**Response to EWG on Aggregated RDS Model**

**Almost Final Draft**

**(To be submitted as NCSG contribution?)**

***Regarding the EWG's suggested Aggregated RDS model, are there additional advantages and disadvantages that should be considered? In such a model, which data repository (ARDS or Registry) should be considered authoritative?***

There is a positive allure toward the organization of information. Without a doubt, a “one stop service” with a centralized database, promises to bring uniformity with possible enhanced search capabilities. While the purpose of streamlining the gathering, organization, and later dissemination of data is intended to be a positive service for the benefit of all users of the Internet, special care must be taken of the registrants’ privacy with a process that guarantees that their information is ultimately utilized only for the intended purpose. In particular, it would be useful if the EWG’s final report would include an introductory explanation describing the original problem with the current system.

There are disadvantages with the proposed RDS model, and all stakeholders should be concerned with a service that gives registries all data control and authoritativeness, while placing the data under one “gate keeper” in one central repository. If the registrant data will now be held by the registrars, the registries and the centralized repository, how will the access be routed to only one source?  We wonder how clear will be to law enforcement and others seeking the data, which of the three places to utilize when requesting that data?

Will the EWG incorporate the WHOIS Review Team's adoption of the concept of "contactability" largely for technical purposes? How will the RDS model protect registrants from the practice of all the information anyone could possible want for any circumstances? What will be the goal of a centralized database, if adopted? Will it be to streamline a minimum amount of information, or the maximum? How will the EWG incorporate the WHOIS PDP Working Group recommendations?

It is not enough to compare this new challenge with existing ones. There are real downsides to a single source of data. In addition, it is not clear what impact this will have with national laws, in particular, and how this will impact directly registrants. These matters should be addressed in light of human right considerations. Regarding which repository should be authoritative (ARDS or Registry), the answer should be the one that offers the maximum privacy protections. Other concerns deserve serious attention before a new model is adopted. How will abuse of data be monitored and controlled? What will be the limits for law enforcement? If the limitation and policing are not done upfront, don't we impose a huge burden on the registrant for that policing or lack thereof? How easy will it be for the registrant to identify who is searching his/her/its data?

***Could the EWG's recommendations for addressing maximum protected registration satisfy both accountability needs and the privacy needs of at-risk individuals? How might a suitable solution be identified and funded?***

This is the key question that the EWG must answer, and one that needs to be solved fully before the success of the RDS Model can be secured. The RDS Model model will require recognition of established privacy and data protection principles before proceeding with any other consideration. This is necessary to protect the data elements in relation to registrants. We recommend robust privacy protections for all registrants, and more so for those considered to be at-risk. The desire for personal data privacy of registrants should not be assumed to be motivated by a registrant's bad faith or intent to perform illegal acts such as infringement or even dilution of intellectual property rights. Personal identifiable data privacy is a basic human right, and in some cases, not recognizing this right could put a registrant’s personal safety and security at risk.

It is not clear to us what process the EWG envisions to evaluate petitions for ‘at risk’ status and the accreditation process to be utilized to grant that status. Will it be an ICANN committee or independent expert group? This accreditation process in and of itself, however, could place such individuals/groups at even greater risk by creating a paper trail that anyone could follow. Accordingly, even putting aside the more difficult question of how exactly this assessment would be carried out, will the proposed solution here may even be worse than the original problem? For example, the credentialing process as proposed by the EWG seems to imply that the same person or law firm or law enforcement agency may get access repeatedly to the centralized database of registrant data. Could this aspect be clarified?

The proposed RDS Model is expected to be managed and monitored, and with restrictions to the access of registrant data; and for legitimate uses. This is the true measure of accountability. The model should allow for tracking tools available to registrants. This proposed new model should not become a system of data retention in order to merely facilitate law enforcement. The transfer process would have to be refined to avoid the imminent danger to consumers. We note that if adopted, this new model must be monitored under strict standards. These standards must be maintained to ensure that all personal information collected is utilized for its intended purpose. There is great risk for potential misuse of this information. In the end, information access and dissemination must reside with the individual registrant. The EWG is encouraged to consult more than one expert in privacy and data protection, preferably for every geographical region of the world.

The EWG may benefit from contacting the DHS Data Privacy and Integrity Advisory Committee (under the Secretary of Homeland Security and the DHS Chief Privacy Officer on programmatic, policy, operational, administrative, and technological issues) for advice in this matter within the US. That committee was established by the Secretary of Homeland Security under the authority of 6 U.S.C. § 451.

***Are the users and purposes identified by the EWG thus far sufficiently representative? Are there any significant gaps in users and purposes that must be addressed?***

Care must be given to gaps in users and purposes that could endanger registrants. For example, will law enforcement and others have unfettered access? The new model should take into consideration the circumstances specific to requests made, and should never grant across-the-board access. For example, how will the bad actor category be defined? For all purposes, bad actors could be identified from nearly every category of user. Bad actors in the WhOIS space could include intellectual property attorneys, individuals and even law enforcement: those that go “fishing” through the data would demonstrate bad acts that clearly go beyond any legitimate purpose. Others may seek to find registrants for the purpose of harassment and intimidation (including to give up domain names they are otherwise entitled to), while seeking disclosures of physical locations (to harass, stalk and intimidate, e.g., for purposes of physical violence or to stop exercise of unpopular free speech positions). What will be the designated parameters? This requires careful analysis, further discussion, and clarification.

***Given the desire for an extensible next-generation RDS that might accommodate the needs of a rapidly-evolving global Internet, how could future new users and purposes be accommodated? Who would decide on permitted purposes, using what criteria?***

The primary consideration should be the protection of registrants. The proposal for the RDS Model should be based on privacy safeguards strengthened by international human right principles and not on the desire by any stakeholder to exercise increased control over data elements. Registrants should not be compelled to disclose personal information. If, in the interest of uniformity and standardization, this model is finally adopted, then standards must be drafted to ensure that all personal information collected is utilized for its intended purpose. How can a centralized-database possibly know what is a valid purpose or invalid purpose?  Won't requesters submit the “correct” information to get the data? Not even a good automated expert system could handle this function. Will a court determine whether a request is non-frivolous within the legal action being pursued? Thus, it will be crucial for the new model what body of individuals will decide the permitted purposes, and above all, the criteria to be utilized for final decisions.

Once a registry or central repository disseminates personal data, it would have to follow standards to protect that data, in particular, against repressive governments. The disclosure of personal data without restraints would threaten individuals’ right to privacy, and thus, their freedom of expression. Without any threshold or required showing of need, doesn't it all amount to a bottom line of -- requestor wants the data and requestor has shown that he/she/it exists? All requestors of information must be required to formally submit an application to access data. All purported beneficial reasons for the creation of the RDS Model should not in itself be the standards to deliver the collected and processed personal data. Most likely the process will require “case by case” evaluations by a judge or panel of some kind (human intervention).  Above all, registrants should be notified of those requests made of their data.

Is the repurposing of data allowed?  As the WHOIS Review Team noted, the only truly agreed upon purpose of the WHOIS is for technical contact when a technical problem is identified within the domain name. That is why many registrants gave this data originally, ten years ago, and continue today.  That scope, further, is consistent with the scope and mission of ICANN. Has the EWG reached out to Data Protection Commissioners to check whether the “repurposing of the personal data” of millions of registrants in the gTLD range of content and purposes is consistent with law - or whether it might be illegal under the last of the EU, Japan, South Korea, Canada, South America and other regions with data protection laws?

Why would EWG recommend a long-abandoned idea to require a domain name registrant to label the “content” of the use of his/her speech for which the domain name may be used as commercial or non-commercial?  This is a) impossible (as use is a spectrum of speech activity very broad and not easily classified, e.g. a link to one's own book on Amazon is not commercial not does fundraising make a non-commercial group commercial, b) it is an idea long abandoned by innumerable task forces because it is unattainable, and c) it puts ICANN into the role of “content gate keeper” far beyond its role as manager and overseer of an infrastructure system. Finally, is this concept a consensus recommendation of the EWG?

The only way to prevent over-use is to flip this entire system around, and provide access ‘only’ on a case by case basis and only ‘after’ a case has been made out for access to the specific data requested. Finally, while some of their envisioned purposes are 'legitimate' (if only in a targeted way), some do not seem legitimate. Neither ‘business research’ nor 'Internet research' seem to be legitimate reasons to force individuals to hand over personal info. These are not remotely related to the operation of the naming system, and are not purposes recognized as having the right to compel access to data under any circumstances (contrast law enforcement and those with legitimate rights claims, which \*can\* compel data access under controlled circumstances). All registrants must have the right to refuse this type of access, by some form of readily accessible consent mechanism (opt in/out).

***Are the registration data elements identified by the EWG thus far sufficiently representative of the data required for each identified purpose? Are there any significant gaps in data elements that must be addressed?***

We note that data elements are also subject to privacy and data protection considerations as noted in the comments to other sections of this questionnaire. Will the new model incur risks associated with new data elements (with no proposal to streamline or limit data), while searching across all gTLDs on a massive scale; something impossible to do today?

***How should public and gated data elements be classified? What criteria should the EWG apply to make initial recommendations in this area?***

The public access to data elements is as important as who controls them. Is the EWG recommending that every element of the existing WHOIS plus much more be included in a centralized database, including name, address and phone, and also never-before-collected data elements such as the purpose of the domain name? The management of billions of entries is hardly a simple matter, which will require restrictions to deter abuse. While this information now relies on the protection of national and regional legal systems, it is this protection that will be lost. Gated access to protected data should be based on legitimate grounds and based on identified purposes, and with the appropriate “gate keeper.”

***What community needs should be considered during the EWG's discussion of registration data storage duration, escrow and access log requirements?***

Should we not logically determine that the centralized storage of information may be problematic? It should be noted that dissemination of additional information for global public distribution may be unnecessary regardless of intent or use. It will contain information that private entities, and government seek – with or without legal rationale, with or with due process. One can envision registries of future gTLDs in Saudi Arabia, Syria, North Korea, China and other jurisdictions, not known for their human rights and due process protection, placing an entire database into the hands of a private company in these countries and others, under enormous pressure to disclose, package, search and sort this data and provide it to law enforcement and a range of private parties for a number of reasons – not necessarily in the best interest of the community. Aggregation then would be a downside, not a benefit. The centralized database will become an enormous target in this regard.

Do privacy protections for the centralized database depend on where it is located? Who would determine the standards - the EWG? ICANN? The GNSO? We want to note that matters of privacy and data protection need further consideration, and should be at the forefront of the discussion. The EWG Report leaves some questions unanswered and we are concerned that a centralized environment may be a challenge; one that will require further analysis within recognized human rights standards. Indeed, further discussion and consideration, not only by the EWG, but by both, the GNSO Council and the ICANN Board will be needed.

***The EWG acknowledges that deploying and operating the suggested RDS will incur costs. In such a system, how could or should those costs be borne?***

These costs should be borne in a manner consistent with the protection of the consumers. While the service is intended to streamline the way data is validated, in the end, this type of data aggregation will require greater accountability by all stakeholders, but in particular, the “sole gate keeper” of this data, and all registries. We note that the costs for registries may not only be technical, but legal, should they face challenges to the transition from local data protection or other authorities, or fines and penalties should such a transition be determined after the fact to be illegal. We ask for clarification in this area. Furthermore, the centralization of data would impose enormous costs on its ultimate “gate keeper” as law enforcement, corporate attorneys, and intellectual property attorneys would seek rapid, easy, unlimited access to its data – and a range of features to assist in the data mining of this data. It could create an enormous target for misuse and abuse from an accessibility perspective. A Risk Analysis seems critical -- and as soon as possible. When would this analysis take place and when would its results become known to the ICANN Community?