Date: 11 November 2017

TO: John Jeffrey, ICANN

FROM: Greg Aaron, iThreat Cyber Group

RE: strawman proposal for WHOIS compliance with GDPR

This document proposes a solution that could be implemented by May 2018. It attempts to deliver the following benefits:

- A proportional solution that provides the protections required by law, while displaying as much data as is allowed by law. This will provide the minimum disruption to current legitimate users of WHOIS, and will provide as much domain contactability as possible.
- Uses existing technology,
- Does not impose unnecessary burdens on registries and registrars, and
- Can be implemented in a uniform way across registries and registrars.

I would like someone to tell me how this proposal stands up to scrutiny. Is anything here unworkable, if so why, and can any challenges be overcome reasonably?

For readability and brevity, this document does not delve into some of the complexities of GDPR, such as heightened requirements for informed consent, etc. Nor does it attempt to address the deployment of RDAP or tiered access. Instead, it provides an expedited, high-level plan for the immediate future, and leaves longer-term problems for a second round of problem-solving.

What Data Must Be Protected under GDPR?

GDPR protects the data of natural persons residing in the European Union. **So, we need a mechanism to protect certain fields of WHOIS contact data for the domains registered by those individuals.** Registrations made by legal persons are not protected by GDPR, and contact data for such domains can be published.[1]

Contact data for all domains should continue to be collected by registrars (and provisioned up into the registries if at all possible). Registries and registrars should continue to publish all of the current field names that are currently published in WHOIS as per contract. This will maintain uniformity of output format.

[1] "This Regulation does not cover the processing of personal data which concerns legal persons and in particular undertakings established as legal persons, including the name and the form of the legal person and the contact details of the legal person." (GDPR, paragraph 14.) If the contact details of natural persons are included in a legal person's domain record, it is assumed that the registrant has gotten permission from their employees/domain contacts. In other words, the registrant is a data controller for its domain record and contacts, and is responsible.

All "thin" data (sponsoring registrar, Create/update/Expiration dates, nameservers, domain statuses, etc.) should *always* be published for *all* domains. Such fields do not contain personally identifiable data. (See below.)

The GDPR might allow the contact data of natural persons to be published in WHOIS by default, *if* natural persons are allowed to opt out of publication. This is the system that Nominet currently uses for .UK.[2] SIDN also uses an opt-out mechanism in .NL.[3] This opt-out question deserved thorough analysis.

If opt-out for natural persons is not possible, then natural persons' data in sensitive fields should be replaced in WHOIS output with placeholder data, such as "Data withheld for privacy law compliance".[4] (See below for an example.) Some contact data fields (such as the Country field) are not sensitive enough to allow the identification of the contact and should continue to be published.

So:

- If the domain is registered by **a natural person residing in the EU:**
    - Personal data should be published in WHOIS, unless the registrant has opted out. If a registrant opts out, then certain contact fields should be masked.
    - If it is not possible to offer an opt-out regime, then data in certain contact fields should be masked by default.
- If the domain is registered by **a natural person outside the EU**, all data should be published as it is now, per ICANN contract.
- If the domain is registered by **a legal person**, all data should be published as it is now, per ICANN contract.
- Proxy and privacy services can continue to be used as they are now, and their use should be identified clearly in WHOIS. Registrars and registries should not provide proxy and privacy services as a way of avoiding the requirements of this plan.

Data collection and display are justified under several lawful bases under GDRP.[5]

Other Necessary Arrangements

The above proposal assumes that registrars can issue revised terms of service that meet the GDPR's requirements. Registrars could start notifying their relevant registrants about GDRP ahead of May 2018,

---

[2] If a natural person registrant does not opt out, Nominet seems to publish at least the registrantt's name and physical address, if not email address and phone address. See http://registrars.nominet.uk/namespace/uk/management/data-quality/whois-opt-out

[3] See https://www.sidn.nl/a/nl-domain-name/sidn-and-privacy?language_id=2 and https://www.sidn.nl/downloads/terms-and-conditions/Data+Protection+Policy+for+nl+Domain+Names.pdf

[4] This is basically what .FRL has been doing.

[5] Including consent, the data is necessary for the performance of a contract to which the data subject is party, and the performance of tasks in the public interest (such as escrowing registration data, anti-abuse activities, access by law enforcement. protection of others' rights, etc.).

offering opt-out for registrants who are natural persons, and notifying how contacts for domains owned by legal persons will be displayed.

There could be common language for a data processing statement, including a purpose statement that includes the reasons why WHOIS exists and its allowable uses. (Contactability, resolution of problems, etc.) These statements can be commonsensical, and do not need to be long. Two notable examples are:

- .UK: https://nominet-prod.s3.amazonaws.com/wp-content/uploads/2015/10/WHOIS_Contract_Terms_1_Sept_2015.pdf
- .NL: https://www.sidn.nl/downloads/terms-and-conditions/Data+Protection+Policy+for+nl+Domain+Names.pdf

After May 2018, there will need to be further policy development and revisions to procedures.

Do Thin WHOIS Fields Contain Personally Identifiable Data?

No. Pieces of "thin WHOIS" data such as Sponsoring Registrar, Create Date, Nameserver, etc. are not enough for a member of public (indeed, even for skilled investigators) to connect them to a registrant, even when combined with other pieces of publicly available data.

The GDRP states that: "30. Natural persons may be associated with online identifiers provided by their devices, applications, tools and protocols, such as internet protocol addresses, cookie identifiers or other identifiers such as radio frequency identification tags. This may leave traces which *in particular when combined with unique identifiers and other information received by the servers, may be used to create profiles of the natural persons and identify them.*" (Emphasis added. GDPR, http://data.consilium.europa.eu/doc/document/ST-5419-2016-INIT/en/pdf)

Paragraph 26 also states: "To ascertain whether means are reasonably likely to be used to identify the natural person, account should be taken of all objective factors, such as the costs of and the amount of time required for identification, ... The principles of data protection should therefore not apply to anonymous information, namely information which does not relate to an identified or identifiable natural person or to personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable."

Stated another way: a piece of data must be protected *only if the party possessing that piece of data may have the ability to connect that piece of data to a natural person.*

The connection between "thin WHOIS data" and a natural person can usually only be made using some other privileged piece of data that ordinary parties will not possess. Usually these privileged pieces of key data are only held by parties such as registrars and ISPs, and they will not publish/disclose that data without legal process.

Note that IP and nameserver data MUST be published in the DNS in order for a domain to resolve.

In its first memorandum to ICANN[6], Hamilton wrote: "In an ICANN context, even thin WHOIS data, IP addresses (including dynamic ones), metadata, etc. *are to be considered personal data* as the identification of an individual by using such data or by combining them with other publicly, easily accessible data is possible." [Emphasis added.] That should not be read to mean that such data must be protected, or that it cannot be published in WHOIS or the DNS.

The Court of Justice of the European Union (CJEU) recently held that IP addresses and the like are "personal data" *only in certain circumstances*. The principle was that if one has no legal means of linking an IP address to the identity of its user, then that IP address is unlikely to be personal data.

The CJEU recently decided that a dynamic IP address will be personal data in the hands of a website operator if:

- there is another party (such as the natural person's ISP) that can link the dynamic IP address to the identity of an individual; *and*
- the website operator has a "legal means" of obtaining access to the information held by the ISP in order to identify the individual.

So in that case, the IP address was "personal data" only in the hands of the ISP (who knows what IP it assigned to its natural person customer), and in the hands of the German government (which could use legal process to compel the ISP to identify the user who used the IP). The ISP and the government would need to take care not to reveal the linkage. But the IP address was not PII for anyone else, since no one else can make the linkage to a natural person.
See: https://www.whitecase.com/publications/alert/court-confirms-ip-addresses-are-personal-data-some-cases

Note that the IP address of a domain name (the A record) just tells us where a domain is hosted. This does not reveal the registrant's physical address or identity. Only the hosting provider may knows the identity of its customer, and the hosting provider should protect that according to the law.

European TLD operators such as Nominet and SIDN have closely reviewed their legal obligations and are publishing thin data. So that's interesting.

The GNSO's RDS WG closely examined the issue in 2017 and generally did not feel that any thin data fields constitute personally identifiable data.

[6] https://www.icann.org/en/system/files/files/gdpr-memorandum-part1-16oct17-en.pdf

WHOIS Output for a Protected Natural Person

Here is what would be published if the domain is registered by a **natural person,** if the natural person has either opted out of the publication of his/her personal data in WHOIS, or if an opt-out regime is not possible. This format is copied from the Base Registry Agreement.[7]

All of the WHOIS fields below should continue to be published in WHOIS. Data in certain of those fields may be masked.

WHOIS FIELD	EXAMPLE DATA PUBLISHED	NOTE
Domain Name	EXAMPLE.TLD	always publish
Domain ID	D1234567-TLD	always publish
WHOIS Server	whois.example.tld	always publish
Referral URL	http://www.example.tld	always publish
Updated Date	2009-05-29T20:13:00Z	always publish
Creation Date	2000-10-08T00:45:00Z	always publish
Registry Expiry Date	2010-10-08T00:44:59Z	always publish
Sponsoring Registrar	EXAMPLE REGISTRAR LLC	always publish
Sponsoring Registrar IANA ID	5555555	always publish
Domain Status	clientDeleteProhibited	always publish
Domain Status	clientRenewProhibited	always publish
Domain Status	clientTransferProhibited	always publish
Domain Status	serverUpdateProhibited	always publish
Registrant ID	5372808-ERL	always publish
Registrant Name	Masked for data privacy compliance	*Mask if natural person*
Registrant Organization	EXAMPLE ORGANIZATION	*Always publish; do not mask; may indicates legal person*
Registrant Street	Masked for data privacy compliance	*Mask if natural person*
Registrant City	Masked for data privacy compliance	*Mask if natural person*
Registrant State/Province	Brittany	*Always publish; do not mask*
Registrant Postal Code	Masked for data privacy compliance	*Mask if natural person*
Registrant Country	FR	*Always publish; do not mask*
Registrant Phone	Masked for data privacy compliance	*Mask if natural person*
Registrant Phone Ext	Masked for data privacy compliance	*Mask if natural person*
Registrant Fax	Masked for data privacy compliance	*Mask if natural person*

[7] https://newgtlds.icann.org/sites/default/files/agreements/agreement-approved-31jul17-en.html

Registrant Fax Ext: 4321	Masked for data privacy compliance	*Mask if natural person*
Registrant Email: EMAIL@EXAMPLE.TLD	Masked for data privacy compliance	*Mask if natural person*
Admin ID		always publish
Admin Name	Masked for data privacy compliance	*Mask if natural person*
Admin Organization		*Always publish; do not mask*
Admin Street	Masked for data privacy compliance	*Mask if natural person*
Admin City		*Mask if natural person*
Admin State/Province	Paris	*Always publish; do not mask*
Admin Postal Code	Masked for data privacy compliance	*Mask if natural person*
Admin Country	FR	*Always publish; do not mask*
Admin Phone	Masked for data privacy compliance	*Mask if natural person*
Admin Phone Ext	Masked for data privacy compliance Masked for data privacy compliance	*Mask if natural person*
Admin Fax	Masked for data privacy compliance	*Mask if natural person*
Admin Fax Ext	Masked for data privacy compliance	*Mask if natural person*
Admin Email	admincontact@example.orf	Always publish[8]
Tech ID:	Masked for data privacy compliance	always publish
Tech Name:	Masked for data privacy compliance	*Mask if natural person*
Tech Organization	Masked for data privacy compliance	*Always publish; do not mask*
Tech Street	Masked for data privacy compliance	*Mask if natural person*
Tech City	Masked for data privacy compliance	*Mask if natural person*
Tech State/Province	CA	*Always publish; do not mask*
Tech Postal Code	Masked for data privacy compliance	*Mask if natural person*
Tech Country	USA	*Always publish; do not mask*

[8] Always publish Admin contact email address and Tech contact email address, to provide contactability and issue resolution. For .NL domains, SIDN publishes the Admin Contact and Tech Contact email address even for domains registered by natural persons: See https://www.sidn.nl/a/nl-domain-name/sidn-and-privacy?language_id=2 and https://www.sidn.nl/downloads/terms-and-conditions/General%20Terms%20and%20Conditions%20for%20.nl%20Registrants%20(with%20changes%20per%201%20March%202016).pdf

Tech Phone	Masked for data privacy compliance	*Mask if natural person*
Tech Phone Ext	Masked for data privacy compliance	*Mask if natural person*
Tech Fax	Masked for data privacy compliance	*Mask if natural person*
Tech Fax Ext	Masked for data privacy compliance	*Mask if natural person*
Tech Email	techcontact@example.ord	Always publish
Name Server	NS01.EXAMPLEREGISTRAR.TLD	always publish
Name Server	NS02.EXAMPLEREGISTRAR.TLD	always publish
DNSSEC	signedDelegation	always publish (if provided)
DNSSEC	signedDelegation	always publish (if provided)