with privacy concerns. In any case, the question of changing access to Whois data is a current and important one.
- **Registrant Awareness of public access to Whois:** The question has also been raised about whether registrants are aware of what Whois data is and how it is displayed and why it is needed.
- **Segregation of registrants into categories presents problems of definition.** There have been discussions about the concept of segregating registrants into different categories and having different requirements for gathering and publishing Whois data, based on the user category. The determination of what category a registrant fits into is not a simple determination, since, for example, individuals may register names for speculation, business development, or for personal use. And the reality is that the problems with consumer fraud, piracy, and trademark infringement are typically perpetrated by individuals, who provide false registration information, in order to avoid pursuit.
- **Differentiated or “tiered” Access by Authenticated Users:** There has been some limited discussion about creating a two tier approach to access and requiring a Whois user to be approved or authenticated to have access all data.
- **Services which offer anonymity for registrants:** Some have raised the issue of providing a mechanism for individual anonymity for legitimate individuals. Such mechanisms exist in telephony, where the telephony provider receives accurate contact information and acts as the point of contact for legitimate requests. Alternatively, anonymous gTLD registrations can be obtained by individuals through several mechanisms such as registration through one’s ISP.
- **Privacy and existing obligations:** Although some entities have raised the question of what privacy laws apply to Whois data, there is not a consistent interpretation of law. A few

countries have established that their privacy laws apply to the display of country-code Whois data. Certain data privacy entities have begun to ask what data privacy protections should apply. Yet many countries require businesses and NGOs to provide accurate information when they apply for services such as a business license, tax exempt status, inclusion in a directory, or trademarks.

- **All data elements are needed.** BC members responding to the questionnaire regarding data elements relied upon by business users indicated that all data elements are used. When some part of the elements are incomplete or inaccurate it is even more important to have access to as many data elements as possible. This enables a thorough effort at contacting the registrant, or in the case of consumer fraud, to support law enforcement.
- **Display of data elements:** All data elements should be displayed, or at a minimum accessible via an easy to use and validated process that would allow access to an authenticated user. However, this needs further and careful examination. It is not acceptable to simply create broad categories of ‘business’ and ‘individual’ without a recognition of the issues involving the misuse of a special category.

The BC therefore proposes:

- **All existing data elements are needed.** The BC recognises the continued need for all the data elements that are available in Whois today.
- **Registrants should be informed:** Fact based, neutral toned information about Whois should be included in the registration process, and specific acknowledgement/consent should be obtained at the time of registration. Registrants should also be renotified when they renew their registration of the importance of accurate and complete data.
- **Assessment of a differentiated access model should be undertaken:** Examination of the broad implications of establishing a differentiated access model, including costs, broad impact on registrants and Whois users, and taking into account CRISP and other emerging standards, should be a community and Council priority. The development of such a change in Whois will require a further PDP process.
- **Updated Information is needed to begin such a consideration:** The Council should be asked to support the briefing by all three TFs by IETF on the status of CRISP and any other emerging and relevant standards.

Task Force 3: Mechanisms to improve quality of contact data

The BC notes:

- **Accuracy because Whois is public communication.** A domain name registration in a TLD is a public form of communication, and as such, requires accurate data for the Whois registry.

- **Accuracy because users need accurate data.** The average Internet user, whether business, government, NGO or individual, has an expectation of accurate Whois information, which they then use to address legitimate issues: verifying the legitimacy of a web site, pursuing a network problem, addressing IP infringement concerns, calling for assistance from law enforcement, etc.
- **Accuracy is important for individuals and organisations.** The same concerns about the need for accurate data are independent of the nature of the registrant. A non-statistical survey of BC members regarding the situations they have experienced with trademark infringements, consumer fraud, and network issues indicates that there are problems with individuals and with organisations. However, none of the consumer fraud incidents encountered by the well-known brand holders involved organisations. The five situations examined all involved individuals who provided false information. Discussions with law enforcement have and continue to evidence similar problems with individuals.
- Some examples of data authentication exist in other industries, including financial services and in some of the ccTLDs.

The BC therefore proposes:

- **Best Practices are available from other sources:** The BC recommends further examination of best practices in authentication in other industries and from selected ccTLDs.
- **Changes to the contracts are needed to ensure there is enforcement.** The requirement to provide accurate data is a part of the Registrar contract, yet it appears that few registrars fulfill this requirement. The BC believes that this must be enforced by ICANN while allowing flexibility in the way registrars carry out this obligation. The previous Whois TF discussed the development of graduated sanctions. They also heard from several ccTLDs with successful data verification practices. The BC calls for the development of policy to evaluate a system of graduated sanctions.

Recommendation: more research is needed, and standards may offer solutions to development of modifications to Whois. Discussion of Whois is limited by a lack of research which would allow fact based policy. The ccTLD registries also have significant experiences which could be better understood and provide useful “understanding” to guide gTLD policy development. The BC encourages the GNSO Council to seek current information on both the CRISP project (on Whois standards undertaken by the Internet Engineering Task Force) and any other relevant standards process, to examine the role of these potential standards in providing a solution. The BC recognizes that the cost of implementing changes in Whois must be analyzed and understood as changes are considered. Changes in Whois should not become an “unfunded mandate” upon registrars.

Footnote: The BC continues to discuss the Whois issues and may provide further comments or modifications to these positions after concluding an ongoing internal process.

GTLD REGISTRIES CONSTITUENCY

This statement is submitted to the ICANN Generic Name Supporting Organization (GNSO) Whois Taskforce 1 on behalf of the gTLD Registry Constituency.

It should be noted that much of what Task Force I does relies on what Task Force II does. If Task Force II makes a recommendation that no data other than non-sensitive data would be displayed, then privacy and data mining become less significant issues. If Whois just shows domain name, IP address, Registrar, creation and expiration date, data mining could be reduced to minimal levels and port 43 concerns could mostly disappear. Because Task Force I and Task Force II are working concurrently, this statement does not assume any particular conclusions from Task Force II.

Process Summary

The gTLD Registry Constituency arrived at the positions described in this statement primarily through email discussions occurring from February through April 2004 supplemented to a small degree by discussions occurring as part of agendas for the in-person constituency meeting in Rome on 2 March 2004 and regular constituency teleconference meetings on 17 and 31 March 2004 and 7 April 2004. All constituency registry members were included in email discussions on the constituency list. Primary contributions were made by the following registry members: DotCoop (.coop), Global Name Registry (.name), Neulevel (.biz), Public Interest Registry (.org), SITA (.aero) and VeriSign (.com & .net). All nine registries participated in voting regarding specific elements of this statement and responses to questions discussed.

Issue Analysis – Impact on the Constituency

Operational Impact

The operational impact of changes to Whois access requirements can be very significant on registries depending on what the nature of the changes are, whether the registry is thick or thin, what implementation time frames are required, available resources, etc. It should also be expected that operational impact can be significant for registrars, possibly even more than registries because the registrars are the custodians of the primary Whois information and are typically the interface with registrants and their contacts.

Registry and registrar Whois systems as they exist today are relied on by millions of users around the world so any changes will potentially affect many if not all of those users. Consequently, it is critical to also consider the operational impact on the various types of Whois users outside of the registry and registrar constituencies.

One specific operational consideration that must be considered is the following: until such time as other means are available for registrars to obtain contact information of registrants associated with other registrars, registrars will need access to Whois data regarding registrants and administrative contacts in order to be able to comply with the new Registrar Transfer Policy;

registries and independent dispute providers will also need access to such data in order to fulfill their roles in the Transfer Dispute Resolution Policy.

Financial Impact

As with operational impact, financial impact to registries of changes to Whois access requirements would vary depending on what the nature of the changes are, whether the registry is thick or thin, what implementation time frames are required, etc. Until specific requirements are defined, it is not possible to quantify financial impact.

Some factors that could lead to increased cost for registries are:

1. The need for manual intervention in providing Whois service
2. Requirements that increase the likelihood of automated Whois queries
3. Complex requirements that cannot be standardized across multiple registries
4. Policies that increase the likelihood of litigation and other forms of dispute resolution
5. Requirements to provide different Whois services for different localities
6. Requirements that conflict with local law and thereby create burden on registries for negotiations and legal fees
7. Changes to the publicly available information - many registrants use Whois for monitoring their registration information and a number of web hosting firms and ISPs use it to confirm registration of domain names; changes to publicly available information could shift additional work to the registry

Any Whois access requirement changes that increase the likelihood of any of these factors occurring can be expected to have financial impact.

Implementation Timeframe Estimates

Registries, large and small, will require full product development cycles to implement any significant changes to Whois systems. These cycles vary by registry but can be longer than six months after final requirements are defined. Registrars also have similar requirements.

Because so many applications rely on Whois information, advance notice must be provided to the community at large to allow sufficient time for such applications to be modified to accommodate changes. Because of the widespread global use of Whois information, it is not unreasonable to expect that at least six months notice should be given to the Internet community for any significant changes to Whois access.

Questions Discussed by the Constituency

The gTLD Registry Constituency specifically raised and discussed six questions relating to the work of Whois Task Force 1. Summaries of the responses to the questions are provided below.

Question 1: *What types of access should be made available for viewing Whois information? (Web-based access, Port 43, Bulk Access, etc.)*

Question 1 Response	% Agree	Comments
Web-based Whois access should be at the discretion of any registry/registrar.	78%	<p>No registries opposed this; two abstained.</p> <p>For web-based Whois, access control is more limited than port 43 or IRIS. Web-based Whois seems most appropriate for a registry's or registrar's customers.</p> <p>Web-based Whois operates on a different port than both the Nicname/Whois protocol (port 43) and the CRISP Working Group's new protocol, IRIS. For web-based Whois, access control is more limited than port 43 or IRIS. Web-based Whois services use the Nicname/Whois protocol (and in the future, possibly IRIS) to gather Whois information from other registrars and registries. It is very difficult for web-based Whois services to gather information from other web-based Whois services. Therefore, at a minimum the Nicname/Whois service on port 43 or a protocol like IRIS must be kept open. However, it should be noted that the Nicname/Whois service does not provide adequate controls for tiered access.</p>
Any implementation of Whois access should permit registries to customize Whois access to applicable law.	100%	
Web-based and port 43 Whois service should not be required of registries and registrars as it is in current agreements with ICANN. (status quo)	100%	
Port 43 Whois access should only be required if it can be implemented to accommodate privacy legislation in the country where the registry operates.	100%	The CRISP IRIS protocol may be able to accommodate this concern.
Bulk access should not be allowed for marketing purposes.	100%	

Question 1 Response	% Agree	Comments
<p>Whois bulk access should not be required as it is under current unsponsored registry agreements.</p>	<p>89%</p>	<p>No registry opposed this; one abstained.</p> <p>Legal restrictions are an important part of an answer to question 1. For example, sponsored registries cannot provide Bulk Access to Whois to anyone except ICANN no matter what the outcome of the task force.</p> <p>Privacy considerations are coming to the fore more and more both on a national and European level and any opinion we volunteer on access to Whois is intimately connected to the legal restrictions of registry jurisdiction.</p> <p>IP community or law enforcement may need bulk access or something like it.</p>
<p>We recognize that certain parties (e.g., law enforcement, IP) may at times need to have better access to Whois. We suggest that a technical solution be identified which allows legitimate parties to search for the information they need, without requiring registries to turn over all data they have in the Whois (i.e., current bulk access). IRIS could be considered as a potential technical solution.</p>	<p>55%</p>	<p>Only five registries voted on this response; all five supported it.</p>
<p>As restrictions are and likely to remain standardized, it would be good to consider standardizing the request format too. With regard to access for registrars, an ICANN-administered registry of authorized IP numbers would be useful.</p>	<p>100%</p>	

Question 1 Response	% Agree	Comments
Non-registry and non-registrar access should be on a need-to-know basis and limited to users that can demonstrate a legitimate need for the information. For example, law enforcement agencies with an appropriate legal basis for a request, e.g., a subpoena, should be able to have access to personal information when necessary for law enforcement purposes. Intellectual property researchers should have access subject to agreements limiting its use.	78%	Only seven registries voted on this response and all of them supported it.

***Question 2:** What has been the effect on registry systems of having to make available Whois information via Port 43 and the web?*

Question 2 Response	% Agree	Comments
The effect on registry systems varies by registry. There has been little or no effect on the thin registry Whois offered for .com and .net. Larger thick registries have experienced operational problems arising from very high rates of requests on port 43, thereby requiring monitoring and maintenance of requisite servers. Smaller registries have not experienced significant negative impact.	89%	One registry, RegistryPro, abstained because it has not yet experienced these problems, but such issues are anticipated after launch.

***Question 3:** Have we noticed a problem with data mining? If so, do we have any facts to support this?*

Question 3 Response	% Agree	Comments
Registry Whois data mining tends to be more significant with larger thick registries. Data is available to support problems incurred. Some registries have received spam complaints from registrants.	89%	One registry, RegistryPro, abstained because it has not yet experienced these problems, but such issues are anticipated after launch.

Question 4: *If the answer to 3 is yes, have we instituted any mechanisms to deal with such mining (i.e., put in speed bumps on Port 43, or a cloudy GIF on web-based access? If yes, what has been the effect of instituting these measures?*

Question 4 Response	% Agree	Comments
Registries have instituted the following types of mechanisms to deal with data mining: 1) limitations on port 43 access; 2) timeouts which temporarily block high-rate users; 3) reduced returns on wildcard queries; 4) system tuning; 5) blocking IP numbers of large-volume abusive requests; and 6) rate controls. Publication of the delete pending list for registrars as required for RGP resulted in reduced mining for some registries.	89%	One registry, RegistryPro, abstained because it has not yet experienced these problems, but such issues are anticipated after launch.
Registries must be allowed to Implement anti-data-mining controls. Because restrictions have unpleasant side-effects for innocent parties, including registries and registrars, standardization of anti-data-mining practices should be considered to minimize undesirable side effects.	100%	

Question 5: *Is it feasible to have tiered access to Whois information (i.e., only some groups being able to use Port 43, while all others using web based access)? If so, how could that be implemented? What are the pros and cons? What issues would still need to be worked out?*

Question 5 Response	% Agree	Comments
Yes, it is feasible to have tiered access to Whois information.	100%	<p>The biggest burden with doing tiered access lies in the administration of authorization and authentication and not within the logistics of writing or running the service itself. IRIS will have specific mechanisms to allow registries/registrar to off-load this burden to policy-management entities (note: the protocol does not mandate the use of these mechanisms). This is important as it allows consistency of tiered access within a policy jurisdiction. Without such consistency, tiered access is much less useful.</p> <p>The two-tier Whois as described would require coordination between registries and registrars to avoid confusion amongst the relevant parties. Any moves toward tiered access would need to take into account the parties and their use of Whois information, i.e., the question of legitimate parties.</p>

Question 5 Response	% Agree	Comments
ICANN should administer an access rights database to Whois information, with appropriate separate treatment for different TLDs where necessary.	100%	<p>The issue of data privacy will inevitably lead to restricting Whois access and eventually create a situation where certain parties will have "better" access than others to Whois data.</p> <p>Providing a centralized administration of access rights will reduce a burden on each individual registry and move the responsibility for granting the access rights to the party which prescribed it.</p> <p>It is not clear that ICANN should administer access to Whois; registries should do that; but it does seem like it might be desirable for ICANN to authenticate access rights based on community input.</p>
Whois policy decisions should be based on the technologies that will be available (e.g., IRIS) not just those that exist today - port 43 Whois and "cloudy gif images".	89%	<p>No registry opposed this; one did not vote.</p> <p>CRISP's protocol documents ("IRIS") have finished last call in the working group and are now being sent to the IESG for their review and comment.</p>
The Whois framework must provide ways for registries and registrars to ensure that they can comply fully with their local legislation requirements. For example registries and registrars operating in Europe must be able to comply with European data regarding personal data processing.	89%	No registry opposed this; one did not vote.

Question 6: *In other words, how can we ensure that legitimate parties (however that is defined) have access to Whois information, but also reduce data mining and the burdens on our systems?*

Question 6 Response	Agree	Comments
The objectives of Whois must be clearly defined before the problem of data mining can be addressed.	100%	
Identification of "legitimate parties" is a core problem.	100%	

Question 6 Response	Agree	Comments
The question for a TLD registry is not just whether it can develop its own side of the IT solution, it must be sure that users (e.g., registrars and registrants) can comfortably follow.	100%	

Concluding Statements

1. It is essential to deal with the paramount concern of personal privacy along with the needs of intellectual property and law enforcement as limited exceptions to the protection of privacy.
2. We recognize that certain parties may at times need to have access to a number of elements listed in the current form of Whois. A technical means of providing this tiered access (i.e., allowing these parties to access the information, while preventing others from getting the information) could be through the IRIS protocol developed by the CRISP working group of the IETF. When finalized, we believe that a comprehensive review of this technical solution be undertaken. We believe a more detailed effort is needed to identify any specific parties that need access to selected elements and what information should be obtained about such access.
3. Cost benefit analysis should be done when considering any significant changes in Whois requirements.
4. Careful consideration should be given to the feasibility of registrars and registries to implement any proposed changes in Whois requirements including but not limited to enforcing such requirements. And sufficient time should be allowed for any associated migration.
5. The Whois framework must provide ways for registries and registrars to ensure that they can comply fully with their local legislation requirements.

INTELLECTUAL PROPERTY CONSTITUENCY

This statement responds to the issue identified in the purpose statement of the terms of reference for Task Force 1, see <http://gnso.icann.org/issues/whois-privacy/tor.shtml>

The purpose of this task force is to determine what contractual changes (if any) are required to allow registrars and registries to protect domain name holder data from data mining for the purposes of marketing. The focus is on the technological means that may be applied to achieve these objectives and whether any contractual changes are needed to accommodate them.

IPC opposes data mining of Whois for the purpose of marketing, although we believe there is strong evidence that Whois data is not a significant source of addresses for spam. Nevertheless, IPC supports, in principle, the use of query volume limitations on Port 43 access in order to discourage such practices. The uses for which trademark and copyright owners need access to domain name Whois do not ordinarily require the extremely high query volume levels that generally would be needed to mine the database for marketing purposes. Being supportive of the debate, the IPC submits that any changes in practice or regulation have to be designed in a manner that does not inadvertently have detrimental effects on the legitimate use of Whois. Based on the work of Task Force 1, we remain confident that this goal is feasible and can be achieved. To this effect, any effective technical/policy solution in the area of discouraging data mining of the domain name Whois database must take a number of points into account, including the following:

- Any provision should maintain and ensure availability of unhampered access to Port 43 for legitimate applications (such as research services) that require high volume access to domain name Whois for use in creating value-added products and services that are of great value to the intellectual property community and to the business community in general. As long as enforcement of the RAA provisions regarding bulk access to Whois remains almost non-existent, availability of port 43 access is essential in assuring the viability of these services.
- Adequate provision must be made for intermediaries which aggregate low-volume requests from end-users into a relatively high volume of queries through Port 43.
- A solution must identify realistic volume break-points between low-volume queries via Port 43 that should remain unrestricted, and a very high volume of queries that could, in principle, require an efficient and workable form of disclosure to registrars (or registries in the thick registry model) of the uses to which query results would be put.
- The solution should also preserve the unrestricted availability of Whois queries through a web-based interface, and the status of Port 43 as a service available free of charge.
- The solution must be accompanied by proactive enforcement of the obligation to make bulk access available.

- Finally, the solution must also address questions of scalability, particularly in the thin registry environment.

IPC does not currently take a position on whether or not the introduction of a solution as described above would require contractual modifications.

IPC would be interested in participating in an ongoing effort to develop such a solution. We propose that this effort be conducted by a small group representing all directly affected interests, on a realistic timeframe, and in a manner that will encourage candid consideration of the technical issues involved, all subject to final review by ICANN.

INTERNET SERVICE AND CONNECTIVITY PROVIDERS

Introduction

The ISPCP Constituency herein provides input to the three Whois Task Forces as required by ICANN by-laws. The ISPCP stresses the need for balanced policy that takes into consideration the interests of all stakeholders, and allows for the effective enforcement of civil and criminal laws while protecting registrant information from marketing or other illegitimate/illegal uses. This goal is the underlying theme running throughout the comments below. It is also consistent with commonly accepted tenets of privacy protections and laws throughout the world.

ISPCP Uses of Whois Data

- to research and verify domain registrants that could vicariously cause liability for ISPs because of illegal, deceptive or infringing content.
- to prevent or detect sources of security attacks of their networks and servers
- to identify sources of consumer fraud, spam and denial of service attacks and incidents
- to effectuate UDRP proceedings
- to support technical operations of ISPs or network administrators

Terms of Reference for Whois Task Forces

Whois Task Force 1

- Focused on restricting access to Whois data for marketing purposes
- Seeks to determine what contractual changes (if any) are needed to protect domain name holder data from data miners.
- What technological means are available to accommodate these possible contractual changes while simultaneously ensuring law enforcement, intellectual property, ISPs, and consumers continue to retrieve information necessary to perform their respective tasks

Whois Task Force 2

- Focused on reviewing Whois data collected and displayed to ensure accurate identification of registrants.
- Seeks to determine the best manner in which to inform registrants of what information is made publicly available when domain names are registered and options for restricting access
- Contemplates the ability of registrants to remove/shield certain parts of required contact information from anonymous, public access
- Furthering this is the need to determine what information may be removed, by whom, and what contractual changes are required to enable this.

Whois Task Force 3

--Focused on developing mechanisms to improve the quality of contact data that must be collected at the time of registration in accordance with the registrar accreditation agreement and the relevant registry agreement

--Related issues:

- Verification of data at time of registration
- Ongoing maintenance of data during registration period
- Protecting against deliberate submission of false information

ISPCP Position

Task Force 1 – Restricting Access to Whois Data

The ISPCP Constituency is in strong favor of limiting access to Whois data in respect of privacy concerns and does not see any legitimate purpose for access to bulk data for marketing purposes. ISPCP members spend tremendous resources to combat spam delivered through their networks and to their subscribers. Even minimal use of Whois data for marketing should be prohibited and further steps should be taken to enforce current policy limiting such use. However, the ISPCP opposes the notion that Whois data is not intended for enforcement purposes and that private parties do not have legitimate need for ready and efficient access to the data.

The ISPCP Constituency proposes that in light of forgoing interests:

- In light of small and regional ISPs' reliance on Port 43 access, the ISPCP Constituency believes its use ought to be preserved at this time. However, its use should be strictly limited by non-technical means such as rate limiting. In the long term, we strongly discourage its continued use.
- A general agreement would be useful on the types of uses that are legitimate and should be continued.
- Any proposed solution should include such legitimate access, including Web based queries and be scalable.
- ICANN staff should undertake development of a uniform access policy that is enforced – in addition, compliance procedures for such a policy should be implemented.
- The ISPCP rejects the notion that the purpose of Whois data is not intended for tracking registrants that are in the business violating laws or deceiving end users and thus, should not be used for any purpose beyond technical reasons.

Task Force 2- Review of Data Collected and Displayed

The ISPCP Constituency is aware of the real and legitimate privacy concerns over the amount and type of data collected and displayed in Whois data. Registrants should be provided with a limited list of needs for which their data may be used, so as to help prevent the possibility of inadequate notice. The ISPCP further notes that for a very small fraction of registrants with legitimate political and free speech concerns, there should continue to be processes in place for proxy registrations where their data will be kept private and provided only upon a limited set of circumstances.

There have been many assertions that the current display of Whois data is not legal or proper under the laws of some regions, namely the EU. However, of the EU member states' ccTLD operators who submitted Task Force 2 responses, all have indicated that they work closely with their respective country's data protection authorities and are in full compliance with their respective privacy laws.

Privacy concerns can further be alleviated by providing proper and adequate notice to all registrants, in a format that is conspicuous and highlights the disclosures within the registrant contract. In many regions it is a common legal requirement that data only be used for the purpose it was originally collected. By itemizing the legitimate needs for which one's data may be used, this requirement can be met.

The ISPCP Constituency proposes:

- That all elements continue to be collected and displayed, for those authorized to obtain access.
- That adequate and full disclosure must be provided regarding the uses of data, at the point of registration, and such requirement should be enforced.
- Anonymous gTLD registrations continue to be made allowed for individuals through current processes.
- The ISPCP supports the concept of tiered access as a principle, but is concerned with cost, enforcement and other practical implementation issues that must be clearly set forth prior to the implementation of such mechanism. The ISPCP will reserve final assessment on this principle until such time that a clearly defined and viable method is proposed.

Task Force 3 – Improving Accuracy of Collected Data

Finally, the ISPCP Constituency is quite concerned about the abundance of inaccurate and incomplete data. Such deficiencies significantly hinder ISPs' ability to identify and contact registrants. Thus, ISPs support ready access to accurate Whois data to facilitate resolution of network problems, sourcing of spam. Further, ready access to accurate data is necessary for the securing our networks and enforcing our acceptable use policies.

Because of the heavy reliance by ISPs on registrants' data to facilitate future contact with the registrant for business issues, security and stability issues, intellectual property infringement and a myriad of other legal issues, accuracy is of the utmost importance.

While automated verification software does exist, its accuracy and therefore its reliability on a global scale is suspect. Registrars should take a multiple steps to ensure that the data they receive is accurate, and there should be some enforcement mechanism to ensure registrars' compliance. In addition, it would be useful for registrars to have a list of best practices that further help verify data and produce an accurate database.

The ISPCP Constituency proposes:

- The creation of a best practices document aimed to improve data verification, with the prospect of a global application.
- Registrars take increased and more uniform measures to verify accurate data. The ISPCP does not advocate removing all flexibility from current or future registrar practices, but some uniformity and compliance with best practices will net a more accurate database.
- ICANN staff should undertake a review of the current registrar contractual terms and determine whether they are adequate or need to be changed in order to encompass improved data accuracy standards and verification practices.

NONCOMMERCIAL DOMAIN NAME HOLDERS CONSTITUENCY

Whois Task Force 1 (TF1) deals with the relatively narrow issue of restricting marketing users' access to Whois data through means other than bulk access under license.

NCUC notes, however, that the results of Whois TF1 may have implications for the other task forces, and vice versa. Our approach to TF1 takes this into account and will be guided by the following principles:

1. First and foremost, NCUC thinks it imperative that ICANN recognize the well-established data protection principle that the purpose of data and data collection processes must be well-defined before policies regarding its use and access can be established. The purpose of Whois originally was identification of domain owners for purposes of solving technical problems. The purpose was not to provide law enforcement or other self-policing interests with a means of circumventing normal due process requirements for access to contact information. None of the current Whois Task Forces are mandated to revise the purpose of the Whois directory. Therefore, the original, technical purpose must be assumed until and unless ICANN initiates a new policy development process to change it.
2. Second, based on input from the community NCUC does not believe it is possible to develop technical mechanisms that can restrict port 43 or port 80 access only to a specific type of purpose; e.g., "nonmarketing uses." Access restrictions imposed by TF1 will inevitably apply to any Whois user regardless of purpose. Moreover, restricting Port 43 access while leaving Port 80 open will only drive the automated processes to Port 80. Therefore we question whether TF1 can achieve anything of value.
3. Third, given the limited scope of TF1, we think it important for the task force to refrain from making judgments about the legitimacy of, justifications for, or "need" for any non-marketing uses. It is outside the scope of TF1 to make any such determinations. Accordingly, we will oppose any access restriction policy based on classification of users.
4. Fourth, we note that automated scripts or programs using port 43 are effectively a substitute for bulk access. According to George Papapavlou of the European Union, under data protection law bulk access is a "disproportionate, privacy infringing step, unless a very convincing, specific case can be made which has to be followed by due process. This applies not only to marketing but to any purpose." Therefore, a policy determination on port 43 access is best made in conjunction with a determination on bulk access, even though this is ruled out of scope by the task force's description of work.
5. Fifth, the best way to stop abuse of ports 43 or 80 is to get data that is valuable to spammers out of the public Whois database. Data that is in Whois will be accessible to lots of people; therefore, privacy concerns require getting data out of Whois or reducing access to it for all. This is, of course, a matter for Whois Task force 2, dealing with data elements.

6. Our participation in the entire Whois process will try to make sure that minor modifications in port 43 (or 80) access do not become an excuse for doing nothing else to protect Internet users' privacy.

Supplemental Statement submitted on May 9, 2004

NCUC opposes on principle the concept of a "White List" of authorized report of TF1, or that the lack of consensus on this idea be noted. If the latter route is taken, we ask that the following analysis of the reasons against the concept be afforded equal treatment in the report with the description of a White list and any reasons advanced for it.

Analysis

As we understand it, a "White List" is intended to give certain approved users the right to access sensitive data via port 43 (or other means). Organizations would apply for approval and once they were placed on the White list they could search, store and download sensitive Whois data, without any further restriction.

This concept is unacceptable to NCUC for the following reasons:

1. The concept is impractical. Creating such a list would add a huge operational burden to ICANN. There are hundreds of millions of Internet users and they come from every geographic region and language group, and involve data use purposes ranging from academic research to IP enforcement. ICANN would in effect be setting up a global certification process that had to be able to respond to all this diversity. If ICANN did this task conscientiously, the administrative burden would be huge. Not only would it have to investigate the legitimacy of each applicant, it should in principle also be able to constantly monitor the behavior of approved entities to make sure that they were not abusing their privileges. It would have to be willing to withdraw the privilege, and handle disputes and appeals relating to that.

If ICANN did not do this task conscientiously, if it simply added entities pro forma to the list whenever they applied, then there is no reason to create the list at all. Anyone and everyone could get the status, which is no different than opening up all Whois information to everyone.

2. The concept is discriminatory. The right to access Whois data must be balanced against the privacy rights of the domain name registrants. Once the proper balance is struck, all Internet users should have the same rights to access Whois data under the same terms and conditions. Intellectual property interests have no greater claim on that information than anyone else. The White List, in our opinion, is designed to create a two-class world of the spied-upon users, who have no rights, and privileged, surveillance- authorized users, who are permitted to spy on registrants.

3. The concept violates international privacy norms. A White List would give any approved user the equivalent of bulk access to Whois zone files. According to George Papapavlou of the European Union, under data protection law bulk access is a "disproportionate, privacy infringing step, unless a very convincing, specific case can be made which has to be followed by due

process. This applies not only to marketing but to any purpose." In other words, no one has the right to fish through sensitive personal data just to see if they can find anything of interest. But a White List would grant this right.

4. The White List concept is unnecessary. Under the proposals supported by registrars, NCUC, and ALAC, the concept of a known user with a known purpose making a request for each individual domain name she wants to investigate can give legitimate users and purposes access to the information they need without creating a centralized administrative entity and without violating privacy.

REGISTRARS STATEMENT

The registrars' policy recommendation for the Restricting Access/Data Mining Whois task force (TF1) has a great dependency on the results of the data collected and displayed (Whois task force (TF2)). If for example, the TF2 determines that the data to be displayed, especially via port-43, is limited to non-sensitive information ("non-sensitive information" defined as the domain itself, name servers, organization-names, and the registrar-of-record) and does not include personally identifiable information, then the information to be mined will be of less value to miners and hence, mining will be reduced. On the other hand, if the TF2 determines that sensitive information ("sensitive information" defined as, but not limited to, person-names, street addresses, phone and fax numbers, and email addresses) is to be displayed, then there will be a great incentive to mine the data because it will be more valuable. There is also a dependency on TF3, because if accuracy requirements are made more exacting, and at the same time, this far more accurate and current data is mandated to be displayed, then it becomes even more valuable, which further increases the motivation for mining. The potential rate of mining is a concern not only to the registrants, whose sensitive data is taken by miners, but also to registrars, for whom this has significant business implications.

Whois data is the registrant's information. It should remain in the control of the data subject as much as possible. As the Whois data moves away from the registrants to the registrars and further, to "thick" registries, and to even more distant (and un-identified) 4th and 5th parties, the registrant loses more and more control. As the public has learned more about how their information is abused, customers have begun to demand more privacy for their information and to object to such loss of control to parties with which they have no relationship or contact. Customers are not happy about their registrars publishing their sensitive Whois data because registrars can not guarantee that the "4th and 5th" parties would treat the data in a manner consistent with the policies and laws under which it was collected.

Requiring registrars to make data available to parties that they can not bind to any standards or restrictions flies in the face of registrars' responsibilities to their customers. Registrars are in the untenable position of having to comply with directly contradictory requirements – from ICANN, and from their customers and national privacy laws. As the Whois information is passed to these other entities, more access policy-control problems are created (because there are geometrically more locations at which to mine the data). Because the registrars are closer to the registrants, their customers, registrars are in the best position of protecting their customers' data, per the permissions provided by the registrants. To protect their customers, registrants strongly advocate for the ability to maintain data control. This means the right to display only non-sensitive information to the public, while providing appropriate limited access to the sensitive information. This also means providing only non-sensitive information at the registry level.

If TF2 determines that sensitive information must be displayed on the Web, the registrars support a policy whereby registrars **may**:

- 1) Shut off port-43 access to the public. This requires a definition of certain issues:
 - a. Who is the "the public"
 - b. Who has access

- i. Registrars must be granted access to port-43 Whois, in standardized format, but only for the purposes of performing transfers and only for so long as all gTLD registries are not EPP (thick or thin) or until another inter-registrar transfer mechanism replaces it.
 - ii. The identities of the non-public requestors must be known to the registrars and may be recorded by the registrars so that it can be communicated to the registrants in appropriate circumstances.
 - iii. The requestor must have a defined, valid purpose for each request and that purpose must be known to the registrars and may be recorded by the registrars so that it can be communicated to the registrants. Some registrars believe a valid purpose exists currently and some do not.
 - iv. The requestor cannot act as a proxy
 - c. Port-43 query rate limiting must be allowed to protect against mining, but the level of the limit must be determined.
- 2) Display the Whois information on a publicly accessible web site, but only in a manner such that the information cannot be easily mined, and consistent with the policies and governmental laws under which it was collected. It is the registrars' real-world experience that CAPTCHA systems (systems that perform checks for humans, such as requesting a person to type in number to access a single Whois record) and other systems (such as tracking the number of queries from a particular IP address), though imperfect, do work to greatly reduce automated data mining of the Whois via the web. Registrars must continue to be allowed to use such systems.
- 3) Continue to provide "identity protection" products to registrants.

The safeguards established for Port 43 access must be put in place for all analogous access points. All of the following access points provide a miner with access to all, or a large portion, of the Whois database of many registrants' sensitive information.

- 1) Mining of registrar's port-43 output
- 2) Mining of fat registry's port-43 output
- 3) Mining a 3rd party's port-43 that proxies access to any registrar's or registry's port-43 output
- 4) Mining the registrar's web-based display of Whois information
- 5) Mining the fat registries web-based display of Whois information
- 6) Bulk access

Therefore, they are the same, and any safe guard policies and controls put in place for one access point must be in place for the others. (For example, if the identity of the requestor (and purpose, lets say) must be known for bulk access, then it also must be known for mining (high query rate) of port-43.)

Whois Task Force 1
Description of Work

Amended: 29 October 2003

Title: Restricting access to Whois data for marketing purposes

Participants:

- 1 representative from each constituency
 - Jeffrey J. Neuman -- Chair
 - David Fares – Commercial and Business Users Constituency
 - Marilyn Cade -- Commercial and Business Users Constituency
 - David Maher – gTLD Registries Constituency
 - Jeremy Banks – IP Constituency
 - John Wolfe – IP Constituency
 - Tony Harris – ISPCP
 - Milton Mueller – Noncommercial Domain Name Holders Constituency
 - Paul Stahura – Registrars Constituency
- ALAC liaison
 - Wendy Seltzer
 - Thomas Roessler
- GAC liaison -- N/A
- ccNSO liaison – N/A
- SECSAC liaison – N/A
- liaisons from other GNSO Whois task forces – N/A
- up to three outside advisors – N/A

Description of Task Force:

=====

In the recent policy recommendations relating to Whois:

it was decided that the use of bulk access Whois data for marketing should not be permitted. However, these recommendations did not directly address the issue of marketing uses of Whois data obtained through either of the other contractually required means of access: Port 43 and web-based. Bulk access under license may be only a minor contributor to the perceived problem of use of Whois data for marketing purposes. A subset of a registrar's Whois database that is sufficiently large for data mining purposes may be obtained through other means, such as a combination of using free zonefile access (via signing a registry zonefile access agreement - the number of these in existence approaches 1000 per major registry) to obtain a list of domains, and then using anonymous (public) access to either port-43 or interactive web pages to retrieve large volumes of contact information. Once the information is initially obtained it can be kept up-to-date by detecting changes in the zonefile, and only retrieving information related to the changed records.

This process is often described as "data mining". The net effect is that large numbers of Whois records are easily available for marketing purposes, and generally on an anonymous basis (the holders of this information are unknown).

The purpose of this task force is to determine what contractual changes (if any) are required to allow registrars and registries to protect domain name holder data from data mining for the purposes of marketing. The focus is on the technological means that may be applied to achieve these objectives and whether any contractual changes are needed to accommodate them.

In-scope

The purpose of this section to clarify the issues should be considered in proposing any policy changes.

The task force should consider the effects of any proposed policy changes on the ability of groups such as law enforcement, intellectual property, internet service providers, and consumers to continue to retrieve information necessary to perform their functions.

The task force should consider the effects of any proposed policy changes on the competitive provision of domain name services including Whois access and transfers, and on the competitive provision of value-added services using Whois information.

Out-of-scope

To ensure that the task force remains narrowly focussed to ensure that its goal is reasonably achievable and within a reasonable time frame, it is necessary to be clear on what is not in scope for the task force.

The task force should not aim to specify a technical solution. This is the role of registries and registrars in a competitive market, and the role of technical standardisation bodies such as the IETF. Note the IETF presently has a working group called CRISP to develop an improved protocol that should be capable of implementing the policy outcomes of this task force. However, the task force should seek to achieve an understanding of the various technological means that could be applied to prevent or inhibit data mining with an eye toward evaluating their impact on other uses and their compatibility with the currently applicable contracts.

The task force should not review the current bulk access agreement Provisions, except to the extent that these can be improved to enhance protection against marketing uses and to facilitate other uses. These were the subject of a recent update in policy in March 2003.

The task force should not study the amount of data available for public (anonymous) access for single queries. Any changes to the data collected or made available will be the subject of a separate policy development process.

Tasks/Milestones

- collect the stated needs and the justification for those needs from non-marketing users of contact information (this could be extracted from the Montreal workshop and also by GNSO constituencies, and should also include accessibility requirements (e.g based on W3C standards)

[milestone 1 date]

- review general approaches to prevent automated electronic data mining and ensure that the requirements for access are met (including accessibility requirements for those that may for example be visually impaired) [milestone 2 date]

- determine whether any changes are required in the contracts to allow the approaches to be used above (for example the contracts require the use of the port-43 Whois protocol and this may not support approaches to prevent data mining) [milestone 3 date]

Each milestone should be subject to development internally by the task force, along with appropriate public comment processes (e.g seeking specific advice from the technical community, or from Whois service operators) to ensure that as much input as possible is taken into account.

Comments for the Whois Task Force 1 Preliminary Report can be submitted to: whois-tf1-report-comments@gns0.icann.org.

The comment archives are available at <http://gns0.icann.org/mailling-lists/archives/whois-tf1-report-comments>.