# ICANN is Seeking Independent Researchers to Respond to Request for Proposal on WHOIS Misuse Studies –

The first area of study focuses on “misuse” of public WHOIS information, and, pursuant to the above request, ICANN is seeking to engage independent research organizations to undertake one or more proposed WHOIS Misuse studies.

As this references different linked articles I will look at the linked articles separately:

Requirements on the documents/web links: Request for Proposals for WHOIS Misuse Studies and Terms of Reference for WHOIS Misuse Studies

WHOIS Misuse Studies - WSMS

[WSMS-D1 & D2-R1]

WHOIS Misuse studies are designed to determine the extent, nature, and impact of WHOIS data misuse (i.e., harmful actions taken using WHOIS contact information) by surveying Registrants about misuse incidents they may have experienced, and conducting experiments to measure misuse and the effectiveness of WHOIS anti-harvesting measures applied by Registrars.

[WSMS-D1 & D2-R2]

Need to eliminate or limit WHOIS misuse refers to harmful acts that exploit contact information obtained from WHOIS. Those harmful acts may include generation of spam, abuse of personal data, intellectual property theft, loss of reputation or identity theft, loss of data, phishing and other cybercrime related exploits, harassment, stalking, or other activity with negative personal or economic consequences.

[WSMS-D1 & D2-R3]

A key objective of these experiments is to identify factors that increase or decrease WHOIS misuse. To enable this, WHOIS query and response practices applied to each test domain should be examined, looking for relationships between anti-harvesting measures and the frequency harmful acts against published addresses. However, these experiments will only examine measures applied to WHOIS queries. Analyzing the impact of Privacy/Proxy registration services that prevent addresses from being published in WHOIS are beyond this study's scope.

[WSMS-D1 & D2-R4]

Different sources are unlikely to supply the same information about each WHOIS misuse incident or represent those data elements consistently. Inputs obtained from all sources must therefore be normalized to enable aggregation, comparison, and statistical analysis.

[WSMS-D1 & D2-R5]

Because all data elements may not be available from every source, a minimum set of elements must be required for any reported incident.

[WSMS-D1 & D2-R6]

It is essential that every incident denote whether misused information was published exclusively by WHOIS – this element is mandatory. Registrants may not reliably know this, Internet search engines should be used to verify that misused addresses are in fact NOT readily available from sources beyond WHOIS.

[WSMS-D1 & D2-R7]

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.

[WSMS-D1 & D2-R8]

WHOIS misuse refers to harmful acts that exploit contact information obtained from WHOIS. Those harmful acts [1] may include generation of spam, abuse of personal data, intellectual property theft, loss of reputation or identity theft, loss of data, phishing and other cybercrime related exploits, harassment, stalking, or other activity with negative personal or economic consequences.

[WSMS-D1 & D2-R9]

The descriptive study documents and analyzes WHOIS misuse incidents (harmful acts) that have already occurred. The experimental study stimulates and records misuse to measure more reliably the impact of making WHOIS data public and WHOIS query filters applied to deter data harvesting.

[WSMS-D1 & D2-R10]

Both kinds of studies are defined here because conducting both (either sequentially or in parallel) is one way to compensate for limitations inherent to each research method. If both kinds of Misuse studies are in fact conducted, terminology, inputs, and outputs for each should be defined consistently to facilitate reuse, integration, and correlation.

[WSMS-D1 & D2-R11]

Inputs will be gathered from a representative set of sources, aggregated, and analyzed to categorize misuse by type (kind of harmful act), severity (impact of act on Registrant), and applicable anti-harvesting measures (WHOIS query/response filters). Input data sources include the following.

a) Registrants: As proposed by, survey a representative sample of Registrants that own domains in the top five gTLDs about specific harmful acts they have experienced which they believe were sent using WHOIS contact information.

b) Registrars and Registries: Pursuant to proposals, survey Registrars and Registries in several regions/countries to obtain contextual information about how WHOIS data can be queried for the above-sampled domains (e.g., supported query vectors, applied anti-harvesting measures, known harvesting attacks).

c) Cybercrime Researchers: To put WHOIS misuse into broader perspective, contact a representative set of independent industry research organizations that track related cybercrime activities (e.g., phishing, spam, identity theft) to gather examples and statistics regarding harmful acts occurring in many different regions/countries.

Examples include the Anti Phishing Working Group, the Privacy Rights Clearing House, and the Online Trust Alliance (AOTA).

d) Consumer Protection, Regulatory, and Law Enforcement Organizations: To further put WHOIS misuse into broader perspective, contact a representative set of organizations that victims contact to report cybercrimes to gather examples and statistics regarding harmful acts occurring in many different regions/countries. Examples include the U.S. Federal Trade Commission, the FBI/NWCC Internet Crime Complaint Center (IC3), and the Identity Theft Assistance Center (ITAC).

[WSMS-D1 & D2-R12]

Sources a) and b) can describe harmful acts attributed to misuse of WHOIS public data and/or the circumstances surrounding incidents first-hand; this primary research allows more qualitative analysis (e.g., actual impact on Registrants, effectiveness of WHOIS anti-harvesting measures).

[WSMS-D1 & D2-R13]

Sources c) and d) can only provide examples and aggregated statistics, but incorporating this secondary research may facilitate more quantitative analysis (e.g., which kinds of harmful acts are most frequently reported worldwide, to what extent are these acts attributed to WHOIS misuse).

[WSMS-D1 & D2-R14]

Due to the diversity and complexity of WHOIS storage and access, survey b) must account for differences between "thick" and "thin" Registries and the impact of resellers, affiliates, and third-party WHOIS operators.

[WSMS-D1 & D2-R15]

Misused data could also have been obtained in bulk form or from a reseller or a third-party WHOIS operator. It is not feasible to identify all possible vectors through which a victimized domain's WHOIS data could have been obtained, but survey b) will at least examine the primary vectors

[WSMS-D1 & D2-R16]

Experimental study should be based on the Harmful acts that attributed to misuse of WHOIS public data can also be measured in a more reliable manner by conducting an experiment that monitors a set of test domains, registered through a representative sample of Registrars, distributed proportionally across the top five gTLDs.

[WSMS-D1 & D2-R17]

These experiments can build upon some of the techniques used by an earlier SSAC WHOIS Spam experiment [8] – notably random generation of names in published / unpublished addresses. However, to meet the Misuse Study's objectives, researchers must start with a broader domain sample and design experiments to measure harmful acts that go beyond spam, including phishing, identity theft, and other cybercrime-related exploits.

[WSMS-D1 & D2-R18]

Email Spam: As proposed by [5], compare the volume of unsolicited bulk email sent to WHOIS-published addresses vs. unpublished addresses. To differentiate between types of misuse, received messages must be divided into at least three categories: solicited email, phishing email (see below), and all other (unsolicited bulk) email – that is, spam.

[WSMS-D1 & D2-R19]

Postal and Telephone Spam: Measure the volume of postal mail delivered to each Registrant's published address and calls received by each Registrant's published telephone number. Here again, unsolicited bulk mail and telemarketing calls would be differentiated from apparent attempts to "phish" for identities and all other postal / telephone communication.

[WSMS-D1 & D2-R20]

Phishing: Categorize a subset of the email, postal, and/or telephone contacts received in the spam test cases as attempted phishing attacks requiring further analysis (e.g., impact assessment). These may include both mass-mail phishing attacks and spear-phishing attacks specifically addressed to the Registrant.

[WSMS-D1 & D2-R21]

Abuse of Personal Data and Identity Theft: To detect and measure abuse of personal data in identity theft attempts, further analyze the content of email, postal, and/or telephone calls addressed to the Registrant. For example, letters from banks or merchants denying a credit application or purchase could signal attempted identity thefts, while email carrying a key-logger trojan (or a URL that leads to one) could represent attempted identity theft.

[WSMS-D1 & D2-R22]

Harassment and stalking: This study will not attempt to solicit or measure these kinds of harmful acts due the extreme difficulty of correlating such behaviour to misuse of WHOIS public data for fictional Registrants.

[WSMS-D1 & D2-R23]

Intellectual property theft and loss of data: Cyber-criminals do not usually attempt to steal intellectual property without an interesting or high-value target in mind; a fictional Registrant is unlikely to be targeted by these harmful acts.

[WSMS-D1 & D2-R24]

A key objective of these experiments is to identify factors that increase or decrease WHOIS misuse. To enable this, WHOIS query and response practices applied to each test domain should be examined, looking for relationships between anti-harvesting measures and the frequency harmful acts against published addresses. However, these experiments will only examine measures applied to WHOIS queries. Analyzing the impact of Privacy/Proxy registration services that prevent addresses from being published in WHOIS.

[WSMS-D1 & D2-R25]

Different sources are unlikely to supply the same information about each WHOIS misuse incident or represent those data elements consistently. Inputs obtained from all sources must therefore be normalized to enable aggregation, comparison, and statistical analysis.