

# Data Analysis

[insert general introduction that covers both consumer trust/safeguards and competition issues; see additional material on consumer trust safeguards issues below]

At various points in this report, we identify analyses that we were unable to conduct because we lacked the needed information. Some of these shortcomings can be overcome in the future if ICANN obtains these data directly from industry participants or if ICANN enters into contractual relationships with parties that collect the data. Others will require improved analyses of the behavior of industry participants, especially analyses that enhance our understanding of the way in which registrants substitute among TLDs. This section discusses these issues in somewhat greater detail. In addition, we believe that ICANN can make better use of publicly available data and that it should develop the capability to analyze both proprietary and public data on an ongoing basis.

The most significant data limitation that we faced was the almost total lack of information about the wholesale prices charged by legacy TLDs. Analysis Group requested wholesale price data directly from both legacy and new registries as part of its study with the understanding that the data would never be provided to ICANN or made public. In addition, Analysis Group provided assurances that the data published in its report would be aggregated and anonymized so as not to compromise confidentiality. Although the Analysis Group obtained some data in response to its request, there were extremely few responses from legacy gTLDs and incomplete data from new gTLDs. We believe that ICANN should acquire this information from all registries on a regular basis and provide assurances that the data would be treated on a confidential basis. The data could then be used for analytic purposes by ICANN staff and by others that execute non-disclosure agreements.

In light of the very high parking rates observed for some gTLDs discussed above, we believe that it is important for ICANN to track this information on a regular basis. Although ntlidstats provides this information on an ongoing basis for new gTLDs, ICANN has had to enter into a contract with them to obtain similar information for legacy gTLDs.

**Commented [1]:** This feels more like an introduction to the competition and choice sections, as it covers limitations related to those analyses.

**Commented [KL2]:** LK: I recommend revision of this paper to either add consumer trust/safeguards issues [see my high-level additions below] or make it a paper that solely relates to Competition and Consumer Choice. If we make this a joint paper, a brief intro should be added that covers both topics. For either approach, I would still advise moving discussions of specific data recommendations on competition and choice issues to those corresponding papers. I view the best use of this paper as a general recommendation to ICANN to be more systematic about how it collects data for the benefit of this and other review teams and a high level summary of where our review team identified data needs. Specific recommendations would be better placed along with their pertinent papers.

**Commented [3]:** Margie's comment: this is a recommendation - what is the purpose/goal for this study? Is it a high/low priority? Is it needed for next round? HIGH PRIORITY IN MY VIEW.

We recommend that ICANN arrange to obtain this information on an ongoing basis in the future.

A third limitation involved our inability to conduct analyses on a regional or country basis. During the course of our work, we learned some of the data that we would need to conduct this analysis had been compiled in the connection with the Latin American and Caribbean DNS Marketplace Study and we are attempting to obtain those data in order to conduct country-specific analysis for that set of countries. We recommend that ICANN collect this information for all countries on an ongoing basis in the future. In this regard, it is important to note that the country-specific analysis would be able to assess the extent to which gTLDs and ccTLDs compete. Some of these data may already be collected, for example by CENTR, and we recommend that ICANN explore the possibility of obtaining the needed data from these sources.

Fourth, it appears that ICANN does not currently make use of retail price data that can be obtained directly from public sources such as <https://tld-list.com/> and <https://namestat.org>. We recommend that ICANN develop the capability of analyzing these data on an ongoing basis.

ICANN may also wish to explore the possibility of obtaining data on prices that prevail in secondary market transactions.

Finally, we note that our ability to define relevant markets has been severely handicapped by the lack of information about how registrants make choices among

TLDs. In this regard, it would have been useful to have answers to the following questions:

- (1) What proportion of the registrants in the new gTLDs were previously registrants in a legacy gTLD but gave up their registrations when they registered in a new gTLD? This will provide some indication of the importance of *switching costs*.
- (2) What proportion of the registrants in the new gTLDs had not previously been registrants in any gTLD? This will provide some indication of the extent to which the introduction of new gTLDs *expanded* the number of individual registrants.
- (3) What proportion of the registrants in the new gTLDs are entities that continued to have registrations in legacy gTLDs? This will provide some indication of whether registrations in legacy and new gTLDs are *complements* as opposed to *substitutes*.

**Commented [4]:** Margie's comment: same questions as above. HIGH PRIORITY IN MY VIEW, ALTHOUGH WE DO NOT DISCUSS THIS ISSUE IN THE COMPETITION DRAFT.

(4) What proportion of the registrants in the new gTLDs registered primarily: (a) for *defensive* reasons, i.e., they felt compelled to register in a new gTLD because they existed but obtained no benefits from doing so and what proportion registered primarily (b) for the *benefits* that they received, perhaps because doing so permitted them to reach users that would have otherwise been inaccessible? This will provide some indication of whether, on balance, the introduction of new gTLDs resulted in net costs or net benefits to registrants.

(5) What are the characteristics of the new gTLDs that attracted registrants primarily because of the benefits that they offered? This will provide some indication of the *sources* of the benefits that the new gTLDs provided, e.g., new allowable characters, service to a specific community, higher levels of security or customer service, ability to offer domain names to non-competing entities.

In order to begin to answer these questions, we recommend that ICANN conduct a survey of new registrants. Among the questions that should be included in the survey are:

1. Did you register a new domain name in the last N months?
2. For each name that you registered, did you register it in a new gTLD or in a legacy gTLD? [Define terms]
3. For each name that you registered in a new gTLD [Check one]
  - a. Was the registration a newly registered name?
  - b. Did the registration replace a registration in a legacy gTLD?
  - c. Did the registration duplicate a registration in a legacy gTLD?
4. For each name that you registered in a new gTLD, was the closest alternative that you considered another gTLD or a legacy gTLD? What was the identity of that gTLD?
5. For each name that you registered in a legacy gTLD, did you consider registering in a new gTLD as an alternative?
6. For each name that duplicated a registration in a legacy gTLD, was the registration intended *primarily* to prevent the name from being used by another registrant?
7. For each name that you registered, indicate whether it is currently parked. [Define parking]

Data Issues Related to Consumer Trust/Safeguards Issues

The review team also faced challenges related to its assessment of the extent to which the expansion of gTLDs promoted consumer trust and the effectiveness of safeguards adopted by new TLDs in mitigating certain risks involved in such expansion.

Two surveys were made available that contained data regarding the extent to which consumer end-users and registrants trusted new gTLDs. However, the review team noted that the surveys did not define consumer trust (and other key terms) and contained few questions that explored the objective behavior of the survey respondents that could serve as a proxy for consumer trust. Moreover, certain responses that identified factors relevant to consumer trust -- such as reputation and familiarity -- were broad concepts that did not lend themselves to providing precise guidance for either future applicants, ICANN, or other community stakeholders. As a result, we would recommend that future review teams consult with survey experts to explore how best to measure consumer trust in a manner that gathers both objective and subjective information, with a goal toward generating more concrete and actionable information.

The review team also lacked sufficient data on how effective safeguards adopted by gTLDs were in mitigating certain risks. For example, although many safeguards for new gTLDs aimed at mitigating DNS abuse, little information was made available to the review team that directly addressed this issue. In response, the review team commissioned a study to establish baseline measures of abuse rates in new and legacy gTLDs that will enable further inquiry into the effectiveness of these safeguards. We hope that future review teams will build on this study and consider how additional studies may shed further light on assessing the effectiveness of new gTLD safeguards both for individual

TLDs and the Domain Name System as a whole. An important and related issue is information about the costs of implementing these safeguards. The review team lacked data regarding the costs to registries and registrars of implementing the safeguards required under the new gTLD program. Such data would be useful to future review teams who may wish to engage in a cost/benefit analysis.

Another challenge faced by the review team was a lack of transparency in the subject matter of complaints submitted to ICANN compliance. Although ICANN makes available information about the general subject matters of the complaints that it receives, such as WHOIS accuracy or DNS abuse, ICANN does not disclose more

specific information about the subject matter of these complaints. For example, regarding complaints about registrars, ICANN compliance reports do not disclose what type of WHOIS accuracy is being complained about (address, email, or identity verification). Similarly, ICANN compliance reports do not identify what types of DNS abuse are the subjects of complaints. Such information would permit review teams to identify more precisely which subject areas generate the most complaints and would enable a better assessment of the effectiveness of current safeguards.

## Recommendations

**Recommendation:** Create a Data Analyst position on the ICANN staff

**Rationale/Related findings:** The lack of data has handicapped attempts both internally and externally to evaluate market trends and the success of policy recommendations.

**To:** ICANN Board

**Must be completed prior to subsequent rounds:** Yes

**Consensus within team:** Yes

**Details:** ICANN should establish a position for a dedicated data scientist to facilitate quantitative analysis, by staff, contractors and the community, of the domain name market and, where possible, the outcomes of policy implementation. This Data Analyst should be directed and empowered to identify and either collect or acquire datasets relevant to the objectives set out in strategic plans, and analysis and recommendations coming from review teams and working groups.

**Success Measures:** The availability of relevant data for use by ICANN staff, contractors and the ICANN community for its work in continuous improvement of ICANN operations.

**Recommendation:** Collection of wholesale pricing for domain marketplace

**Rationale/Related findings:** The lack of wholesale data will continue to frustrate future CCT review teams efforts to analyze competition between registries and TLDs in the domain marketplace.

**To:** ICANN Board

**Must be completed prior to subsequent rounds:** Yes

**Consensus within team:** Yes

**Details:** ICANN, or an outside contractor, should acquire wholesale price information from all registries on a regular basis and provide necessary assurances that the data would be treated on a confidential basis. The data could then be used for analytic purposes by ICANN staff and by others that execute non-disclosure agreements. This may require amendment to the Base Registry Agreement.

**Success Measures:** The availability of relevant data for use by ICANN staff, contractors and the ICANN community for its work in evaluating competition in the DNS space.

**Recommendation:** Collection of retail pricing for domain marketplace

**Rationale/Related findings:** The lack of retail data will continue to frustrate future CCT review teams efforts to analyze competition between registries and TLDs in the domain marketplace.

**To:** ICANN Board

**Must be completed prior to subsequent rounds:** Yes

**Consensus within team:** Yes

**Details:** ICANN does not currently make use of retail price data that can be obtained directly from public sources such as <https://tld-list.com/> and <https://namestat.org>. We recommend that ICANN develop the capability of analyzing these data on an ongoing basis. Alternatively, an amendment to the Registrar Accreditation Agreement to ensure the availability of this data with all due diligence to protect competitive information.

**Success Measures:** The availability of relevant data for use by ICANN staff, contractors and the ICANN community for its work in evaluating competition in the DNS space.

**Recommendation:** Collection of parking data

**Rationale/Related findings:** The high incidence of parked domains suggests an impact on the competitive landscape but insufficient data frustrates efforts to analyze this impact.

**To:** ICANN Board

**Must be completed prior to subsequent rounds:** Yes

**Consensus within team:** Yes

**Details:** ICANN should regularly track proportion of TLDs which are parked with sufficient granularity to identify trends on a regional and global basis.

**Success Measures:** The availability of relevant data for use by ICANN staff, contractors and the ICANN community for its work in evaluating competition in the DNS space.

**Recommendation:** Collection TLD sales at a country by country level

**Rationale/Related findings:** The lack of country level data will continue to frustrate future CCT review teams efforts to analyze competition between registries and TLDs in the domain marketplace. In particular, lack of country specific data frustrates efforts to understand the competition between gTLDs and ccTLDs.

**To:** ICANN Board

**Must be completed prior to subsequent rounds:** Yes

**Consensus within team:** Yes

**Details:** Some of this data is collected by 3<sup>rd</sup> parties such as CENTR so it is possible that ICANN can arrange to acquire it.

**Success Measures:** The availability of relevant data for use by ICANN staff, contractors and the ICANN community for its work in evaluating competition in the DNS space.

**Recommendation:** Periodic Surveys of Registrants

**Rationale/Related findings:** The inability to determine registrant motivations and behavior frustrates efforts to study competition and choice in the TLD marketplace.

**To:** ICANN Board

**Must be completed prior to subsequent rounds:** Yes

**Consensus within team:** Yes

**Details:** The survey should be designed and continuously improved to collect registrant trends. Some initial thoughts on potential questions is in appendix X.

**Success Measures:** The availability of relevant data for use by ICANN staff, contractors and the ICANN community for its work in evaluating competition in the DNS space.

**Recommendation:** Periodic Surveys of Consumers

**Rationale/Related findings:** While there are some proxies for consumer trust which can be measured by analyzing behavior, it's still imperative to survey consumers to gain an understanding of consumer motivations, apprehensions and the effectiveness of ICANN policy on consumer awareness.

**To:** ICANN Board

**Must be completed prior to subsequent rounds:** Yes

**Consensus within team:** Yes

**Details:** While two surveys were conducted, greater effort at specificity in question is necessary for valid analysis of consumer trends.

**Success Measures:** The availability of relevant data for use by ICANN staff, contractors and the ICANN community for its work in evaluating consumer trust in the DNS space.

**Recommendation:** Ongoing tracking of DNS Abuse

**Rationale/Related findings:** In order to continuously evaluate the effectiveness of safeguards, it is important to understand DNS abuse trends over time.

**To:** ICANN Board

**Must be completed prior to subsequent rounds:** Yes

**Consensus within team:** Yes

**Details:** While the CCTRT commissioned a backwards looking study of DNS abuse, it will be constrained by time and available data. Ongoing analysis of DNS abuse is necessary. .

**Success Measures:** The availability of relevant data for use by ICANN staff, contractors and the ICANN community for its work in evaluating consumer trust in, and the trustworthiness of, the DNS space.

**Recommendation:** Track Costs of Safeguard Implementation

**Rationale/Related findings:** For future CCT review teams to evaluate the costs benefit of various safeguards, fairly granular data on safeguard implementation costs from registries and registrars will be necessary.

**To:** ICANN Board

**Must be completed prior to subsequent rounds:** Yes

**Consensus within team:** Yes

**Details:** ICANN should regularly survey registries and registrars on the cost of safeguard implementation. Alternatively, this may require amendments to the Base Registry Agreement and the Registrar Accreditation Agreement.

**Success Measures:** The availability of relevant data for use by ICANN staff, contractors and the ICANN community for its work in evaluating consumer trust in, and the trustworthiness of, the DNS space.

**Recommendation:** More granular data collection by ICANN compliance

**Rationale/Related findings:** Insufficiently granular data from ICANN compliance frustrates analysis in a number of areas including DNS abuse and the effectiveness of safeguards.

**To:** ICANN Board

**Must be completed prior to subsequent rounds:** Yes

**Consensus within team:** Yes

**Details:** The CCTRT will work with ICANN staff to better understand the types of fields necessary for capture by the compliance team to enable better analysis of complaints and the effectiveness of organizational reform.

**Success Measures:** The availability of relevant data for use by ICANN staff, contractors and the ICANN community for its work in evaluating consumer trust in, and the trustworthiness of, the DNS space.