Inventory of WHOIS Service Requirements – Final Report

STATUS OF THIS DOCUMENT

This is an Inventory of WHOIS Service Requirements, requested by the GNSO Council. As requested by the GNSO in its resolution of May, 2009, staff shared the initial draft with the GNSO Council and with the other SOs (ccNSO, ASO) and ACs (SSAC, ALAC, GAC) for their input. This document is an updated report following those consultations as well as GNSO Council and community discussions in Brussels.

SUMMARY

THIS FINAL REPORT IS SUBMITTED TO THE GNSO COUNCIL IN RESPONSE TO A REQUEST RECEIVED FROM THE COUNCIL ON 7 MAY 2009.

TABLE OF CONTENTS

1 Executive Summary

For many years, the GNSO community has discussed a variety of concerns about WHOIS, and there is recognition that the current WHOIS service might decrease in reliability and usefulness over time. Recognizing these concerns, in May 2009 the GNSO Council asked staff to compile a comprehensive set of requirements for WHOIS that includes known deficiencies in the current service and "any possible requirements that may be needed to support various policy initiatives that have been suggested in the past" [21]. In compiling this report, staff has attempted to distil from current requirements and policy discussions the technical requirements that would be necessary to implement to correct deficiencies and implement various policy proposals. The report is a technical inventory and does not intend to define or suggest the policies or operational rules that should apply. As requested by the GNSO in its resolution of May, 2009, staff shared the initial draft with the GNSO Council and with the other SOs (ccNSO, ASO) and ACs (SSAC, ALAC and GAC) for their input. This updated report reflects input received during those consultations as well as GNSO Council and community discussions in Brussels.

2 Introduction

Created in the 1980s, WHOIS began as a service used by Internet operators to identify and contact individuals or entities responsible for the operation of a network resource on the Internet. The WHOIS service has since evolved into a tool used for many purposes, such as determining whether a domain name is available for registration, identifying the registrant of a domain name that has been associated with malicious activities, contacting domain name registrants on matters related to trademark protection, and verifying online merchants.

As usage of WHOIS evolved, few changes have been made to the protocol or the services that make use of the protocol. There are increasing community concerns that the current WHOIS service is deficient in a number of ways, ranging from data accuracy and reliability, to other technical areas, such as accessibility and readability of WHOIS contact information by users whose local languages cannot be represented in US-ASCII7. These are noted in recent reports from ICANN's Security and Stability Advisory Committee (SSAC), in reports of other ICANN supporting organizations and advisory committees and by external sources.

Acknowledging these concerns, in May 2009, the Generic Names Supporting Organization (GNSO) Council asked the staff to compile a comprehensive set of requirements for WHOIS service based on current requirements and a review of previous GNSO WHOIS policy work.

Based on Council discussions leading up to Council approval of this request and following staff has interpreted the term *requirements* as *technical requirements*. These requirements came from three areas: current service features that have been identified as needing improvement, features needed to support policy proposals that have been discussed in the past, and features recommended by ICANN supporting organizations and advisory committees.

This report is organized as follows. In section 2, we present background information on WHOIS service as well as the GNSO Council resolution. In section 3, we list current WHOIS service requirements specified in various registrar and registry contracts. We then enumerate requirements and cite their origins in sections 4, and in section 5 we present a straw-man draft of requirements for community consultation.

As requested by the GNSO, staff shared the initial draft of this report on March 25 with the GNSO Council and with the other SOs (ccNSO, ASO) and ACs (SSAC, ALAC, GAC) for their input. In addition, we held two webinars in April and May 2010 to brief the broader community on the initial report and to solicit feedback. As of May 31 2010, staff have received formal comments from the ALAC and Registry Stakeholder Group (RySG), and from the following community members: Tom Vest, Luis Diego Espinoza S, Marco d'Itri, and the following group of technical experts: Jaap Akkerhuis, Patrik Fältström, James Galvin, Warren Kumari, and Doron Shikmoni, who provided a joint set of comments. We thank them for their input, and this report has been updated based on their comments. We include the formal comments from the ALAC and Registry Stakeholder Group in Annex III and IV of this report.

3 Background and Terminology

The WHOIS service and protocol were originally developed and deployed in 1982 as a transaction-based service that “provides online directory look-up equivalent to the ARPANET Directory” [5]. The original WHOIS protocol, in simple terms, is a client-server, query-response protocol. The client query 1) connects to the service host (SRI-NIC) at TCP or NCP service port 43; 2) sends a single “command line”, and 3) signals the end of the command line with a <CRLF> (carriage-return and line-feed) character sequence. The server listening to port 43 4) accepts and parses the query, 5) composes a response, again using a <CRLF> to signal end of response, and 6) returns the response to the client. The server closes its connection as soon as the output is finished [5]. The client displays the response to the standard output or processes the response as otherwise indicated by the user.

Although the protocol was subsequently modified in 1985 (RFC 954 [6]), and again in 2004 (RFC 3912, [2]) to remove historical references to protocols and authorities and to generalize the applicability of WHOIS to the Internet community, the core functionality remains the same. The only notable change in the WHOIS service is that with the advent of the World Wide Web, looking up WHOIS information via the web has become a popular alternative to the original command-line client-server WHOIS application. At present, ICANN accredited registrars, gTLD registries, some country code TLD registries (ccTLDs) and regional Internet registries offer Web-enabled and client-server WHOIS services.

3.1 Components of the WHOIS service

The WHOIS service includes WHOIS clients, WHOIS servers, WHOIS data stores, and WHOIS data (domain name registration records). Figure 1 illustrates a simple interaction of the WHOIS system components in a simple WHOIS query “icann.org”. We also explain each component in detail below.

It is critical to understand that the WHOIS service is *not* a single centrally managed data store. Rather, registration data are held in disparate locations and administered by multiple, autonomous parties who set their own standards and conventions for WHOIS service.

When people refer to WHOIS, they may mean several different things. Some mean the WHOIS protocol that specifies the network exchange between a WHOIS client and server, whereas others refer to a perceived or conceptual WHOIS database that registries or registrars support, or to the WHOIS data that registrants provide and ICANN accredited registrars are obliged to make public according to the terms of the Registrar Accreditation Agreement (RAA). In this document, unless otherwise specified, when we refer to WHOIS, we mean the WHOIS services that include the WHOIS clients of all kinds and WHOIS servers that implement the protocol as well as the databases that store the domain registration data (every component in Figure 1 below).

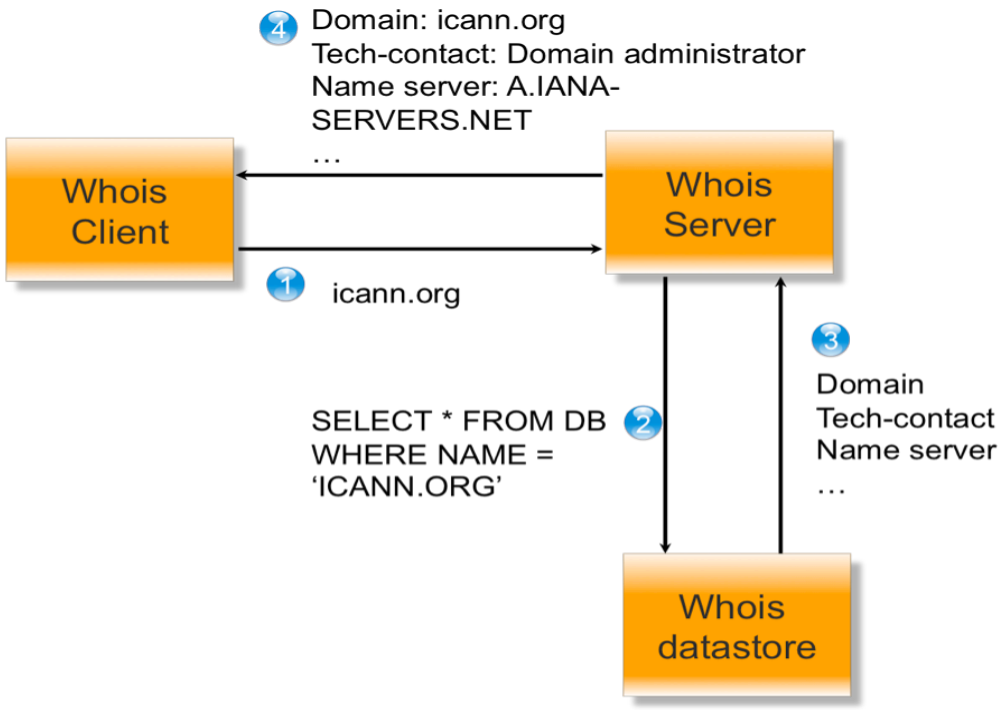

Figure 1: A simple WHOIS query. 1) A WHOIS client makes a request to the authoritative server that maintains the WHOIS data for "icann.org", 2) the WHOIS server translates the request to a database query and queries the data store, 3) the data store returns the query results, 4) the WHOIS server formats the output and sends it back to the client, and 5) the client displays the response to the user.

WHOIS clients can be categorized into command-line clients, web clients and automated client applications. Originally, the only method to contact a WHOIS server was to use a text-based, command line interface client available from a Unix or Unix-like operating system. A WHOIS command line client typically has options to choose which host to connect to for WHOIS queries; often, a default WHOIS server list is preconfigured for the client.[1] Like most TCP/IP client/server applications, a WHOIS client takes the user input, opens an IP socket to the destination server on port 43, sends the query conforming to the WHOIS protocol, and waits for a response from the server. The client either displays the response to the end-user or uses the data to make additional queries [39].

With the advent of the World Wide Web, looking up WHOIS information via the web has become more accessible. At present, web-based WHOIS queries may be performed through the websites of ICANN accredited registrars and registries (pursuant to RAA 3.3), most regional Internet registries, and third

[1] Additional options may allow control of what port to connect on, displaying additional debugging data, or changing recursion/referral behavior.

party WHOIS client providers. Early web-based WHOIS clients were simple front-end processes that invoked a command-line client application, and the WHOIS responses were displayed on a webpage with little, if any formatting. More recently developed web based WHOIS clients usually perform the WHOIS queries directly and then format the results for display [39].

In recent years, to accommodate automatic processing of WHOIS information, many automated WHOIS client applications have been written. These programs not only provide WHOIS query functionality, but often provide the ability to parse the WHOIS results into individual components as well. One example of such client is the NET::WHOIS module maintained by the Comprehensive Perl Archive Network (CPAN), a large collection of Perl software.[2]

WHOIS servers are usually maintained by registrars and registries that offer the WHOIS service. The server software listens to requests on port number 43. When a request comes in, it processes the request, connects to the registrar or registry data store, retrieves the WHOIS data based on the query, composes and sends a response to the client, and immediately closes the connection. To avoid abuse, registrars and registries usually implement technologies and policies to limit the rate of queries a client can make (for example, one query per second). Rules regarding access and query limits vary across registrars and registries.

WHOIS data stores are databases where the WHOIS data resides. Normally, these are relational database(s), although this does not have to be the case. Some registrars may simply put the WHOIS data in a text file.

WHOIS data refers to the registration data that registrants provide and registrars or registries disclose. The Registrar Accreditation Agreement (RAA 3.3.1) specifies the following data elements that must be provided by registrars in response to a query:

3.3.1.1 The Registered Name;
3.3.1.2 The names of the primary nameserver and secondary nameserver(s) for the Registered Name;
3.3.1.3 The identity of Registrar (which may be provided through Registrar's website);
3.3.1.4 The original creation date of the registration;
3.3.1.5 The expiration date of the registration;
3.3.1.6 The name and postal address of the Registered Name Holder;

[2] http://search.cpan.org/search?m=module&q=WHOIS

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

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

Bulk Whois Access: Registries and registrars also provide bulk WHOIS access to third parties.

Regarding the definition of WHOIS, the ALAC noted the following:

> "As noted in the report under 3.1, Components of the WHOIS service, the name "WHOIS" refers to multiple concepts and it is important to distinguish between them. The At-Large suggests it might be necessary to come up with another name to refer to the "WHOIS service", to avoid confusion with the WHOIS protocol. This is especially true if the service itself might be running over other protocols in the future."
>
> "We define the WHOIS service as an interaction between the client and the server, running on TCP port 43, and implementing the protocol defined in RFC3912. We disagree that web-based interfaces that query a database can be considered "WHOIS clients". They do not suffer from the same limitations as the text-based clients, and can easily handle authentication, internationalization and anti-abuse features."

3.2 Usage of WHOIS

Internet operators originally use WHOIS to identify individuals or entities responsible for the operation of a network resource on the Internet. Over time, it evolved to serve the need of different stakeholders such as registrants, law enforcement agents, intellectual property and trademark owners, businesses and individual users. In addition to accessing domain information, it is also used to access other resources such as autonomous system (AS) numbers, IP addresses, registrars, nameservers, contact handles. In this section, we list some of the known uses or abuses of WHOIS services. This list was originally compiled in SSAC 23 [25], and is extended here.

WHOIS service today is used:

- To determine whether or not a given domain is available.

- To contact network administrators for resolution of technical matters related to networks associated with a domain name (e.g., DNS or routing matter, origin and path analysis of DoS and other network-based attacks).
- To diagnose registration difficulties. WHOIS queries provide information that is often useful in resolving a registration issue, such as the creation and expiration dates and the identity of the registrar.
- To contact web administrators for resolution of technical matters associated with a domain name.
- To obtain the real world identity, business location and contact information of an online merchant or business, or generally, any organization that has an online presence.
- To associate a company, organization, or individual with a domain name, and to identify the party that is operating a web or other publicly accessible service using a domain name, for commercial or other purposes.
- To contact a domain name registrant for the purpose of discussing and negotiating a secondary market transaction related to a registered domain name.
- To notify a domain name registrant of the registrant's obligation to maintain accurate registration information.
- To contact a domain name registrant on matters related to the protection and enforcement of intellectual property rights.
- To establish or look into an identity in cyberspace, and as part of an incident response following an Internet or computer attack. (Security professionals and law enforcement agents use WHOIS to identify points of contact for a domain name.)
- To gather investigative leads (i.e., to identify parties from whom additional information might be obtained). Law enforcement agents use WHOIS to find email addresses and attempt to identify the location of an alleged perpetrator of a crime involving fraud.
- To investigate spam, law enforcement agents look to the WHOIS database to collect information on the website advertised in the spam.

Miscreants use WHOIS:

- To collect or "farm" email addresses for the purpose of delivering unsolicited, bulk electronic mail.
- To gather information about a company, organization, or individual as part of the footprinting and target acquisition phase of an *Internet attack*. Internet footprinting involves searches and queries of available publicly accessible databases, including web pages, the U.S. Securities Exchange Commission's Electronic Data Gathering, Analysis, and Retrieval (EDGAR) database, WHOIS, and DNS.

- To gather information about a company, organization, or individual as part of the footprinting and target acquisition phase of a *social engineering attack*. For example, a miscreant may query the WHOIS of a targeted organization, contact the sponsoring registrar, impersonate the registrant to a registrar's customer care, and convince the registrar's customer care representative to grant him access to the registration account for the targeted domain. The miscreant may also attempt to impersonate one of the registrant's contacts to deceive a different contact in this manner as well.
- To acquire information about registrants for other abusive purposes.

3.3 May 2009 GNSO council Resolution and Scope of Work

As noted above, although WHOIS has evolved to serve many stakeholders, few changes are made to the protocol. As a result, there are increasing community concerns that the current WHOIS service is deficient in a number of ways. Recognizing these concerns, in May 2009, the GNSO Council asked the staff to compile a comprehensive set of requirements for an improved WHOIS service.

Text of Council Resolution

"Whereas there have been discussions for several years on the adequacy of the current set of WHOIS tools to provide the necessary functions to support existing and proposed WHOIS service policy requirements,

and, there have been questions as to the adequacy of these tools for use in an IDN environment,

and, that there have been extensive discussions about the requirements of the WHOIS service with respect to Registry and registrar operations, and, new architectures and tools have been developed and suggested by the technical community,

and, the GNSO accepted the recommendation of the IRT-A Working Group to encourage staff to explore further assessment of whether IRIS would be a viable option for the exchange of registrant email address data between registrars and conduct an analysis of IRIS' costs, time of implementation and appropriateness for IRTP purposes,

Resolved,

The GNSO Council requests that Policy Staff, with the assistance of technical staff and GNSO Council members as required, collect and organize a comprehensive set of requirements for the WHOIS service policy tools.

These requirements should reflect not only the known deficiencies in the current service but should include any possible requirements that may be needed to support various policy initiatives that have been suggested in the past.

The synthesis of requirements should be done in consultation with the SSAC, ALAC, GAC, the ccNSO and the GNSO and a straw-man proposal should be prepared for these consultations. The Staff is asked to come back with an estimate of when this would be possible." ([21])

Scope of the Work

Staff has developed this report based on the GNSO Council's request to collect a comprehensive set of requirements for the service policy tools. These requirements have been derived from three sources: current features identified as needing improvement (or known deficiencies); features to support policy initiatives that have been suggested in the past; and features recommended by the ICANN community, such as various supporting organizations and advisory committees. Following release of this report, staff will gather input from the supporting organizations and advisory committees and produce a final report.

For a given WHOIS policy, there are typically both technical and operational requirements. Technical requirements are concerned with the capability of the supporting technology while the operational dimension is concerned with making a given policy operational. To illustrate this distinction, consider the tiered access proposal [18], which envisioned different levels of access to non-public WHOIS data depending on who was requesting the data. If a proposal of this kind were to be adopted as a consensus policy, the technical dimension would be concerned with finding the right technology that is capable of authenticating users to the WHOIS system and granting proper access to different resources based on each requester's identity or role. The operational dimension would be concerned with putting such an access policy into operation. To further illustrate, if a consensus policy required, for example, that law enforcement should be given access to certain information, operationally we would need a working definition of "law enforcement," and a means to determine which individuals would be included in such a "group," and to determine what credentials those entities would need to submit to demonstrate membership in such a group. Based on discussions with the Council when this request was initiated, staff has interpreted the term *requirements* to mean *technical requirements*. Thus, staff has prepared this report with the understanding that the goals are *neither* to gather policy requirements, *nor* to recommend policy, *nor* to consider operational requirements. In this report staff has attempted to gather a set of technical requirements (including the data and supporting technology) that would enable the deployment of policies that may be developed via consensus processes. The report seeks not to recommend or attempt to specify policy but to speak only to the capabilities that are required by underlying technology to *support* a policy. For instance, this document may say that a protocol MUST support a certain feature. This only means that it is necessary for the software implementing the protocol to provide the feature; in practice, a service provider may voluntarily implement the feature but would only be obligated to support the feature if a consensus policy were developed.

Today, entities operate WHOIS service include registrars, gTLD registries, ccTLDs and Regional Internet Registries (RIRs). The staff recognizes that ICANN's contractual relationships are limited to gTLDs registries and accredited registrars, however in this document we try to be more inclusive and we also include potential improvements to the WHOIS service that could be adopted by ccTLDs and RIRs.

Finally, the terms “must” and “should,…” are used in this document in compliance with RFC2119.

4 Current WHOIS service requirements

We begin our requirements compilation by examining existing WHOIS service requirements. Through its contracts and agreements, ICANN requires gTLD registries and registrars to offer WHOIS services.

4.1 WHOIS Service at the Registry Level

For the generic top-level domain (gTLD) registries, ICANN specifies WHOIS service requirements through the registry agreements (ICANN 2009 Registry Agreements).

Registries satisfy their WHOIS obligations using different services. The two common models are often characterized as “thin” and “thick” WHOIS registries. A thin registry only includes data sufficient to identify the sponsoring registrar, status of the registration, creation and expiration dates for each registration, name server data and last time the record is updated in its WHOIS data store. Registrars maintain the complete set of registration data as required by RAA 3.3 for those domains they sponsor. .COM and .NET are examples of thin registries. Thick registries maintain the registrant’s contact information and designated administrative and technical contact information, in addition to the sponsoring registrar and registration status information supplied by a thin registry. .INFO and .BIZ are examples of thick registries.

As an example of thick and thin WHOIS, consider the WHOIS for cnn.com and cnn.org. Both domains are registered by CNN, but one of them is in thin registry (.COM), the other is in thick registry (.ORG). If we query .COM’s WHOIS server for cnn.com, we get the following results:

```
Domain Name: CNN.COM
Registrar: CSC CORPORATE DOMAINS, INC.
WHOIS Server: whois.corporatedomains.com
Referral URL: http://www.cscglobal.com
Name Server: NS1.TIMEWARNER.NET
```

```
   Name Server: NS3.TIMEWARNER.NET
   Name Server: NS5.TIMEWARNER.NET
   Status: clientTransferProhibited
   Updated Date: 04-feb-2010
   Creation Date: 22-sep-1993
   Expiration Date: 21-sep-2018
```

However, if we query the .org's whois server, we get the full result of WHOIS information:

```
Domain ID:D5353343-LROR
Domain Name:CNN.ORG
Created On:16-Apr-1999 04:00:00 UTC
Last Updated On:04-Feb-2010 22:48:15 UTC
Expiration Date:16-Apr-2011 04:00:00 UTC
Sponsoring Registrar:CSC Corporate Domains, Inc. (R24-LROR)
Status:CLIENT TRANSFER PROHIBITED
Registrant ID:1451705371f82308
Registrant Name:Domain Name Manager
Registrant Organization:Turner Broadcasting System, Inc.
Registrant Street1:One CNN Center
Registrant Street2:13N
Registrant Street3:
Registrant City:Atlanta
Registrant State/Province:GA
Registrant Postal Code:30303
Registrant Country:US
Registrant Phone:+1.4048273470
Registrant Phone Ext.:
Registrant FAX:+1.4048271995
Registrant FAX Ext.:
Registrant Email:tmgroup@turner.com
…
```

The content of registration data provided via WHOIS may differ across TLDs registries. Some gTLD registry agreements, such as .tel, have provisions in place that in certain circumstances exclude personal information from the public WHOIS. For example, .tel WHOIS output for individuals may only mention registrant's name with no other contact information.

The ALAC noted a similar point, that "the WHOIS protocol and associated servers and clients are being used outside the gTLD space. ccTLDs use them in a way similar to gTLDs, but often need to implement variations on the server side to comply with local laws on privacy."

4.2 WHOIS Service at the Registrar Level

ICANN requires registrars to provide public access to data on registered names through the Registrar Accreditation Agreement (RAA) (ICANN 2009 RAA). Specifically, the 21 May 2009 RAA requires that "At its expense, Registrar shall provide an interactive web page and a port 43 WHOIS service providing free public query-based access to up-to-date (i.e., updated at least daily) data concerning all active Registered Names sponsored by Registrar for each TLD in which it is accredited." (ICANN 2009 RAA 3.3.1). ICANN has implemented policies and measures to improve the accuracy and availability of domain name registration records, including

- the WHOIS Data Reminder Policy (WDRP) [7],
- the WHOIS Data Problem Reporting System (WDPRS) [8] , a problem reporting system that allows parties to report allegedly inaccurate WHOIS data and requires that registrars verify the data with the registrant, and
- annual WDRP compliance audits, and ICANN commenced a WHOIS data accuracy audit in 2007.

In addition to these requirements, registrars are required to provide third-party bulk access to WHOIS data. The RAA requires that:

> "Registrar shall make a complete electronic copy of the data available at least one (1) time per week for download by third parties who have entered into a bulk access agreement with Registrar." (ICANN 2009 RAA 3.3.6)

The RAA requires registrars to investigate claims of WHOIS inaccuracy:

> "Registrar shall, upon notification by any person of an inaccuracy in the contact information associated with a Registered Name sponsored by Registrar, take reasonable steps to investigate that claimed inaccuracy. In the event Registrar learns of inaccurate contact information associated with a Registered Name it sponsors, it shall take reasonable steps to correct that inaccuracy." (ICANN 2009 RAA 3.7.8)

There are several other important WHOIS requirements set forth in the RAA and in various registry agreements:

Authorized License Terms: Section 3.3.5 of the RAA prohibits use of WHOIS for marketing purposes. Registrars are required to permit the use of WHOIS data for any lawful purposes except to "(a) allow, enable, or otherwise support the transmission by e-mail, telephone, or facsimile of mass, unsolicited, commercial advertising or solicitations to entities other than the data recipient's own existing customers; or (b) enable high volume, automated, electronic processes that send queries or data to the systems of any Registry Operator or ICANN-Accredited registrar, except as reasonably necessary to register domain names or modify existing registrations." (ICANN 2009 RAA 3.3.5). In actuality, many registrars and registries implement rate limiting of WHOIS queries.

Data Escrow: Section 3.6 of the RAA specifies that registrars should submit an electronic copy of the registrar's database to an escrow agent on a schedule, under the terms, and in the format specified by ICANN. This database includes the public WHOIS data and other data that registrars hold such as the billing data for the customer. In addition, the database includes information that the Registrar chooses to escrow related to customers of any privacy service or licensee of any proxy registration service offered by a Registrar. If the Registrar decided not to escrow such customer data, it is required to display a conspicuous notice to such customers at the time an election is made to utilize such privacy or proxy service that their data is not being escrowed. It is important to note that the RAA does not require any specific information to be included in the WHOIS Service with respect to these privacy or proxy services.

Centralized WHOIS: Section 3.3.4 of the RAA specifies that if the WHOIS service implemented by registrars "does not provide reasonably robust, reliable and convenient access to accurate and up-to-date data, and ICANN deem it necessary (or through the consensus policy process), the registrars may need to supply data from Registrar's data store to facilitate the development of a centralized WHOIS database for the purpose of providing comprehensive Registrar WHOIS search capability." (ICANN 2009 RAA 3.3.4) It is worth noting that **Section 3.1 of the 2006 .COM registry agreement also specifies that "**Registry Operator shall develop and deploy a centralized WHOIS for the .com TLD if mandated by ICANN insofar as reasonably feasible, particularly in view of Registry Operator's dependence on cooperation of third parties [10]."

5 Potential New Requirements

In this section, we identify a list of potential new requirements. These "requirements" come from three sources: current features identified as needing improvement (or known deficiencies); features to support all policy initiatives that have been suggested in the past; and features recommended by the ICANN community, such as various supporting organizations and advisory committees.

This list was further complemented by reviews of academic and industry research on WHOIS, reviews of opinions and statements from various ICANN supporting organizations and constituencies (e.g. intellectual property constituency), and advisory committees (e.g. SSAC), and interviews with developers of WHOIS software and other experts. For each of the proposed requirements, we cite the origin of the requirement.

5.1 Mechanism to find Authoritative WHOIS servers

Currently each registry, and for thin registries, each registrar, also maintains its own authoritative WHOIS server. There is no easy way to find out the domain names and IP addresses of the WHOIS servers for a given TLD[3] or registrar, although whois.nic.TLD is a common host naming convention for WHOIS servers. Clients attempt to derive the host name of the proper WHOIS server through a combination of heuristics, hardwired tables, DNS SRV records, etc.

With the advent of the new gTLD program [12], the number of registries could significantly increase, thereby making it even more difficult for WHOIS clients to constantly maintain its hardwired tables.

To solve this problem, registrars and registries could make this information available for ICANN to publish. Alternatively, they could agree to a uniform naming convention, or publish the host name at a well-known location (e.g. on their website or in their DNS as an unique resource record [34]). Whatever mechanism is used, the host name information should be both machine and human-accessible.

Inventory item R-1:

- Provide a publicly accessible and machine parsable list of domain names or IP locations of WHOIS servers operated by ICANN accredited registrars, gTLD registry operators, ccTLDs operators, and regional internet registries (RIRs) [23, 31].

[3] It should be noted that the IANA maintains a root zone website that includes the WHOIS server location data for some TLDs.

5.2 Structured Queries

The WHOIS server applications vary with respect to how they expect clients to format query data. WHOIS client applications also vary with respect to how they have compensated for the server variability. These factors affect user experience in a negative way. Because the query syntax can change from server to server, that also adversely affects automation. For example, to query the ARIN Database for information about an AS number, you would use:

```
whois –h whois.arin.net a 6.
```

However, to query the RIPE Database for information, you would use:

```
whois –h whois.ripe.net –Taut-num as7.
```
[4]

Users would benefit from a standard query structure. For example, a user may wish to submit a list of domain names to WHOIS to check for the creation date of the domain names. With a standard query structure, he or she can do so without checking out the specific syntax for each WHOIS servers.

Inventory item R-2:

- Define a standard query structure that clients can implement and that all gTLD registries and ICANN accredited registrars will support.

5.3 Well-defined schema for replies

The WHOIS protocol (RFC 3912) does not specify a format the WHOIS server must use when it composes a data response. Registrars and registries return the WHOIS query results differently. For example, a query of cnn.com using command line WHOIS client (version 4.7.3 ubuntu2) yields the following results:

```
Registrant:
      Turner Broadcasting System, Inc.
      Domain Name Manager
      One CNN Center 13N
      Atlanta, GA 30303
```

[4] Staff notes that in the context of RIR-allocated resources then there is an existing standard (RPSL) which is used by RIPE, AFRINIC and APNIC and partially mimicked by LACNIC.

```
      US
      Email: tmgroup@turner.com

   Registrar Name....: CORPORATE DOMAINS, INC.
   Registrar WHOIS...: whois.corporatedomains.com
   Registrar Homepage: www.cscprotectsbrands.com

….

   Domain Name: cnn.com

      Created on..............: Thu, Apr 02, 2009
      Expires on..............: Fri, Sep 21, 2018
      Record last updated on..: Thu, Jun 11, 2009

   Administrative Contact:
      Turner Broadcasting System, Inc.
      Domain Name Manager
      One CNN Center 13N
      Atlanta, GA 30303
      US
      Phone: +1.4048273470
      Email: tmgroup@turner.com
```

A similar WHOIS query of cnn.org that was also registered by CNN yields the following results:

```
Domain ID:D5353343-LROR
Domain Name:CNN.ORG
Created On:16-Apr-1999 04:00:00 UTC
Last Updated On:27-Jun-2009 00:30:50 UTC
Expiration Date:16-Apr-2011 04:00:00 UTC
Sponsoring Registrar:CSC Corporate Domains, Inc. (R24-LROR)
Status:CLIENT TRANSFER PROHIBITED
Registrant ID:145170537lf82308
Registrant Name:Domain Name Manager
Registrant Organization:Turner Broadcasting System, Inc.
```

```
Registrant Street1:One CNN Center
Registrant Street2:13N
Registrant Street3:
Registrant City:Atlanta
…
```

Generally speaking, responses today are commonly unstructured USASCII7 or unstructured UTF-8 for gTLD registry and registrars. Some ccTLDs respond using a Latin-1 format, rather than answering in UTF-8, and many more have no regard to charsets. Lack of standard format or data structure for responses makes it difficult for humans to interpret the result and also for programs to parse the results.

Currently, most structured data formats are based on XML format. Therefore, SSAC, along with others [38], have proposed that the community define an XML or similar formal, extensible data structure and schema [27]. The structured data would contain and uniquely identify the data elements that must be returned in a manner that assures that there is no ambiguity across elements, and that the data elements are syntactically and semantically correct.

Inventory item R-3:

- Define a standard data structure for WHOIS responses. The data structure would contain and uniquely identify the data elements that must be returned in a manner that assures there is no ambiguity across elements, correct syntax, and correct semantics.

Regarding this inventory item, the group of technical experts noted the following:

> "The use of a structured data model would allow for easier localization of the client software. This would be most welcome by those who do not have English as one of their languages and do not understand what "tech-c" may mean."

5.4 Standardized errors

No standard set of error messages is defined for WHOIS servers, and WHOIS servers may handle errors differently. For example, if a WHOIS client exceeds the query limit set by the WHOIS server, some WHOIS servers will not return the queried domain result (silent discard), others will return an error message that is specific to the server application or service provider, and still others will simply close the connection. The lack of uniform error handling and error messages introduces ambiguity and confusion: users and applications cannot determine unequivocally whether the limit was exceeded or an exception was encountered.

Inventory item R-4:

- Define a set of standardized error messages and standard handling of error conditions. Examples of useful error messages include: queries exceeding the limit; no records found; unable to process query; etc.

5.5 Standardized Set of Query Capabilities

Currently, no standardized set of query capabilities are implemented across all WHOIS services. Some registries are required to offer a search capability that is not limited to using a domain name as a search argument, but also other registration information as well. For example, the .MOBI registry agreement requires that WHOIS data:

> "be searchable by domain name, registrant's name, registrant's postal address, contacts' names, Registrars Contact IDs and Internet Protocol address without arbitrary limit. In order to provide an effective WHOIS database, Boolean search capabilities may be offered." (.MOBI Registry Agreement [9])

Such features are valuable to parties that investigate spam and other criminal activity on a given domain and typically try to determine whether other domains have been registered by the same registrant. Past SSAC reports have also noted this problem and recommended that the WHOIS support searching [23]. In addition, a 2002 report by the WHOIS task force [16] recommended that WHOIS permit users to submit not only domain names as arguments to search functions but other registration data elements as well.

Inventory item R-5:

- Allow users to submit not only domain names as arguments to search functions but other registration data elements as well [16, 23].

Staff notes that while supporting this feature is not particularly challenging technically for thick registries, it does pose a major technical challenge for a thin registry. A key issue is that in its current form and definition, the WHOIS protocol would not be able to support searching the architecture of distributed indices.

On this point, the Registry Stakeholder Group (RySG) noted that drawbacks and challenges should be mentioned as well. These include:

- "The paper is incorrect in saying that such contact searches are "not particularly challenging technically for thick registries." Such searches do pose significant technical issues, and indeed it

might not be possible to deliver such searches under the contractual SLAs that gTLD registries are obligated to deliver to ICANN.

- Contact searches may facilitate malicious activities.
- Current practice in both gTLDs and ccTLDs is to not provide such contact searches. There are probably both social and technical reasons for this current state of the industry, and it may be worth noting in the Initial Report that this should be explored further."
- Searches by registrant name, contact postal address, etc. were what Bulk WHOIS access was designed to provide."

The group of technical experts also noted potential privacy invasions. They said:

"It can be argued that the requested capability -- obtaining a cumulative report on all of one's registered objects -- would constitute a non-proportional invasion of that party's privacy, on one hand, with questionable value for law enforcement etc., on the other. There are TLD registries who "disabled" this capability for this reason."

5.6 Quality of domain registration data

Registration data associates individuals or organizations with the domain names they register. It is important that these data are accurate and that the registration data provide correct contact information for all of the positive uses of WHOIS identified in section 2.2. There are several metrics that measure the quality of the registration data, including *accuracy*, *applicability*, *availability*, and *currency*.

Accuracy: Accuracy is the central metric of the quality of domain registration data. Do the data accurately identify and provide information sufficient to reliably contact the registrant? Various studies have assessed the quality and accuracy of domain name registration information; these studies have shown that the quality of the domain registration data needs to be improved [27, 35, 32].

Many reasons have been offered for the current extent of inaccurate WHOIS information:

1. *Privacy considerations*. People intentionally submit false information. They do not wish to disclose personal contact information that can be accessed publicly and anonymously [23, 32].
2. *Stealth, intentional deception*. Miscreants intentionally provide false information to obfuscate identification by law enforcement or parties that investigate malicious use of domains [3].
3. *Little or no corroboration of submitted data*. Current registration requirement take a minimalist approach to verifying identity. Unless credit verification measures are stringently applied for all

levels of payment, little or no additional proof of identity and verification of contact information are required when a user registers a domain name.

4. *User error*. Users may mistype when registering domains. The current verification processes can overlook errors.
5. *"Scofflaw" effect*. Users may not understand the consequences of the WHOIS data accuracy program and annual obligation to maintain accurate and complete registration data, or refuse to take time to check that their contact information is current, or reject the notion that they will forfeit a domain registration simply because some registration data are inaccurate.

These reasons, if not addressed, will continue to contribute to WHOIS inaccuracy. From a technical perspective, certain measures can be taken to reduce unintentional errors by registrants; for example, a formal data structure and strong typing of data (e.g., this field must be Arabic numbers only, this field must be alphabetical characters only) can reduce certain typographical errors. Enforcing mandatory submission of data for all fields may reduce cases where users omit information.

Applicability is a soft characteristic. It describes whether or not the data collected and displayed are useful or relevant to the user or querying applications. To keep WHOIS data relevant and useful, it is important that a future WHOIS data model accommodate extensibility and changeability. Certain registration data are not as useful today as they were 20 years ago; fax, for example, is an obligatory field but fewer registrants use fax today than ten years ago [24]. Fax number is a possible example of data that may no longer be useful and thus could be deprecated. Acting to deprecate a fax number is an example of changeability.

Next, consider how society has embraced short and instant messaging and texting services, and whether certain of these might be practical ways to contact registrants. AIM handles, Jabber IDs, Twitter handles, etc. might be more timely and useful contact methods for certain registrants. Expanding WHOIS contact information to reflect and accommodate changes in messaging is an example of extensibility. Extensible, structured data facilitates such additions.

Another possible example of a benefit from having structured data is to provide the ability to suppress the display of certain registration data based on the applicability of the "Conflict of Law Policy[5]." A "contact" for example, could have an attribute "subject to Conflict of Law..."; if the attribute value is YES, then the server would not return contact information when queried by a client.

[5] ICANN Procedure For Handling WHOIS Conflicts with Privacy Law" http://www.icann.org/en/processes/icann-procedure-17jan08.htm

Defining extensible, structured data lays the foundation for future, additional information. For example, a would-be registrant currently has no simple way to determine whether a domain name has been associated with some malicious activity or abuse. It may be useful to provide some indication in WHOIS that a domain name was suspended, black listed, or deleted for such reasons. It may also be useful to provide some form of identification of or contact information for the reseller that processes registrations as an agent of an ICANN accredited registrar.

Finally, **Currency** is a database attribute. Are the collected registration data current? Are these data maintained in an appropriate cycle to provide the most accurate picture of the registrant possible? Sometimes the WHOIS data can be quite outdated. From a technical perspective, one way to inform the currency of the WHOIS data is to add a time stamp in the WHOIS data that shows when the field was last verified or updated.

Inventory item R-6a:

- Adopt a structured data model for WHOIS data that provides extensibility and changeability properties. Employ a formal data schema language such as XML to describe the characteristics of the structured data.

Inventory item R-6b:

- Consider extending the currently defined set of registration data elements to include: alternative forms of contact than the contacts currently collected; information that discloses the history or "pedigree" of a domain; and additional registration service provider contact information.

Inventory item R-6c:

- Add a time stamp in the WHOIS data that shows when the field was last verified or updated.

On inventory R-6a, ALAC noted that

> "The introduction of a structured data format would also be an excellent opportunity to require the use of internationally agreed standards on the display of postal addresses and phone numbers. The use of a machine-parsable output would certainly be beneficial for legitimate uses of the WHOIS information, allowing to automate processes. On the other hand, it will also make the life of those with malicious intents much easier, too. There should be mechanisms put in place to prevent large scale harvesting of data for malicious use."

5.7 Internationalization

The current WHOIS protocol has not been internationalized. It has no mechanism for indicating the character set in use. Originally, the predominant text encoding in use was US-ASCII. In practice, some WHOIS servers, particularly those outside the USA, might use another character set either for requests, replies, or both. This inability to predict or express text encoding has adversely impacted the interoperability (and, therefore, usefulness) of the WHOIS protocol [2].

SSAC investigated this issue in a recent report [28]. The SSAC report examines how the use of characters from local scripts affects the Internet user experience with respect to domain name registration data submission, usage and display. The report presents examples of what users may encounter today when they access registration data via WHOIS or via the web. The report examines the issues related to supporting characters from local scripts in the context of current and future applications that various parties (e.g., registrars, registries, third parties) provide for the submission, usage and display of domain names and registration data.

Currently, IDN guidelines are sufficient for recording and displaying domain names; however, no standards or conventions currently exist that would make WHOIS service more accessible to users whose local languages cannot be represented in USASCII7.

At SSAC's recommendation, ICANN's Board tasked GNSO and SSAC to form an Internationalized Registration Data Working Group (IRD WG) to study the feasibility and suitability of introducing display specifications or standards to deal with the internationalization of Registration Data. ICANN staff is currently supporting an IRD working group [22] composed of members who have considerable expertise with IDNs and issues associated with supporting local languages in web and other Internet applications.

As the IRD WG is deliberating in parallel with staff's preparation of this inventory, staff has elected to defer consideration of internationalized registration data issues while those deliberations continue. Staff will coordinate with the IRD WG to see that its recommendations are included in an updated inventory when those recommendations are made available.

On the issue of internationalization, the ALAC noted that:

> "We understand that the requirement 7, which does not appear in this document, has been submitted to a specialized working group on the internationalization of WHOIS data. On that matter, the At-

Large is of the opinion that the data should be displayed both in native script and in latin characters. Domain names should be displayed both in native script and punycode."

5.8 Security

5.8.1 Authentication

Current WHOIS services offer public access to registration data. The services are largely anonymous, requiring no identity assertion, credentialing or authentication.

SSAC has discussed the utility of an authentication framework for WHOIS in SAC 33 [27], where it concluded "[An authentication] framework allows organizations to employ stronger authentication methods to sensitive data and simpler authentication methods for access to public or less sensitive data. The value of authenticating users, even when they access public data, is to allow an organization to audit user activity." Adopting an authentication framework merely establishes the underlying capabilities to define and implement a range of verification methods and credential services. What methods are needed and how they are applied are matters of policy development.

In addition to the benefits mentioned in the SSAC document, client authentication mechanisms are also useful for setting rate limits for performance. Currently most operators use the client IP address as an authenticator, but this is not useful for clients with temporary addresses, and there are easy ways for such mechanisms to be circumvented.[6] A client authentication mechanism would also provide a means to protect the privacy of the WHOIS data, should policy be developed to distinguish registrations by natural persons and restrict access to certain personally identifiable information of those registrants.

Authentication is considered to be a requirement for enabling authorization services: you must be able to distinguish identity A from identity B before you can grant different permissions to A and B. This same principle holds true for groups of people. For example, should policy be developed to prohibit anonymous access to non-public information included in the registrations of natural persons, some ICANN stakeholders [18] will maintain that certain groups (e.g., law enforcement) should still have access to that information. Future WHOIS services should be able to support any future policy that may be developed to

[6] http://www.dnforum.com/f17/pir-limits-org-whois-thread-110375.html. The circumvention involves using an extra dynamic IP DSL line on another box and writing a timed loop script that resets the router, which will renew the IP address from the ISP. If the query is refused on the primary box, route the query to the secondary box as above. Add more lines/boxes according to load. The loop time will depend on the amount of queries being refused.

control whether individuals or members or a particular group are authorized to access certain registration data and an authentication framework is necessary to enable such support.

Inventory item R-8.1:

- Define an authentication framework for WHOIS service that is able to accommodate anonymous access as well as verification of identities using a range of authentication methods and credential services.

5.8.2 Access Control (authorization)

Various past GNSO activities have discussed differentiated access to the WHOIS data, including the operational point of contact (OPOC) proposal [18], the "special circumstances" proposal [17], and the "financial services" proposal.

Furthermore, the Inter-Registrar Transfers PDP A (IRTP A) discussed the potential need for WHOIS of the future to allow for the exchange of registrant email addresses for purposes of inter-registrar transfers. One of the ways to solve this problem, which was discussed in the report, is that "registrars implement a tiered access approach to providing WHOIS information that would permit the private provision of Registrant e-mail address."

Besides GNSO activities, the SSAC has issued the following recommendations in SAC 003 [23]:

1. WHOIS services must provide mechanisms to protect the privacy of registrants.
2. A WHOIS service must discourage the harvesting and mining of its data. [23].

In addition, SSAC 33 and 40 have recommended that a successor to the WHOIS protocol be considered that can support future consensus policies requiring identification and validation of users, control access (for privacy or other reasons), etc.

Section 4.8.1 SSAC 33 considers a scenario where policy may be developed to control whether individuals or members of a particular group are permitted access to registration data. Similarly, policy may be developed to prohibit certain groups from accessing all or particular elements of registration data. An authorization model that accommodates granular (per object) access controls is needed to support policy flexibility for both cases. For example, policy might be developed that would prohibit anonymous access to all data of a registration held by a natural person (who has provided satisfactory proof of identity to a registrar, in accordance with a "natural persons identification" policy). Alternatively, a policy

might be developed that would allow public access to certain registration data that are determined not to be personally identifying data.

Inventory item R-8.2:

- Implement an authorization framework that is capable of providing granular (per registration data object) permissions (access controls).

On inventory item R-8.1, 8.2, ALAC noted that:

> "The authentication framework, coupled with granular access to data for the WHOIS service should not be an option or a nice to have feature, but is a fundamental prerequisite **to allow for the protection of the privacy of individuals. It should be sufficiently flexible to allow those outside the gTLD community, notably ccTLDs, to implement access policies required by their local laws."

The group of technical experts noted similarly, that

> "The requirements mention several recommendations the SSAC has provided in the past regarding authentication and granular access to information, which has been a major request of the At-Large Advisory Committee (ALAC) over the years. It is also a technical necessity for some registrars and registries that need to comply with local privacy laws. For example, Telnic had several problems implementing a WHOIS service that would comply with the United Kingdom's laws on privacy."

5.8.3 Access Auditing

SAC 033 describes how directory service applications typically provide auditing functions, i.e., methods to monitor, record, and report data object access activities. Auditing of WHOIS access has several possible beneficial applications. For example, collecting statistics that reveal the history or frequency of access to registration records and the locations or identities of users performing the access, and then correlating these against the types of access (view, change, delete) is a variation on methods used by credit card fraud and network intrusion detection systems to detect anomalous or suspicious access activity. In the context of domain registrations, such monitoring might provide registrars or registrants opportunities to block subsequent activity by an attacker and thus prevent a domain hijacking or other type of attack.

Many applications support considerable granularity with respect to auditing, and are able to audit not only access to a record but given elements of a record. Specifically, monitoring access to DNS configuration elements of a registration record(s) might provide registrars with opportunities to detect and block an

attacker's attempt to hijack a domain's name service (a precursor to a web site defacement or hijacking of other Internet services such as mail).

Registrars or registries could conceivably implement and offer WHOIS auditing services today, but no framework or taxonomy of metrics to audit exists to accommodate registrants who register domains under multiple TLDs and through multiple registrars. Based on the initial recommendation from SAC 033, it would be useful to consider a framework and baseline set of metrics that can accommodate future policy development for WHOIS auditing.

Inventory item R-8.3

- Define a framework and baseline set of metrics that can accommodate future policy development for auditing of WHOIS access.

5.9 Thick vs. Thin WHOIS

There have been considerable debates on the merits of thin WHOIS versus thick WHOIS,[7]. In the following section, we summarize some of the technical arguments made in favor of each type.

From a technical perspective, a thick WHOIS model is essentially a centralized database. Historically, centralized databases are operated under a single administrator that sets conventions and standards for submission and display, archival/restoration and security have proven easier to manage. By contrast, a thin WHOIS model is a decentralized database. Registrars set their own conventions and standards for submission and display, archival/restoration and security registrant information. Today, for example, WHOIS data submission and display conventions vary among registrars. The thin model is thus criticized for introducing variability among WHOIS services, which can be problematic for legitimate forms of automation.

Like other centralized databases, a thick WHOIS model offers attractive archival and restoration properties. If a registrar were to go out of business or experience long-term technical failures rendering them unable to provide service, registries maintaining thick WHOIS have all the registrant information at hand and could transfer the registrations to a different (or temporary) registrar so that registrants could continue to manage their domain names.

[7] See GNSO discussions outlined by this thread: http://gnso.icann.org/mailing-lists/archives/registrars/thrd35.html#02038

From a technical perspective, some argue that the thin WHOIS model has its benefits as well. For example, they comment that the extensible provisioning protocol (EPP) was not designed to handle the extensive updates every time a registrar makes changes to the WHOIS record.[8]

After carefully considering the strengths and merits of both models in the context of the new gTLD program, ICANN staff recommended that new gTLD applicants must implement a "thick WHOIS" [12]. The recommendation cites two scenarios in which the additional option of retrieving the data at the registry would be valuable:

- Where the registrar WHOIS service might be experiencing a short- or long-term outage (in violation of the registrar's accreditation agreement), and
- Where the registrar has implemented strong (or sometimes overly-defensive) measures to prevent large-scale automated harvesting of registrar data.

Inventory item R-9:

- Adopt a thick WHOIS for all new gTLDs. Consistent with these recommendations for future WHOIS services, new or legacy registries could consider evolving to a thick WHOIS [12].

Regarding this requirement, RySG commented that "It seems like it would be useful to point out that this would be a significant undertaking for gTLDs like .com with well over 80 million registrations. In that regard, on the technical side, what would be the impact to EPP commands and service level requirements? On the service side, what would be the impact on registrants and registrars? Some discussion of these issues would probably be a good idea."

On this point, ALAC also commented that "The At-Large believes that the thick vs thin WHOIS debate is outside the scope of this document and that its implementation is a policy decision that is not dependent on the underlying protocol. We disagree that "new or legacy registries should consider evolving to a thick WHOIS". Irrespective of the policy decision taken, all gTLD registries should behave the same way. It should not be an option for the registry to consider or not."

On this point, the group of technical experts noted that:

> "In the current WHOIS, there is no difference between WHOIS services at thick and thin registries when searching by domain name. The WHOIS service at a thin registry will respond

[8] See above thread.

with a referral to the WHOIS service with the complete information. It is straightforward to follow the referral for the complete information.

If it is desirable to be able to search by tokens other than the domain name, then it is true these tokens must be present in the WHOIS service at the registry. A thick registry WHOIS service could have all possible tokens available, although it is not clear that searching by all possible tokens is the desired level of service. For example, it may be sufficient to specify the minimum information that must be present in the registry WHOIS service in order to support more general queries, which could still be quite a bit less than the complete WHOIS information."

5.10 Domain WhoWas service

Once a domain name deletion request has been completed, the domain name is removed from the registry WHOIS service. This deletion occurs at the time the deletion transaction is processed for a domain name within the Add Grace Period, or for domain names that are deleted outside of the Add Grace Period, upon expiry of the Pending Delete period. The WHOIS data associated with the domain name is no longer readily available through WHOIS. This means that users in general cannot ascertain from WHOIS information whether domains that have been registered were in the past black or block listed because they were used in support of a malicious or criminal activity (or to prevent criminal activity, as in the case for Conficker). A user could register a domain only to find that it is not usable in the manner the user anticipated.

Recognizing these deficiencies, certain third parties offer fee-based access to domain name "histories". In 2009, VeriSign proposed to launch a similar service, "domain name WhoWas," through ICANN's Registry Service Evaluation (RSEP) Process [36]. VeriSign's WhoWas Service provides an automated capability for a customer (which may be either a registrar or non-registrar) to look up a domain name and receive a response with the registration history for the entire life of that domain name which includes the domain name, registration dates and registrar of record for each period of time. The RSEP request was approved by ICANN in July 2009.[9] While VeriSign is thus far the only registry to consider a domain history service, it is possible that other registries might also offer such a service.

Inventory item 4-10:

- A WhoWas service could be provided by all registries. This is another example of data that could complement existing registration data as we described in section 4.6.

[9] http://www.icann.org/en/registries/rsep/verisign-whowas-16jul09-en.pdf

5.11 Registrar Abuse Point of Contact

In recent years, abuse on the Internet has been on the rise [1]. Internet miscreants use compromised infrastructure to launch various attacks such as phishing, distributed denial of Service (DDoS) attacks, spam, malware, etc. Domain names are an important weapon in the miscreant/criminal arsenal [1]. To deal with this growing threat, law enforcement, corporations and other anticrime and antiphishing groups have a growing need to contact registrars and registries about certain domains. However, publicly available contact information for registries and registrars does not always lead to an individual or group capable or qualified to handle abuse complaints. Finding the right person to contact often requires special institutional knowledge; as a result, investigation processes may be impeded while responders seek out this information.

Recognizing these difficulties, SSAC first studied this issue and recommended that registrars and registries publish abuse point of contact information [29]. The new gTLD malicious conduct report [13] follows up on SSAC's recommendation and requires that "Registry Operator shall provide a single abuse point of contact for all domains within the TLD. This abuse contact would be responsible for addressing and providing timely response to abuse complaints received from recognized parties, such as other registries, registrars, law enforcement organizations and recognized members of the anti-abuse community. Registries must also provide a description of their policies to combat abuse." Furthermore, the "Registry Operator may require of all registrars with whom they contract for services that they provide an abuse point of contact."

Inventory item R-11:

- Registrars and registries could provide and publish abuse point of contact information as an element of a domain registration record [13,29]. There are several ways this could be supported; for example, registrars could populate the current sponsoring registrar contact information with an abuse point of contact rather than a general purpose business contact; alternatively, an abuse identifier that serves as an index into a publicly accessible table of abuse points of contact could be added to a registration record. These are further examples that demonstrate the utility of adopting an extensible data structure and formal schema.

6 Draft Compilation of Potential New Requirements

In this section, we briefly list the possible new requirements, examined in this report, that have been suggested in the past.

- R-1: Provide a publicly accessible and machine parsable list of domain names or IP locations of WHOIS servers operated by ICANN accredited registrars and gTLD registry operators and ccTLDs operators.
- R-2: Define a standard query structure that clients can implement and that all gTLD registries and ICANN accredited registrars will support.
- R-3: Define a standard data structure for WHOIS responses. The data structure would contain and uniquely identify the data elements that must be returned in a manner that assures there is no ambiguity across elements, correct syntax, and correct semantics.
- R-4: Define a set of standardized error messages and standard handling of error conditions. Examples of useful error messages includes queries exceeding the limit, no records found, unable to process query, etc.
- R-5: Allow users to submit not only domain names as arguments to search functions but other registration data elements as well.
- R-6a: Adopt a structured data model for WHOIS data that provides extensibility and changeability properties. Employ a formal data schema language such as XML to describe the characteristics of the structured data.
- R-6b: Consider extending the currently defined set of registration data elements to include: alternative forms of contact than the contacts currently collected; information that discloses the history or “pedigree” of a domain; and additional registration service provider contact information.
- R-8.1. Define an authentication framework for WHOIS that is able to accommodate anonymous access as well as verification of identities using a range of authentication methods and credential services.
- R-8.2: Implement an authorization framework that is capable of providing granular (per registration data object) permissions (access controls).
- R-8.3: Define a framework and baseline set of metrics that can accommodate future policy development for auditing of WHOIS access.
- R-9: All new TLDs should operate a thick WHOIS. Consistent with these recommendations for future whois, new or legacy registries should consider evolving to a thick WHOIS.
- R-11: Registrars and registries should provide and publish abuse point of contact information as an element of a domain registration record. There are several ways this could be supported; for example, registrars could populate the current sponsoring registrar contact information with an abuse point of contact rather than a general purpose business contact; alternatively, an abuse identifier that serves as an index into a publicly accessible table of abuse points of contact could

be added to a registration record. These are further examples that demonstrate the utility of adopting an extensible data structure and formal schema.

Regarding the compilation, RySG added the following requirements:

- Ensuring consistency of data between registries and registrars (for thin registries).
- Accommodating privacy services in a manner that effectively provides access to information
- Mitigating impacts to SLAs (Service Level Agreements) and EPP (Extensible Provisioning Protocol) commands in migrations from thin to thick WHOIS data.

7 General Comments and Next Steps

In this section we listed some of the general comments we received on the initial report.

ALAC commend that "Most of the issues we face today are due to the lack of features of the protocol. The WHOIS, as defined in RFC3912 is rudimentary. It does not define a format neither for the query nor for the data being returned", and that "The At-Large supports all the requirements expressed in the document, and believes there is a consensus in the community on these."

The RySG "expresses appreciation for what we believe is very constructive report. We believe that it provides an excellent basis for additional definition of WHOIS service requirements for the future."

With regard to the next steps, the Registry stakeholders Group (RySG) made the following comments:

> "As the community moves forward with regard to new WHOIS requirements an important question for inclusion in the Initial Report is which of the proposed requirements in this section involve Internet standards issues that are the responsibility of the Internet Engineering Task Force (IETF). It is possible that some of them have already been dealt with by the IETF. It could also be the case that additional standards work needs to be done regarding some of the requirements. Whatever the case, we recommend that any standards work that may be needed be identified and steps taken to initiate the any needed standards development work as soon as possible so as to avoid possible delays later when additional WHOIS policy work may occur."

The ALAC provided the following comments for next steps:

> "The discussion over the WHOIS has been going on for several years. The At-Large would like to see a clear roadmap and a timeline with milestones for the implementation of the above requirements. Obviously, the At-Large Community and the Committee is willing to work with the GNSO, the staff and other parts of the ICANN community in helping to move the process forward."

The group of technical experts provided the following comments regarding next steps:

- Whatever new solution is chosen / changes are made, we need some sort of backwards compatibility / phased introduction / transition plan.
- It is not clear if the intention is to update the WHOIS protocol to match the new requirements, in which case it should go through the Internet Engineering Task Force (IETF) standards process, or if ICANN intends to develop its own WHOIS protocol-like service. In any case, because the WHOIS protocol is being used outside the generic top level domain (gTLD) in country code TLDs (ccTLDs), regional Internet Registries (RIRs) and some Local Internet Registries (LIRs), we need to avoid having different dialects of WHOIS, which would share a similar name, but different interfaces and output.
- As a next step, we recommend the community discuss what services / protocols would satisfy these requirements and how to move forward to make these changes.

References

1. Anti-Phishing Working Group. (2009) Global Phishing Survey: Trends and Domain Name Use in 2H2008. Retrieved January 10, 2010, from http://www.apwg.org/reports/APWG_GlobalPhishingSurvey2H2008.pdf

2. Daigle, L. (2004) WHOIS Protocol Specification, RFC 3912.

3. Edelman, B. (2002) *Large-Scale Intentional Invalid WHOIS Data: A Case Study of "NicGod Productions" / "Domains For Sale"*. Cambridge, MA: Harvard University. Retrieved October 29, 2009, from http://cyber.law.harvard.edu/archived_content/people/edelman/invalid-whois/

4. Gasster, L. (2007) *Staff Overview of Recent GNSO WHOIS Activities.* Marina Del Rey, CA: Internet Corporation for Assigned Names and Numbers (ICANN). Retrieved October 21, 2009, from http://gnso.icann.org/drafts/icann-staff-overview-of-whois11oct07.pdf

5. Harrenstien, K. and White, V. (1982) "NICNAME/WHOIS", RFC 812.

6. Harrenstien, K., Stahl, M. and E. Feinler. (1985) "NICNAME/WHOIS", RFC 954.

7. Internet Corporation for Assigned Names and Numbers (ICANN). (2003) *WHOIS Data Reminder Policy.* Marina Del Rey, CA: ICANN. Retrieved November 10, 2009, from http://www.icann.org/en/registrars/wdrp.htm.

8. Internet Corporation for Assigned Names and Numbers (ICANN). (2005) *Community Experiences with the InterNIC WHOIS Data Problem Reports System.* Marina Del Rey, CA: ICANN. Retrieved November 10, 2009, from http://www.icann.org/en/whois/wdprs-report-final-31mar05.htm.

9. Internet Corporation for Assigned Names and Numbers (ICANN). (2005) *.MOBI Agreement Appendix S.* Marina Del Rey, CA: ICANN. Retrieved November 10, 2009, from http://www.icann.org/en/tlds/agreements/mobi/mobi-appendixS-23nov05.htm.

10. Internet Corporation for Assigned Names and Numbers (ICANN). (2006) *.COM Agreement Appendix 5:WHOIS Specifications.* Marina Del Rey, CA: ICANN. Retrieved November 10, 2009, from http://www.icann.org/en/tlds/agreements/verisign/appendix-05-01mar06.htm

11. Internet Corporation for Assigned Names and Numbers (ICANN). (2009a) *New gTLD draft Applicant Guidebook v3.* Marina Del Rey, CA: ICANN. Retrieved October 21, 2009, from http://www.icann.org/en/topics/new-gtlds/draft-rfp-clean-04oct09-en.pdf

12. Internet Corporation for Assigned Names and Numbers (ICANN). (2009b) *New gTLD Program Explanatory Memorandum. Thick vs. Thin WHOIS for New gTLDs.* Marina Del Rey, CA: ICANN. Retrieved January 19, 2010, from http://www.icann.org/en/topics/new-gtlds/thick-thin-whois-30may09-en.pdf

13. Internet Corporation for Assigned Names and Numbers (ICANN). (2009c) *New gTLD Program Explanatory Memorandum. Mitigating Malicious Conduct.* Marina Del Rey, CA: ICANN. Retrieved January 19, 2010, from http://www.icann.org/en/topics/new-gtlds/mitigating-malicious-conduct-04oct09-en.pdf

14. Internet Corporation for Assigned Names and Numbers (ICANN). (2009d) *Registrar Accreditation Agreement.* Marina Del Rey, CA: ICANN. Retrieved November 10, 2009, from http://www.icann.org/en/registrars/ra-agreement-21may09-en.htm#2

15. Internet Corporation for Assigned Names and Numbers (ICANN). (2009e) *Registry Agreements.* Marina Del Rey, CA: ICANN. Retrieved November 10, 2009, from http://www.icann.org/en/registries/agreements.htm

16. ICANN Generic Names Supporting Organization (GNSO). (2002) *WHOIS Task Force Interim Report.* Shanghai, China. Retrieved October 21, 2009, from http://gnso.icann.org/issues/whois-privacy/whois-shanghai-oct02.pdf

17. ICANN Generic Names Supporting Organization (GNSO). (2007a) *Final Task Force Report on WHOIS Services.* Marina Del Rey, CA: ICANN. Retrieved October 21, 2009, from http://gnso.icann.org/issues/whois-privacy/whois-services-final-tf-report-12mar07.htm

18. ICANN Generic Names Supporting Organization (GNSO). (2007b) *Final Outcomes Report of the WHOIS Working Group 2007.* Marina Del Rey, CA: ICANN. Retrieved January 15, 2010, from http://gnso.icann.org/drafts/icann-whois-wg-report-final-1-9.pdf

19. ICANN Generic Names Supporting Organization (GNSO). (2009a) *Terms of Reference for WHOIS Misuse Studies.* Marina Del Rey, CA: ICANN. Retrieved October 21, 2009, from http://gnso.icann.org/issues/whois/tor-whois-misuse-studies-25sep09-en.pdf

20. ICANN Generic Names Supporting Organization (GNSO). (2009b) *Terms of Reference for WHOIS Registrant Identification Studies.* Marina Del Rey, CA: ICANN. Retrieved October 25, 2009, from http://gnso.icann.org/issues/whois/tor-whois-registrant-identification-studies-23oct09-en.pdf

21. ICANN Generic Names Supporting Organization (GNSO). (2009c) *Council Resolutions May 2009.* Marina Del Rey, CA: ICANN. Retrieved October 25, 2009, from http://gnso.icann.org/resolutions/#200905

22. ICANN Generic Names Supporting Organization (GNSO). (2009d) *Internationalized Registration Data Working Group Draft Charter.* Marina Del Rey, CA: ICANN. Retrieved February 10, 2010, from http://gnso.icann.org/issues/ird/ird-wg-charter-24sep09.htm

23. ICANN Security and Stability Advisory Committee (SSAC). (2003) *WHOIS Recommendation of the Security and Stability Advisory Committee* (SSAC publication No. 003). Retrieved from http://www.icann.org/en/committees/security/sac003.pdf

24. ICANN Security and Stability Advisory Committee (SSAC). (2006) *Information Gathering Using Domain Name Registration Records* (SSAC publication No. 014). Retrieved from http://www.icann.org/en/committees/security/information-gathering-28Sep2006.pdf

25. ICANN Security and Stability Advisory Committee (SSAC). (2007) *Is the WHOIS Service a Source for email Addresses for Spammers?* (SSAC publication No. 023). Retrieved from http://www.icann.org/en/committees/security/sac023.pdf

26.ICANN Security and Stability Advisory Committee (SSAC). (2008a) *SSAC Comment to GNSO regarding WHOIS studies* (SSAC publication No. 027). Retrieved from http://www.icann.org/en/committees/security/sac027.pdf

27.ICANN Security and Stability Advisory Committee (SSAC). (2008b) *Domain Name Registration Information and Directory Services* (SSAC publication No. 033). Retrieved from http://www.icann.org/en/committees/security/sac033.pdf

28.ICANN Security and Stability Advisory Committee (SSAC). (2009a) *Display and usage of Internationalized Registration Data: Support for characters from local languages or scripts* (SSAC publication No. 037). Retrieved from http://www.icann.org/en/committees/security/sac037.pdf

29.ICANN Security and Stability Advisory Committee (SSAC). (2009b) *Registrar Abuse Contacts* (SSAC publication No. 038). Retrieved from http://www.icann.org/en/committees/security/sac038.pdf

30.ICANN Security and Stability Advisory Committee (SSAC). (2009c) *Measures to Protect Domain Registration Services Against Exploitation or Misuse* (SSAC publication No. 040). Retrieved from http://www.icann.org/en/committees/security/sac040.pdf

31.Marco d'Itri. "Re: questions about whois client." Email to Steve Sheng 21 January. 2010. (Marco is the author of the open source WHOIS client included in most Linux distributions)

32.National Opinion Research Center. (2010). *Draft Report for the Study of the Accuracy of WHOIS Registrant Contact Information*. Retrieved March 18, 2010, from http://www.icann.org/en/compliance/reports/whois-accuracy-study-17jan10-en.pdf

33.Newton, A. (2006) *Replacing the WHOIS Protocol: IRIS and the IETF's CRISP Working Group.* Internet Computing, IEEE Volume: 10 Issue: 4 July-Aug. 2006 Page(s): 79-84

34.Sanz, M. (2004) Using DNS SRV records to locate whois servers, Internet Draft. IETF.

35.U.S. Government Accountability Office (GAO). (2005) *Internet Management: Prevalence of False Contact Information for Registered Domain Names.* (GAO publication No. GAO-06-165). Washington, DC: Author. Retrieved from http://www.gao.gov/products/GAO-06-165

36.VeriSign, Inc. (2009a) *Domain Name Who Was - com/net (*Applications for New Registry Services)*.* Marina Del Rey, CA: ICANN. Retrieved October 21, 2009, from http://www.icann.org/registries/rsep/verisign-whowas-01jul09-en.pdf

37. VeriSign, Inc. (2009b) *Registry-Registrar Two-Factor Authentication Service (*Applications for New Registry Services)*.* Marina Del Rey, CA: ICANN. Retrieved October 21, 2009, from http://www.icann.org/registries/rsep/verisign-auth-request-25jun09.pdf

38.Wesson. R (2001). WHOIS Export and Exchange Format. Internet draft. http://xml.coverpages.org/draft-wesson-whois-export-03.txt

39.WHOIS. (2010, January 27). In *Wikipedia, the free encyclopedia*. Retrieved February 1, 2010, from http://en.wikipedia.org/wiki/WHOIS

Annex 1 – GNSO Request for Inventory of WHOIS Service Requirements

Excerpt from Council Resolutions May 2009. Marina Del Rey, CA: ICANN. Retrieved October 25, 2009, from http://gnso.icann.org/resolutions/#200905

"Whereas there have been discussions for several years on the adequacy of the current set of WHOIS tools to provide the necessary functions to support existing and proposed WHOIS service policy requirements,
and, there have been questions as to the adequacy of these tools for use in an IDN environment,
and, that there have been extensive discussions about the requirements of the WHOIS service with respect to Registry and registrar operations, and, new architectures and tools have been developed and suggested by the technical community,
and, the GNSO accepted the recommendation of the IRT-A Working Group to encourage staff to explore further assessment of whether IRIS would be a viable option for the exchange of registrant email address data between registrars and conduct an analysis of IRIS' costs, time of implementation and appropriateness for IRTP purposes,
Resolved,
The GNSO Council requests that Policy Staff, with the assistance of technical staff and GNSO Council members as required, collect and organize a comprehensive set of requirements for the WHOIS service policy tools.
These requirements should reflect not only the known deficiencies in the current service but should include any possible requirements that may be needed to support various policy initiatives that have been suggested in the past.
The synthesis of requirements should be done in consultation with the SSAC, ALAC, GAC, the ccNSO and the GNSO and a straw-man proposal should be prepared for these consultations. The Staff is asked to come back with an estimate of when this would be possible."

Annex II: ICANN 2009 RAA provisions that are relevant to the provisioning of WHOIS Service

Excerpt from: http://www.icann.org/en/registrars/ra-agreement-21may09-en.htm#3

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

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

3.3.1.1 The name of the Registered Name;
3.3.1.2 The names of the primary nameserver and secondary nameserver(s) for the Registered Name;
3.3.1.3 The identity of Registrar (which may be provided through Registrar's website);
3.3.1.4 The original creation date of the registration;
3.3.1.5 The expiration date of the registration;
3.3.1.6 The name and postal address of the Registered Name Holder;
3.3.1.7 The name, postal address, e-mail address, voice telephone number, and (where available) fax number of the technical contact for the Registered Name; and
3.3.1.8 The name, postal address, e-mail address, voice telephone number, and (where available) fax number of the administrative contact for the Registered Name.

The appendix to this Agreement for a particular TLD may state substitute language for Subsections 3.3.1.1 through 3.3.1.8 as applicable to that TLD; in that event the substitute language shall replace and supersede Subsections 3.3.1.1 through 3.3.1.8 stated above for all purposes under this Agreement but only with respect to that particular TLD.

3.3.2 Upon receiving any updates to the data elements listed in Subsections 3.3.1.2, 3.3.1.3, and 3.3.1.5 through 3.3.1.8 from the Registered Name Holder, Registrar shall promptly update its database used to provide the public access described in Subsection 3.3.1.

3.3.3 Registrar may subcontract its obligation to provide the public access described in Subsection 3.3.1 and the updating described in Subsection 3.3.2, provided that Registrar shall remain fully responsible for the proper provision of the access and updating.

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

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

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

3.3.6.1 Registrar shall make a complete electronic copy of the data available at least one (1) time per week for download by third parties who have entered into a bulk access agreement with Registrar.

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

3.3.6.3 Registrar's access agreement shall require the third party to agree not to use the data to allow, enable, or otherwise support any marketing activities, regardless of the medium used. Such media include but are not limited to e-mail, telephone, facsimile, postal mail, SMS, and wireless alerts.

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

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

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

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

3.6 Data Escrow. During the Term of this Agreement, on a schedule, under the terms, and in the format specified by ICANN, Registrar shall submit an electronic copy of the database described in Subsection 3.4.1 to ICANN or, at Registrar's election and at its expense, to a reputable escrow agent mutually approved by Registrar and ICANN, such approval also not to be unreasonably withheld by either party. The data shall be held under an agreement among Registrar, ICANN, and the escrow agent (if any) providing that (1) the data shall be received and held in escrow, with no use other than verification that the deposited data is complete, consistent, and in proper format, until released to ICANN; (2) the data shall be released from escrow upon expiration without renewal or termination of this Agreement; and (3) ICANN's rights under the escrow agreement shall be assigned with any assignment of this Agreement. The escrow shall provide that in the event the escrow is released under this Subsection, ICANN (or its assignee) shall have a non-exclusive, irrevocable, royalty-free license to exercise (only for transitional purposes) or have exercised all rights necessary to provide Registrar Services

Data of Registered Name Holders- this data is the WHOIS plus the billing data, and the customer data for a privacy/proxy service.

3.4 Retention of Registered Name Holder and Registration Data.

3.4.1 During the Term of this Agreement, Registrar shall maintain its own electronic database, as updated from time to time, containing data for each active Registered Name sponsored by it within each TLD for which it is accredited. The data for each such registration shall include the elements listed in Subsections 3.3.1.1 through 3.3.1.8; the name and (where available) postal address, e-mail address, voice telephone number, and fax number of the billing contact; and any other Registry Data that Registrar has submitted to the Registry Operator or placed in the Registry Database under Subsection 3.2. Also, Registrar shall either (1) include in the database the name and postal address, e-mail address, and voice telephone number provided by the customer of any privacy service or licensee of any proxy registration service offered or made available by Registrar or its affiliate companies in connection with each registration or (2) display a conspicuous notice to such customers at the time an election is made to utilize such privacy or proxy service that their data is not being escrowed.

3.4.2 During the Term of this Agreement and for three (3) years thereafter, Registrar (itself or by its agent(s)) shall maintain the following records relating to its dealings with the Registry Operator(s) and Registered Name Holders:

3.4.2.1 In electronic form, the submission date and time, and the content, of all registration data (including updates) submitted in electronic form to the Registry Operator(s);

3.4.2.2 In electronic, paper, or microfilm form, all written communications constituting registration applications, confirmations, modifications, or terminations and related correspondence with Registered Name Holders, including registration contracts; and

3.4.2.3 In electronic form, records of the accounts of all Registered Name Holders with Registrar, including dates and amounts of all payments and refunds.

3.4.3 During the Term of this Agreement and for three (3) years thereafter, Registrar shall make these records available for inspection and copying by ICANN upon reasonable notice. ICANN shall not disclose the content of such records except as expressly permitted by an ICANN specification or policy.

3.4.4 Notwithstanding any other requirement in this Agreement, Registrar shall not be obligated to maintain records relating to a domain registration beginning on the date three (3) years following the domain registration's deletion or transfer away to a different registrar.

3.7.7 Registrar shall require all Registered Name Holders to enter into an electronic or paper registration

agreement with Registrar including at least the following provisions (except for domains registered by the Registrar for the purpose of conducting its Registrar Services where the Registrar is also the Registered Name Holder, in which case the Registrar shall submit to the following provisions and shall be responsible to ICANN for compliance with all obligations of the Registered Name Holder as set forth in this Agreement and ICANN policies established according to this Agreement):

Annex III: Comments from Registry Stakeholder's Group

GNSO gTLD Registries Stakeholder Group Statement

Issue: Inventory of WHOIS Service Requirements Initial Report

Date: May 17, 2010

Issue Document URL: http://gnso.icann.org/issues/ (See WHOIS)

Regarding the issue noted above, the following statement represents the views of the ICANN GNSO gTLD Registries Stakeholder Group (RySG) as indicated. The RySG statement was arrived at through a combination of RySG email list discussion and RySG meetings (including teleconference meetings).

The RySG expresses appreciation for what we believe is very constructive report. We believe that it provides an excellent basis for additional definition of WHOIS service requirements for the future.

We submit the following comments with the intent of identifying possible areas of improvement in future versions of the report.

RySG Comments

The Initial Report contains "Inventory Items" (the first is R1 on page 16). It would help to further clarify the purpose of these for readers, and to standardize their wording in a consistent fashion. Some are stated as suggestions or options, while others are phrased as policy directives or conclusions (such as Items R-8.2 and 4-10). At this time none are "requirements" – rather they are possible or proposed requirements, and the actual requirements will be determined later. Inventory Items phrased as policy directives and conclusions seem confusing since "the report is a technical inventory and does not intend to define or suggest the policies or operational rules that should apply."

The Initial Report says "In this document, unless otherwise specified, when we refer to WHOIS, we mean the WHOIS services that include the WHOIS clients of all kinds and WHOIS servers that implement the protocol as well as the databases that store the domain registration data" (section 3.1). We note that the veracity of "truthfulness" of data is often completely separate from and non-dependent upon the WHOIS service and WHOIS protocol.

Bulk WHOIS is a "WHOIS service" as per the RAA and registry contracts, and should be mentioned in the Initial Report (such as in "4.1: WHOIS Service at the Registry Level" and "4.2: WHOIS Service at the Registrar Level"). The Initial Report should not assume that port 43 WHOIS should be the only means to provide WHOIS information, or that port 43 (or a successor protocol) is the one or best tool for solving all business or technical needs.

Content of registration data provided via WHOIS may differ across TLDs. Some gTLD registry agreements, such as .tel, have provisions in place that in certain circumstances exclude personal information from the public WHOIS. For example, .tel WHOIS output for individuals may only mention registrant's name with no other contact information. We recommend that these sorts of special provisions be mentioned in the report.

3. Background and Terminology

Bulk WHOIS is a "WHOIS service" as per the RAA and registry contracts, and should be mentioned in the Initial Report (including in sections 3: "Background and Terminology," "4.1: WHOIS Service at the Registry Level," and "4.2: WHOIS Service at the Registrar Level"). The Initial Report should not assume that port 43 WHOIS should be the only means to provide WHOIS information, or that port 43 (or a successor protocol) is the one or best tool for solving all business or technical needs.

3.2 Usage of Whois

The third bullet on page 8 says, ". . . WHOIS queries provide information that is often useful in resolving a registration ownership issue". We suggest deleting the word "ownership" because it has implications that do not apply to domain name registrations.

Section 3.2 says "Miscreants allegedly use WHOIS….." This should be changed to "Miscreants use WHOIS….." These uses are not theoretical and have been well-substantiated by SSAC 023, security firms, law enforcement, and registries and registrars.

4.1 Whois Service at the Registry Level

The third sentence of the first paragraph on page 12 says, "A thin registry only includes data sufficient to identify the sponsoring registrar, status of the registration, and creation and expiration dates for each registration in its WHOIS data store." Note that thin registries also include "name server data, and may provide other fields such as "Last Updated." This should be corrected in this sentence as well as in the next-to-last sentence in the same paragraph.

5, New Requirements

As the community moves forward with regard to new WHOIS requirements an important question for inclusion in the Initial Report is which of the proposed requirements in this section involve Internet standards issues that are the responsibility of the Internet Engineering Task Force (IETF). It is possible that some of them have already been

dealt with by the IETF. It could also be the case that additional standards work needs to be done regarding some of the requirements. Whatever the case, we recommend that any standards work that may be needed be identified and steps taken to initiate the any needed standards development work as soon as possible so as to avoid possible delays later when additional WHOIS policy work may occur.

5.1 Mechanism to find Authoritative Whois Servers

The first sentence of this section on page 15 says, "Currently each registry, and for thin registries each registrar also, maintains its own WHOIS server." Should this be changed to "Currently each registry, and for thin registries, each registrar also maintains its own authoritative WHOIS server?"

The second sentence says, "There is no easy way to find out the domain names and IP addresses of the WHOIS servers for a given TLD or registrar, although whois.nic.TLD is a common host naming convention for WHOIS servers." It should be noted that the IANA maintains a root zone database that includes the WHOIS server location data for each TLD. As an example, see the following URL for the .com delegation record: http://www.iana.org/domains/root/db/com.html.

We are not sure of the significance of the IP addresses of WHOIS servers. It is generally inadvisable for users to go to an IP rather than a server URI because service providers occasionally need to change their IPs.

5.5 Standardized Set of Query Capabilities

Searches by registrant name, contact postal address, etc. were what Bulk WHOIS access was designed to provide. This should be noted.

The Initial Report mentions benefits/beneficiaries of contact searches. Drawbacks and challenges should be mentioned as well. These include:

- The paper is incorrect in saying that such contact searches are "not particularly challenging technically for thick registries." Such searches do pose significant technical issues, and indeed it might not be possible to deliver such searches under the contractual SLAs that gTLD registries are obligated to deliver to ICANN.
- Contact searches may facilitate malicious activities.
- Current practice in both gTLDs and ccTLDs is to not provide such contact searches. There are probably both social and technical reasons for this current state of the industry, and it may be worth noting in the Initial Report that this should be explored further.

We note that the papers cited in 5.5 are eight years old – ancient in ICANN and Internet terms. The question is whether the observations and recommendations in them are still relevant. We assume that such questions will be explored in the processes to come.

5.6 Quality of domain registration data

The third full paragraph on page 21 says, "Finally, **Currency** is a database attribute. Are the collected registration data current? Are these data maintained in an appropriate cycle to provide the most accurate picture of the registrant possible? Sometimes the WHOIS data can be quite outdated. From a technical perspective, one way to inform the

currency of the WHOIS data is to add a time stamp in the WHOIS data that shows when the field was last verified or updated." We note that an 'Inventory Item' was not included for this. Was that intentional? If so, why?

We note that the currency and the accuracy of WHOIS data are not necessary equivalent.

5.8.1 Authentication

A current example of a "WHOIS-WHOIS" authentication service is the .name premium name Whois service.

5.8.2 Access Control (authorization)

The first paragraph on page 24 refers to the "financial services" proposal but there is no reference information. What is this?

5.9 Thick vs. thin Whois

At the bottom of page 25, the Initial report says, "Registrars set their own conventions and standards for submission and display, archival/restoration and security (of) registrant information….. Today, for example, WHOIS data submission and display conventions vary among registrars." It should be noted that the RAA is clear about registrar escrow and what WHOIS fields that registrars must display.

That section also brings up matters separate from WHOIS display – such as how registrars operate their Web sites ("standards for submission"), disaster recovery ("archival/restoration"), and how registrars execute their IT security ("security of registrant information"). These seem out of scope.

The first full paragraph on page 26 says the following: "Like other centralized databases, a thick WHOIS model offers attractive archival and restoration properties. If a registrar were to go out of business or experience long-term technical failures rendering them unable to provide service, registries maintaining thick WHOIS have all the registrant information at hand and could transfer the registrations to a different (or temporary) registrar so that registrants could continue to manage their domain names." It should be noted that data escrow is the primary means for dealing with such a situation. The need to fall back on registry data occurs if the registrar has failed to escrow its data, which is a contractual compliance problem.

Inventory item R-9 on page 26 says, "All new TLDs should operate a thick WHOIS. Consistent with these recommendations for future WHOIS services, new or legacy registries should consider evolving to a thick WHOIS [12]." It seems like it would be useful to point out that this would be a significant undertaking for gTLDs like .com with well over 80 million registrations. In that regard, on the technical side, what would be the impact to EPP commands and service level requirements? On the service side, what would be the impact on registrants and registrars? Some discussion of these issues would probably be a good idea.

5.11 Registrar Abuse Point of Contact

The Issues Report says: "The new gTLD malicious conduct report follows up on SSAC's recommendation and requires that a Registry Operator shall provide a single abuse point of contact for all domains with the TLD [etc.]." That malicious conduct report contains a variety of assertions, but it is ancillary/extra-contractual and "requires" nothing. Staff needs to re-examine the second paragraph of 5.11.

6. Draft Compilation of Requirements

Here are some possible additional requirements that could be considered:

- Ensuring consistency of data between registries and registrars (for thin registries)
- Accommodating privacy services in a manner that effectively provides access to information
- Mitigating impacts to SLAs and EPP commands in migrations from thin to thick WHOIS data.

RySG Level of Support

1. **Level of Support of Active Members**: Supermajority
   1.1. # of Members in Favor: 10
   1.2. # of Members Opposed: 0
   1.3. # of Members that Abstained: 0
   1.4. # of Members that did not vote: 3
2. **Minority Position**(s): N/A

General RySG Information

- Total # of eligible RySG Members[10]: 14
- Total # of RySG Members: 13
- Total # of Active RySG Members[11]: 13
- Minimum requirement for supermajority of Active Members: 9
- Minimum requirement for majority of Active Members: 7
- # of Members that participated in this process: 13
- Names of Members that participated in this process: 13
  1. Afilias (.info & .mobi)
  2. DotAsia Organisation (.asia)
  3. DotCooperation (.coop)
  4. Employ Media (.jobs)
  5. Fundació puntCAT (.cat)

[10] All top-level domain sponsors or registry operators that have agreements with ICANN to provide Registry Services in support of one or more gTLDs are eligible for membership upon the "effective date" set forth in the operator's or sponsor's agreement (RySG Articles of Operation, Article III, Membership, ¶ 1). The RySG Articles of Operation can be found at <http://gnso.icann.org/files/gnso/en/improvements/registries-sg-proposed-charter-30jul09-en.pdf>. The Universal Postal Union recently concluded the .POST agreement with ICANN, but as of this writing the UPU has not applied for RySG membership.

[11] Per the RySG Articles of Operation, Article III, Membership, ¶ 6: Members shall be classified as "Active" or "Inactive". A member shall be classified as "Active" unless it is classified as "Inactive" pursuant to the provisions of this paragraph. Members become Inactive by failing to participate in a RySG meeting or voting process for a total of three consecutive meetings or voting processes or both. An Inactive member shall have all rights and duties of membership other than being counted as present or absent in the determination of a quorum. An Inactive member may resume Active status at any time by participating in a RySG meeting or by voting.

6. Museum Domain Management Association – MuseDoma (.museum)
7. NeuStar (.biz)
8. Public Interest Registry - PIR (.org)
9. RegistryPro (.pro)
10. Societe Internationale de Telecommunication Aeronautiques – SITA (.aero)
11. Telnic (.tel)
12. Tralliance Registry Management Company (TRMC) (.travel)
13. VeriSign (.com, .name, & .net)

- Names & email addresses for points of contact
  - Chair: David Maher, dmaher@pir.org
  - Vice Chair: Jeff Neuman, Jeff.Neuman@Neustar.us
  - Secretariat: Cherie Stubbs, Cherstubbs@aol.com
  - RySG representative for this statement: Chuck Gomes, cgomes@verisign.com

Annex IV: Comments from ALAC

Introduction

By the Staff of ICANN

1.1.1.1 The attached statement on Initial WHOIS Service Requirements Report was originally drafted by Patrick Vande Walle, member of the At-Large Advisory Committee (ALAC) and sent to the members of the At-Large Whois working group for review on April 5th 2010.

The first revision of the statement (the attached document) was published by Patrick Vande Walle on April 23rd and discussed during the monthly conference call of the ALAC on April 27th. Please click here for a comparison of the two documents.

On April 29th, the Chair of the ALAC asked the Staff to start a five-day online vote on the ALAC Statement on Initial WHOIS Service Requirements Report.

The online vote resulted in the ALAC endorsing the statement with 14-0. You may review the result independently under: https://www.bigpulse.com/pollresults?code=2AzcTXhB8MGuGtJCA9Ru

On May 10th 2010, the statement was officially transmitted to Liz Gasster, the Staff person assisting the GNSO on Whois-related work.

[End of Introduction]

1.2 At-Large comments on the Initial WHOIS Service Requirements Report

The At-Large community thanks the GSNO and the ICANN staff for this opportunity to comment on the Initial WHOIS Service Requirements Report.

As noted in the report under 3.1, Components of the WHOIS service, the name "WHOIS" refers to multiple concepts and it is important to distinguish between them. The At-Large suggests it might be necessary to come up with another name to refer to the "WHOIS service", to avoid confusion with the WHOIS protocol. This is especially true if the service itself might be running over other protocols in the future.

1.2.1 Technical discussion

We define the WHOIS service as an interaction between the client and the server, running on TCP port 43, and implementing the protocol defined in RFC3912. We disagree that web-based interfaces that query a database can be considered "WHOIS clients". They do not suffer from the same limitations as the text-based clients, and can easily handle authentication, internationalization and anti-abuse features.

Most of the issues we face today are due to the lack of features of the protocol. The WHOIS, as defined in RFC3912 is rudimentary. It does not define a format neither for the query nor for the data being returned.

We note also that the WHOIS protocol and associated servers and clients are being used outside the gTLD space. ccTLDs use them in a way similar to gTLDs, but often need to implement variations on the server side to comply with local laws on privacy.

Regional Internet Registries have WHOIS services as an essential part of their work with regard to the allocation of IP addresses, autonomous system numbers, as well as in-addr.arpa and ipv6.arpa PTR delegations. This is why we suggest that the ASO should be consulted in the framework of this process. The last sentence of the executive summary does not indicate the ASO as one of the parties to be consulted, and neither did the original GNSO resolution.

Given that WHOIS clients are included in most operating systems today, and are being used outside of the gTLD space, it is of upmost importance that, whatever new requirements are implemented do not break the existing installed base. We need to avoid having different dialects of WHOIS, which would share a similar name, but different interfaces and output.

We note that the requirements mention several recommendations the SSAC has done in the past regarding authentication and granular access to information. The At-Large obviously supports these, as it has done multiple times over past years.

1.2.2 Requirements discussion

The At-Large supports all the requirements expressed in the document, and believes there is a consensus in the community on these. We add the following additional comments:

- R-4: Standardized error messages will make the localization of the client software much easier. This would be most welcome by those who do not have English as one of their languages and do not understand what "tech-c" may mean.
- R-6a: The introduction of a structured data format would also be an excellent opportunity to require the use of internationally agreed standards on the display of postal addresses and phone numbers. The use of a machine-parseable output would certainly be beneficial for legitimate uses of the WHOIS information, allowing to automate processes. On the other hand, it will also make the life of those with malicious intents much easier, too. There should be mechanisms put in place to prevent large scale harvesting of data for malicious use.
- R-8.1 and 8.2: The authentication framework, coupled with granular access to data for the WHOIS service should not be an option or a nice to have feature, but is a fundamental prerequisite **to allow for the protection of the privacy of individuals. It should be sufficiently flexible to allow those outside the gTLD community, notably ccTLDs, to implement access policies required by their locals laws.
- R-9: The At-Large believes that the thick vs thin WHOIS debate is outside the scope of this document and that its implementation is a policy decision that is not dependent on the underlying protocol. We disagree that "new or legacy registries should consider evolving to a thick WHOIS". Irrespective of the policy decision taken, all gTLD registries should behave the same way. It should not be an option for the registry to consider or not.

We understand that the requirement 7, which does not appear in this document, has been submitted to a specialized working group on the internationalization of WHOIS data. On that matter, the At-Large is of the opinion that the data should be displayed both in native script and in latin characters. Domain names should be displayed both in native script and punycode.

1.2.3 Next steps

The discussion over the WHOIS has been going on for several years. The At-Large would like to see a clear roadmap and a timeline with milestones for the implementation of the above requirements.

Obviously, the At-Large Community and the Committee is willing to work with the GNSO, the staff and other parts of the ICANN community in helping to move the process forward.