Packet Clearing House
572 B Ruger Street, Box 29920
The Presidio of San Francisco
San Francisco, California
94129-0920 USA
+1 415 831 3100 main
+1 415 831 3101 fax

Board of Directors
ICANN

Wednesday, January 22, 2020

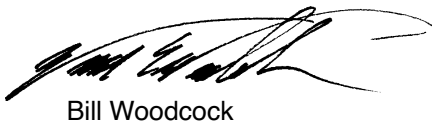
Ladies and Gentlemen of the Board:

In our letter of December 11, 2019, we addressed one of the three critically harmful consequences of allowing Ethos Capital to pay to circumvent the discipline of competition and the rigor of a multistakeholder process. Here we address a second harm: the harvesting and selling of personal information and communications metadata.

Packet Clearing House, a 501(c)(3) not-for-profit public-benefit organization, has provided domain name service for the .ORG domain for the past sixteen years. We provide the same service to hundreds of other top-level domains and have operated the largest DNS network in the world for several decades. Our centrality in the domain name industry has given us insight into the effects that the surveillance economy has wrought upon the industry, the harms it visits upon the public, and the consequent diminution of human rights and chilling of speech. A conclusion we draw from that insight is that allowing unqualified and deeply indebted parties to purchase control of the .ORG domain constitutes a grave danger to its registrants, and society at large, which is dependent upon them.

Ethos Capital, the would-be purchaser of control of the .ORG domain, is preparing to engage in practices most regard as ethically suspect and which are, in many countries, simply illegal: all four of its other acquisitions are companies that monetize personal data. Ethos is thus preparing a pipeline for the exploitation of .ORG registrants' privacy and the sale of the personal information of everyone who communicates with any .ORG registrant. Attached is a more detailed explanation of the issue, its mechanisms, and its dangers.

ICANN has two duties: to the stability of the domain name system, and to the public interest. Allowing circumvention of both the multistakeholder process and the evaluation of qualifications ensures bad outcomes and would be an abandonment of the duty for which ICANN exists. If ICANN reduces itself to a rubber-stamp and a fee-collector, it has no remaining purpose. We encourage you to live up to ICANN's responsibility and enact a multistakeholder evaluation of competitive proposals for the operation of .ORG.



Bill Woodcock
Executive Director
Packet Clearing House

cc: Maarten Botterman, maarten.botterman@board.icann.org
Goran Marby, goran.marby@icann.org
John Jeffrey, john.jeffrey@icann.org
Cyrus Namazi, cyrus.namazi@icann.org

Regarding the Privacy and Integrity of Registrant Communications

Financial Constraints

Ethos has stated that it was able to secure only $360M in debt financing from real banks, and that the remainder of its $1.135B purchase price for PIR had to be met from private equity. Assuming an average bank lending interest rate of 5%, and the current average PE return rate of 22.6%, financing costs from the two sources amount to $18M/year and $175M/year, respectively, a total of $193M/year. ISOC's current profit margin is $70M/year, but this includes a $29M/year subsidy that would not exist after a transfer to for-profit ownership. So, if .ORG were to be operated sustainably with no other changes, Ethos would be left with a net $152M annual shortfall.

Ethos is presumably planning to trade a portion of this insolvency for instability (in the form of increased communications downtime for .ORG registrants); a portion for increased rent extraction (in the form of higher prices, less value for more money); and a portion for extraction of value from registrants via other mechanisms. Here, we discuss those other mechanisms.

The Ethos website lists four other companies it has acquired. All are startups in the data-brokering business, each exploring different methods of monetizing personally identifiable information (PII). Together, they form a pipeline of sorts, which when filled with PII would produce money. The .ORG registry would fill the head of Ethos's pipeline.

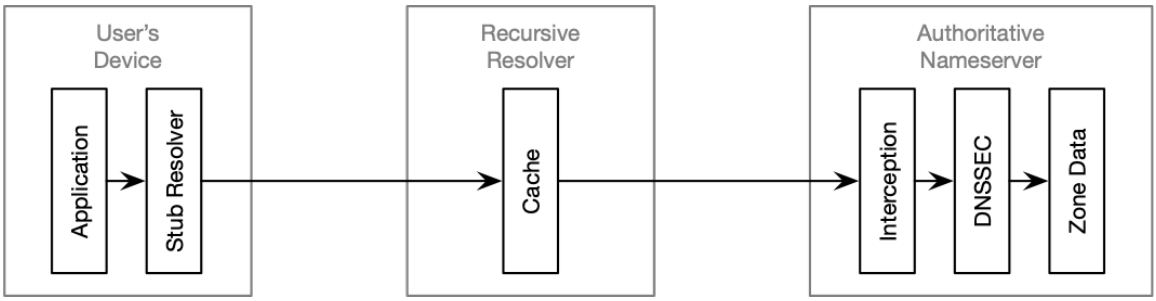
So, how do you turn .ORG registrants' private information into cash? The contract under which the .ORG domain is delegated by ICANN precludes selling the contact information of registrants, and there are already plenty of mailing lists of non-profits available for less than $0.0013 per organization. So even if Ethos were to risk the delegation by flouting this restriction, it would have to sell each .ORG registrant's identity nearly 12,000 times each year to reach $15 (each domain's share of Ethos's annual budget deficit)—an unlikely prospect. It does, however, put into perspective the degree of shortfall Ethos would have to make up, and the value of the information that its pipeline is built to sell relative to that of traditional mailing lists.

Operation of .ORG would give Ethos the communications metadata about every transaction that anyone in the world engages in with any .ORG registrant. If you send email to someone on a .ORG email address, that information will be available for Ethos to sell. If you visit a .ORG website, that information will be available for Ethos to sell. If someone with a .ORG address purchases an advertisement and you see it, that information will be available for Ethos to sell. That information can be recombined and resold in a multitude of ways. Want the identity of individuals in the state of Alabama who sent email to a family-planning clinic in the last year? No problem. Want a list of individuals in China who visited the Falun Gong website? No problem. Want a list of all of individuals in Russia who saw advertisements purchased by alternative political parties in the months before an election? No problem.

Those who would collect individuals' personal information and monetize it must first collect as much as possible. The following are a few of the methods used by unscrupulous DNS operators to maximize their harvesting of users' personal information.

Attacks Against the Privacy of DNS Communications

In this diagram, a DNS query originates in an application within a user's device. The application forwards its query to the stub resolver resident on the device, which in turn forwards it to a recursive resolver, perhaps encrypting that leg of the query. By intention, the recursive resolver maintains a cache of DNS answers, only performing onward queries to authoritative servers as necessary to refresh its cache, and using its own IP address as the origin of such queries. And upon receipt at an authoritative nameserver, by intention, all queries receive the same answer, regardless of who originated them.



But these intentions have been undermined. The rise of the surveillance economy and monetization of the public's personal information has created perverse incentives and eroded the integrity and performance of the DNS by introducing "dark design patterns" into its protocols and usage.

Using Shortening Time-To-Live and Synthetic Labels to "Cache-Bust"

By design and intention, the resilience of the DNS is served by each party to a transaction caching the result, so as to avoid overloading servers above them in the hierarchy of zone data publication with unnecessary repetitive queries, and so that brief failures of servers or the network will not affect most name resolution. Caching also has the privacy benefit of avoiding the creation of any direct association between users' actions and interceptable queries. But it doesn't serve those who would profit from the taking of users' privacy.

Consequently, "cache-busting" techniques have arisen, to destroy the utility and efficiency of the layers of caching hierarchy in the Internet, by causing each individual to send a unique query all the way up the chain each time they act, or even each time their unattended device displays a new advertisement. Two mechanisms are principally used to bust caches: short TTLs and unique synthetic labels.

A simple method of cache-busting is to decrease the Time-To-Live (TTL) of a DNS record. This is a number associated with each DNS answer that is used by the cache as a signal of the answer's "freshness." When Mariko sends an email to a user at Hotmail, her organization's recursive resolver caches the MX records for hotmail.com, such that when Abdul subsequently sends email to Hotmail, he receives a cached answer more quickly, and without unnecessarily using the organization's Internet connection to ask the same question again. This also has the benefit of working equally well whether Hotmail's authoritative DNS servers are reachable and functioning or not. Hotmail can decide how long Mariko and Abdul's recursive resolver retains the answer by setting the TTL of the answer to, for instance, the common default value of 86,400 seconds, one day. Thus, if Mariko was the first person to send email to Hotmail, the MX record would be returned with a value of 86,400, and the recursive resolver would begin decrementing that value by one each second. Two hours later, when Abdul sends an email, the cached answer he receives would have a TTL of 79,200. Another twenty-two hours later, when the counter reached

zero, the record would be expunged from the cache, and the next user to send email to Hotmail would receive a new answer that would, in all likelihood, be exactly the same as the previous answer, yet would have a fresh new TTL of 86,400. The TTL is necessary because Hotmail may periodically wish to rebuild its network fundamentally, bringing up servers on new addresses and deprecating servers on old addresses; without a TTL, users would continue trying to contact the old addresses and would not learn of new addresses. Allowing the domain owner to set the TTL allows it to time such infrastructural cut-overs precisely. But this power can and is abused, not only by domain owners but by those who would surveil users anywhere in the DNS answer distribution hierarchy: if a TTL is set to zero, all downstream caches are told to delete the answer immediately after use and retransmit the query again each time they see it. Zeroing TTLs has the effect of unnecessarily revealing each instance of each user's DNS queries to external parties.

A more complex method of cache-busting is the creation of unique synthetic DNS labels. When users Mariko and Abdul seek to view the same resource online, a tiny hidden beacon on a web page that Mariko's browser accessed will typically be caused by the content distribution network to resolve to the unique synthetic DNS label "mariko37682.content-distribution-network.org", and a similar beacon on Abdul's machine would reference "abdul92783.content-distribution-network.org". This ensures that each time either of them views that page, a unique DNS query will be passed to their stub resolver, which would have its "cache busted" and be unable to resolve it; it would then pass that privacy-destroying unique identifier along to the recursive resolver, which would likewise have its "cache busted," and the recursive resolver would have to, in turn, perform a new unique query to the authoritative server, making Mariko's and Abdul's identities and actions visible to third parties. When these synthetic labels are within the .ORG hierarchy, .ORG authoritative nameservers will see queries to them. When a content distribution network operator has not already done this work for them, the registry can create the opportunity to interject such labels by creating a referral chain, in which a zero-TTL label chains to something that is uniquely assigned and cacheable, which in turn chains to something that has a zero-TTL, creating an entirely DNS-based cookie scheme. Like zeroing TTLs, unique synthetic labels reveal each instance of each user's DNS queries to external parties.

Extended Client Subnet

Extended Client Subnet (ECS) is another mechanism that can be abused to undermine users' privacy. As noted above, by intention and design recursive resolvers shield the identity of individuals performing DNS queries by breaking the one-to-one relationship between user-originated queries and those visible at the authoritative nameserver, and by removing the user's IP address from the query the authoritative nameserver sees. ECS directly and intentionally subverts this barrier by adding the user's IP address or enclosing subnet back into the upstream query as an additional field of data for collection by the authoritative nameserver operator. Because this additional information is monetizable, there is growing evidence that the more rapacious content delivery networks have begun sporadically penalizing users whose recursive resolvers do not betray their identities. By artificially making those recursive resolvers seem unreliable, they steer users toward recursive resolvers that pass their PII along rather than protecting it. This form of net-neutrality-violating punishment of users is one of the tools available to any operator of a top-level domain.

Decryption

Users have come to believe that the encryption offered by DNS-over-TLS (DoT) and DNS-over-HTTPs (DoH) protects them against interception of their traffic. In reality, these provide hop-by-hop privacy, but not end-to-end privacy, and no data integrity. Although this basic encryption may protect users against incidental spies in the network, it does not protect them against spying done by DNS operators, who must be able to read the clear-text of a DNS query in order to provide a response. Recent legal requests by Internet Service Providers in the United States against DoT and DoH operators have exposed the degree of monetization of DNS traffic. DoH will result in fewer data providers servicing this illicit market, thereby driving up the price of DNS data captured by a registry operator.

Attacks Against the Integrity of Communications

Even more dangerous than collecting data about people communicating with .ORGs would be Ethos's capacity to redirect those communications, transparently hijacking connections, providing false cryptographic signatures, and enabling difficult- or impossible-to-detect man-in-the-middle attacks against people who believe they are communicating with authentic .ORG registrants.

DNS Hijacking

Ethos is able to facilitate the hijacking any .ORG domain name due to their position in the domain name resolution process. The substitution of a domain name's name server (NS) records, perhaps via proxy name servers, can allow for the monitoring of end-users communication with targeted domains, and the modification of records assumed to be published by authentic .ORG registrants. DNS hijacking can be targeted to affect only users in specific geographic regions, decreasing the likelihood of detection. Zone files, published via CZDS or other mechanisms, are not guaranteed to match the zone data answered by all authoritative name servers under all conditions.

DNSSEC Requires Trustworthy Parent Domains

The method by which end-to-end data integrity is protected in the DNS is called DNSSEC. It is a suite of cryptographic signing protocols and processes that operate hierarchically. As a domain owner, you cryptographically sign the domain including all of the records within it. These signatures are trusted when Delegation Signer (DS) records are published in the parent zone, just as the domain's name server (NS) records are published to delegate the domain. If they were given authority over the .ORG domain, Ethos could either remove or add DS records to facilitate the hijacking of .ORG domain names. On-the-fly signing provides for difficult-to-detect targeted attacks, literally not visible to anyone other than the victim.

Conclusion

Each of the mechanisms described above is already in common use in the Internet today, though not by reputable operators of top-level domains. It requires no stretch of the imagination to recognize that a newly minted TLD operator, facing a shortfall of $152M/year, would be strongly incentivized to put all of these tools into service immediately, to feed the waiting PII monetization pipeline it had already constructed.