The ICANN GNSO "Business Constituency"

Comment on the WHOIS Accuracy Pilot Study Report

Status: FINAL

Version: 2

27-Feb-2015

Business Constituency Submission

GNSO//CSG//BC

Background

This document is the response of the ICANN Business Constituency (BC), from the perspective of business users and registrants, as defined in our Charter:

> The mission of the Business Constituency is to ensure that ICANN policy positions are consistent with the development of an Internet that:
> 1. promotes end-user confidence because it is a safe place to conduct business
> 2. is competitive in the supply of registry and registrar and related services
> 3. is technically stable, secure and reliable.

The BC appreciates the opportunity to provide its input to this effort.   The BC recognizes the amount of effort across multiple parties required to complete the pilot study, and sees it as an important initial step toward complete implementation of the WHOIS Accuracy Reporting System (ARS).

Below are BC reactions to the 5 aspects of the report on which ICANN is seeking feedback.

1. Sample Design & Methodology

- The design appropriately accounts for the most important contact methods in domain registration records. It would be helpful for FAX, and its elimination from the validation set, to be addressed.

- The ARS should address accuracy of all contact addresses, not just the registrant's address.

- We would appreciate detail on treatment of Admin, Billing, Tech, and Registrant contacts, if they are treated in different ways.

- The pilot appears to have resulted in useful information that will be incorporated into future WHOIS accuracy testing and ARS. The report does not, however, include detail of issues that are likely to arise with larger volumes of records.

- It may be useful for the design to include the validator's Testing Server policy for email address testing.

- Understanding resources anticipated to be required for ARS will be helpful in setting timeline expectations.

- It is not clear why the study focuses on new gTLDs to the exclusion of legacy gTLDs (.com;.net;.org) where inaccurate data is known to be an issue.

2. Types of Accuracy Reports to be published through the ARS

- The delineation of syntactic accuracy, operational accuracy, and registrant validation designed for the reports are appropriate and helpful. Likewise, the degrees of validity per contact method appear to be appropriate.

- Because the pilot did not include identity confirmation, it may be useful to organize a pilot for just this scope, if it is determined that this will be included in the scope of the ARS.

- The ARS should include a scorecard of registrar responsiveness, records corrected, and tracking of challenges to assuring WHOIS accuracy.

3. Should ICANN conduct Identity Validation in subsequent phases of the ARS Development?

- Ultimately, identity validation should be included in the ARS.

- In the meantime, the BC believes that the task and costs of acquiring Identity Validation should be quantified and shared, and that the level of assurance should be quantified and shared, so that the community is able to make a qualified judgment about moving forward with it. Also, we prefer to see these data as the result of a pilot, prior to its incorporation in ARS development.

- Performing the Syntactical evaluation and the Operational Evaluation without the Identity evaluation is leaving out a major component as to the relevancy of the data. This study does not look at fraudulent use of a third party's information. The data if examined would appear to be accurate as it is authentic and even validated contact information, but it is not information that would allow you to contact the individual that registered the domain name. This contravenes the objective of WHOIS. We urge the study of including identity confirmation. Fraudulent use of contact data or alleged identity theft should be treated differently when reported through the WHOIS ARS. It does not make sense to ask a registrant who has intentionally provided contact information fraudulently, if the information is correct.

- The core objective of this effort is to improve the WHOIS data so that responsible parties can be contacted. Therefore, when inaccuracies are found, even in this pilot study phase, a process should be established to forward those inaccuracies to registrars for appropriate corrective action.

4. Should the methodology treat registrations under privacy or proxy services differently, and if so, how?

- The email address methodology should be different for privacy or proxy services, since previous studies have shown that these registrations comprise at least 25% of the domain name registrations – and growing. The work of the PPSAI has shown that many of the email addresses provided by a privacy or proxy service are technically operational, but do not serve the purpose of WHOIS because emails may not be delivered to the actual registrant and operator of a domain name.   Many privacy and proxy vendors allow their customer to choose not to receive any email sent to the proxy privacy specified email address for their domain name. The email address may be shown to be operational but not relevant. To avoid overstating the quality of the WHOIS data for meeting the core objective of WHOIS, these emails should be treated separately in any accuracy study.

5. Any other aspect of the ARS

- Because WHOIS pertains to domain names, IP addresses, and autonomous system numbers, a more appropriate formal name for WHOIS Accuracy should include the words "domain name".

- Should list safeguards included in the policy, setup, testing, and reporting. For example, with testing, larger validation data sets increase the possibility of validators being blocked at sites like gmail.com, yahoo.com, and MSN, among others.

- The report does not include details about how recipient validity is detected for domains that do not give a definitive response through an SMTP conversation (e.g., Yahoo, Hotmail, AOL, etc.).

- Q&A check with multiple requests using known invalid addresses.

- According to the report, a record "received a scoring of "Full Failure (-2)" if the username or mailbox was rejected as a valid email address, the email was invalid according to RFC specifications, and the email was identified as a spam trap." In the future, having a breakdown of the occurrence of these individual conditions may produce useful information.

- Presently, an email address "Full Failure (-2)" may result from other conditions not mentioned in the report, such as a server temporarily down or server timeouts.

- If email is intended to be used in the future as part of identity validation methods, a defined policy and method as it relates to WHOIS validation will be helpful in guarding against false negatives. For example, these methods might be useful toward guaranteeing the most accurate results:

    o testing via DMARC (Domain-based Message Authentication, Reporting & Conformance);

    o testing via VBR (Vouch by Reference);

    o use of IANA Email Authentication Parameters and SPF and DKIM records when sending;

    o accounting for milters/MTA's, receivers and relays using iprev

- A documented process should exist for processing rejected or bounced mail from servers running SPF with HELO (or EHLO) identity checks, and rejected or bounced mail from servers running SenderID or other controls in the receiver's Administrative Management Domain (ADMD).

--

This comment was drafted by Angie Graves, with input from J Scott Evans, Susan Kawaguchi, and Ellen Blackler.

It was approved in accordance with the BC charter.