At-Large Whois Registrant Identification Study, Draft Report Workspace

Comment Close Date	Statement Name	Status	Assignee(s) and RALO(s)	Call for Comments	Call for Comments Close	Vote Announcement	Vote Open	Vote Reminder	Vote Close	Date of Submission	Staff Contact and Email	Statement Number
09.03.2013	Whois Registrant Identification Study, Draft Report	Adopted 11Y, 0N, 1A	Salanieta Tamanikaiwaimaro (APRALO)	07.03.2013	08.03.2013 20:00 UTC	08.03.2012	08.03.2012	14.03.2013	15.03.2013	09.03.2013	Barbara Roseman Policy-Staff@icann.org	AL/ALAC/ST /0313/2

Comment/Reply Periods (*)		Important Information Links
Comment Open:	15 February 2013	
Comment Close:	9 March 2013	
Close Time (UTC):	01:00	Public Comment Announcement
Reply Open:		To Submit Your Comments (Forum)
Reply Close:		View Comments Submitted
Close Time (UTC):		Report of Public Comments

Brief Overview	
Originating Organization:	GNSO
Categories /Tags:	Policy Processes
Purpose (Brief):	This study, conducted by NORC at the University of Chicago, uses Whois to classify entities that register gTLD domain names, including natural persons, legal persons, and Privacy/Proxy service providers. Using associated Internet content, the study classifies entities using those domains, and observed potentially commercial activities. Findings will help the community understand how Registrants identify themselves in Whois.
Current Status:	This Public Comment solicitation represents an opportunity for the community to consider the study results detailed in this draft report, ask questions, and request clarifications. In parallel, ICANN and NORC will conduct Webinars to facilitate feedback by summarizing this study's purpose, methodology, key findings, and conclusions.
Next Steps:	NORC will consider all comments submitted to this Public Comment forum during the comment period, incorporate any needed clarifications, and then publish a final version of this WHOIS Registrant Identification study report. Afterwards, the GNSO Council will use this report to inform future Whois policy-making.

Staff Contact:	Barbara Roseman	**Email:**	Policy-staff@icann.org

Detailed Information
Section I: Description, Explanation, and Purpose
The WHOIS Registrant Identification Study uses Whois to classify entities that register gTLD domain names, including natural persons, legal persons, and Privacy/Proxy service providers. Using associated Internet content, it then classifies entities using those domains and potentially commercial activities. NORC at the University of Chicago has been selected to conduct this study and has issued a draft report, which is now available for public comment.
Section II: Background

As part of its effort to develop a comprehensive understanding of the gTLD Whois system, the GNSO Council expressed an interest in conducting an in-depth study of how entities that register and use gTLD domain names identify themselves in Whois. At the GNSO's request, ICANN issued an RFP in October 2009 describing a study to examine the extent to which domains used by legal persons or for commercial purposes (1) are not clearly identified as such in Whois Registrant data and (2) are correlated to use of Privacy and Proxy services.

After considering RFP responses received in late 2009 from researchers willing to undertake that Registrant Identification study, along with significant concerns raised by GNSO Council members regarding the above-stated study hypothesis, theGNSO Council decided to revamp the study's goals and approach.

In May 2010, a revised study was approved by the GNSO Council and awarded to NORC at University of Chicago. This exploratory study – detailed by this draft study report – seeks a more foundational understanding of the types of entities and kinds of activities observed in gTLD domains, including (but not exclusively focused on) those registered using Privacy or Proxy services. Accordingly, the categories of entities and activities to be studied were not pre-determined, but rather generated as researchers examined representative samples of active websites and related Whois data.

Study findings are intended to provide raw data needed to understand how entities that register and use gTLD domain names identify themselves in Whois, including (but not limited to) domains registered Privacy/Proxy services and domains engaged in potentially commercial activity. This empirical data will not only enable Council to respond to GAC questions, but will also create a baseline for evaluating potential Whois policy changes.

Section III: Document and Resource Links

- WHOIS Registrant Identification Study Revised Terms of Reference [PDF, 719 KB]
- WHOIS Studies Report for GNSO [PDF, 487 KB]
- WHOIS Registrant Identification Study – Rationale for Changes [PDF, 121 KB]
- NORC Selected to Conduct a gTLD Whois Registrant Identification Study
- WHOIS Registrant Identification Study | Draft Project Summary Report [PDF, 505 KB]
- ICANN WHOIS Registrant Identification Project | Appendix A: Exploratory Analysis Report [PDF, 1.1 MB]
- WHOIS Registrant Identification Study | Appendix B: Variable Glossary [PDF, 250 KB]

**Section IV: Additional Information**

Additional WHOIS studies are now being conducted at the request of the GNSO Council, as summarized by:http://gnso.icann.org/issues/whois/

(*) Comments submitted after the posted Close Date/Time are not guaranteed to be considered in any final summary, analysis, reporting, or decision-making that takes place once this period lapses.

FINAL VERSION TO BE SUBMITTED IF RATIFIED

Please click here to download a copy of the PDF below.
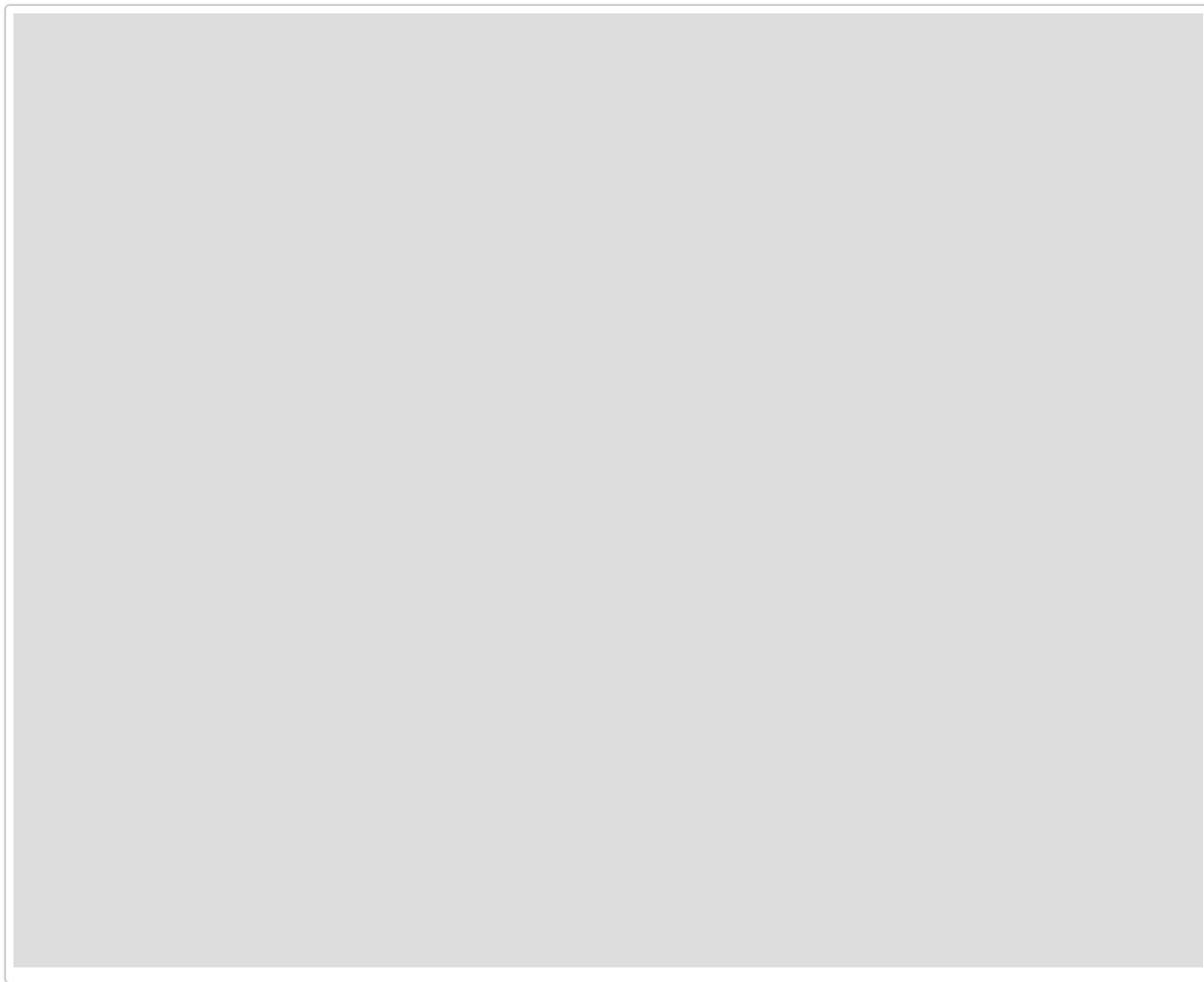
FINAL DRAFT VERSION TO BE VOTED UPON BY THE ALAC

The At-Large Advisory Committee (ALAC) would like to thank the GNSO and particularly NORC at the University of Chicago for its Draft Project Summary Report.

The WHOIS Registrant Identification Study was useful in identifying the types of registrants, in adding context to issues raised within previous WHOIS Reports and helped to inform the discussions on WHOIS Misuse, Abuse and Privacy related issues. The findings are very useful and will serve as an important factor in dealing with the development of Policies.

The breakdown of Apparent Domain Name users was interesting where:

- 37% were "legal person" users;
- 26% were tagged as "no usable online content";
- 21% were "parked domains";
- 12% were of an "unknown user type"; and
- 5% were "natural person" users.

As the second largest group, it is surprising to find the "non-usable online content" has such importance and we would like to see the breakdown of this figure from a geographical perspective.

On the methodology, we do not understand why NORC checked DNSBL listings, which are mostly used for fighting e-mail spam. This is even more of a paradox because the focus appeared to be on reviewing web sites and FTP servers. Today, the latter is seldom used. Domain names that may be used solely for e-mail do not appear to have been verified. Given the small sample, it would have been possible to send out an e-mail to the domain registrant, asking if they were an individual or a business. This might have reduced the large number of unknowns in the research.

It was interesting to discover the relative percentage of Privacy and Proxy use among legal persons and that of the 586 sampled domains, 55 % were legal person registrants, 25% were natural person registrants, 15% were privacy and proxy registrants and 5% were unknown.

A closer inspection of the study reveals that of the 320 domain registrants coded as Privacy/Proxy service providers, only 10 were determined to be privacy service providers. Therefore, 3 % of domain registrants appear to be using a privacy or proxy provider. The ALAC notes that this figure is significantly different from ICANN's *Study on the Prevalence of Domain Names Registered using a Privacy or Proxy Service* which had found a much larger percentage: approximately 15 percent of domain registrants. The ALAC would be interested in receiving an explanation about this discrepancy.

We are concerned with the size of the "unknown country" data as per Table F.1 on pages 83-84 of the Draft Report. This appears to be very significant. Also, the study appeared to be using a data sample that was very US Centric. The ALAC would have preferred to see more sampling from the Asian and Eastern European regions since these are two areas with high Internet service growth.

The ALAC is grateful both for the Briefing and the Report and may revert with requests for clarifications on various aspects of the Study once we have completed un-packaging the Draft Report.

FIRST DRAFT SUBMITTED

The At-Large Advisory Committee (ALAC) would like to thank the GNSO and particularly NORC at the University of Chicago for its Draft Project Summary Report.

The Whois Registrant Identification Study was useful in identifying the types of registrants, in adding context to issues raised within previous Whois Reports and helps to inform the discussions on Whois Misuse, Abuse and Privacy related issues. The findings are very useful and will serve as an important factor in dealing with the development of Policies. We are still in the process of unpackaging this study which took two years to complete.

The breakdown of Apparent Domain Name users was interesting where legal person users were at 37%, no usable online content at 26%, parked domain was at 21%, unknown user type was at 12% and natural person user at 5%. The non-usable online content is alarming and we would like to see the breakdown of this from a geographical perspective.

It was interesting to discover the relative percentage of Privacy and Proxy use among legal persons and that of the 586 sampled domains, 55 % were legal person registrants, 25% were natural person registrants, 15% were privacy and proxy registrants and 5% were unknown.

The fact that the study reveals, is that of the 320 domain registrants coded as Privacy/Proxy service providers, only 10 were determined to be privacy service providers. In other words, about 3 % of domain registrants are using a privacy or proxy. The ALAC notes that this is significantly different from ICANN's *Study on the Prevalence of Domain Names Registered using a Privacy or Proxy Service* which had found a much larger percentage: approximately 15 percent of domain registrants.

We are concerned with the unknown data as per Table F.1 on pages 83-84 of the Draft Report. It was also interesting to see the breakdown of these unknown types by countries. However, we would have liked to see more categories in the area of illegal activities as well as see more samples from the Asian and Eastern European regions.

The ALAC is grateful both for the Briefing and the Report and will revert with requests for clarifications on various aspects of the Study once we have completed our unpackaging the Draft Report.