Report of Public Comments

Title:	Whois Misuse Study Draft Report
Publication Date:	
Prepared By:	ICANN Staff

Comment Period:	
Comment Open Date:	27 November 2013
Comment Close Date:	27 December 2013
Reply Close Date:	18 January 2014
Time (UTC):	23:59 UTC

Important Information Links
Announcement
Public Comment Box
View Comments Submitted
Report of Public Comments

Staff Contact:	Mary Wong	**Email:**	policy-staff@icann.org

Section I: General Overview and Next Steps

As part of its efforts to obtain a comprehensive, objective and quantifiable understanding of various aspects of the Whois gTLD registration data system, the GNSO Council approved this study in September 2010. Carnegie Mellon University's Cylab (CMU) was selected to perform the study in April 2011.

The purpose of this study was to attempt to prove or disprove the following hypothesis: ***Public access to Whois data leads to a measurable degree of misuse – that is, to actions that cause actual harm, are illegal or illegitimate, or otherwise contrary to the stated legitimate purpose.*** From the start of the study, the GNSO Council had recognized that it would not be possible to either quantitatively or qualitatively assess the extent to which Whois misuse is "significant", although it was possible to measure and categorize many different types of harmful acts often attributed to the use of Whois data.

The overall study consisted of two related studies. First, the research team surveyed (1) registrants of a representative sample of domain names registered in the top five gTLDs – .biz, .com, .info, .net and .org; (2) registries and registrars associated with registration of the surveyed domain names to identify Whois anti-harvesting mechanisms they employ; and (3) cybercrime researchers and law enforcement organizations to gather examples and statistics related to harmful acts attributed to Whois misuse. Secondly, the research team designed and conducted an experiment to measure Whois misuse by registering 400 domains across 16 registrars, associating unique, synthetic Whois contact information with test domains and monitoring incidents of misuse for six months.

The body of public comment received has been analyzed by ICANN staff and CMU as described below,

and the comments and analysis report will be forwarded to the GNSO Council for its consideration and review.

Section II: Contributors

At the time this report was prepared, a total of fourteen (14) community submissions had been posted to the Forum. The contributors, both individuals and organizations/groups, are listed below in chronological order by posting date with initials noted. To the extent that quotations are used in the foregoing narrative (Section III), such citations will reference the contributor's initials.

Organizations and Groups:

Name	Submitted by	Initials
Messaging, Malware & Mobile Anti-Abuse Working Group	Jerry Upton	M3AAWG
ICANN At Large Advisory Committee	ALAC Staff	ALAC
GNSO Registries Stakeholder Group	Keith Drazek	RySG
GNSO Business Constituency	Steve DelBianco	BC
International Anti-Counterfeiting Coalition	Travis Johnson	IACC
General Electric	Sean Merrill	GE
GNSO Intellectual Property Constituency	Steve Metalitz	IPC
Dotgay LLC	Scott Seitz	DG
GNSO Non Commercial Users Constituency	Kathy Kleiman	KK
GNSO Non Commercial Stakeholder Group	Avri Doria	NCSG

Individuals:

Name	Affiliation (if provided)	Initials
Net Gremlin		NG
Stephen		SS
Nacer Adamou Saidou		NAS
Greg Aaron		GA

Section III: Summary of Comments

General Disclaimer: This section is intended to broadly and comprehensively summarize the comments submitted to this Forum, but not to address every specific position stated by each contributor. Staff recommends that readers interested in specific aspects of any of the summarized comments, or the full context of others, refer directly to the specific contributions at the link referenced above (View Comments Submitted).

Usefulness of the Study

A number of commentators noted that the study findings (1) corroborate individual user and anecdotal experiences of Whois misuse (ALAC; NCUC; NCSG); (2) confirm that information published

in Whois is misused to a measurable degree (ALAC; GA; NCUC; NCSG; RySG); and (3) demonstrate that anti-harvesting techniques may reduce Whois email misuse (RySG). One commentator stated that the findings are at least illustrative, if not necessarily definitive (BC). Two commentators congratulated the research team for an excellent and thorough study (NAS; NCUC), although a number of commentators expressed concerns about study limitations (detailed further below).

Two commentators urged ICANN to reject the study's findings, questioning its design and execution —most notably, the small sample size of the descriptive survey (IACC; M3AAWG). Other commentators either cautioned against basing ICANN policy development on the study findings, given its limitations (IPC) or questioned its utility in relation to the question of balancing the benefits and drawbacks of publicly-available Whois information (GA).

Methodology of the Study

Several commentators expressed concern at the small sample size and low response rate of the descriptive survey conducted by the research team (GA; IACC; IPC; M3AAWG; RySG). Two commentators noted that these limitations had been acknowledged by the study team in its report (BC; RySG). One commentator suggested that the low response rate argues for the continuation of research into creating a more accurate Whois system as well as registrant education (BC).

Some commentators were concerned that the registrant survey relied upon self-reporting of Whois misuse by the respondents and that eighty percent of the spam emails measured by the experimental study were associated with domains issued by a single registrar (GA; IACC; IPC). GA suggested that these factors could affect a conclusion that a registrant's experience of misuse can be attributed to the public availability of Whois contact information.

The limited geographical scope of respondents for the registrant and registry surveys was noted by two commentators (IPC; M3AAWG), although a third commentator commended the research team for its attempts to encourage study participation from other geographical regions such as Latin America (NCUC).

Comments about Specific Findings of the Study

One commentator thought that CMU's finding that domain name price is negatively correlated to misuse was not necessarily supported by the data available, suggesting that this might be due more to a registrant's choice of registrar rather than price (GA). Another commentator suggested that the possible correlation between pricing and misuse could be something that individual registry operators could take into account as one means of reducing Whois misuse (RySG).

Commentators also noted CMU's finding that Whois anti-harvesting techniques are statistically-significant in alleviating misuse (GA; RySG), suggesting that this demonstrates the value of rate-limiting port 43 Whois access (GA) and that mitigating actions can be taken to limit Whois misuse (ALAC; BC). In particular, GA thought that the problem with registrars having weak (or no) anti-harvesting measures on their port 43 servers may be alleviated should registries all move to a "thick"

rather than a "thin" Whois model, increasing the role played by registry anti-harvesting.

Several commentators were not surprised by CMU's finding that Whois publication of email addresses could be correlated with a high incidence of spam emails (GA; GE). One commentator (GA) cautioned that spam emails can be received even at email addresses that are not published in Whois, due to tactics such as guessing on the part of spammers, despite methods used to deter this when setting up test domains for the experimental study.

One commentator questioned whether there was a statistically significant link between junk mail received and addresses published in Whois (GE). Another commentator requested clarification regarding the study duration, the finding that the gTLD itself is the sole statistically significant characteristic affecting phone number misuse, and the possible effect of the availability of privacy or proxy services (NAS).

A commentator suggested that the team's finding that experimental domain names representing natural persons attracted less misuse than other types of names could be the subject of further consideration by ICANN (RySG).

Other Comments and Suggestions

One commentator described examples of Whois misuse known to its members but outside the scope of this study, including cases of stalking, political persecution and anti-competitive activity (NCUC). However, another commentator noted that some other conceptually similar forms of Whois misuse - such as harassment, stalking, and identity or intellectual property theft – had not been reported by survey respondents to occur at a significant rate (GE). GE also suggested that ICANN consider cooperating with the United States Federal Trade Commission and other similar bodies to develop more direct ways of reducing spam.

GA suggested that the study did not answer some important questions raised by its methodology and findings, e.g. lack of malware received at the published email addresses for its experimental domains; the relatively high number of voicemails received from just two callers to a small number of "registrants" of the experimental domains; the distinction between "targeted Whois spam" and generic postal spam; and the absence of content analysis for emails received by experimental domains preventing measurement of possible phishing attacks or drive-by malware. GA also noted that Whois misuse may not be evenly distributed across the registrant population; that it is sometimes localized around certain registrars, resellers or TLDs, and just a few perpetrators might be responsible for a significant amount of Whois misuse.

Section IV: Analysis of Comments

General Disclaimer: This section is intended to provide an analysis and evaluation of the comments received along with explanations regarding the basis for any recommendations provided within the analysis.

A number of commentators expressed support for ICANN's efforts to study the Whois system, including the development of policies to mitigate misuse (ALAC; IACC; IPC) and to ensure Whois accuracy (IACC). One commentator (BC) recommended that registries and registrars adopt a fully-integrated set of mitigating actions, and another (ALAC) suggested that a useful beginning for a coordinated response from registries and registrars would be the adoption of best practices from every domain that has proven and useful anti-harvesting implementation measures.

The limitations of the study were the focus of some commentators, as noted above. While several commentators thought ICANN should reject the study's findings as a result of its limitations (IACC; M3AAWG), others suggested that certain findings could be used with appropriate context and caution (RySG). On the other hand, one commentator (IPC) suggested that any further study of Whois misuse should not be based on the findings of this study but rather a new study that addresses these limitations.

In view of the study's findings, NCUC recommended that measures be taken to protect registrants from Whois misuse. Other commentators (BC; IPC) noted the benefits of ensuring public access to reliable and accurate Whois data where there is evidence of actionable harm. GE commented that making Whois more rigorous should help in addressing the most severe form of Whois misuse and SS stated that Whois contact information must be readily available with transparency as a default.

ICANN and CMU will consider all public comments received in preparing a final report for publication, which will be forwarded to the GNSO Council for further review and action. In particular, CMU expects to clarify areas of the report that in their view may not have been fully understood by commentators (include further detail about methods and findings where available) and address questions raised about sample design, the statistical significance of findings, the impact of acknowledged limitations on findings, and the technical validity of suggested interpretations.