Contents

1. Objective 1
1.1 Background 1
1.2 Challenges 3
1.3 Objective 3
2. Approach 4
3. Inputs 6
3.1 Primary Sources 6
3.2 Secondary Sources 9
3.3 Data Collection 10
3.4 Data Elements 11
4. Outputs 13
5. References 15

WHOIS Proxy / Privacy Relay & Reveal Studies – Draft Definition

This exploratory study will analyze a sample of relay and reveal requests sent for Privacy/Proxy-registered domain names to document how they are processed and identify factors that may promote or impede timely communication and resolution.

1. Objective

This exploratory study is based on several proposals [8][9][10][11][12] by members of the ICANN community. While proposed study approaches varied, all sought empirical data about communication relay and identity reveal requests sent for Privacy/Proxy-registered domain names. Absent such data, the community has been unable to agree whether policy changes are needed to make those requests more efficient and reliable.

Currently, each Proxy or Privacy service provider has its own independently-developed practices for handling such requests. There is no common format for submitting these requests and no central repository for tracking them. The highly diverse and distributed nature of these practices has made it difficult to even assess the effectiveness of related ICANN policies. The objective of this study is therefore to help the ICANN community better understand how communication relay and identity reveal requests sent for Privacy/Proxy-registered domain names are actually being handled today.

1.1 Background

Individuals, businesses, law enforcement, and other parties may wish to identify and contact domain name registrants for a wide variety of reasons. WHOIS services are often used to locate that contact information.

Section 3.7.7.1 of the ICANN Registrar Accreditation Agreement (RAA) [4] requires that all Registered Name Holders provide accurate and reliable contact details including the name of an authorized person for contact purposes. Section 3.3.1 of the RAA further requires Registrars to provide an interactive web page and a port 43 WHOIS service to enable free access to up-to-date data concerning all active registered domain names. This WHOIS service can be used to obtain the name and address of the Registered Name Holder and technical and administrative contacts.

Some domain names use Proxy or Privacy registration services [1] to provide anonymity or privacy protection for domain name users. *Privacy* services offer alternate WHOIS contact information and mail forwarding services while not actually shielding the Registered Name Holder's identity. *Proxy* services register domain names on a third party's behalf and then license their use so that the provider's identity and contact information (and not the licensee's) is published in WHOIS.

When any party attempts to identify or contact the user of a Proxy/Privacy-registered domain for any reason, the associated Proxy/Privacy provider could relay, acknowledge, respond with an explicit accept/reject, or otherwise act upon those requests. For example, a Privacy service might auto-forward all emailed communication relay requests, including domain name purchase offers, registration service renewal offers, or other commercial messages. A Proxy service might respond to a reveal request with the licensee's identity or notify law enforcement, given reasonable evidence of actionable harm. In some cases (such as email bounces), parties may also send these requests to the domain name's Registrar.

Although there is no explicit RAA requirement that Proxy/Privacy services or Registrars handle communication relay or identity reveal requests, many providers have developed their own policies and procedures for doing so. These are business practices that may also be constrained by local and national data protection and privacy laws and influenced by section 3.7.7.3 of the RAA, which states:

> "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 current contact information provided by the licensee and the identity of the licensee to a party providing the Registered Name Holder reasonable evidence of actionable harm."

See also Draft Advisory [13] which seeks to clarify what this clause means.

1.2 Challenges

Proposals [8][9][10][11][12] made by various members of the ICANN community have asserted that some individuals, businesses, first responders, and law enforcement officials who have tried to identify and contact Privacy/Proxy-registered domain users have encountered significant challenges.

For example, in proposal [9], members of the Anti-Phishing Working Group (APWG) hypothesized that Privacy/Proxy registrations lengthen phishing website take-down times in two ways:

- By preventing direct contact with legitimate Registered Name Holders of Privacy-registered domains that have been hacked by phishers, and
- By waiting for providers to assess claims about actions allegedly taken by third parties that license Proxy-registered domains to for phishing attacks.

Similar concerns have been expressed by brand owners and their representatives who investigate trademark and copyright infringement. In proposal [12], members of the Intellectual Property Constituency asserted that some providers have not responded to reveal requests at all, while other responses have denied any knowledge of or relationship to the domain name while refusing to disclose the licensee's identity without a subpoena.

1.3 Objective

This study is intended to provide the ICANN community with empirical data to evaluate such concerns. However, the diverse and distributed nature of today's request handling prevents studying a statistical microcosm; there is no recognized or comprehensive repository from which to pull a random sample. Furthermore, so little is commonly understood about communication relay or identity reveal requests that even identifying concrete message flows, measurable factors, and testable hypotheses has proven difficult.

As a result, this study will establish a foundation for future research by **exploring a large, broad sample of actual relay and reveal requests**. Individuals, businesses, first responders, complaint centers, and law enforcement agencies across the globe that volunteer to participate in this study will be asked to supply a complete and accurate set of requests sent during a study period. For each request, researchers will then solicit secondary input from the associated Privacy/Proxy service provider and Registrar to determine if the request was received, relayed, responded to, or otherwise acted upon.

By documenting real-world experiences, this study will attempt to characterize ways in which requests are commonly handled, enumerate potential outcomes (e.g., request bounced, request explicitly denied, problem resolved), and isolate measurable factors that appear to promote or impede timely communication or resolution. Ultimately, this study may find that request practices are too diverse to permit meaningful statistical analysis. Alternatively, this study may document flows and isolate factors suitable for empirical measurement by future studies, designed to prove or disprove testable hypotheses.

2. Approach

This study will explore a large, broad sample of actual relay and reveal requests sent for Privacy/Proxy-registered domains within the top five gTLDs (.biz, .com, .info, .net, .org), pertaining to Registered Name Holders or third party licensees that participants have tried to contact or identify during normal business activities.

This study focuses solely on Privacy/Proxy-registered domain names to better understand the challenges encountered by anyone wishing to identify or contact these Registered Name Holders or third party licensees, no matter why they want to do so. By definition, Privacy/Proxy services insert some level of indirection into these request flows:

- For many domains, Registered Name Holders can be reached directly at addresses obtained from WHOIS. However, for Privacy/Proxy-registered domains, Registered Name Holders or third party licensees cannot be reached directly via WHOIS-published addresses. Instead, **communication relay requests** may be sent to the Privacy/Proxy service provider published in WHOIS, or attempted using addresses obtained from other sources, websites or communications associated with the domain.

  This study will explore factors that could potentially impact relay request delivery, such as address accuracy, communication method (email, fax, postal), message size, and sender. Because Registered Name Holders and third party licensees may reasonably opt not to respond to relayed requests, this study will only try to isolate factors that appear to promote or impede **timely delivery** to those intended recipients.

- For many domains (including those registered via Privacy services), the Registered Name Holder's identity is published directly in WHOIS. However, for domains registered via Proxy services, the name of the licensee is not published in WHOIS; third party licensees can typically only be identified by asking the Proxy to **reveal the licensee's identity**, given reasonable evidence of actionable harm [4][13].

  This study will explore factors that could potentially impact reveal request outcome, such as the request's source, content, and stated reason, the licensee's geographic location and documented activities, the domain's Proxy and Registration service state, a Registrar's obligation to maintain confidentiality during an active law enforcement investigation, and Proxy and Registrar policies and practices for handling such requests. Because Proxies and Registrars may take action without explicitly responding to request senders, this study must do more than isolate factors that appear to promote or impede **timely response**. It should also attempt to determine how each request was actually handled by the Proxy and Registrar, documenting multi-step message flows, potential outcomes, and factors that appear to promote or impede **timely resolution** of the actionable harm (if any) behind the request.

This study is framed as an exploratory study rather than a hypothesis-driven study because Privacy/Proxy and Registrar request handling processes vary so widely. For example, according to study proposal [12], some Proxies routinely reject all reveal requests, requiring a subpoena or lawsuit before revealing licensee identity. While Proxies may have sound reasons for rejecting some requests, senders may also have sound reasons for not instigating legal action before the licensee's identity is known. Even when requests are accepted, some Proxies do not explicitly respond to request senders but instead terminate the licensee's Proxy service, thereby publishing their identity in WHOIS unless/until domain name registration service is also terminated.

Cases like these have been documented, but no empirical data is available to quantify how often requests do not elicit a response or contact information, how often requests result in remedial action without explicit response, common reasons that requests are rejected or accepted, or how much delay is typical between request generation and resolution. Absent this kind of data – or even a uniform understanding of request handling policies and processes – the ICANN community has been unable to assess the overall effectiveness of related policies or agree upon any needed improvements.

To advance policy debate, this study will use actual relay and reveal requests, responses, and outcomes, recorded by volunteers who agree to submit all requests (and supporting documentation) generated during normal business over a 12 month study period. This approach reflects feedback on earlier drafts and related proposals. In particular:

- This study does **not** use test domains or simulated requests because doing so would not document real-world experiences for real-world incidents.
- This study does **not** originate requests about fictitious incidents because fictitious requests could not reasonably claim actionable harm.
- This study explores **both** relay and reveal because many reveal requests are preceded by and/or trigger relay requests.
- This study examines **both** Privacy and Proxy-registered domains because request sources do not always or consistently differentiate between them.
- This study does **not** measure response time alone because that metric is not very meaningful without broader context (i.e., what actually happened and why?)
- This study does **not** exclude any requests based on the sender's reason or recipient's actions; all submitted requests will be characterized to establish a complete picture.

Reasonable concerns have been raised about the statistical validity of a volunteer-based study. To validate primary inputs and examine more of the underlying process, this study also enlists the help of Privacy/Proxy providers and Registrars. Specifically, the Privacy/Proxy provider and Registrar associated with each relay or reveal request will be given an opportunity to supply secondary input about how each request was handled, as well as their published policies and practices for handling such requests.

Nonetheless, because input from both primary and secondary sources is still voluntary, sampled data will inevitably yield an incomplete view of current practices. For example, participating senders may be skewed towards those experiencing more significant or

frequent challenges, while participating providers may be skewed towards those with more rigorous policies and case tracking. This proposal incorporates many steps intended to improve the completeness and usefulness of sampled data; researchers are encouraged to suggest further steps. By assembling all available pieces of each sampled flow, it is hoped that this study can shed light on how requests are often handled, establishing a foundation for future empirical studies. However, it must be understood that this exploratory study's approach will not yield a statistically random sample.

Note that other WHOIS studies [3][6][7] have been defined to measure the overall frequency of Privacy/Proxy registrations, the types of entities that commonly use Privacy/Proxy-registered domains and for what apparent purpose, and whether Privacy/Proxy-registered domains are often abused by parties engaged in harmful/illegal Internet activity. Those questions are therefore outside the scope of this study. However, to explore a representative slice of the top 5 gTLD Privacy/Proxy-registered domain population, this study must consider gTLD distribution (see study [2]), frequency of Privacy/Proxy registrations (see study [3]), and geographic location of study participants.

3. Inputs

The first step in conducting this study is to generate a sufficiently large and broad sample of relay and reveal requests for Privacy/Proxy-registered domain names. As noted above, this primary input will be gathered from volunteers who routinely send relay and/or reveal requests during normal business activities.

- **Communication relay requests** are sent in a wide variety of situations in which a sender wishing to contact a Registered Name Holder or third party licensee does so using the alternate WHOIS contact data supplied by a Privacy/Proxy service.

- **Identity reveal requests** tend to be sent to the WHOIS-published Proxy service and/or Registrar associated with domain names allegedly involved in (or affected by) illegal or harmful Internet activities [1].

This study will consider *all* submitted relay and reveal requests, no matter why they were sent or how recipients chose to handle them. Secondary inputs from Privacy/Proxy providers and Registrars will then be used to categorize requests and provide context.

3.1 Primary Sources

Researchers are expected to identify and reach out to possible primary input sources during the first phase of this study. In particular, this study requires volunteers that are together capable of generating a sufficiently large, broad sample of relay/reveal requests which ensures that primary input data:

- Includes domains registered within the top 5 gTLDs [2],
- Includes all major Privacy/Proxy registration service providers [3],
- Includes requests made by geographically-diverse participants, and
- Includes requests that were likely generated for a wide variety of reasons.

Possible sources of primary study input are suggested below; additional volunteers are welcome – *especially sources that might supply global input*. Note that, although many possible sources investigate allegedly illegal or harmful activities, request recipients may not have done anything wrong. For example, hacked servers or botnet hosts are often used without the associated domain name Registered Name Holder's permission.

- Some domain name resellers register names with perceived market value on speculation, hoping to sell or license them to third parties. These resellers may use relay requests to market their domain name inventory to potential customers.

- More generally, many businesses use targeted email to communicate with customers and advertise products or services. These businesses may use relay requests to reach customers that have used Privacy or Proxy services to protect their identity or contact information against WHOIS data misuse.

- Organizations that maintain real-time Domain Name System Blacklists (DNSBLs) might possibly use relay requests to investigate (possibly hacked) domain names associated with spam sender IPs. Possible sources include Spamhaus Blocklist, Mailshell Live-Feed, SURBL, URIBL, and DNSBL.

- Organizations that maintain phishing website live-feeds might possibly use relay or reveal requests to investigate (possibly hacked) domain names associated with phishing URLs. Possible sources include OpenDNS, Internet Identity, and the Anti Phishing Working Group (APWG).

- Malware researchers and/or Internet security vendors might possibly use relay or reveal requests to investigate (possibly hacked) domain names associated with malware dissemination. Possible sources include SRI Malware Threat Center, FireEye Malware Analysis & Exchange, and Malware Domains.

- First responders that investigate major DoS and DNS attacks might possibly use relay or reveal requests to investigate (possibly hacked) domain names associated with attack originators or command and control centers. Potential sources include the IMPACT Global Response Centre NEWS feed and FIRST-member response teams.

- Organizations like the International Trademark Association (INTA) might be able to identify possible study participants who send relay or reveal requests about domain names cited in alleged cybersquatting incidents.

- Organizations like the UK Alliance Against IP Theft or the International Intellectual Property Rights (IPR) Advisory Program might be able to identify possible study participants who send relay or reveal requests about domain names cited in intellectual property theft complaints.

- The International Federation of the Phonographic Industry (IFPI), the Motion Picture Association of America (MPAA), the Recording Industry Association of America (RIAA), and their international counterparts might possibly send relay or reveal requests about domain names associated with servers alleged to illegally share copyrighted movies and/or music.

- Software vendors like Microsoft and Adobe or anti-piracy organizations like the Business Software Alliance (BSA), Software and Information Industry Association (SIIA) or Entertainment Software Association (ESA) might possibly send relay or reveal requests about names associated with servers alleged to illegally distribute copyrighted software or circumvent access controls on copyrighted materials.

- Members of organizations like the International Trademark Association (INTA) or commercial first-responders like Mark Monitor might possibly send relay or reveal requests about domain names alleged to infringe upon registered trademarks.

- Agencies like the International Anti-Counterfeiting Coalition (IACC) or US National Intellectual Property Rights Coordination Center Cyber Crimes Section (CCS) might possibly send relay or reveal requests about domain names associated with online sale of counterfeit merchandise and illegal pharmaceuticals.

- Legitimate job recruitment websites like Monster and HotJobs might possibly send relay or reveal requests about domain names associated with fraudulent online money laundering scams.

- Bodies that handle Internet fraud complaints such as the FBI/NWCC Internet Crime Complaint Center (IC3) might possibly send relay or reveal requests about domain names associated with advanced fee fraud email scams, such as those documented by Artists Against 419.

- Agencies like the FBI/NWCC Internet Crime Complaint Center (IC3), the National Data Protection Commissions within the EU and Canada, or the US National Intellectual Property Rights Coordination Center Identity Fraud Initiative might possibly send relay or reveal requests about domain names associated with online identity thefts.

- Agencies like the US National Intellectual Property Rights Coordination Center Cybercrimes Child Exploitation Section (CES) and Operation Predator might possibly send relay or reveal requests about domain names associated with online distribution of child pornography.

All primary input sources participating in this study must agree to record and submit all[1] relay and reveal requests sent as part of normal operation during a defined period.

[1] This requirement applies to requests generated by a single source. In it understood that a sender submitting input on behalf of many sources may not be able to secure the permission of *all* sources.

Participants shall be asked to report *all* requests initiated by a given source, no matter what their outcome. They will also be asked to supply *all* intermediate replies (including bounced emails) and final responses received from the Registered Name Holder, third party licensee, Privacy/Proxy provider, and/or Registrar. Finally, participants shall be asked to describe whether, when, and how the request was resolved (e.g., relayed message delivered, request denied, identity supplied, identity published in WHOIS, associated domain name suspended or site(s) taken down).

Some challenges that researchers may face when gathering primary input:

1. When many requests are sent about the same domain, providers may respond faster, having already completed initial investigation. Requesting secondary input for pending requests could potentially affect their outcome. Study duration must take temporal factors such as this into consideration.

2. Law enforcement agencies and first responders who generate frequent requests may have different experiences than individuals or businesses generating singular requests. Requests triggered by illegal/harmful activities also tend to be handled differently depending on type of activity. Inputs must include potentially diverse experiences.

3. Cases in which domain registrants or licensees are easily contacted may be under-reported – for example, only challenging cases may be handled by an outside counsel participating in this study. Sources should be asked about potential under-reporting.

4. Participants could also generate spurious requests, accidentally or intentionally. Collection methods must deter over-reporting – for example, by collecting time-stamped copies of requests and matching them to secondary inputs where available, or cross-referencing requests from different sources that involve the same domain.

3.2 Secondary Sources

Researchers are expected to contact secondary input sources to corroborate primary inputs where possible and better understand message flows, policies and practices, and identify factors that could potentially promote or impede request delivery and resolution.

Specifically, every Privacy or Proxy service provider and Registrar associated with a relay or reveal request sent by a primary input source shall be afforded an opportunity to supply the following background information:

- Published Terms of Service (ToS) governing Domain Name registration services, Privacy registration services, and/or Proxy registration services.

- Description of communication relay request policies and practices, including all supported forwarding methods, filtering criteria, and limits applied to those requests.

- Description of identity reveal request policies and practices, including any forms used to submit requests, any data protection or privacy laws that constrain when and how

responses are returned, and any actions routinely taken in response to specific kinds of alleged harmful or illegal Internet activity.

Additionally, for each individual request gathered from a primary input source, the associated Privacy or Proxy service provider and Registrar shall be invited to indicate whether and when the request was received. For received requests, providers and Registrars shall be invited to describe when, how, and why each request was resolved (e.g., relayed message delivered, request denied, identity supplied, identity published in WHOIS, associated domain name suspended or site(s) taken down). Where possible, participants are encouraged to accompany resolutions with explanations – for example, reason for denying request or for taking action without returning an explicit response.

Some providers may choose not to supply this information, may not keep records of this information, or may not be able to supply details for certain cases due to data protection and privacy laws. However, in cases where this secondary input can be obtained, this study will be able to shed more light on how requests are actually handled and steps that might have improved request efficiency and effectiveness.

3.3 Data Collection

An input collection and submission process must be designed to minimize participant effort while promoting accurate reporting. Researchers must develop a short, simple reporting form that all participants can use to consistently record over time and then submit input describing each relay/reveal request, outcome, and supporting documentation.

- Defining reporting requirements prior to study start is essential to ensure that all participants record necessary input data in a timely manner, including the WHOIS Registrant Name, Organization, Address, and Registrar for associated domain names when the request was generated. Note that "associated domain name" depends upon the reason for the request and may in some cases be obtained by reverse DNS lookup. To promote completeness and accuracy, data should be recorded by participants throughout the request handling process, for submission when that process ends.

- Obtaining the content of each relay/reveal request (including any attached supporting documentation) is essential for this study to examine factors that promote or impede timely communication or resolution, but care must be taken to avoid influencing that content. For example, this study should not define how participants should formulate or send requests, even though doing so might yield more consistent results. To promote completeness and accuracy, copies of requests and replies shall be supplied.

- Data collection forms and submission processes need to address reasonable confidentiality concerns. For example, primary input sources must supply the actual WHOIS record used to obtain the contact information used to address each request – including the full domain name associated with the request. However, redacted copies of actual requests and responses may be supplied to preserve the source's privacy. Similarly, secondary sources may briefly describe request resolutions without

supplying details that conflict with privacy policies or laws. Submitted data must be kept confidential; no results should be published about any individual request source or domain.

Ultimately, sources must supply enough data to meet study goals (e.g., input validation, result classification) without prohibiting collection of a sufficiently large and broad sample. In particular, some sources may not be able to participate in this study because confidentiality concerns stop them from supplying full domain names. This element has been required to enable primary/secondary input correlation and independent WHOIS data validation, but researchers are asked to comment on the utility of studying requests that do not include this mandatory data element.

3.4 Data Elements

After a sufficiently large, broad set of requests have been reported, researchers will clean, code, and classify this sample to capture the following input data for each request:

Primary Inputs to be recorded and submitted by Request Generators

- Type of Request being reported: Relay or Reveal[2]
- Reason for generating this Request
- Description of Request Source: Type and Location
- Description of Request Sender [if not Source] : Type and Location
- Associated Domain Name(s)
- Actual WHOIS record(s) used to formulate Request, including:
    - Registered Name Holder's Contact Name, Organization, Addresses
    - Apparent Privacy/Proxy service provider
    - Apparent Registrar

- Request Sent Date
- Request Method: Email, Postal, Fax, Phone
- Request Format and Size (e.g., X MBs, Y Pages)
- Request Destination (address, obtained from WHOIS)
- Copy of Request & Attachments (may be redacted to preserve source's privacy)

- Initial Reply Type & Date (e.g., none, bounced, forwarded, acknowledged)
- Initial Reply Sender: Privacy/Proxy, Registrar, Registered Name Holder, licensee
- Copy of Initial Reply & Attachments (may be redacted)

[2] Relay refers to any request to simply have communication forwarded to the registrant or licensee. Reveal refers to any request for the registrant or licensee's identity and direct contact information.

- ALL Follow-Up Communication(s) (e.g., none, status update, info requested) each including Type, Date, Sender, and Copy of Follow-up (may be redacted).

  Note: Senders may not be copied on all communication between a Privacy/Proxy service or Registrar and Registered Name Holder or licensee, so only communications actually received by the Sender can be documented here.

- Final Resolution, from Sender's perspective
  (e.g., not resolved, relayed message delivered, request denied, identity supplied, identity published in WHOIS, domain name suspended or site(s) taken down)

Secondary Inputs to be requested from affected Registrars

- Registrar's Published Terms and Price for Domain Name Registration services
- Registrar's Relay and Reveal Request handling policies and practices
- For each Request associated with a customer's Domain Name:
  - Request First-Received and Final-Resolution Dates
  - Domain Name's current Registration status
  - Description of Action(s) taken by Registrar to resolve Request
  - Final Resolution, from Registrar's perspective
    (e.g., not resolved, relayed message delivered, request denied, identity supplied, identity published in WHOIS, domain name suspended or site(s) taken down) and stated reason

Secondary Inputs to be requested from Privacy/Proxy Providers, for each Request

- Provider's Published Terms and Price for Privacy or Proxy services
- Provider's Relay and Reveal Request handling policies and practices
- For each Request associated with a customer's Domain Name:
  - Request First-Received and Final-Resolution Dates
  - Domain Name's current Privacy/Proxy service status
  - Description of Action(s) taken by Provider to resolve Request
  - Final Resolution, from Provider's perspective
    (e.g., not resolved, relayed message delivered, request denied, identity supplied, identity published in WHOIS, domain name suspended or site(s) taken down) and stated reason

Note: Some Registrars offer Privacy/Proxy services to their Domain Name Registration customers. In such cases, Registrars may supply both Privacy/Proxy and Domain Name Registration inputs, but to enable consistent handling, these should not be merged.

Additional Data Elements to be supplied by researchers

- Is the input data complete enough to enable analysis?
  If not, can missing essential details be obtained from sources?
- Request Source and Sender Country Codes
- Current WHOIS record for Domain Name, including:
  - Registered Name Holder's Contact Name, Organization, Addresses
  - Confirmed Privacy/Proxy service provider,
    classified using the methodology specified by study [3]
  - Confirmed Registrar

WHOIS data used to send a request may differ from current WHOIS data for many reasons. Privacy/Proxy service may have been terminated, Domain Name Registration service may have been terminated, the Registered Name Holder may have updated their WHOIS contact information, the Domain Name may have been sold to another unrelated Registrant, etc. Comparing old and new WHOIS data may shed light on why requests were not delivered or how requests were resolved. Researchers will therefore query WHOIS data associated with each request to:

(1) independently validate inputs, and
(2) detect cases where senders may not be aware of WHOIS changes.

The input data elements listed above are a starter list, drafted to solicit possible participant feedback on data availability and feasibility. Entity, message, and common resolution types must be refined during the first phase of this study by using a small pilot to gather example inputs from a few diverse participants. During the pilot, particular attention must be paid to confidentiality concerns and redaction, resulting in participant guidelines that will help to ensure supplied data is sufficient for meaningful analysis, and that input element definitions are sufficiently unambiguous. Guidelines should also take steps to promote collection of a uniform data set (e.g., requesting specific empirical or enumerated input elements, not just anecdotal information of varying form or narrative description).

4. Outputs

Study outputs will illustrate relay and reveal request process flows, enumerate reported outcomes, and used open-ended data analysis to look for factors that appear to promote or impede timely communication or resolution.

As an exploratory study, input data may not be sufficiently random to permit statistical analysis. Nonetheless, it may be useful to categorize study outputs in various ways, looking for possible trends and hypotheses that could warrant future study. For example:

- Inputs might be categorized by type of request (relay or reveal), reason for request, type of source and sender (e.g., individual, large business, legal counsel, first responder, law enforcement agency), type of provider (Privacy or Proxy service), and associated gTLD.

- Request content, frequency, and sender might be examined to isolate factors that could impede timely delivery or resolution. Are there any characteristics that quickly delivered or resolved requests appear to share? Do there appear to be common (possibly measurable) reasons why requests are bounced or delayed?

- Initial replies may be examined to identify trends that could reflect Registrar or Proxy/Provider policies and practices. For example, how often do initial replies appear to be automated acknowledgements or routine rejections? Are there possible patterns among requests that are quickly denied, such as location or type of sender, or inclusion of supporting documentation? Have other sources made requests pertaining to the same domain? How well do initial replies dovetail with Registrar and Proxy/Provider published policies and practices?

- Follow-up communications should be charted to isolate common message flows between affected parties, how many interactions are needed to reach resolution, and additional information often requested from senders. For example, how often are requests sent directly to Registrars? Are senders frequently copied on communication between Registrars, Privacy/Proxy providers, and other parties involved in resolution? Here, the goal is to shed light on the request handling blind spots that have made studying this process difficult.

- Finally, resolutions may be examined to characterize the most common outcomes, reasons given for them, and the range of times required to reach this state. For example, are many requests resolved without senders realizing that action has been taken? Do requests unresolved requests have any (possibly measurable) characteristics in common? How useful is the identity information provided in reveal responses?

The above list is provided as a starting point for further discussion. Researchers are asked to propose feasible data analysis and verification methods to generate useful findings that can help the ICANN community advance related policy debates.

This study may well conclude that some aspects of request handling cannot be analyzed in a statistically-meaningful way. However, it is hoped that this study will be able to identify factors that can be measured and hypotheses that can be tested by future studies.

5. References

[1] Working Definitions for Key Terms that May be Used in Future WHOIS Studies, GNSO Drafting Team, 18 February 2009

[2] Proposed Design for a Study of the Accuracy of Whois Registrant Contact Information (6558,6636), NORC, June 3, 2009

[3] ICANN's Study on the Prevalence of Domain Names Registered using a Privacy or Proxy Service among the top 5 gTLDs, ICANN, September 28, 2009

[4] Registrar Accreditation Agreement (RAA), ICANN, 21 May 2009

[5] Terms of Reference for WHOIS Misuse Studies, ICANN, September 2009

[6] Terms of Reference for WHOIS Registrant Identification Studies, ICANN, Oct 2009

[7] Terms of Reference for WHOIS Privacy/Proxy Abuse Studies, ICANN, May 2010

[8] Study Suggestion Number Study 3a/b, Analysis of compliance by registrars operating proxy services, Steve Del Bianco

[9] Study Suggestion Number 13b/c, Measure growth of proxy/privacy services vis-à-vis all registrations, Laura Mather

[10] Study Suggestion Number Study 20, Timeliness of proxy services in relaying communications to registrants, Claudio DiGangi

[11] GAC Data Set 1, To what extent are legitimate uses of WHOIS data curtailed by use of proxy or privacy services, GAC Recommendations for WHOIS Studies, 16 April 2008

[12] Re: Circumvention of Registrar Accreditation Agreement Section 3.7.7.3, Intellectual Property Constituency (IPC), April 2009

[13] DRAFT Advisory re: RAA Subsection 3.7.7.3, ICANN, 14 May 2010