

## **ICANN Generic Names Supporting Organisation**

### **WHOIS Study Hypotheses Group Report to the GNSO Council**

**Prepared by the WHOIS Study Hypothesis Group**

**26 August 2008**

# WHOIS Study Hypotheses Group Report to the GNSO Council

1. OVERVIEW AND BACKGROUND .....	3
2. WHOIS STUDY HYPOTHESES TABLE .....	5

## 1. Overview and background

On 26 June 2008, the GNSO Council voted to convene a Whois study hypothesis group. The group was charged with reviewing the study recommendations offered through the previous public comment period and the studies requested by the GAC in its letter of 16 April 2008, and, based on those recommendations and that request, prepare a concise list of hypotheses. The group was asked to deliver a report containing the above with any supporting rationale to the Council.

The Council will then decide whether any potential studies should be further considered, and if so, determine cost, feasibility, potential methodology, and estimated time frames for testing. The text of the resolution can be found at: <http://gns0.icann.org/resolutions/#200806>.

Following is the result of the work of that Whois hypotheses study group. Volunteers for the Whois study hypotheses group are listed in the table below; key participants are marked with an asterisk (\*).

Chuck Gomes*	GNSO Council vice chair and chair of group
Jordi Iparraquirre	gTLD Registry C
Ken Stubbs	gTLD Registry C
David Maher*	gTLD Registry C
Adam Palmer	PIR
Steve Metalitz*	IPC
Lee Eulgen*	IPC
Steve DelBianco*	CBUC
Tony Harris*	ISPC
Tim Ruiz*	Registrar C
Paul Stahura	Registrar C
James Bladel*	Registrar C
Stéphane Van Gelder	Registrar C
Eric Brunner-Williams*	Registrar C
Olga Cavalli*	NomCom appointee to GNSO Council
Avri Doria	GNSO chair
Bertrand de la Chapelle	GAC
Danny Younger	
Beau Brendler	

Wendy Seltzer	ALAC Liaison on the ICANN Board
Alan Greenberg*	ALAC Liaison on the GNSO Council
Liz Gasster*	staff
Patrick Jones	staff
Glen de Saint G�ry*	GNSO Secretariat

## 2. Whois Study Hypotheses Table

### Hypotheses for Whois Studies as developed by the Whois Study Hypotheses WG

Notes regarding the hypotheses:

Note (1): throughout this document the term "registrant" or "registrant data" refers to what is sometimes called the "beneficial user" or customer of a proxy/privacy service. In that regard, note the following from the Registrar Accreditation Agreement (RAA) 3.7.7.3: "Any Registered Name Holder that intends to license use of a domain name to a third party is nonetheless the Registered Name Holder of record and is responsible for providing its own full contact information and for providing and updating accurate technical and administrative contact information adequate to facilitate timely resolution of any problems that arise in connection with the Registered Name. A Registered Name Holder licensing use of a Registered Name according to this provision shall accept liability for harm caused by wrongful use of the Registered Name, unless it promptly discloses the identity of the licensee to a party providing the Registered Name Holder reasonable evidence of actionable harm." See <http://www.icann.org/registrars/ra-agreement-17may01.htm>.

Note (2): The Hypotheses Table below is intended to categorize, consolidate, and add relevant detail to the hypotheses originally submitted. In some cases, as with the GAC recommendations, the hypotheses needed to be inferred from the information submitted. As Council considers which of these studies should be pursued, it will be helpful to refer to the original study submissions (posted at <http://forum.icann.org/lists/whois-comments-2008/> ). These original submissions include statements of how study results could lead to an improvement in WHOIS policy. Many submitters also described the type of survey/study needed, including data elements, data sources, population to be surveyed, and sample size. These original submissions should be used by council and its consultants in designing studies and deciding which are worthwhile to pursue. The GAC suggestions can be found at: <http://www.icann.org/correspondence/karlins-to-thrush-16apr08.pdf>.

Note (3): Further work regarding some of the proposed studies should include consultation with ICANN contract compliance staff to minimize overlap or duplication with their work.

Note (4): The GAC has suggested that we collect two data sets, as follows:

- the amount and source of traffic accessing WHOIS servers and the types and numbers of different groups of users and what those users are using WHOIS data for; and
- the types and extent of misuses of WHOIS data and what harm is caused by each type of misuse, including economic, use of WHOIS data in SPAM generation, abuse of personal data, loss of reputation or identity theft, security costs and loss of data.

Note (5): In cases where the original hypothesis offered by a submitter was modified by the group, effort was made to contact the submitter for input and feedback, and to incorporate their views where possible. The Whois study group also provided the GAC with draft hypotheses for the proposals they recommended, but given the abbreviated time frame and the fact that the GAC does not meet on an intercessional basis, no substantive response was received by the deadline for this report.

Study	Hypotheses
Area 1	<p><b>WHOIS misuse studies</b></p> <p>The hypotheses in Area 1 generally regard "public access to Whois", but there are distinct aspects of public access that should be measured separately in any studies designed:</p> <ol style="list-style-type: none"> <li>1) some registrars prevent automated email harvesting by allowing public web-based access to Whois registrant data only after the user deciphers a "captcha" image.</li> <li>2) registrants who use proxy registration or other privacy services should be measured separately from those registrants whose actual information is open for public access.</li> </ol>
1	<p>Public access to WHOIS data is responsible for a material number of cases of misuse that have caused harm to natural persons whose registrations do not have a commercial purpose.</p> <p><a href="http://forum.icann.org/lists/whois-comments-2008/msg00001.html">http://forum.icann.org/lists/whois-comments-2008/msg00001.html</a></p> <p>Note: In any analysis of misuse, it is critical to determine whether the data was, or could easily have been obtained from a source other than Whois.</p>

14	<p>The Whois database is used only to a minor extent to generate spam and other such illegal or undesirable activities.  <a href="http://forum.icann.org/lists/whois-comments-2008/msg00017.html">http://forum.icann.org/lists/whois-comments-2008/msg00017.html</a></p> <p>Note: The methods employed by previous studies of Whois and the results of those studies should be considered when designing subsequent studies in this area. For example, ICANN's Security and Stability Advisory Committee (SSAC) has already studied email spam arising from Whois data, including an analysis of data protection measures used by ICANN-accredited registrars. See "Is the WHOIS service a source for email addresses for spammers?" at <a href="http://gns0.icann.org/correspondence/ssac-whois-study-27oct07.pdf">http://gns0.icann.org/correspondence/ssac-whois-study-27oct07.pdf</a> . Other studies of Whois misuse can be found at: <a href="http://gns0.icann.org/drafts/whois-available-data-points-04oct07.pdf">http://gns0.icann.org/drafts/whois-available-data-points-04oct07.pdf</a>.</p>
15	<p>Those using Whois data to facilitate illegal or undesirable activities (such as spam) depend on port 43 access to Whois to obtain Whois data. <a href="http://forum.icann.org/lists/whois-comments-2008/msg00018.html">http://forum.icann.org/lists/whois-comments-2008/msg00018.html</a></p>
21 & GAC data set 2	<p>There are significant abuses caused by public display of Whois. Significant abuses would include use of WHOIS data in spam generation, abuse of personal data, loss of reputation or identity theft, security costs and loss of data (note – definition is from GAC recommendation 2). <a href="http://forum.icann.org/lists/whois-comments-2008/msg00026.html">http://forum.icann.org/lists/whois-comments-2008/msg00026.html</a></p> <p>Note: As an example of such abuses, the original submitter noted that public Whois databases are being used and mined regularly by direct mail and related companies for their commercial benefit to compile personal data which they then use, combine, sell and distribute as part of massive lists and databases.</p>
GAC 3	<p>There are technical measures available that would effectively curtail misuse of data published on WHOIS databases while preserving legitimate use and open access to the databases.</p>
Area 2	<p><b>Compliance with data protection laws and the Registrar Accreditation Agreement</b></p> <p>NOTE: GAC #s 12, 13 and 14 are all interdependent and their hypotheses are also interdependent.</p>

16	<p>Two hypotheses:</p> <ol style="list-style-type: none"> <li>1. Registrars do not have a uniform method of disclosing or obtaining consent for collection of data for WHOIS purposes.</li> <li>2. The methods employed by registrars to disclose and obtain consent have not been adjudicated with regard to their consistency with national law.</li> </ol> <p><a href="http://forum.icann.org/lists/whois-comments-2008/msg00019.html">http://forum.icann.org/lists/whois-comments-2008/msg00019.html</a></p> <p>Note: Because there may be significant variations in consent in different jurisdictions, the analysis should be segmented by common legal consent regimes.</p>
22	<p>(a) More restrictive Whois policies than the general ICANN Whois requirements have been adopted by some of the 30 top ccTLDs.</p> <p>(b) ccTLD operators report that Whois policies have been adopted in order to become compliant with the data protection laws of the territory.</p> <p>(c) ccTLDs are moving towards more restrictive WHOIS policies motivated by national data protection laws.</p> <p><a href="http://forum.icann.org/lists/whois-comments-2008/msg00024.html">http://forum.icann.org/lists/whois-comments-2008/msg00024.html</a></p>
23	<p>Some national data protection laws explicitly apply, or have been adjudicated to apply, to information submitted by gTLD registrants and made available via Whois. <a href="http://forum.icann.org/lists/whois-comments-2008/msg00025.html">http://forum.icann.org/lists/whois-comments-2008/msg00025.html</a></p>
GAC 12, GAC 13, GAC 14 & GAC 15	<p>GAC 12 - As reported by gTLD registries or registrars, as reflected in their contractual documents, or as adjudicated in relevant fora, the WHOIS contractual obligations of gTLD registries and registrars are governed by:</p> <ul style="list-style-type: none"> <li>• the laws of their local jurisdiction, or</li> <li>• the laws of the jurisdictions of their Registrants, or</li> <li>• the laws of ICANN (California, U.S.), or</li> <li>• some other jurisdiction.</li> </ul> <p>GAC 13 - Those gTLD registries or registrars that are governed by a local jurisdiction provide a contractual mechanism (or have had a mechanism imposed upon them by law or binding decision) to resolve any conflicts between the law applicable to their WHOIS requirements and the law of any other jurisdiction.</p> <p>GAC 14 - Incorporated into GAC 12.</p> <p>GAC 15 - Out of scope for proposed studies of “key factual issues”</p>



24	Some Registrars are not obtaining agreement to terms required under section 3.7.7 of the RAA. <a href="http://forum.icann.org/lists/whois-comments-2008/msg00013.html">http://forum.icann.org/lists/whois-comments-2008/msg00013.html</a>
<b>Area 3</b>	<b>Availability of privacy services</b>
2	The cost of proxy services precludes some registrants from using them. <a href="http://forum.icann.org/lists/whois-comments-2008/msg00002.html">http://forum.icann.org/lists/whois-comments-2008/msg00002.html</a>
5	Whois at present allows resellers and registrars to offer privacy services to differentiate themselves on value. <a href="http://forum.icann.org/lists/whois-comments-2008/msg00005.html">http://forum.icann.org/lists/whois-comments-2008/msg00005.html</a>
GAC 7	A growing share of registrants is protecting the privacy of their Whois data by using proxy registrations and/or privacy services.
GAC 8	A growing share of registrars and affiliates are offering proxy registration and/or privacy services.
<b>Area 4</b>	<b>Demand and motivation for use of privacy services</b>
17	The majority of domain names registered by proxy/privacy services are used for abusive and/or illegal purposes. <a href="http://forum.icann.org/lists/whois-comments-2008/msg00020.html">http://forum.icann.org/lists/whois-comments-2008/msg00020.html</a>
18, 19, GAC 9 & GAC 10	<p>18 - The majority of domain names registered by proxy/privacy services are used for commercial purposes and not for use by natural persons. <a href="http://forum.icann.org/lists/whois-comments-2008/msg00021.html">http://forum.icann.org/lists/whois-comments-2008/msg00021.html</a></p> <p>19 - A disproportionate share of requests to reveal the identity of registrants who use proxy services is directed toward registrations made by natural persons. <a href="http://forum.icann.org/lists/whois-comments-2008/msg00022.html">http://forum.icann.org/lists/whois-comments-2008/msg00022.html</a></p> <p>GAC 9 - A growing and significant share of proxy/privacy service users are legal persons.</p> <p>GAC 10 - A growing and significant share of domains that are registered using proxy/privacy services are used for commercial purposes.</p>
<b>Area 5</b>	<b>Impact of WHOIS data protection on crime and abuse</b>
6	There is a statistically significant correlation between more restrictive ccTLD Whois policies and levels of cybercrime in a domain. <a href="http://forum.icann.org/lists/whois-comments-2008/msg00006.html">http://forum.icann.org/lists/whois-comments-2008/msg00006.html</a>
GAC 1	The legitimate use of gTLD WHOIS data is curtailed or prevented by the use of proxy and privacy registration services.

13 & GAC 11	<p>13. <a href="http://forum.icann.org/lists/whois-comments-2008/msg00016.html">http://forum.icann.org/lists/whois-comments-2008/msg00016.html</a></p> <p>a) The number of proxy registrations is increasing when compared with the total number of registrations.</p> <p>b) Proxy and private WHOIS records complicate the investigation and disabling of phishing sites, sites that host malware, and other sites perpetrating electronic crime as compared with non-proxy registrations and non-private registrations.</p> <p>c) Domain names registered using proxy or privacy services are disproportionately associated with phishing, malware, and other electronic crime as compared with non-proxy registrations or non-private registrations.</p> <p>d) (GAC 11) Domain names registered using proxy or privacy services are disproportionately associated with fraud and other illegal activities as compared with non-proxy registrations.</p>
GAC 2	Restrictions on some or all of the legitimate uses of WHOIS have a negative economic impact.
<b>Area 6</b>	<b>Proxy registrar compliance with law enforcement and dispute resolution requests</b>
3	<p>Some registrars are not revealing registrant data that is shielded by proxy services when presented with requests that provide reasonable evidence of actionable harm, as required under RAA 3.7.7.3.</p> <p><a href="http://forum.icann.org/lists/whois-comments-2008/msg00003.html">http://forum.icann.org/lists/whois-comments-2008/msg00003.html</a></p>
Metalitz Comment	<p>a. Some registrars operating proxy/privacy services are not revealing registrant data when requested in a UDRP proceeding.</p> <p>b. A party's use of a proxy/privacy registration service reduces the party's ability to respond to a UDRP proceeding.</p> <p><a href="http://forum.icann.org/lists/whois-comments-2008/msg00012.html">http://forum.icann.org/lists/whois-comments-2008/msg00012.html</a></p>
20	<p>a. Some proxy and privacy services do not promptly and reliably relay information requests to and from actual registrants.</p> <p>b. Some proxy and privacy services are failing to adhere to RAA 3.7.7.3 – Suggest that this be consolidated with study suggestion #3.</p> <p><a href="http://forum.icann.org/lists/whois-comments-2008/msg00023.html">http://forum.icann.org/lists/whois-comments-2008/msg00023.html</a></p>
12	<p>Registrants would be less likely to falsify their Whois data if the sensitive information of private persons can be secured while giving law enforcement access. <a href="http://forum.icann.org/lists/whois-comments-2008/msg00015.html">http://forum.icann.org/lists/whois-comments-2008/msg00015.html</a></p>
<b>Area 7</b>	<b>WHOIS data accuracy</b>
8	<p>Some Registrars knowingly tolerate inaccurate or falsified Whois data so as to attract and retain registrations by spammers and other bad actors, and do not face deterrent consequences for doing so.</p> <p><a href="http://forum.icann.org/lists/whois-comments-2008/msg00008.html">http://forum.icann.org/lists/whois-comments-2008/msg00008.html</a></p>

11	<p>The use of non-ASCII character sets in Whois records will detract from data accuracy and readability.</p> <p>Note: The hypothesis should be considered in light of the fact that this is a proposed technical analysis and not a study. The original submission for this item suggests a technical analysis in lieu of a survey or statistical study. That is, a technical analysis of how the use of non-ASCII characters in Whois data elements might increase risks of inaccurate data, particularly through use of client-side software that fails to properly check the syntax of fields that contain both ASCII and non-ASCII strings. This analysis should examine and recommend methods for web display and Port 43 retrieval of non-ASCII Whois data, such that those accessing Whois can effectively read, recognize, and reliably use the information to reach registrant contacts and name server resources. <a href="http://forum.icann.org/lists/whois-comments-2008/msg00014.html">http://forum.icann.org/lists/whois-comments-2008/msg00014.html</a></p>
GAC 4	<p>A significant number of Registrars do not apply effective methods to detect fraudulent domain name registrations, and do not take adequate corrective measures when fraudulent information is detected.</p>
GAC 5	<p>A significant percentage of registrants who are legal entities are providing inaccurate Whois data that implies they are natural persons. Furthermore the percentage of registrants with such inaccuracies will vary significantly depending upon the nation or continent of registration. (These hypotheses could be combined with GAC 6.)</p>
GAC 6	<p>A significant percentage of registrants who are operating domains with a commercial purpose are providing inaccurate Whois data that implies they are acting without commercial purposes. Furthermore the percentage of registrants with such inaccuracies will vary significantly depending upon the nation or continent of registration. (These hypotheses could be combined with GAC 5.)</p>