Complete deliberation on the charter question: *What steps should be taken to control "thin data" access?*

These WG Agreements are limited to “Thin Data” elements, for example:

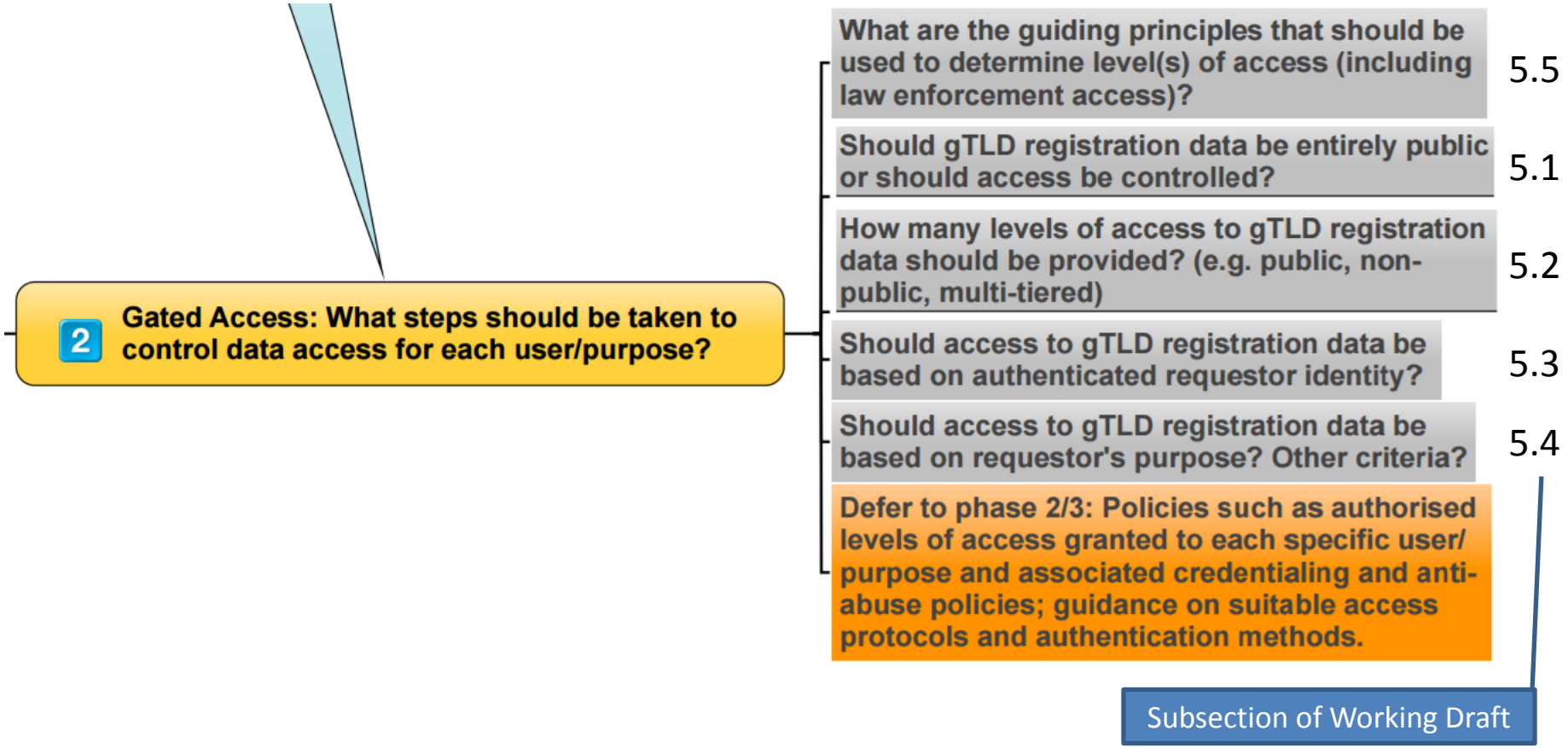
Domain Name: ANVILWALRUSDEN.COM
Registrar: TUCOWS DOMAINS INC.
Sponsoring Registrar IANA ID: 69
Whois Server: whois.tucows.com
Referral URL: http://www.tucowsdomains.com
Name Server: NS1.SYSTEMDNS.COM
Name Server: NS2.SYSTEMDNS.COM
Name Server: NS3.SYSTEMDNS.COM
Status: clientTransferProhibited https://icann.org/epp#clientTransferProhibited
Status: clientUpdateProhibited https://icann.org/epp#clientUpdateProhibited
Updated Date: 17-jan-2017
Creation Date: 30-jun-2010
Expiration Date: 30-jun-2017

Rough consensus WG Agreements reached thus far:

20. *“gTLD registration "thin data" should be accessible without requestor identification, authentication, or stated purpose.”* (9 May call, poll question 2)
21. *“There must be no RDS policies that prevent RDS operators from applying operational controls such as rate limiting and CAPTCHA, provided that they do not unreasonably restrict legitimate access.“* (2 May call, poll question 3*)

* Action item to define underlined text

Review all of this charter question's subquestions to complete first pass deliberation on "thin data" access



Source: KeyConceptsDeliberation-WorkingDraft-9May2017.pdf, Page 19

Review all of this charter question's subquestions to complete first pass deliberation on "thin data" access

5.1) *Should gTLD registration "thin data" be entirely public?*
Possible answer: Yes, see WG Agreement #20

5.2) *How many levels of access to gTLD registration "thin data" should be provided?*
Possible answer: One, see WG Agreement #20

5.3) *Should access to gTLD registration "thin data" data be based on authenticated requestor identity?*
Possible answer: No, see WG Agreement #20

5.4) *Should access to gTLD registration "thin data" be based on*
- Requestor's purpose?
- Other criteria?

5.5) *What guiding principles should be applied to determine level(s) of access to "thin data"?*

No.	Data Access Principles – Excerpted from EWG Report, pages 58-60, as starting point for WG consideration **WHICH OF THESE PRINCIPLES APPLY TO "THIN DATA" ACCESS?**
41.	A minimum set of data elements, at least in line with the most stringent privacy regime, must be accessible by unauthenticated RDS users.
42.	Multiple levels of authenticated data access must be supported, consistent with stated permissible purposes.
43.	RDS user access credentials must be tied to an auditable accreditation process, as further defined in Section IV(c), RDS User Accreditation.
44.	Access must be non-discriminatory (i.e., the process must create a level playing field for all requestors, within the same purpose).
45.	To deter misuse and promote accountability: • All data element access must be based on a stated purpose; • Access to gated data elements must be limited to authenticated requestors that assert a permissible purpose; and • Requestors must be able to apply for and receive credentials for use in future authenticated data access queries.
46.	Some type of accreditation must be applied to requestors of gated access: • When accredited Requestors query data, their purpose must be stated every time a request is made. • Different terms and conditions may be applied to different purposes. • If accredited requestors violate terms and conditions, penalties must apply.

Resume deliberation on the charter questions on Purpose and Data Elements for "thin data" only

As a reminder, initial rough consensus WG Agreements for “thin data” purpose include:

Should gTLD registration thin data elements be accessible for any purpose or only for specific purposes?

1. *The WG should continue deliberation on the purpose(s) of "thin data."*
2. *Every "thin data" element should have at least one legitimate purpose.*
3. *Every existing "thin data" element does have at least one legitimate purpose for collection.*

For what specific (legitimate) purposes should gTLD registration thin data elements be collected?

4. *EWG-identified purposes apply to at least one "thin data" element.*
5. *Domain name control is a legitimate purpose for “thin data” collection.*
6. *Technical Issue Resolution is a legitimate purpose for “thin data” collection.*
7. *Domain Name Certification is a legitimate purpose for "thin data" collection.*
8. *Business Domain Name Purchase or Sale is a legitimate purpose for "thin data" collection.*
9. *Academic / Public Interest DNS Research is a legitimate purpose for "thin data" collection.*
10. *Regulatory and Contractual Enforcement is a legitimate purpose for "thin data" collection.*
11. *Criminal Investigation & DNS Abuse Mitigation is a legitimate purpose for "thin data" collection.*
12. *Legal Actions is a legitimate purpose for "thin data" collection.*
13. *Individual Internet Use is a legitimate purpose for "thin data" collection.*

Resume deliberation on this subquestion:
What is the purpose of each "thin data" element?

Merged text from Sullivan's Purposes in Detail & EWG Report Annex D

Thin Data Element	EWG Purposes	Collection Rationale	Publication Rationale
Domain Name	All	The domain name is required to be collected under the Statement of Purpose, purpose 1. Without this, there is no domain name, so it is literally impossible to have anything to collect or publish.	The domain name is required to be published under purpose 1, because it is a key by which data is accessed. If you wish to look up the current data about a particular name, you use the name as the key by which you query. (This is not the only possible key. For instance, in an EPP registry you could in principle use the ROID to look up a particular name object. But that does not give you the current data for the thing so named; it just gives you the data about that Repository Object. Two different versions of the same name -- like if example.com is registered by Alice then deleted and later registered by Bob -- have different ROIDs.)
Registrar	Domain Name Control Business Domain Name Purchase/Sale Academic/Public Interest DNS Research Regulatory/Contractual Enforcement Criminal Investigation/ DNS Abuse Mitigation DNS Transparency	IANA has a registry of registrar IDs (https://www.iana.org/assignments/registrar-ids/registrar-ids.xhtml#registrar-ids-1), and that contains their (iii) names. This is a protocol parameter registry, but it appears to be managed by ICANN so it is probably appropriate for this PDP to make the policy about how that is to be managed. Data (iii) needs to be collected in order to give (i) Registrar ID meaning, because it is the only way to know whether two IANA ids are bound to the same organization or person.	See data (i) Registrar ID?

See https://community.icann.org/download/attachments/64078512/Merged-ThinDataPurposes-v1.pdf

Resume deliberation on this subquestion:
What is the purpose of each "thin data" element?

Merged text from Sullivan's Purposes in Detail & EWG Report Annex D

Thin Data Element	EWG Purposes	Collection Rationale	Publication Rationale
Sponsoring Registrar IANA ID (aka Registrar IANA Number)	Domain Name Control Business Domain Name Purchase/Sale Academic/Public Interest DNS Research Regulatory/Contractual Enforcement Criminal Investigation/ DNS Abuse Mitigation DNS Transparency	(i) Registrar ID provides data about the entity that created the entry in the registry (formally, in EPP, "repository"). Data (i) is required to be collected under RDS purposes 1 and 2. Without this data it is not possible to know the source of the data and it is not possible to trace it further in the system.	Data (i) are possibly required to be published under purpose 1. This largely depends on whether we think the identity of who is managing an object in the registry is part of the "lifecycle of a domain name". My feeling is "yes". Also, this information is likely to be disclosed anyway; owing to the way these work, publication of these is likely to "leak" information about (i) and (iii)
Whois Server and Referral URL (aka Registrar URL)	Domain Name Control Business Domain Name Purchase/Sale Academic/Public Interest DNS Research Regulatory/Contractual Enforcement Criminal Investigation/ DNS Abuse Mitigation DNS Transparency	(ii) Whois Server and Referral URL both provide metadata necessary for the operation of the distributed database that makes up the RDS (in systems other than whois, approximately the same data with the same relation to identity would be in place, but the details might be different. I think we can treat this as a class anyway) Data (ii) is required to be collected under purposes 1 and 2 (dissemination of registration data). Without this data it is not possible to know the source of the data and it is not possible to trace it further in the system.	Data (ii) are required to be published under purposes 1 and 2, as long as there is at least one data element that is required under some purpose and is not available from the registry. (Since the actual registration life cycle is controlled by the registrar and not the registry, this appears likely.)

See https://community.icann.org/download/attachments/64078512/Merged-ThinDataPurposes-v1.pdf

Resume deliberation on this subquestion:
What is the purpose of each "thin data" element?

Merged text from Sullivan's Purposes in Detail & EWG Report Annex D

Thin Data Element	EWG Purposes	Collection Rationale	Publication Rationale
Name Servers	Domain Name Control Technical Issue Resolution Domain Name Certification Business Domain Name Purchase/Sale Academic/Public Interest DNS Research Regulatory/Contractual Enforcement Criminal Investigation/ DNS Abuse Mitigation	Without collecting the name servers, domain names cannot function on the Internet, so this is required under purposes 1 and 2. (Given that the registration of the name itself and the collection of the name servers are both required for the basic functioning of the Internet Domain Name System, it strikes me that we may be missing a more obvious purpose in our list, but I guess (1) and (2) will be enough and we're already so late that I am loathe to suggest something more.)	Whenever a name is available on the Internet, the name server data is already available in the DNS, so this data is necessarily published. Under either purpose 1 or 2 (or both), the data about nameservers in the RDS provides an avenue for troubleshooting issues in the DNS, and so it is required for those purposes.
Statuses (aka Registration Status, Client Status (Registrar) Server Status (Registry))	Domain Name Control Business Domain Name Purchase/Sale Academic/Public Interest DNS Research Regulatory/Contractual Enforcement Criminal Investigation/ DNS Abuse Mitigation	The status values are not exactly "collected", but are at least in part the result of various actions by the sponsoring registrar and registry on the name. (Some can be set directly.) These govern the disposition of the name in question, and are a necessary condition for having a shared registration system, so they are required under purpose 1.	The status values govern the possible things that could be done to a name, and therefore the data must be published under purpose 1.

See https://community.icann.org/download/attachments/64078512/Merged-ThinDataPurposes-v1.pdf

Resume deliberation on this subquestion: *What is the purpose of each "thin data" element?*

Merged text from Sullivan's Purposes in Detail & EWG Report Annex D

Thin Data Element	EWG Purposes	Collection Rationale	Publication Rationale
Updated Date and Creation Date and Expiration Date (aka Registrar Expiration Date)	Domain Name Control Business Domain Name Purchase/Sale Academic/Public Interest DNS Research Regulatory/Contractual Enforcement Criminal Investigation/ DNS Abuse Mitigation	While the dates might appear to be different kinds, they aren't, since for our purposes they all have at least one common utility (see below). The dates, like status values, are not exactly "collected": they're a consequence of certain activities. They're necessary for the workings of the shared registration systems using the current fee-for-term model that (approximately?) all gTLD registries use today, so they're required under purpose 1.	The dates are required under purpose 1 or 2 in order to aid troubleshooting of resolution. (If a name worked yesterday and not today, it is helpful to know that it was just created -- meaning the old one was deleted -- or that it is expired, or that someone updated the name only last night.)

In addition, Sullivan-SuggestionForPurposeInDetail.pdf provides rationale for "Maximal Audience," noting: *I use the "maximal audience" because I think that if there is any "whole public" use then there's no point considering more restrictive uses. (For instance, if we need the domain name to be published to everyone on the Internet because it won't work otherwise, then it makes no difference if LEOs want that data under some sort of authorized-access protocol, because they'll just get it under the wide-open rules instead. So we don't need to care about the LEO purpose in that case.) "Maximal audience" might not work for cases where two different classes have different needs both of which require some restrictions, but it's handy here because we're talking about thin data.*

This concept has not yet been included in the above table but can be added in a second pass.

See https://community.icann.org/download/attachments/64078512/Merged-ThinDataPurposes-v1.pdf

EXCERPTS FROM INPUT MATERIALS FOR REFERENCE AS-NEEDED

Related Input Materials

Final Issue Report [PDF] (7 October 2015), especially

- Section 4.2, Gated Access
- Annex C, Charter – Gated Access Question, Phase 1 Goals (Page 70)

EWG Recommendations for a Next-Generation RDS, especially

- Section 4b, Principles for Unauthenticated and Gated Data Access
- Annex E, Unauthenticated and Gated Access Examples
- Video FAQs "Does the RDS eliminate free public access to data?" and "What would I need to do to access gated RDS data?
- EWG Tutorial Pages 15-21, 42-60

Question 5: https://community.icann.org/download/attachments/64078601/ICANN58-DataProtectionExpert-Responses-7April2017-plus-Intro.pdf

https://community.icann.org/display/gTLDRDS/Phase+1+Documents

- KeyConceptsDeliberation-WorkingDraft-9May2017.pdf, Section 5
- All WG decisions are added to this document during deliberation

Relevant Question/Answer from ICANN58 Data Protection Experts

5. Below is an example of "thin data" elements made publicly accessible in today's WHOIS system for every registered gTLD domain name. Do you believe that any of the following data elements are considered personal information under the General Data Protection Directive, and why?

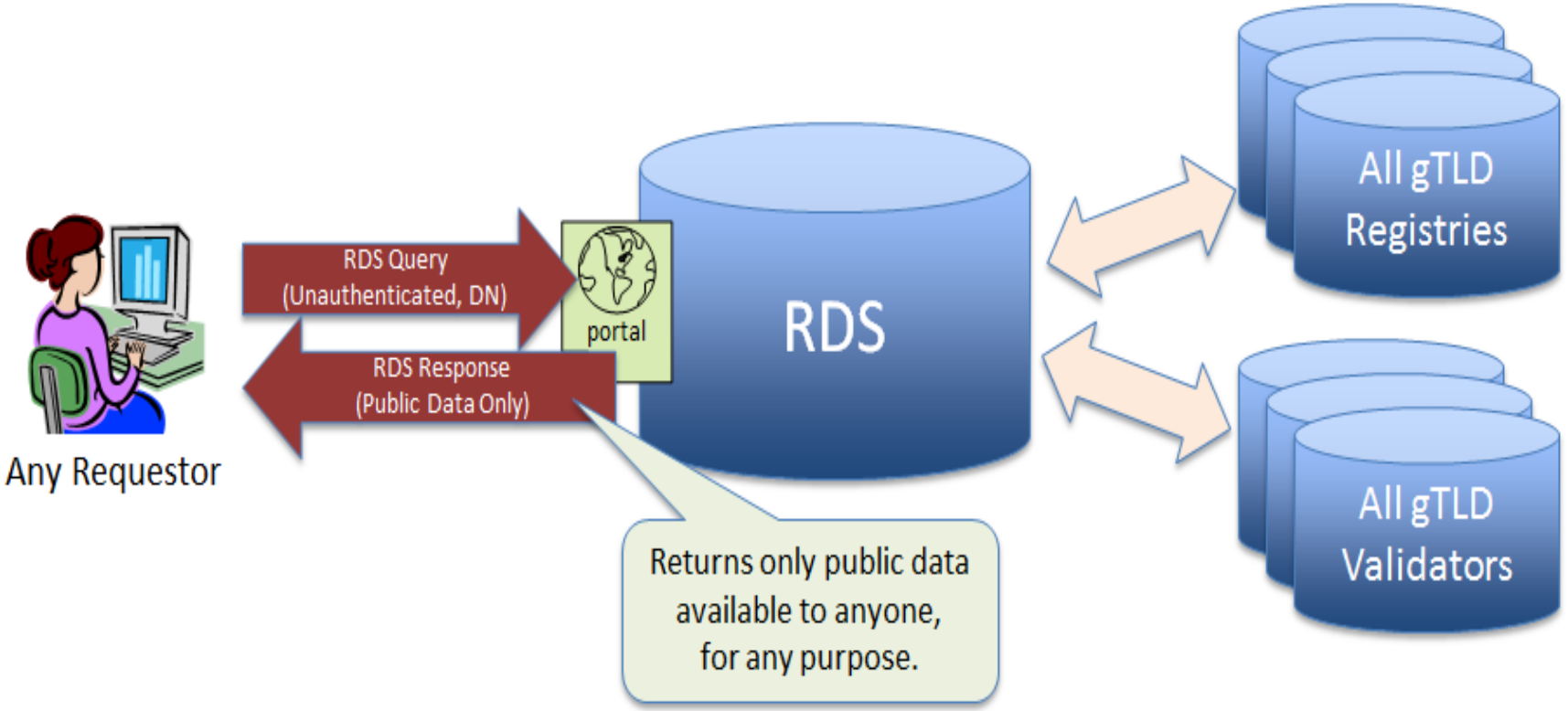
Domain Name: CNN.COM
Registrar: CSC CORPORATE DOMAINS, INC.
Sponsoring Registrar IANA ID: 299
Whois Server: whois.corporatedomains.com
Referral URL: http://www.cscglobal.com/global/web/csc/digital-brand-services.html
Name Server: NS-1086.AWSDNS-07.ORG
Name Server: NS-1630.AWSDNS-11.CO.UK
Status: clientTransferProhibited https://icann.org/epp#clientTransferProhibited
Status: serverDeleteProhibited https://icann.org/epp#serverDeleteProhibited
Status: serverTransferProhibited https://icann.org/epp#serverTransferProhibited
Status: serverUpdateProhibited https://icann.org/epp#serverUpdateProhibited
Updated Date: 15-feb-2017
Creation Date: 22-sep-1993
Expiration Date: 21-sep-2018

This information can be easily combined with other data sets freely or easily accessible, then yes, it is "personal data". Google itself is offering look up services, reverse look up services (for free). Besides there are websites which are harvesting data from whois.corporatedomains.com and making them accessible freely with personal data as on WHOIS Servers there is personal data. (see: www.who.is for instance). As long as the identification of a person behind this information and numbers is possible, it is considered as personal data.

Source: https://community.icann.org/download/attachments/64078601/ICANN58-DataProtectionExpert-Responses-7April2017-plus-Intro.pdf

Answer to 5.1 given by the EWG Report, Pages 61-62:

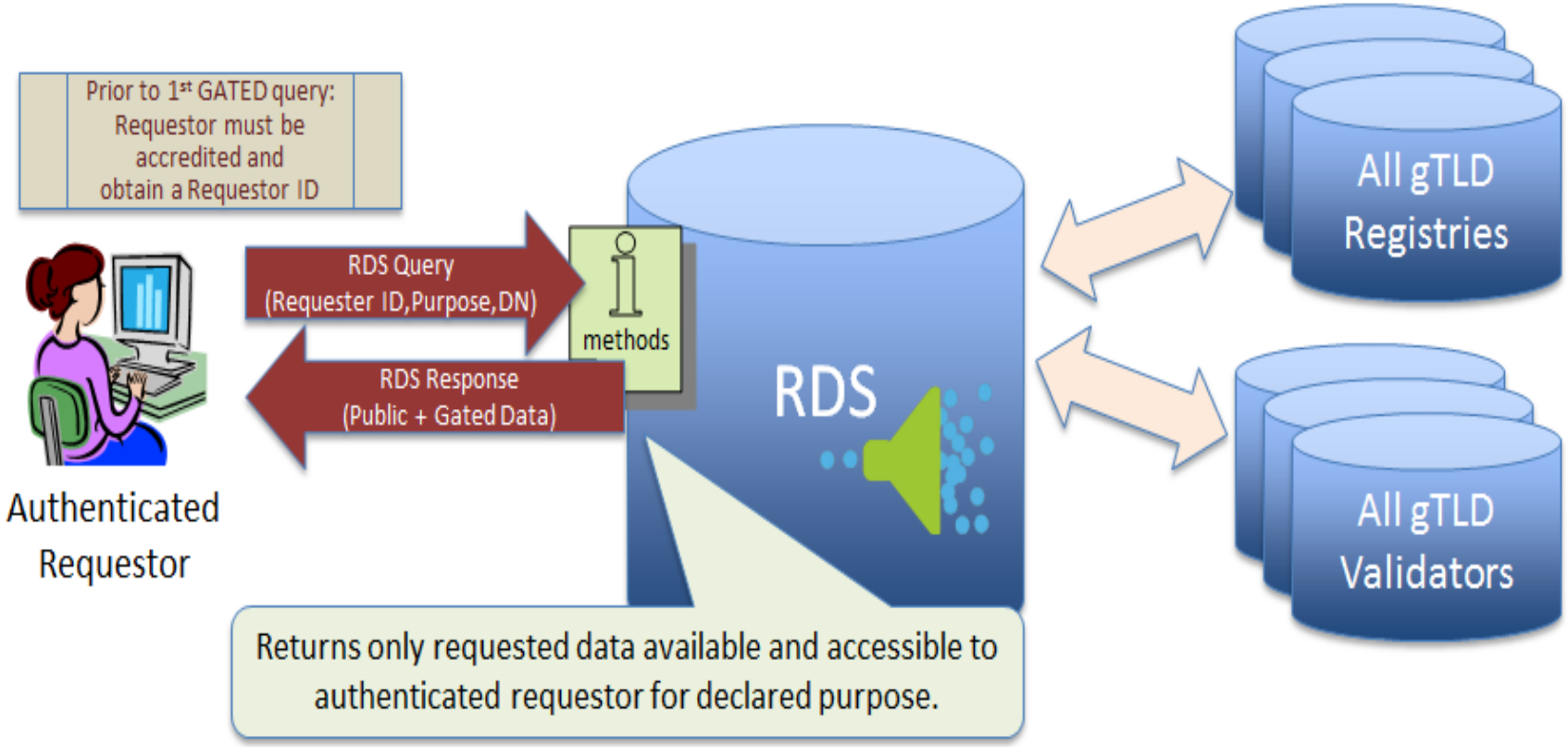
As depicted in the following figure, public data elements can still be requested from the RDS by anyone, with or without authentication.



A minimum set of data elements, at least in line with the most stringent privacy regime, must be accessible by unauthenticated RDS users.

Answer to 5.1 given by the EWG Report, Pages 61-62:

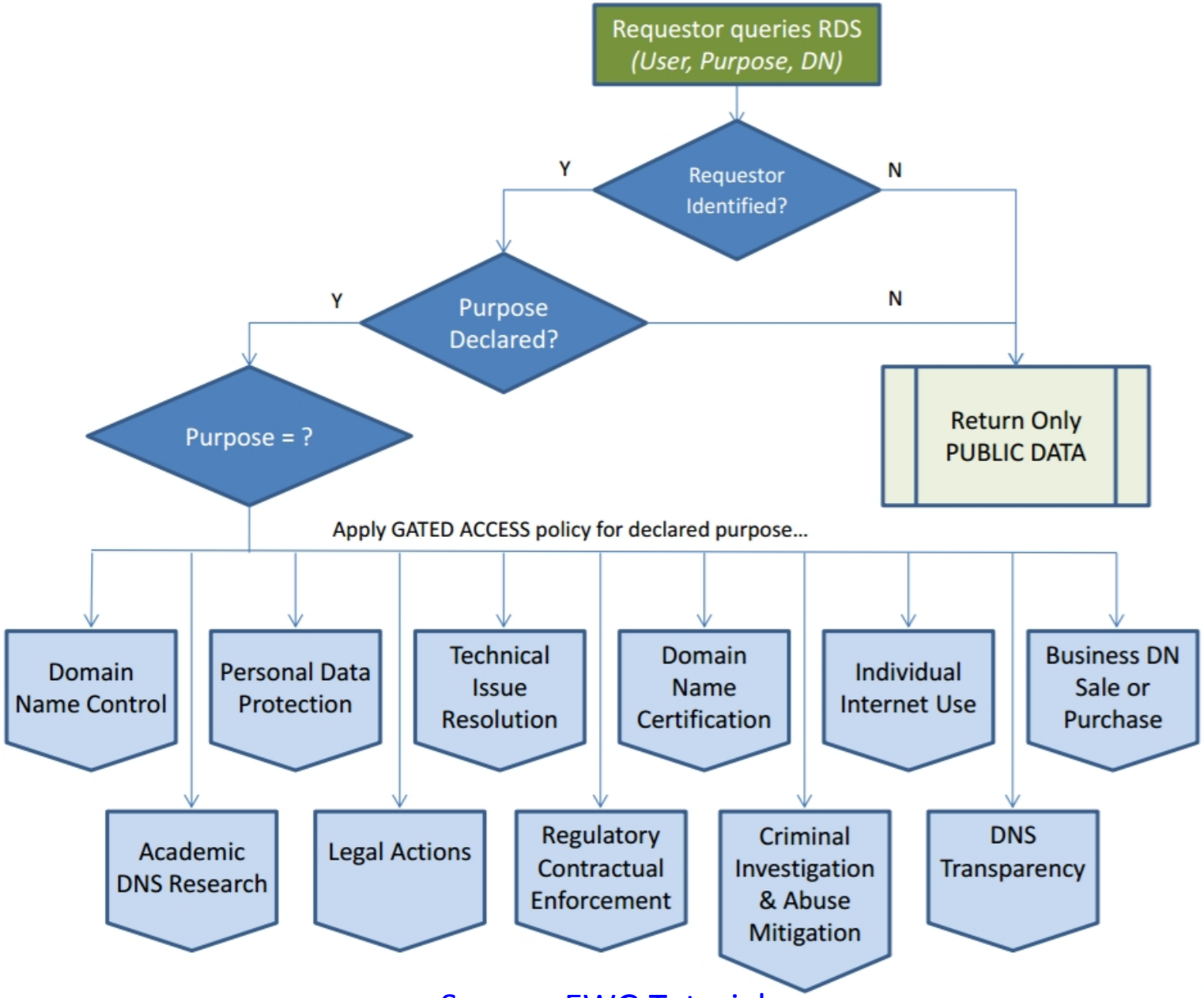
As depicted in the following figure, gated data elements can also be requested via the RDS. To do so, requestors must first be accredited. Thereafter, requestors may submit authenticated queries requesting data elements for a stated purpose.



Multiple levels of authenticated data access must be supported, consistent with stated permissible purposes.

How does this differ from WHOIS?

- The EWG Report split RDS Data Elements into categories
  - Minimum Public Data Set (includes today's "thin data" elements)
  - Gated Data (includes most of today's "thick data" elements)
  - See EWG Report for definitions and criteria
- Requestors optionally IDENTIFY and AUTHENTICATE themselves
  - Anonymous RDS queries return ONLY Minimum Public Data Set
  - Authenticated RDS queries may or may not return Gated Data Subset
- Requestors optionally state a PURPOSE
  - Users can be ACCREDITED for one or more purposes
  - Accredited users are AUTHORIZED to access Minimum Public Data Set + Gated Data Subset AS NEEDED BY PURPOSE & LIMITED BY APPLICABLE LAW
- Requestors optionally ACCREDITED for RDS access, for example
  - Self-accreditation for purposes authorized to access to low-risk data
  - Third-party accreditation for purposes with access to higher-risk data
- No identification or authentication? No purpose?
  - Such queries return ONLY Minimum Public Data Set
- Anti-abuse measures such as RATE LIMITING apply to all kinds of RDS access



Source: EWG Tutorial

Purposes for collection? Purposes for providing access?

Purposes identified for Thin Data	Includes tasks such as… (Note: may involve more than thin data)	Related Thin Data Elements	Example Use Cases developed by PDP WG (Note: may involve more than thin data)
Domain Name Control	Creating, managing and monitoring a Registrant's own domain name (DN), including creating the DN, updating information about the DN, transferring the DN, renewing the DN, deleting the DN, maintaining a DN portfolio, and detecting fraudulent use of the Registrant's own contact information.	Domain Name [Name Servers] Sponsoring Registrar Registrar's RDS/WHOIS URL Registration Status(es) Registration Creation Date Registration Expiration Date RDS/WHOIS Last Updated Date	DN maintenance - Transfer DN maintenance - Deletions DN maintenance - DNS Changes DN maintenance - Renewal
Technical Issue Resolution	Working to resolve technical issues associated with domain name use, including email delivery issues, DNS resolution failures, and website functional issues, by contacting technical staff responsible for handling these issues.	Domain Name [Name Servers] Sponsoring Registrar Registrar's RDS/WHOIS URL Registration Status(es) Registration Creation Date Registration Expiration Date RDS/WHOIS Last Updated Date	Technical Issue Resolution Technical Issue Resolution (specific examples)
Domain Name Certification	Certification Authority (CA) issuing an X.509 certificate to a subject identified by a domain name needing to confirm that the DN is registered to the certificate subject.	Domain Name Name Servers	Certification Authority
Business Domain Name Purchase or Sale	Making purchase queries about a DN, acquiring a DN from another Registrant, and enabling due diligence research.	Domain Name Name Servers Sponsoring Registrar Registrar's RDS/WHOIS URL Registration Status(es) Registration Creation Date Registration Expiration Date RDS/WHOIS Last Updated Date	Business DNs - Bankruptcy Asset Purchase Business DNs - Mergers and Acquisitions Business Intelligence

Source: ICANN58 RDS PDP WG F2F Meeting Slides

Purposes for collection? Purposes for providing access?

Purposes identified for Thin Data	Includes tasks such as… (Note: may involve more than thin data)	Related Thin Data Elements	Example Use Cases developed by PDP WG (Note: may involve more than thin data)
Academic/ Public Interest DNS Research	Academic public-interest research studies about domain names published in the RDS, including public information about the Registrant and designated contacts, the domain name's history and status, and DNs registered by a given Registrant.	Domain Name Name Servers Sponsoring Registrar Registrar's RDS/WHOIS URL Registration Status(es) Registration Creation Date Registration Expiration Date RDS/WHOIS Last Updated Date	None developed by PDP WG EWG example cases include: DN Registration History DNs for Specified Contact Survey DN Registrant or Contact
Regulatory and Contractual Enforcement	Tax authority investigation of businesses with online presence, UDRP [and URS] investigation, contractual compliance investigation, and registration data escrow audits.	Domain Name Name Servers Sponsoring Registrar Registrar's RDS/WHOIS URL Registration Status(es) Registration Creation Date Registration Expiration Date RDS/WHOIS Last Updated Date	Services required by Registry Agreement
Criminal Investigation & DNS Abuse Mitigation	Reporting abuse to someone who can investigate and address that abuse, or contacting entities associated with a domain name during an offline criminal investigation.	Domain Name Name Servers Sponsoring Registrar Registrar's RDS/WHOIS URL Registration Status(es) Registration Creation Date Registration Expiration Date RDS/WHOIS Last Updated Date	Investigate Abusive Domain Find Domains Registered by Miscreant Reputation Services Law Enforcement - Compromised websites WHOIS queries for compliance purposes

Source: ICANN58 RDS PDP WG F2F Meeting Slides

Purposes for collection? Purposes for providing access?

Purposes identified for Thin Data	Includes tasks such as… (Note: may involve more than thin data)	Related Thin Data Elements	Example Use Cases developed by PDP WG (Note: may involve more than thin data)
Legal Actions	Investigating possible fraudulent use of a Registrant's name or address by other domain names, investigating possible trademark infringement, contacting a Registrant/Licensee's legal representative prior to taking legal action and then taking a legal action if the concern is not satisfactorily addressed.	Domain Name Other Thin Data Elements?	Obtain DN holder details for legal action Fraudulent contact information Trademark Infringement
Individual Internet Use	Identifying the organization using a domain name to instil consumer trust, or contacting that organization to raise a customer complaint to them or file a complaint about them.	Domain Name Other Thin Data Elements?	Real-World Contact

Source: ICANN58 RDS PDP WG F2F Meeting Slides