This is a response to the request for feedback on the three proposed discussion models for collecting registration data and implementing registration directory services published by ICANN on 12th January 2018.[1] It is written in our capacity as internet privacy researchers. In particular, we focus on user tracking via first and third parties in mobile and web apps. WHOIS has an important role in our attempts to identify and map information flows that may impact user privacy. Before commenting on the specific proposed models, we would like to point out the positive role that WHOIS (and any future iterations of RDS) plays in protecting the privacy of data subjects.

WHOIS provides transparency for data subjects

In debates about whether WHOIS data should be publicly available, the privacy of registrants is usually pitted against the public benefits of WHOIS. These benefits include facilitating the commercial market for domain names, enabling administrators associated with domains to be contacted about problems and held accountable, as well as for research for cyber-security or spam detection. However, in addition to these public benefits, WHOIS can also play a crucial role in *protecting* privacy. To the extent that personal data is collected via the internet, and most internet services are associated with a domain name, WHOIS provides a single, consistent and reliable means through which data subjects can establish the identity of various online data controllers.

Most non-technical users do not make use of this functionality directly, but many benefit from it indirectly. Over 1 in 10 people use browser plugins or built-in browser features which block certain forms of third party tracking. Such tools are reliant on the ability of contributors (mostly volunteers) to compile comprehensive information about domain names which end-users may wish to blacklist (as used in **tracker protection lists** such as e.g. Easylist),[2] for which WHOIS is an invaluable resource (despite efforts by some trackers to evade detection through anonymous registrations). Similarly, privacy and security researchers frequently use WHOIS to trace the entity behind any third party servers that a web service or app is sending data to.

WHOIS can promote GDPR compliance for online data controllers

Article 13(1)(a) of the General Data Protection Regulation requires data controllers to provide data subjects with their identity and contact details. The Article 29 Working Party has produced guidance on the transparency requirements of the GDPR.[3] Regarding Article 13(1), it notes that the identity and contact details of the controller should be “easily accessible” to the data subject, "taking account of the device used (if applicable), the nature of the user interfaces/interactions with the data controller (the user “journey”) and the limitations that those factors entail". The Article 29 Working Party recommends that such information be made available in a prominent place, e.g. on a website or app, where the user is likely to see it. In this regard, some of the DNRD that is currently listed in WHOIS is in fact already required to be made publicly available by data controllers.

---

[1] https://www.icann.org/news/blog/data-protection-and-privacy-update-seeking-community-feedback-on-proposed-compliance-models

[2] https://easylist.to/

[3] ec.europa.eu/newsroom/just/document.cfm?doc_id=48850

However, while domain owners are often identifiable from information contained on website(s) associated with their domains, not all domains have corresponding websites, and even if they do there is no standard protocol for the provision of data controller identity and contact details within websites or apps. In particular, third party tracking generally does not involve any direct visual interaction with data subjects via a GUI. Instead, one of the only reliable ways to detect the presence of such third party trackers is via their domain name. Once an organisation has been identified as a tracker, it is very helpful to be able to easily find other domain names associated with it to provide deeper insights into online tracking which benefits scholars, practitioners, and regulators.

The ability to perform a WHOIS lookup on such names is therefore an invaluable feature for greater transparency. End-user privacy tools such as tracker blockers are an increasingly common way for data subjects to identify and block entities attempting to track or otherwise collect data about them online. Data controllers who collect data in these ways who are not listed on WHOIS actively inhibit the effectiveness of such tools. While opting to have their identity and contact details listed on WHOIS is probably insufficient for a data controller to demonstrate compliance with Article 13(1)(a), the converse - actively opting out of public WHOIS records - runs counter to these transparency requirements, especially if suitable alternative forms of identification are absent. For many data controllers operating online, WHOIS is a consistent and effective way for them to make themselves easily identifiable to the privacy protection tools that millions of end-users rely on today.

It is no surprise then, that certain third party trackers have been observed using existing WHOIS workarounds which hide such contact information, such as privacy-protection services offered by registrars, presumably in order to evade identification and blacklisting by tracking protection tools. The debate about data protection law and the ICANN WHOIS service has centred on protecting the rights of domain registrants as data subjects. This focus ignores the fact that many - perhaps most - registrant domain owners are also data controllers, and as such are required to provide information about their identity and contact details to data subjects in a variety of formats to facilitate transparency.

Comments on the three models:

We hope that whatever model is adopted will not inhibit the transparency of data controllers that is currently indirectly promoted by WHOIS. With these issues in mind, we believe that models 1 or 3 would best enable academic, commercial, and community efforts to create tracker protection lists.

Model 1 would allow the continued investigation of domains that are observed to be tracking users via websites and apps without further need for formal certification. Although this model also seems to be at odds with ICANN's own data protection obligations raised by the Article 29 Working Party.

Under Model 2, tracker identification could become effectively impossible if those working in this space have to gain accreditation with each relevant registree / registrar. The administrative burden would likely mean that WHOIS data ceases to be a useful source for such work, unless an effective, cross-registree system can be put in place. This burden is

also likely to exacerbate inequalities of access to WHOIS data, especially for those facing economic and/or language barriers relative to the registry / registrar country.

Model 3 may be the most viable measure through which ICANN's own data protection obligations can be met, based on our understanding of the situation, and would also in principle enable the identification of tracking domains as described above. One downside under Model 3 is that domain owners who intend to use their domains to track users could evade their transparency obligation by simply including personal data in the relevant registrant fields. However, providing there were adequate measures in place to prevent such abuse, model 3 would still enable WHOIS to be a valuable resource for privacy researchers to identify tracking domains.

Given our support for models 1 / 3, it is important to note that the approaches suggested in the consultation document are flawed in their current form. The document describes certain measures applying 'to personal data included in registrations of natural persons'. While the distinction between legal and natural persons is important in data protection law (where only the latter is afforded protection), we would argue that whether the *registrant* is a natural or legal person is not the most relevant distinction. Rather, a better distinction is whether the *information displayed* in a field identifies a legal or natural person, whether or not the registrant is a legal or natural person. For instance, a legal person registrant might supply the email address of a particular employee, e.g. firstname.lastname@example.com (which would probably be personal data), while a natural person registrant might supply a generic email address, e.g. contact@example.com (which might not be personal data). So for any registrant (whether natural or legal person), there are various fields which may or may not be personal data, depending on whether they identify someone or not. The public display of registrant data should depend on whether a *field* contains data that identifies a natural person, rather than on whether the registrant is a natural or legal person.

- Reuben Binns, Department of Computer Science, University of Oxford
- Max Van Kleek, Department of Computer Science, University of Oxford
- Jun Zhao, Department of Computer Science, University of Oxford
- Nigel Shadbolt, Department of Computer Science, University of Oxford
- Tim Berners-Lee, Department of Computer Science, University of Oxford
- Tim Libert, Reuters Institute for the Study of Journalism, University of Oxford
- Huw Davies, Oxford Internet Institute, University of Oxford
- Meredydd Williams, Department of Computer Science, University of Oxford