COALITION FOR ONLINE ACCOUNTABILITY
WWW.ONLINEACCOUNTABILITY.NET

C/O MITCHELL SILBERBERG & KNUPP LLP ● 1818 N STREET N.W., 8TH FLOOR ● WASHINGTON, D.C. 20036-2406
TEL: (202) 355-7900 ● FAX: (202) 355-7899 ● E-MAIL: INFO@ONLINEACCOUNTABILITY.NET

August 8, 2013

The Coalition for Online Accountability (COA) appreciates this opportunity to comment on the Initial Report of the Expert Working Group on gTLD Directory Services (EWG). See http://www.icann.org/en/groups/other/gtld-directory-services/initial-report-24jun13-en.pdf.[1]

COA consists of eight leading copyright industry companies, trade associations and member organizations of copyright owners (listed below). COA and its participants have engaged actively in many aspects of ICANN's work since the inception of the organization, including through the Intellectual Property Constituency of the GNSO. COA's stated goal is to enhance online transparency and accountability by working to ensure that domain name Whois databases remain publicly accessible, accurate and reliable, as key tools against online infringement of copyright, as well as to combat trademark infringement, cybersquatting, and other forms of fraudulent or criminal misconduct. See http://www.onlineaccountability.net/. Accordingly, the EWG Initial Report is of particular interest to us.

COA commends the EWG on its work and supports in principle its proposals for a new Registration Directory Service (RDS) model. Its recommendations could provide a pathway to a more accessible, accurate, and usable RDS for the gTLD space than the Whois model that exists now. However, the Initial Report leaves many important issues unresolved, and in some cases its recommendations are more restrictive of access than is warranted. COA welcomes the chance to provide input at this stage, but reserves its ultimate position until a more detailed and revised proposal is available.

I. Comments on Selected Main Features

While it is not feasible for COA to comment on all aspects of the Initial Report, we offer the following observations on three of the main features of the proposal.

[1] COA submits these comments despite a public comment announcement that does not appear designed to encourage any comments other than responses to pre-set survey questions. Unlike any other ICANN public comment announcements that COA has reviewed over the years, the public comment page (http://www.icann.org/en/groups/other/gtld-directory-services/share-24jun13-en.htm) includes neither a link to the document on which comment is sought, nor the e-mail address to which comments (other than survey responses) should be sent.

A. Aggregated Registration Data System (ARDS) Model

COA strongly supports the proposal for an aggregated (although not necessarily physically centralized) architecture for the new RDS. This architecture enables or greatly facilitates a number of vital positive features. For instance, it is difficult to see how the "gated access" model, with the need to authenticate requesters, issue credentials, and audit uses, could feasibly be implemented in a consistent manner unless something like the ARDS model is employed. Similarly, the provision of important services such as registration history (or "WhoWas"), and the ability to search by registrant or other data elements (sometime referred to as "reverse Whois"), appears much more feasible in an environment in which there is a single point of access to all RDS data.[2] Furthermore, it makes sense that the operation of the ARDS service be combined with other key functions of the model, including not only credentialing and auditing of requesters, but also the standard validation service contemplated by section 4.10 of the Initial Report, as well as some of the data validation and revalidation functions laid out in section 4.9. COA agrees with EWG that "the current distributed system introduces inefficiencies and additional costs," and notes that this problem is likely to worsen with the impending addition of many hundreds of new gTLD registries, even though all these will be operating in a thick Whois structure. Having reviewed the discussion in section V of the Initial Report, COA agrees that the advantages of the ARDS structure outweigh the possible disadvantages.

B. Enhanced Access for Authenticated Requesters

Although the structure proposed in the Initial Report differs from the unfettered public access model that has been a feature of Whois since its inception, and that COA and many others have striven to defend over the years against numerous proposals to reduce public access, COA is prepared to support, at least in principle, the new "gated access" model proposed. However, this support is conditional on satisfactory resolution of a number of issues in the proposal. These include the following:

1. Better definition of "purposes". One of the main ways that COA participants currently use Whois data is to investigate (and hopefully to identify, locate and enable contact with) the entity responsible for a web site or similar resource associated with a domain name on which activity that infringes copyright or trademark is occurring. This includes websites offering services that constitute or facilitate copyright piracy, that offer counterfeit goods or services, or that are associated with domain names confusingly similar to trademarks of COA participants (or their member companies). Of the ten purposes listed on page 3 of the Initial Report, and that are associated with "example use cases" in Table 1 on pages 12-14, the one that appears to overlap most closely with this use is called "legal actions."[3] However, the investigative use summarized above also seems to overlap to some extent with other categories, notably "abuse mitigation." Furthermore, COA participants (or their member companies) also make other uses of Whois data today (and would make other uses of registration data in the

[2] As noted below, COA has some concerns about possible impediments to accessing such services under the Initial Report's proposal, but regardless about how that issue is resolved, an aggregated architecture will facilitate provision of these and similar services to whoever is qualified to access them.

[3] If this "legal actions" label is ultimately used to describe this "purpose," it should be made clear it is not limited to requests associated with a formal lawsuit or similar proceeding. In the vast majority of cases, incidents of the kind of online misconduct with which COA participants are most concerned are resolved without the need for formal litigation.

future), and some of these uses more closely resemble listed "purposes" such as "business domain name purchase or sale," or "domain name research." Clearly it will be necessary to have much clearer definitions of the various permissible purposes in order to implement any system of "gated access," in which a requester's ability to access specific data elements depends on an accurate (and ultimately an auditable) representation of the specific "purpose" for which access to non-public registration data is requested.

2. Better alignment of data elements with "purposes." Proceeding on the assumption that the core use of registration data summarized above fits best under the label of "legal actions," COA participants reviewed in detail Annex C of the Initial Report, which seeks to align permissible purposes with specific data elements, and concluded that many of the data elements not associated with "legal actions" in Annex C could in fact be quite important in carrying out this use.[4] For example:

- Original registration date/creation date: historical information such as this can be critical in determining whether a domain name was registered before or after a trademark was registered, and also (in conjunction with a "WhoWas" service) can sometimes provide leads to contact information that was submitted before a registration was taken behind a proxy curtain.
- Client status/Server status: This information is important to COA participants for their investigations, as it may indicate whether there is a pending UDRP case involving the domain name, or whether the registrant is seeking to transfer the registration to a registrar that is less cooperative with third party requestors or more tolerant of incomplete or inaccurate Whois data. Furthermore, since this information is generally currently available today via Whois, and since it implicates no personal privacy considerations, COA is not aware of any justification for removing it from public access, let alone access by authenticated requestors.
- Expiration date: This information can be critical to making tactical decisions about how best to proceed with an investigation. For instance, a trademark owner who learns that an infringing domain name is on the brink of expiration may choose to claim it upon expiration, rather than focusing on tracking down the current registrant.
- DNS servers: This information is crucial to the extent that it provides IP addresses that can be associated with those of other suspect sites whose registrants are already known, and also is invaluable in geographically locating the servers involved in infringing activity (in some cases, this determines whether or not a particular requester even has jurisdiction to investigate further). As with client/server status, this data element is listed as currently available to the public via Whois[5], and no justification appears in the

[4] In some cases we were unclear what the data element actually meant, particularly for those that do not currently appear in Whois results. It is critical for any new RDS model that data elements be clearly defined and that those definitions are readily available to registrants, registrars, registries and the general public. This is certainly not the case currently with Whois.

[5] We assume that the "DNS Servers" data element is meant to be the same as the "Name Servers" data element in today's Whois. This is another case in which a clear definition would be useful.

EWG report for suppressing this public access, let alone access by authenticated requesters.

While there might be some data elements listed in Annex C that are of limited importance for "legal actions" requesters[6], COA suggests that the rule of thumb ought to be that once such requesters are authenticated, credentialed and made subject to audit, they ought to be able to access the full range of registration data for such requests, absent a compelling justification for suppressing their access.

3. "Premium" services: Although at the time it came under ICANN's stewardship Whois included the ability to search for all registrations made by a particular registrant, ICANN never insisted that this functionality be preserved, and today it is generally available only from private sector vendors whose compilation of Whois records may be incomplete. As noted above, the ability of an ARDS architecture to facilitate the restoration of such "reverse Whois" capabilities, as well as other services such as "WhoWas" historical records of past registration data, is a positive feature of the Initial Report's proposal. However, COA questions the statement in section 4.8.5 of the Initial Report that such so-called "premium data access services" should be "subject to some type of accreditation regime." We await further explanation about why these services should not be available to all authenticated requesters, at least in the "Legal Actions" arena, and what sort of additional "accreditation regime" the EWG considers necessary.

4. Authentication process: The Initial Report does not provide any details about how a requester would become qualified as a "legal actions" requester (or whatever other "purpose(s)" are most relevant to the uses of registration data of most importance to COA participants). Whatever process is put into place needs to be expeditious, functional, persistent (i.e., credentials would be valid for a reasonable time period before any re-qualification were required), and applicable not only to rights owners themselves, but also to associations of rights owners and authorized third parties of such rights owners or associations. While in some sense these are implementation questions, how they are resolved could spell the difference between whether the new RDS model is acceptable or unacceptable for rights owners.

5. Appropriate level for public access: While COA accepts the concept that authenticated requestors could receive a broader level of access than members of the general public, the EWG needs to bear in mind that the category of "individual Internet user" could embrace people that have most of the other "purposes" listed when they seek access to registration data. In particular, as individuals today use Whois in order to get a better idea of who they (or their children) are dealing with online, their ability to continue to do so must be preserved under the new model. Of course, many copyright owners are in fact "individual Internet users," and their ability to investigate online infringements of their rights through access to relevant registration data must be safeguarded. Annex C recommends that "individual Internet users" be denied access to at least half a dozen data elements that they can readily access today via Whois. To the extent that the transition from Whois to the new RDS model is viewed as a mechanism for reducing public access to data that has been publicly accessible since the birth of the DNS, ICANN will face a significant burden to explain why such suppression of public access is justified. The Initial Report does not (and understandably does not fully attempt to) shoulder this burden, but such an explanation will be essential for public acceptance of any new model.

[6] Examples could include Registration Agreement Language or EPP Transfer Key.

C. Data Accuracy Improvements

The new RDS model will only succeed if it is perceived as curing the main problem with Whois that the Initial Report accurately characterizes: "giving every user the same anonymous public access to (too often inaccurate) gTLD registration data." Initial Report at 3. If the new model is to significantly restrict anonymous public access, then it must in return deliver significant data accuracy improvements. Those improvements must be measured against the baseline that exists today (or that will be in force for much of the gTLD space as early as January 1, 2014). In other words, the baseline includes –

- the provisions of the 2013 version of the Registrar Accreditation Agreement, which include new obligations for registrars to verify and re-verify certain Whois data[7];
- implementation of the recommendations of the Whois Policy Review Team, including for example an ICANN-provided tool for flagging potentially inaccurate Whois data and funneling registrations associated with this data into compliance processes[8]; and
- (for new gTLDs) the implementation of the GAC Safeguards Advice with respect to Whois accuracy.[9]

To reiterate, all of these data accuracy improvements are already coming into force in the Whois environment of unrestricted public access to much registration data. The new RDS model must add to these mechanisms if the proposal to move from public access to more restricted "gated access" is to be credible.

Sections 4.9 and 4.10 of the Initial Report provide a useful list of some of the additional improvements needed. Although not all of these are really "plus" elements to the existing baseline, some important ones are, such as the proposed requirement (in section 4.9.7) that "operational validation methods should not rely exclusively upon a single contact method" (in contrast to the RAA 2013 provision that allows registrars to comply with the validation requirements by confirming solely the e-mail address OR the phone number of the registrant).[10]

However, further "plus" elements are needed. One example might be a requirement that action in response to evidently inaccurate data flagged by the standard validation service in section 4.10, or by the registrar itself in compliance with the requirements of section 4.9, must proceed significantly faster than with respect to third party Whois inaccuracy complaints under the 2013 RAA. Since presumably there is much less risk of abuse if an inaccuracy is discovered by the registrar itself, or by the accredited standard validation service, it should not be acceptable

[7] See http://www.icann.org/en/resources/registrars/raa/approved-with-specs-27jun13-en.htm , especially Whois Accuracy Program Specification.

[8] See http://www.icann.org/en/groups/board/documents/briefing-materials-1-08nov12-en.pdf (Appendix to Board paper on "WHOIS Policy Review Team Report Recommendations), as referenced in Board Resolution 2012.11.08.02.

[9] See http://www.icann.org/en/groups/board/documents/resolutions-new-gtld-annex-i-agenda-2b-25jun13-en.pdf, section 1, approved by Board Resolution 2013.06.25.NG02.

[10] See http://www.icann.org/en/resources/registrars/raa/approved-with-specs-27jun13-en.htm, Whois Accuracy Program Specification, paragraph 1(f).

for an active registration to remain in the system with inaccurate and unvalidatable data for longer than 3-5 days.

COA also applauds the requirement in section 4.9.9 that the date of validation of each data element should be recorded in the system. The method of validation should also be recorded through a code that would help explain how the validating authority (registrar, registry or standard validation service provider) confirmed the validity of the data. COA urges that both the date and the method of validation be included with the data made available either to the general public, or to credentialed requestors, depending on the data element in question. In other words, this meta-data regarding validation would become part of the RDS record and would be accessible on the same terms as the data element to which it refers. Certainly with respect to those data elements made available to authenticated "legal actions" requesters, it would be very useful to have this indication, both of the freshness of the data, and of the degree to which it has been validated through a method that best ensures its reliability.

II. Major Unresolved Issues

COA offers the following observations on several critical issues that are flagged in the Initial Report, but for which no specific proposal is presented. Satisfactory resolution of these issues will be essential to the successful adoption of a new RDS model.

A. Proxy services

The Initial Report proposes that "enhanced protected registration" be available to all gTLD registrants without distinction as to type of registrant or type of use of the domain name. At one level, this is the status quo: today, anyone may make a proxy registration in a gTLD regardless of the purpose for which the domain name will be used, and between one-fifth and one-quarter of all registrants have availed themselves of this option, effectively removing their contact data from the publicly accessible Whois. If the new RDS model does nothing to solve this problem, then it will be a complete failure.

The solution to this problem has at least two dimensions: who can use a proxy service, and what is the significance of using it? COA believes that the answer to the second question is probably more important than the first. While there is some appeal to the idea that proxy registration would be open only to individuals, and/or only to those making non-commercial uses of the domain name, there are huge practical problems to implementing either requirement. Some of the complexities are discussed in the most recent Whois study commissioned by ICANN, the *Whois Registrant Identification Study*.[11] In particular, the problem of filtering commercial from non-commercial registrants at the time of registration would appear almost insurmountable, for at that moment no use whatever has been made of the domain name in question. It is also hard to envision a practical means for ICANN (or an agent of ICANN) to determine whether a registrant is or is not an individual, and to adjudicate whether or not specific uses of the domain name are commercial.

It seems preferable to focus on the second question: what are the ground rules regarding proxy registration, regardless of who makes one? The answer, in COA's view, must address three elements:

[11] See http://gnso.icann.org/en/issues/whois/registrant-identification-summary-23may13-en.pdf .

- Standards for "reveal." Because proxy services in effect draw a veil of anonymity across critical registration data elements, the most important task is to clearly define when that veil can be pierced and by whom. These rules must be clear and predictable, and must accommodate the needs of the various requesters who will seek to pierce the veil, including but by no means limited to intellectual property owners seeking to identify suspected infringers. A "gated access" system may be well adapted to this task, since it can readily distinguish among different requesters and thus set different thresholds at which the "real" registrant contact data would be revealed. (It might also be possible to incorporate in the "reveal" thresholds the differentiation between commercial and non-commercial use which it would be impractical to implement at the time of registration. In other words, if a requester shows that the domain name is associated with commercial activity, a "reveal" could be required for all or specified categories of requesters, including individual Internet users.) It should go without saying that once the thresholds are set, they need to be scrupulously enforced, and services that refuse to reveal the data when presented with a qualifying request should be effectively penalized.

- Validated data. While for the vast majority of proxy registrants, the contact data for the "real registrant" may never be revealed, the potential for revelation exists for all proxy registrants, and therefore the contact data of the customers of proxy services must be subject to the same verification and re-verification requirements as registration data generally. (This is another "plus element" for registration data accuracy compared with the status quo under the 2013 RAA, in which ICANN ultimately declined to require the same level of verification for data on proxy registrants.)

- Prompt action. Unlike the other aspects of the new RDS model, for which we can and should take whatever time is necessary to get them right, work on a proxy accreditation system needs to be a top priority for ICANN. The limited new requirements imposed by the 2013 RAA expire by their own terms on January 1, 2017 unless a Privacy and Proxy Accreditation Program is in place by then.[12]

B. "Maximum protected registration"

COA agrees that, for a very small percentage of gTLD registrants who face physical danger if their contact information is made public, and who have no practical alternatives for engaging in online speech and expression other than gTLD domain name registration at the second level, a much more robust type of identity masking service would be appropriate and may be necessary. However, the discussion of this issue in the Initial Report leaves many questions unresolved, notably what entity would be accredited as a Trusted Agent and under what criteria. We are also concerned by the imprecision of the language in section VI of the Initial Report, which could be read to confer this extraordinary status upon any registrant who "wish[es] to exercise rights of free speech on the Internet which are widely regarded as protected," even in the absence of any evidence of threat, coercion or vulnerability. Since virtually every use that is made of a domain name could plausibly be brought within the ambit of

[12] See http://www.icann.org/en/resources/registrars/raa/approved-with-specs-27jun13-en.htm, Specification on Privacy and Proxy Registrations.

protected free speech at some level, this criterion must be applied as an "and" factor, rather than as an "or" in the current version of section VI.

C. Cost

A sound and transparent RDS is an essential feature of the Domain Name System. ICANN was accorded stewardship of the current version of the RDS (Whois) more than a decade ago. It has allowed the reliability of this critical resource to become significantly degraded on its watch, and it is now (quite properly) shouldering the lion's share of the burden of designing and initially implementing an improved system. Recovering the costs incurred in ongoing operations of a more viable, useful, accessible and accurate RDS should be built into the system and treated as a cost of doing business. In particular, requestors of data should not be expected to bear the bulk of the costs of operating the new system; these should be taken on primarily by registrars and registries as a cost of doing business to which all of them must contribute as responsible players under contract to ICANN.

III. Next steps

COA commends EWG members for their dedication and hard work, which have produced a document that contributes significantly to informed debate about this critical topic. We look forward to further dialogue with EWG, and ultimately to reviewing the final EWG report later this year. As the expectation now appears to be that the output of EWG will simply become one input to an ordinary Policy Development Process, with no indication that ICANN plans to expedite the process in any significant way, all participants should acknowledge that, even assuming that consensus can be achieved on a new RDS model, it will be a number of years before it can be implemented. In the meantime, efforts to improve the existing RDS – Whois – must be accorded high priority and should be undertaken as expeditiously as possible. The recommendations of the Whois Policy Review Team, which the ICANN Board chair states have been approved by the Board in their entirety, should provide the framework for these efforts. In particular, as noted above, it is urgent to put in place strong standards for accreditation of proxy services (and, to the extent they still exist, privacy services)[13] well before the January 2017 expiration of the existing interim specification on this topic in the 2013 RAA.

Respectfully submitted,

Steven J. Metalitz, Counsel
Coalition for Online Accountability

[13] The Whois Registrant Identification Study found that proxy registrations (94% of the total of proxy/privacy registrations) vastly outnumber privacy registrations (6% of the total, revised from 3% in the study draft). See http://gnso.icann.org/en/issues/whois/registrant-identification-summary-23may13-en.pdf , at Appendix C.