



WHOIS Contact Data Availability and Registrant Classification Study

A Study of the Effects of GDPR and ICANN Policy

by

Greg Aaron

Lyman Chapin

David Piscitello

Dr. Colin Strutt

Interisle Consulting Group, LLC

25 January 2021



Executive Summary

Domain names are one of the essential components of the Internet. *Domain name registration data* identifies who registered and controls the domain name, and provides technical information that makes the domain name function. This data has always been vital for a variety of legitimate purposes. For those reasons, registration data has always been available in WHOIS, a well-known “registration data directory service” (RDDS) publication system.

The General Data Protection Regulation (GDPR) was adopted by the European Union (EU) and took full effect on 25 May 2018. This important data protection and privacy law had significant impact on how domain name registration data could be published. In response, the Internet Corporation for Assigned Names and Numbers (ICANN) established a new policy, allowing registrars and registry operators to comply with GDPR by redacting (withholding) personally identifiable data from publication in WHOIS. This policy created a tension between ICANN’s two competing goals: allowing compliance with GDPR while also preserving the publication of non-protected data (and the utility of the WHOIS system itself) “to the greatest extent possible.”

To date, ICANN has not performed work to determine key facts about the effects of its policy, such as how much data has been redacted, or under what circumstances. Therefore, there has not been a factual basis for determining if or how the policy is having its intended effects. The purpose of the current study is to establish answers to those questions. This study uses previously established methodologies, so its results can be compared to the results of previous studies and provide historical continuity. It is our hope that this data can be used for fact-based discussion about what effects the GDPR and ICANN’s resulting policy have had.

The major findings of this study are:

ICANN’s GDPR-driven policy has led to the redaction of the contact data for most gTLD domains and has allowed registrars and registry operators to redact (withhold) much more contact data than is required by GDPR—perhaps five times as much as is necessary. While ICANN’s policy has generally protected the data that must be protected per GDPR, ICANN’s policy has also been used to conceal a much larger set of contact data that is *not* subject to GDPR.

1. Before the GDPR went into effect and ICANN changed its registration data policy, the actual identities of about 75.7% of gTLD domain registrants were available in WHOIS. The other 24.3% of domains were under privacy/proxy protection.
2. At present, only 13.5% of domains have an actual registrant identified in WHOIS. Registrars and registry operators have used ICANN’s post-GDPR policy to redact contact data from 57.3% of all domains. Adding proxy-protected domains, this means that 86.5% of registrants cannot be identified via WHOIS.
3. The use of privacy/proxy protection has increased over time—from 20.1% of all domains in 2013 to about 29.2% in 2020.
4. About 23.1% of gTLD domains are covered by the GDPR’s jurisdictional reach. These are the domains for which the registrant, registrar, registry operator, or registry back-end provider is located in the EU.
5. However, the data of legal persons is not protected under GDPR. The data suggests that only around 11.5% of domains may belong to natural persons who are subject to GDPR. This 11.5%

may be the percentage of domains that is *necessary to protect* under GDPR. In contrast, registrars and registry operators have redacted contact data from 57.3% of all domains, or five times the amount that may be necessary.

6. Different registrars are making very different choices about how much contact data they redact. Some display the contact data of registrants outside of GDPR's jurisdictional reach. Others redact the data of all their registrants, whether GDPR applies to them or not.
7. The data indicates that domain registrars have significant control over (or effective influence on) whether their registrants' contact data is displayed in WHOIS.
8. More than half of the gTLD namespace—51.7%—is now controlled by unidentifiable parties. These are domains that cannot be attributed to a registrant or user, either via WHOIS or by examining their web site content. This is in contrast to the time before GDPR and ICANN's policy, where only 18% of domains were controlled by unidentifiable parties.
9. While 23.1% of all domains fall within the GDPR's jurisdictional reach, only 12.5% of domains had a registrant that resided in the EEA. By protecting personal data *processed within* the EU, the GDPR's reach extends protection beyond EU residents to a larger set of domains and registrants.

Sponsorship

This industry study was funded by contributions from the following organizations. All are participants in the ICANN community. This report is the work of Interisle Consulting Group, and Interisle is wholly responsible for its content. The sponsors had no input into the design of the study, and had no input into or influence on the findings or the content of this report.

Anti-Phishing Working Group (apwg.org)

CAUCE, the Coalition Against Unsolicited Commercial Email (cauce.org)

DomainTools (domaintools.com)

Facebook (facebook.com)

The ICANN Business Constituency (bizconst.org)

Microsoft (microsoft.com)

RiskIQ (riskiq.com)