

This document has been prepared by ARI Registry Services in consultation with Neustar, Verisign and Demand Media. This document has also been reviewed by the TMCH-Tech working group and is now offered to the broader community for review and comment.

Should you have any questions or comments about this document please send an email to:

- Chris Wright, Chief Technology Officer at ARI Registry Services - [chris.wright@ariservices.com](mailto:chris.wright@ariservices.com), or
- Post to the ICANN TMCH-Tech working group - to post a message to all the list members, send email to [tmch-tech@icann.org](mailto:tmch-tech@icann.org).

## 1 Issues with the current ICANN proposed TMCH solution

The following is a summary of the issues we perceive with the Trademark Clearinghouse (TMCH) model as is currently proposed by ICANN. The proposed model can be found at the following link:

<http://mm.icann.org/pipermail/tmch-tech/attachments/20120629/92b95395/TMCH-Model-28jun2012-clean-0001.pdf>

In this paper will explore the issues we perceive in the proposed model, starting with sunrise and followed by Trademark Claims.

## 2 Current proposed sunrise model

As registrars, registry operators and backend registry service providers, we have significant concerns about the ICANN proposed sunrise model. In summary the issues we have with this model are as follows:

- Obscure codes, that do not have any relation to the actual domain label they represent  
*This makes diagnosing errors and providing customer support difficult for both registries and registrars.*
- No access to trademark information for registries, meaning this information is unable to be used in applying further eligibility rules, for hierarchical allocation models or displayed in the WHOIS  
This means registries will need to independently request and re-validate trademark information increasing costs for sunrise registrations.
- The unnecessary burden placed on the:
  - TMCH to generate codes for; and
  - Registries to replicate codes for every eligible (for some definition of eligible) mark in the TMCH  
*This is despite only a small percentage of those marks will participating in each sunrise.*
- The unnecessary burden of replicating data to all the different registries  
*There are alternative models that meet all the objectives of all parties and don't suffer from any of the issues raised here without requiring replication of TMCH database data.*

We will now explore each of these issues in detail.

## 2.1 Participation in sunrise

In previous sunrise scenarios registries accepted sunrise registrations based on trademark registration dates, rather than an arbitrary time that one's mark was entered into a clearinghouse.

Some may believe that this may be partially mitigated by giving mark holders several months to validate their mark in the TMCH before any TLDs are able to launch. However as discussed in this paper, we believe that, especially during the early stages of TLD launch, most mark holders' first interaction will be with TLD registries (via registrars, or in most cases by smaller 'resellers' that specialize in online brand management). Only then would they be directed to the clearinghouse—not the other way around.

ICANN have indicated that they can address this concern by providing registries with 'deltas', list of codes that have been added or removed since the original list (or the previous delta) was received. This now introduces an overhead of registries needing to receive and process those deltas; issues with this are discussed in following sections in more detail. Further, depending on the timing of delta generation, and subsequent receipt and processing by registries, a mark holder will have to wait until those deltas have been processed by the TLD they are applying to, before they will be able to successfully submit their application.

The pre-generation of all sunrise codes also introduces a cost that does not consider registration volumes. This will greatly disadvantage small registry operators such as 'single-registrant, single-user' registries (e.g. for .brands), that may not even register any names during the compulsory sunrise period) but may have to pay fees similar to an 'open' TLD that is expecting large registration volumes.

This can be potentially mitigated by allowing each TLD to specify specific criteria, and only require the generation of codes for marks that match that criteria. However, in the case of an open TLD, the potential is for there to be millions of marks in the TMCH, but at most only 100,000 registrations may be received; it is a lot of effort to generate and distribute codes for millions of marks for each registry when less than 10% of them are actually used. We believe that this is very inefficient and introduces unnecessary.

## 2.2 Obscure Codes not Explicitly Related to the Mark Name

The obfuscation/encryption used to generate codes in the ICANN model makes it impossible for a registry operator and/or a registrar to provide any level of support to the rights holders.

In the case of a rights holder supplying an incorrect code, neither registry nor registrar will be able to inform the rights holder whether:

- The code is not valid for the TLD,
- The code is not for the provided name,
- The code was copied incorrectly,
- The registry has incorrect data for the provided name, or
- There was some other error.

Neither the registry operator nor registrar will be able to determine the error or examine the code to determine the TLD and mark name. The rights holder will have no choice but to receive support through the TMCH. As a result of the number of parties involved the TMCH will have difficulty resolving issues.

If the current proposed sunrise model were adopted in its current form, the TMCH would be required to provide 24/7 support to registrars, registries and mark holders and stringent SLAs for resolution of issues, which will undoubtedly increase the cost of providing the service.

## **2.3 Access to trademark data**

The current proposed sunrise model does not provide registries with access to the trademark data being used for the registration as the reason they are eligible to participate in the sunrise. This is problematic for a number of reasons which are discussed in the following sections.

### **2.3.1 Eligibility for rights holders**

Registry operators may restrict eligibility to rights holders that have a mark in a particular jurisdiction or class of goods (or indeed some other criteria applied to the mark). In order to validate such eligibility, under the current proposed sunrise model, a registry will require the rights holder to submit additional information about the mark they are using during the registration process. This additional information will then need to be validated by the registry, the cost of which will be passed on to the potential registrant. As such information is the same information that has already been supplied and validated by the clearinghouse, thus this increase in cost and effort is unnecessary.

It has been suggested that requesting that the clearinghouse only provide sunrise codes for marks that meet a certain criteria can mitigate this issue. In this instance registries, without having means to verify would rely on the clearinghouse to have done this correctly. Registries would seek legal indemnifications from the clearinghouse to protect them against any issues arising from incorrect allocation of sunrise codes to non-eligible parties; something that is likely to only increase the TMCH costs, thus preferable to avoid. This also does not leave registries 'in control of their own destinies' and leaves them having to rely on a third party whom they had no hand in choosing, or have no power to dismiss if their performance is not acceptable.

### **2.3.2 Allocation methods**

Registry operators may implement an allocation method that preferences jurisdiction of mark or class of goods. An example is a geographic TLD giving preference to marks registered within its locality over marks from other jurisdictions when resolving contention between applications.

This differs from the issue identified in section 2.3.1 above in that whilst a particular set of marks is eligible, smaller subsets of those marks are given higher preference than others.

In the current proposed sunrise model, as previously described, additional data will also be required to be collected to meet this requirement and, as above, this will result in increasing costs for all and the requirement of further indemnifications from the TMCH.

Simply separating the codes into hierarchal lists may not be sufficient to meet this requirement, because sometime the allocation techniques may not be 100% objective and may be open to subjective decisions being made by the registry operator, whom, without access to the information will be unable to make such a determination.

Creating unique set of codes based for different eligibility policies will result in creating sub-phases for sunrise that increases the complexity for all parties in the ecosystem:

- TMCH in having to generated different sets of codes,

- Registry for sub-dividing their sunrise into smaller phases and handling multiple sets of code that can't be validated,
- Registrar for having to deal with multiple sunrise sub-phases with different supported codes, and
- The registrant who would also need to be aware of the sunrise sub-phases and what marks they can use in each sub-phase).

Further to this, registries will need access to trademark data used in registration of domain names to facilitate investigations of abuse, or trademark disputes, this is critical to the proper executions of those functions.

### 2.3.3 Transparency

Lack of trademark data reduces transparency of domain names allocated through the sunrise process. Previous sunrises have included trademark data in the WHOIS responses of domain names registered during sunrise processes. An example is observable by performing a query for microsoft.info, taking notice of the trademark information in the beginning of the response.

The inclusion of such data in the WHOIS response provides visibility for those external to the sunrise process, including other rights holders, and helps when organizations are making a decision to challenge the eligibility of a registrant to the given domain name, or when other mark holders are considering making UDRP or URS complaints. Note that one of the steps in URS and UDRP is to provide WHOIS details to the UDRP/URS provider, by making the complainant perform a WHOIS, and then seeing the trademark information, the complainant may then decide to not, unnecessarily, make a complaint.

Not publishing this information in the registry WHOIS, will increase the difficulty for entities to decide if they have a case against a domain name registrant, thus most likely increasing the number of redundant cases being raised through UDRP and other processes.

## 2.4 Generation and replication of unrequired data

The current proposed model involves the replication of codes for **every entry** in the TMCH to all TLDs. Each TLD has its own distinct replicated copy of the marks, using different cryptography keys, thus registry service providers, managing multiple (in some cases hundreds) of TLDs, will be required to maintain (download, store and update) hundreds of distinct copies of the same data for no clear reason.

This process is unnecessarily cumbersome and costly. The current proposed model requires registries to store data that grows linearly to the number of TLDs. This is further exacerbated given that expected data set sizes and update frequencies have not been approximated.

Registries are also required to keep a large data set of which they only really need access to a small portion under the current proposed model. If there are millions of entries in the clearinghouse, however only 10,000 domains are registered during the sunrise period, the registry unnecessarily had to replicate codes for (and the TMCH had to generate codes for), more than 99% of the marks in the clearinghouse. In a .brand scenario it is highly likely that 99.99% of the codes are useless to the registry and will never be used.

### 3 Current proposed claims model

As registrars, registry operators and backend registry service providers, we have significant concerns about the ICANN proposed claims model. In summary the issues we have with this model are as follows:

- The unnecessary burden placed on the TMCH and registries to replicate data for every entry in the TMCH even though only a small percentage of those marks will match domain names registered during the claims period (first 60 days).
- The TMCH outage problem.
- The burden for registries to have to protect and securely manage trademark data.
- Unnecessary and obscure encryption of data, with no relation to the actual domain label they represent, make diagnosing errors and providing customer support difficult for both registries and registrars.
- The risks and issues associated with registries and registrars utilizing stale data
- Trademark data is not protected since it will be easy for any party (registry, registrar, registrant) to mine the data in the encrypted files – replicated copies of the data broaden the data mining attack vector.
- Replication delay increases the risk that registrations will go through without the required claims notice being displayed - why would the mark holders accept a long delay (days to a week) where their trademarks don't result in the required claims notice being displayed to registrants and them not getting notified of any matching registrations?
- The issue of the time lag between when a potential registrant accepts the claims notice and the actual registration of the domain name takes place. Situations such as when registrars accept pre-registration requests in anticipation of the opening of first come – first served or in a land rush scenario where applications for names are submitted over a period of time and then allocated at the closing of that period need to be considered with respect to the exact 'how' trademark claims notices are supposed to work. We certainly want to try to avoid (where possible) the situation where a potential registrant is not notified about matching claims in the TMCH, yet the trademark holders are notified of the proceeding registration.

We will now explore each of these issues in detail.

#### 3.1 Replication of data

The current proposed claims model involves the **entire** TMCH database being replicated to each registry. This process is subject to a number of concerns as follows.

##### 3.1.1 Maintenance

Each TLD has its own distinct replicated copy of the database, using different cryptography keys, so registry service providers, managing registry services for multiple (in some cases hundreds of TLDs), will have to maintain (download, store and update) hundreds of distinct copies of the same data for no clear reason. This is unnecessary and costly.

The current proposed claims model requires registries to store claims data that grows linearly to the number of TLDs. This is further exacerbated given that expected data set sizes and update frequencies have not been approximated.

### 3.1.2 Security

Registries are required to keep secure and contractually protect a large data set of which they only really need access to a small portion of.

If there are millions of entries in the clearinghouse, however only 100,000 domain names are registered during the claims period, the registry unnecessarily had access to, and had to protect, more than 90% of the TMCH data. In a .brand scenario it is highly likely that 99.99% of the data is useless to the registry and will never be used.

The current proposed claims model is fundamentally flawed, goes against basic security principles, and by burdening registries with unnecessarily taking on this responsibility, most registries will seek indemnification against the exposure of this data; whether by mining by a third-party, security breaches or any other method.

### 3.1.3 Unknown requirements

Registries are also concerned with the unknown nature of the problem space. Currently the expected volume of data (both initially and updates), frequency of updates, and other metrics have not been provided nor discussed, affecting the ability of a registry to plan for the deployment and operations of the critical launch phases of new TLDs.

## 3.2 The TMCH 'Outage' problem

It is **extremely** important to note that the proposed model does not solve the 'TMCH' outage problem; it simply changes its form.

Previous proposals have revolved around the TMCH maintaining a 'query' interface of some sort for the TMCH data. This solution has its issues when we consider the situation that arises when that query interface is no longer available due to an issue.

The current proposed solution does not eliminate this situation, consider the following scenario:

*What if during the rsync process with the TMCH, the TMCH has a malfunction of some sort that results in the incorrect data being rsynced to the registry?*

This is indeed extremely feasible, especially with the TMCH being asked to managed one, encrypted, obfuscated, not easily inspected or verifiable model of the data for **each registry**. In addition given the encryption, in some case the registry would have a hard time even knowing that the data that they have is incorrect (or gibberish). Depending on the encryption used, they will either not be able to decrypt the data or, decrypt it and get garbage as the response.

Should this situation arise, the registry cannot have an unplanned outage and still needs to continue taking registrations - What are is the registry to do? This is exactly the same issue as when an online query interface is not available, there still **must** be an acceptable action plan for registries when the TMCH data is not available and this **cannot** be 'stop taking registrations' (unless the TMCH will to financially compensate registries for the outage time and ICANN will relax SLA commitments for the period the data is unavailable - this however is still extremely unfavorable).

## 3.3 Use of encryption

In the view of the authors the current proposed encryption model is meaningless. The encryption only exists to stop registries reading the entire TMCH database; it serves no other purpose.

We recommend that should data still be provisioned with a registry, that the data is provided without encryption. Most registries do not have any interest in this data, however encrypting it assumes that registries are bad actors. However any registry that is a bad actor can easily get access to the data anyway, as described in the following section.

As was discussed during the call on 14 May 2012, encryption is not enough to restrict access from the registry and contractual agreements will need to be put in place. If there is reliance on contractual provisions, then encryption should be removed to allow registries to better meet their requirements.

### **3.3.1 Assumption of a bad-actor registry**

This use of encryption assumes registries are non-trusted parties who are 'bad-actors'. Assuming a bad-actor registry, it is relatively trivial to gain access to most (if not all) of the database. This is because the encryption key is made up of something the registry already has (the registry part), and something that is determined from each domain name; all that is needed to decrypt the data is a list of domain names.

The zone file access service which provides access to the .com and .net zone files, gives access to 100,000 million domain names to check against the database. Thus running the following 'pseudo code' would likely result in the decryption of most, if not all data.

```
for each name in source of domain names (.com & .net zone file)  
  determine if name is subject to trademark claims  
  if subject to trademark claims then decrypt claims data  
done
```

Given that most marks will be registered in .com (whether by the mark holder or not) the probability of getting greater than 90% of the database with this method is high.

### **3.3.2 Round-trip time for queries**

The use of encryption does not align with the goals of a registry operator, whose interest is to provide the lowest round-trip time for queries due to contractual penalties applying if round-trip-times exceed a specific threshold.

It should be expected that a pre-computation of the TMC data would be considered by registries, to remove the cost of decryption from the round-trip-time of claims lookups. Having access to the decrypted data also allows a registry operator to verify the integrity of claims data, and to have knowledge of the data that the registry will be transmitting (maximum length of claims notices, average length of claims notice)—required for testing of software and capacity planning. Such upfront analysis removes the possibility of run-time failures (such as failure to decrypt data) affecting domain name registration.

In short, registries, especially back end registry service providers, may have legitimate reasons for using the process described above to decrypt the information provided by the clearinghouse.

### **3.3.3 Data mining concerns**

The proposed encryption does not prevent data mining by parties other than registries. Instead, distribution of claims data to multiple registries will actually increase the ability of calculated attempts to mine claims data.

Malicious users can distribute their queries amongst a combination of registrars and registries, never providing enough information to any one party to determine mining. Should

mining be detected, damage would be done before the application of remedies across each of the registries and/or registrars.

#### **3.3.4 Variants**

Encryption also prevents registries from calculating mark name variants. The issues are identical to those described in response to the sunrise process above, thus the recommendation is to remove the encryption of mark names.

### **3.4 Allocation**

The proposed model appears to only consider allocation based on a first-come first-served model. Consider a land-rush process where potential registrants submit applications for domain names and resolve contention via auction. This would create a separation between the submission of the application, and subsequent allocation (if any) of the requested domain name.

The concern is that the submission of notices to rights holders before allocation could be considered front-running if the rights holder subsequently filed an application for the said domain name. In response to this, applications might be submitted in the closing hours of the application window, negating the purpose of reducing the rush of applications in a short period.

At a minimum, we expect clarification of the process for auction-based allocation methods, which should involve the removal or otherwise definition of the term 'registration'.

There is also the need to consider issues around timing between display of notices and when the actual registration goes ahead. These issues are not specific to any one implementation model (at least based on all implementation possibilities presented to date), however are more generic issues with the TMCH in general which need to be resolved.

## **4 Conclusion**

We believe that this current proposed solution is complex and does not meet the goals of Registrars and Registries.

The current proposed solution increases the cost (time, effort and expense) for all actors participating in it, and remains open to significant flaws. We strongly urge ICANN to consider the issues raised in this document and consult with the affected parties in their resolution.