KEEPING THE INTERNET
OPEN • INNOVATIVE • FREE

www.cdt.org

CENTER FOR DEMOCRACY
& TECHNOLOGY

1634 I Street, NW
Suite 1100
Washington, DC 20006

P +1-202-637-9800
F +1-202-637-0968
E info@cdt.org

PUBLIC COMMENT ON "INITIAL REPORT FROM THE [ICANN] EXPERT WORKING GROUP ON GTLD DIRECTORY SERVICES: A NEXT GENERATION REGISTRATION DIRECTORY SERVICE"

12 August 2013

In a report published on May 11, 2012, the ICANN WHOIS Policy Review Team described several shortcomings of the current WHOIS directory service, including inaccurate data, privacy risks, lack of a common query interface, and lack of support for internationalized domain names.[1]

On December 13, 2012, ICANN CEO Fadi Chehadé announced the creation of an Expert Working Group on gTLD Directory Services ("EWG") to propose reforms to the WHOIS system.[2] On June 24, 2013, the EWG published a draft report proposing a complete overhaul, replacing the current distributed system with a centralized database called the Aggregated Registration Data Service ("ARDS").[3]

ICANN has requested public input on the EWG proposal by August 12, 2013.

The Center for Democracy & Technology welcomes ICANN's efforts to improve the security, privacy, and accessibility of the WHOIS system. However, to achieve these goals, certain aspects of the EWG proposal should be extensively reconsidered. Our questions and suggestions are below.

I. The current WHOIS system raises privacy and free expression concerns by requiring registrants to disclose sensitive information.

ICANN's registrar agreements for top-level domains ("TLDs") require registrars to collect data from registrants and serve that data in response to WHOIS queries.[4] There are two types of WHOIS data fields:

1. **Technical:** This information is necessary for DNS functionality. The fields are the domain name, primary and secondary nameservers, the registrar's identity, and the domain name's creation and expiration dates.

[1] "Whois Policy Review Team Final Report," May 11, 2012, http://www.icann.org/en/about/aoc-review/whois/final-report-11may12-en.pdf.

[2] "Expert Working Group on gTLD Directory Services Launched," December 14, 2012, http://www.icann.org/en/news/announcements/announcement-2-14dec12-en.htm.

[3] "Initial Report from the Expert Working Group on gTLD Directory Services: A Next Generation Registration Directory Service," June 24, 2013, http://www.icann.org/en/groups/other/gtld-directory-services/initial-report-24jun13-en.pdf.

[4] See "Registrar Accreditation Agreement," May 21, 2009 (updated August 2, 2012), http://www.icann.org/en/resources/registrars/raa/ra-agreement-21may09-en.htm. § 3.3 describes the policy for public access to registrant data.

2. **Personal:** This information identifies the domain's owner, but it is not necessary for DNS operations. These fields include the registrant's name and postal address, as well as the name, postal address, email address, telephone number, and fax number of the administrative and technical contacts.

A simple WHOIS query that anyone can make anonymously returns both this technical and personal information. This system raises serious privacy concerns by revealing sensitive personal information to the public. According to the OECD's privacy guidelines, personal data should be relevant to its intended purpose and should be protected from unreasonable or unauthorized disclosure.[5] Similarly, to guard against identity theft, the U.S. Department of Justice encourages individuals to adopt a "need to know" approach in distributing personal information.[6] But the WHOIS system needlessly exposes registrants' sensitive data to anonymous queries, granting easy access to malicious users.

Registrants concerned with safeguarding their personal information often choose to provide less than accurate information or use proxy services. Although proxy services can limit the exposure of private information, typically only more savvy registrants use them. This highlights a systemic problem with the WHOIS system: users who are unfamiliar with the WHOIS system are more likely to unwittingly expose sensitive data, making these less-savvy users all the more susceptible to harm. As we explain below, we think the WHOIS system should protect individuals and noncommercial entities (hereinafter "individuals") registrants' privacy by default.

The current WHOIS system also raises free expression concerns. As the Security and Stability Advisory Committee ("SSAC") explained, the WHOIS system was developed in 1982 for the needs of a small, close-knit network of technologists.[7] Since then, the Internet has become a vital platform for free expression around the world. Widespread Internet censorship and surveillance have prompted the development of new technologies to secure the key freedom of anonymous speech.[8] The WHOIS system risks undermining free expression – for example, controversial political and artistic uses of the Internet – by requiring individual registrants to either forgo anonymity or contribute to the inaccuracy of the WHOIS directory by providing incorrect information.

The EWG report does not address a number of key questions underlying these concerns. For instance, what registrant information *must* be collected for the domain name system to function and what information *must* be published? Should ICANN require individual registrants to disclose information beyond what is necessary for DNS operations, and why? What is the proper function of the WHOIS system? Is it to provide access to sensitive personal information

[5] "OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data," http://www.oecd.org/internet/ieconomy/oecdguidelinesontheprotectionofprivacyandtransborderflowsofpersonaldata.htm. The "Data Quality Principle" reads: "Personal data should be relevant to the purposes for which they are to be used, and, to the extent necessary for those purposes, should be accurate, complete and kept up-to-date." The "Security Safeguards Principle" reads: "Personal data should be protected by reasonable security safeguards against such risks as loss or unauthorised access, destruction, use, modification or disclosure of data."

[6] "Identity Theft and Identity Fraud," http://www.justice.gov/criminal/fraud/websites/idtheft.html.

[7] "WHOIS: Blind Men And An Elephant," September 14, 2012, http://www.icann.org/en/groups/ssac/documents/sac-055-en.pdf, at 7. "Historically, WHOIS was created to provide a means to make contact information available for both sites and prominent individuals of what was then a very small (and essentially homogeneous in terms of user community) Internet compared to what exists today."

[8] See, e.g., Ronald Deibert et al., *Access Controlled: The Shaping of Power, Rights, and Rule in Cyberspace* (Cambridge, MA: MIT Press, 2010).

to any party with a "use case," or are some data needs more appropriately addressed by direct contact with registrars?

The EWG report does a good job of outlining use cases for access to *currently available* registrant data, but we think it should also reexaminine what data must be available today, in light of the vastly more complex modern Internet environment. We hope ICANN will facilitate a dialogue among DNS stakeholders – including those who would likely want access to more detailed registrant data, such as law enforcement and intellectual property holders – around what data registrants should have to disclose to registrars to facilitate legitimate inquiries. It should then address the question of whether that information should be published as a matter of course or only available pursuant to local legal process. We question whether registering a domain should automatically publish the registrant's personal data in the equivalent of an "Internet phone book." We suggest that individual registrants should be able opt-out entirely of having this information included in WHOIS, effectively allowing anonymous domain registration (at least with respect to the WHOIS directory).

The policy Nominet has adopted for .uk domain names offers a sensible model: individual, non-commercial registrants can choose to keep sensitive information private, while commercial registrants cannot.[9] This policy properly balances the interests and obligations of commercial and non-commercial entities in the internet ecosystem: entities offering services or engaged in trade should necessarily disclose more contact information as part of WHOIS, such that the public can access details needed for commercial and legal activities. Nominet also employs a simple but clever method of dealing with those that abuse this distinction: if Nominet determines that a commercial entity has improperly self-identified as an individual, they can change the setting on that registry entry such that more detailed commercial-entity contact information is publicly shared through WHOIS. Anti-abuse teams that fight spam, phishing and other sorts of malicious behavior can report to Nominet obvious misclassifications and then act legally or procedurally with the more detailed registrant data. We encourage ICANN to recognize the differences in rights and obligations between commercial and non-commercial registrants; individuals should not have sensitive information exposed by default.

II. The proposed privacy scheme and validation of registrants is unnecessary and unworkable. Instead, ICANN should protect registrants' privacy by default.

The EWG report proposes three tiers of privacy protections for domain registrants:[10]

1. **Ordinary registration**: registrants receive no additional privacy protections.
2. **"Enhanced Protected Registration"**: registrants can use privacy or proxy services to conceal their personal information.
3. **"Maximum Protected Registration"**: registrants using domains for "At-Risk, Free-Speech uses"[11] would receive additional protections by "using blind credentials issued by a trusted third party."

[9] "WHOIS opt-out," http://registrars.nominet.org.uk/registration-and-domain-management/query-tools/whois/opt-out.

[10] EWG Report, fn. 3, at 12.

[11] EWG Report, fn. 3, at 21.

This proposed system is arbitrary and would not improve protections for user privacy. The first level mirrors the current WHOIS system. The second level would provide the same functionality as current privacy and proxy registration services, although it is not clear how ICANN would determine how registrants would be eligible to use these services.[12] The third level raises two significant concerns for user privacy and free expression.

First, the proposed system would rely on an unspecified third party to determine whether registrants are "at risk" or are exercising their free speech rights. This is a poorly defined concept that puts users' fundamental rights to freedom of expression in the hands of an unidentified, unaccountable arbiter who will use unspecified criteria to make potentially unreviewable decisions in an undisclosed process. How would these determinations be made? What sort of third party would be capable of making them in an international context with inconsistent norms and laws concerning free speech and privacy? How would international human rights case law, resolutions, and norms factor into the decision?

Assuming that a third party could reliably discern which registrants are speaking freely, there remains the problem of legal jurisdiction. Individuals exercising free speech or those at risk due to their beliefs or activities in one jurisdiction might be subject to law enforcement sanctions from another. How will the third party decide which jurisdiction's law governs?

Second, while this system aims to protect vulnerable registrants, it would also identify them to bad actors. Imagine two lists of IP addresses: one associated with registrants who have freely disclosed their identifying information, and one associated with registrants designated as exercising free speech and engaging in "at-risk" activities. This system would create a vector for attacks designed to access sensitive data about the most vulnerable users. Moreover, this system may create false security by giving registrants the sense that their identities are safely concealed, while providing no protection against a host of other attacks.

The EWG report proposes disclosure of registrant information as the default, allowing increased privacy protections only under certain circumstances. We strongly encourage ICANN to consider the inverse approach: privacy as the default for individual and noncommercial registrants, with disclosure only when necessary.

Further, there is little hope that the ARDS or registrars will be able to strongly validate registrant data. The SSL certificate system serves as an instructive example to demonstrate how difficult it will be to verify registrant data submitted to registrars. At the lowest level, anybody can self-sign an SSL certificate, enabling encryption of traffic but offering no verification of the server's identity to the client. For vendors and customers in need of better identity verification, an active market for trusted certificate authorities has emerged, offering varying levels of validation. The most basic level of validation is "domain validation" where the certificate authority verifies the requestor has control over the domain (typically through an email challenge/response). The highest level of validation is the "Extended Validation Certificate" system, in which the certificate authority takes elaborate steps to verify the identity of the certificate's owner.[13]

[12] EWG Report, fn. 3, at 12. The report says that "enhanced protected registration" would "[e]nable use of accredited privacy or proxy registration services by any registrant seeking to minimize public access to personal names and addresses."

[13] "Guidelines For The Issuance And Management Of Extended Validation Certificates," May 29, 2012, http://www.cabforum.org/Guidelines_v1_4.pdf.

The EWG proposal calls for "initial and ongoing validation of domain name registration data (e.g., designated contacts, addresses)."[14] This proposal would be akin to mandating the Extended Validation Certificate system for all SSL certificates, rather than allowing market forces to develop a variety of levels of verification.[15] Mandating the most intensive form of verification for all domain registrations would add significant expense and complexity to the domain registration process. In cases where such verification is necessary, market solutions providing a range of authentication levels, including a high-assurance option like the Extended Validation Certificate, would be more effective and efficient than a sweeping mandate for high-level authentication in all cases.

III. A centralized system is unnecessary and unstable.

The current WHOIS system is distributed across a global network of registrars. The EWG report advocates replacing this system with a centralized database under the control of an unspecified third party. There are several reasons why a centralized system is unnecessary and more problematic than a distributed system. Although centralized systems have operated successfully in other contexts – for instance, Verisign's operations as the .com registry – the EWG report does not explain why a centralized WHOIS database would be superior to a distributed system.

For instance, ICANN can address the WHOIS issues it identified – data accuracy, privacy, the lack of a common interface, and internationalization – by modifying its contracts with registrars. Since 2003, the SSAC has urged ICANN to adopt contractual solutions to problems with the WHOIS system.[16] It is not clear how centralizing the WHOIS system would address these issues more effectively than modifying contracts with registrars.

It is also unclear how the WHOIS use cases the EWG identifies require a centralized system. The report notes that the ARDS would "[enable] validation/accreditation of requestors qualifying for special purposes (i.e., law enforcement)." For example, if law enforcement sought information about a specific registrant, it could turn to the ARDS gatekeeper instead of seeking data from the registrar. But law enforcement can use existing legal processes to require registrars to disclose information about registrants, and it is not clear why ICANN should provide law enforcement an alternative to circumvent these processes. The ARDS gatekeeper would be a poor substitute for existing legal processes because the ARDS operator would likely lack the capacity to identify illegitimate or overbroad requests originating from national and local jurisdictions across the globe.

Diverting law enforcement requests to the ARDS would also circumvent any state-specific procedural safeguards on law enforcement access. While today a registrar may use local law and process to evaluate law enforcement requests, the ARDS gatekeeper would be in no position to challenge law enforcement requests. As an extension of ICANN, it would need to be "extra-jurisdictional" and not default to a given body of law. The proposed approach to law

[14] EWG Report, fn. 3, at 12.

[15] Extended Validation Guidelines, fn. 13.

[16] "Whois Recommendation of the Security and Stability Advisory Committee," February 7, 2003, http://www.icann.org/en/committees/security/sac003.pdf. After making several recommendations on WHOIS privacy, data accuracy, and other issues, the SSAC wrote, "ICANN should modify the Registry and Registrar contracts to require the recommendations as described in the previous section."

enforcement requests might also circumvent the data protection laws of the registrant's jurisdiction.

Finally, as the EWG report recognizes, the ARDS would create a single point of attack and a single point of failure.[17] A centralized database would be more vulnerable to accidental or intentional data breaches. A technical failure in the ARDS would implicate the safety and security of all registrants and malicious attackers would only have to compromise a single system to access this information. The centralized system would also be susceptible to insider abuse, particularly as registrant data becomes more accurate.[18]

IV. Conclusions

The EWG report attempts to address shortcomings in the current WHOIS system. We agree that an overhaul is necessary. However, we are highly skeptical that a centralized database like the ARDS is the right solution. The ARDS exacerbates serious concerns about user privacy and free expression. We believe that ICANN can fix the WHOIS system while avoiding these pitfalls by modifying its contracts with registrars.

In addition, we are concerned that the EWG has focused on a single model for a new registrant database. We encourage ICANN to consider multiple solutions to this complicated problem and the EWG should be explicitly tasked with recommending a number of candidate models, not just the one current flawed ARDS proposal.

As a next step, ICANN should seek input from all DNS stakeholders about (1) what information registrars must collect from domain name registrants, consistent with the OECD's privacy principles and the Internet of 2013 instead of 1982, and (2) how to restructure the WHOIS system to provide access to registrant data only when strictly necessary, while also facilitating access for legitimate DNS purposes.

For further information, contact CDT's Senior Staff Technologist, Joseph Lorenzo Hall (202-407-8825, joe@cdt.org) or CDT Summer Legal Intern, Joseph Mornin (joseph@mornin.org).

[17] EWG Report, fn. 3, at 6.

[18] The National Opinion Research Council at the University of Chicago found in 2010 that only 23% of WHOIS data is fully accurate. EWG Report, at 11. The EWG proposal aims to increase this level of accuracy through better data validation.