

**Priority:** N5

**Issue:** Matching Rules

**Description:** The Trademark Clearinghouse will perform numerous comparisons between strings, with parameters for identifying an “identical match” as defined in the Applicant Guidebook. Because of the plethora of languages and character sets used in trademarks, and the limited set of DNS-permissible characters, clear definition of both the principles and rules by which matches are to be permitted must be established to ensure transparency of Clearinghouse operations.

**Business**

**Requirements:** (1) Ensure predictability and reproducibility of matches  
 (2) Avoid divergence from the legally recognized protection of trademark rights  
 (3) Provide transparency into rules and processes used in Clearinghouse processes

### **Issue 1: Applying the Identical Match rules**

As noted in the published Trademark Clearinghouse model, “Identical Match” means that the domain name consists of the complete and identical textual elements of the mark. In this regard:

- a) spaces contained within a mark that are either replaced by hyphens (and vice versa) or omitted;
- b) only certain special characters contained within a trademark are spelled out with appropriate words describing it (@ and &);
- c) punctuation or special characters contained within a mark that are unable to be used in a second-level domain name may either be (i) omitted or (ii) replaced by spaces, hyphens or underscores and still be considered identical matches; and
- d) no plural and no “marks contained” would qualify for inclusion.

Some interpretation is required of these rules, as described below.

#### Rule B.

When translation of a “rule b” special character into appropriate words describing it (in the case of ‘@’ and ‘&’) must occur, should it occur into a given language (for example: the UN languages, the intended language(s) for the domain name registration, the national language(s) of the jurisdiction where the mark is registered)? How could the appropriate language(s) be determined for each case? Alternatively, should a single set of languages be established as a standard to apply in all cases? How would this be determined?

Consider the fictional mark “X & Y.” Looking only at English, French and Spanish, the ampersand could be spelled out as “and,” “et,” or “y,” causing the mark to match to “X and Y,” “X et Y” or “X y Y.” What circumstances will dictate which language(s) are to be used?

For a Sunrise period, these issues will have an impact on the number of names a trademark holder is eligible to register.

For a Trademark Claims service, this issue will have an impact on the number of Claims notices generated and displayed to prospective domain name registrants.

### Rule C.

When a mark contains more than one “rule c” special character that is to be dropped or transformed into a dash (to be a DNS-permissible character), an expanded number of relevant strings will be generated (i.e., the rule applies more than once for the same string).

If a mark contains, for instance, three “rule c” characters, this means that there will be 8 possible matches. Consider the fictional mark “a’b:c,d” -- how it might be presented. This would be an identical match to “abcd,” “a-bcd,” “a-b-cd,” “a-b-c-d,” “ab-cd,” “abc-d,” “a-bc-d,” “ab-c-d.”

### **Issue 2: Registry Character Mappings**

The Unicode standard provides a repertoire of code points used in world scripts, including various classifications of character properties, and normalization rules. The Internationalizing Domain Names in Applications (IDNA) protocol specifies rules for determining whether a code point is a candidate for inclusion in domain names. An internationalized domain name label can be represented as a Unicode string or an ASCII string. Taking the IDN test TLD string in Cyrillic script as an example, the U-label is <испытание> and the A-label is <xn--80akhbyknj4f>.

It is expected that no characters outside the Unicode standard would be included in trademark records in the Clearinghouse. This should not pose a difficulty as the vast majority of characters in common use are included in this standard. Additionally, except as provided in the matching rules above, a trademark containing characters that may not appear in a domain name would not be a basis for a Sunrise registration or a Trademark Claims notice.

- How should the Clearinghouse service the different character mapping rules that could occur in different new TLD registries?

Registries may establish rules and policies for characters to be allowed in the TLD, including, in some cases, “variant” characters or characters that are mapped to one or more other characters in some way. At the present time, registry practices differ, and so it is likely that a number of character mappings may apply in some registries and not others. As is the case presently, this results in a different experience across various TLDs.

In consideration of registry character mappings and the Clearinghouse services, there are two broad approaches:

- (a) One is that the registry maintains the responsibility for integrating its character mapping policies into the Sunrise and Claims services.

For example, in the Trademark Claims service, where a registry policy maps “e” to “é” and a domain name applicant attempts to register the string “xyz” – the registry could query both “xyz” and “éxyz” to determine whether there is a match to a Clearinghouse record, and provide the results for both queries to the domain name applicant. However, the registry should also ensure that the domain name applicant is provided

an explanation that this is occurring according to the registry policy, to aid in the interpretation of the Claims notice information.

In the case of a Sunrise period, the registry could take the information on Sunrise name eligibility generated from a Clearinghouse record and apply the registry rules to generate the additional eligible names based on that record, as appropriate.

- (b) In the other approach, the Clearinghouse would be provided the mapping rules for each TLD, and would need to have a mechanism for applying the correct set of rules in a given case. Essentially, the definition of "Identical Match" would be supplemented with the registry-specific matching rules to be applied.

Under this approach, the results returned for a Clearinghouse query would differ according to the TLD and its rules, and the same string might trigger a Claims notice in one TLD but not another.

For a Sunrise period, the Clearinghouse would be able to generate the eligible strings for a Clearinghouse record according to both the Identical Match rules, and the registry mapping rules.

In this case, the issue arises of what process requirements are desirable with respect to adding, removing or changing the Clearinghouse mapping rules used for a given TLD. For example, is it acceptable to change the mapping rules applied to a given TLD during a Sunrise or Claims period? If not, how long prior to the new TLD's Sunrise should the rules be established?