

Follow-up to WHOIS Studies discussion on 1 April 2010
Prepared by Liz Gasster, 5 April 2010

Note: This is a working document and does not constitute or reflect the views of ICANN on any pending or proposed WHOIS study.

On the GNSO Council's 1 April 2010 call, Council members began to discuss staff's recent report analyzing information gathered to-date on two proposed studies of WHOIS. That report is posted here: <http://gnso.icann.org/issues/whois/whois-studies-report-for-gnso-23mar10-en.pdf>, summarized by slides: <http://gnso.icann.org/correspondence/whois-studies-presentation-01apr10-en.pdf>.

On the call, Council members suggested that it would be helpful to identify the WHOIS policy issues that each proposed study is intended to inform. Recognizing that there is a long history of debate, Council members also wanted to understand better the nature of the concerns and viewpoints about these studies that have been expressed to-date. The following is an initial response to that request. Most of the content was extracted from the original study proposals themselves, since study proposers were asked to specifically state the "utility" of each study being proposed. Some additional ideas were gleaned from past council and working group discussions. Council members are encouraged to amplify, modify or correct anything offered below.

While I'm at it, I also offer some thoughts and insights shared by others on the potential policy "relevance" of other pending WHOIS studies.

General Background – a useful document

As I mentioned on the call, there were many steps that led to the Council decision to pursue staff investigation of certain studies of WHOIS. In 20071031-3, the Council concluded that "a comprehensive, objective and quantifiable understanding of key factual issues regarding the gTLD WHOIS system will benefit future GNSO policy development efforts. In a subsequent Council working group, constituency representatives came together to try to agree on which studies to pursue. The WG went through an extended effort to rank each study in priority as a way of guiding further decision, but ultimately disagreed as to whether studies should be conducted at all. Here is a link to the report. <http://gnso.icann.org/issues/whois/gnso-whois-study-group-report-to-council-22may08.pdf>. A chart appended to the report identifies how WG participants "rated" or viewed various studies at that time.

Of course this WG Report long predates my recent analysis, but it may be helpful to the Council because constituencies stated in their own words exactly what they thought the merits and drawbacks were of pursuing various studies at the time.

WHOIS Misuse studies – Potential value of results for future policy discussions (draft for discussion):

The Misuse studies will assess the extent, nature, and impact of harmful actions like spam, phishing, and identity theft that exploit WHOIS contact information.

Note – these must be read in conjunction with feasibility analysis provided in staff's 23 March 2010 report, cited above.

- Studies may shed some light on the extent and type of misuse of WHOIS data and the harm caused by each type of misuse (including economic) by counting and categorizing many different kinds of harmful acts often attributed to misuse of WHOIS data.
- Studies may help identify where (meaning in which countries and gTLDs) harmful acts caused by WHOIS data misuse are more frequent.
- Studies may result in useful qualitative data about the nature of misuse and provide a rough quantitative estimate of cases of misuse.
- Some have suggested that people suffer harm and harassment because their data is publicly displayed in WHOIS. Factual data produced by these studies may help assess how big an issue this is by documenting the impact that misuse has had on those registrants.
- Studies may help to understand how different modalities of WHOIS access impact misuse by documenting how different registrars and registries provide access and deter harvesting.
- Studies may loosely correlate a variety of anti-harvesting measures used by registrars to frequency of misuse to help the Council decide if anti-harvesting measures should be encouraged to reduce misuse, and if so which might be most effective.
- Studies may help to assess the impact of WHOIS public data on personal privacy abuse.

WHOIS Registrant Identification Studies -- Potential value of results for future policy discussions (draft for discussion):

The WHOIS “Registrant Identification” study will explore the extent to which domains used by legal persons (commercial entities) or for commercial purposes are: 1) not clearly identified as such in WHOIS; and 2) correlated to use of Privacy and Proxy services.

Note – these must be read in conjunction with feasibility analysis provided in staff’s 23 March 2010 report, cited above.

- Study will provide factual information about registrants by classifying how natural persons and legal persons identify (or fail to clearly identify) themselves in WHOIS.
- Study may provide information to help the community understand why certain registrants do not clearly identify themselves in WHOIS (e.g., misunderstanding of WHOIS data or purpose, third-party registration of domains used by others, apparent personal privacy concern).
- Study will look at the proportion of individuals that appear to use privacy and proxy services for non-commercial purposes.
- Study will measure how often proxy and privacy services are used by commercial registrants, providing real-world examples that may help the community better understand the demand and motivation of such registrants for proxy and privacy services.
- Study will create a straw-man definition for how registrants can be classified as businesses and what constitutes commercial use, thus helping the community to identify cases where there is ready agreement on these definitions, even if some cases elude easy classification.
- Study results may be used to consider more specific WHOIS data requirements to improve usefulness and accuracy, such as how businesses should be identified in WHOIS, requiring registrants to indicate if domain names are to be used for commercial purpose, or limiting privacy/proxy registration to non-commercial domain names only.
- If this study shows that the majority of registrations by proxy are used for commercial enterprise, this may diminish the policy arguments that proxy registrations are primarily used to preserve the privacy of individuals who register domains for non-commercial purposes.

Privacy and Proxy Abuse Study

The Privacy and Proxy Abuse study will look at the extent to which domain names used to conduct illegal or harmful Internet activities are registered via Privacy and Proxy services.

- Study will quantify the frequency of and illustrate abusive uses of privacy and proxy services.
- Study may show whether privacy and proxy services tend to be abused more often for certain kinds of illegal/harmful activity.
- If this study shows that a high proportion of registrants use proxy/privacy services to shield cyber-squatters, such services could be eliminated or limited in accordance with the extent of legitimate use.

Privacy and Proxy Reveal Study

The Privacy and Proxy Reveal studies will look at the degree to which Privacy and Proxy registrant reveal handling impedes timely identification of parties allegedly involved in such activities.

- Study will quantify the extent to which privacy and proxy service operators reveal registrant information when presented with evidence of actionable harm.
- If this study demonstrated that the use of proxy and privacy services in WHOIS registrations impeded timely identification of criminals, results might support a policy to allow certain organizations (like those affiliated with the APWG and others) to obtain more immediate access to non-public data about registrants who use privacy services.
- Study results may support the need to establish more efficient ICANN domain registration dispute mechanisms in cases where private and proxy domain names are determined to have been used for phishing or other harmful/illegal purposes.
- Study data may support the need for ICANN to strengthen its contractual compliance requirements or efforts for registrars offering proxy services.

WHOIS Service Requirements Study Initial Report

- In calling for this report, the Council recognized that current WHOIS is deficient in a number of ways and is inadequate to support features that have been recommended to address these deficiencies. The Council viewed this inventory of current and potential future requirements as necessary foundation to help guide how WHOIS would need to change in the future. The report is a technical inventory and does not intend to define or suggest the policies or operational rules that should apply. The Council should consider how to make best use of this report in considering next steps.

Study of use of non-ASCII character sets in WHOIS records (study deferred for now pending work of the SSAC-GNSO Internationalized Registration Data Working Group)

This is a technical analysis of how entry of non-US-ASCII WHOIS contact information could increase risks of inaccurate data, particularly through use of client-side software that fails to properly check the syntax of fields that contain both IDN and ASCII strings. This analysis should examine and recommend

methods to display IDN-based WHOIS information such that those accessing WHOIS can effectively read, recognize, and reliably use the information to reach registrant contacts and name server resources.

- If analysis supports this hypothesis, ICANN could undertake policy development to amend requirements for WHOIS data collection and display. For instance, a new policy could require contracted parties to perform syntax checking when users enter any WHOIS data that contains non US-ASCII characters. A new policy might require contracted parties to display WHOIS information in multiple formats to increase usability for those accessing information (e.g. display punycode, Unicode, and font-rendered forms of domains, email addresses, and name servers).

ICANN Compliance Studies that may also be of interest:

NORC WHOIS Data Accuracy Study

- The National Opinion Research Center (NORC) recently completed a study commissioned by ICANN to determine the percentage of domain names registered under the top 5 gTLDs (i.e., .com, .net, .org, .biz, and .info) that contain accurate WHOIS data. See: <http://www.icann.org/en/announcements/announcement-3-15feb10-en.htm>
- The draft report contains a number of useful findings, but in particular it identifies two barriers to accuracy that seem relevant from a policy perspective. First, the report suggests that requiring basic proof of identity at the point of registration would improve accuracy at the point of data entry. Second, the report notes the challenges and further efforts that might be made to keep data accurate and current. The report also notes the costs associated with addressing both sets of challenges.

ICANN Study on the Prevalence of Registrations using Privacy or Proxy Services

The objective of this study was to establish baseline information to inform the community on the prevalence of the use of privacy and proxy services when registering domain names.

<http://www.icann.org/en/compliance/reports/privacy-proxy-registration-services-study-28sep09-en.pdf>

An updated report will be published shortly.