

2014年7月18日
第49回JPドメイン名諮問委員会
資料5

ICANNにおける国際化ドメイン名 (IDN) TLDに関する検討状況

2014年7月18日(金)
株式会社日本レジストリサービス

IDNのルール

- IDNの技術規格
 - IETFが制定
 - ドメイン名登録管理の規則はこれに従う
 - IDNを扱うアプリケーションはこれに従う
- ドメイン名登録管理でのIDN文字列のルール
 - ルールとして次を定義
 - 文字列の中で使用可能な文字の集合
 - その集合の中での異体字(後述)の定義
 - TLD(「.△△」の△△部分)の場合
 - ICANNが統ルールを制定
 - 2LD, 3LD, ... (例えば「〇〇.JP」の〇〇部分)の場合
 - 各ドメイン名レジストリが制定
 - 例えば、「〇〇.JP」のルールはJPRSが制定
 - 使用可能な文字は、漢字+ひらがな+カタカナ+ASCII
 - 英大小文字は異体字であるが、それ以外の異体字は定義せずすべて異なる文字とみなす

TLDにおけるIDN文字列のルール

- ICANNが統一ルールを制定
 - TLDは世界中のインターネットユーザーが使うので、TLDにおけるIDNにはさまざまな言語の文字が使われる
 - 各TLDは、どの言語の文字列であるかがTLD申請時に決められる
 - ただし、TLDを構成する文字列は複数の言語の文字を組み合わせてはいけない
 - 文字列の組み合わせはその言語で使用可能と定義された文字の中からのみ
 - ある言語のTLDを構成する文字列のルール(言語ルール)は、その言語のコミュニティがICANNに提案し、ICANNが提案全体を統合して統一ルールを作成する
 - 言語コミュニティの代表: 生成パネル(Generation Panel)
 - 統一ルール作成: 統合パネル(Integration Panel)
 - 漢字など、複数の言語で共有される文字は、その文字を共有する言語コミュニティが協力して統一ルールの基礎を作成する
 - 漢字の他には、アラビア文字、インド系文字、ラテン系文字などがある

IDN TLDに関して解くべき課題

- 各言語で使用する文字
 - TLD文字列の中で使用可能な文字の集合
 - その集合の中での異体字の定義
- 複数の言語が文字を共有する場合の異体字の扱い
 - 異体字とは
 - 同じ発音・意味を持つが異なる字体の文字のこと
 - 漢字の「国」と「國」など
 - 異体字の定義は言語、国、文化、使用される文脈で異なる
 - 課題
 - 各言語ルールにおいて異体字をどう定義するか、どう扱うか
 - TLD(ルートゾーン)用のルールで、異体字を持つ言語ルールをどのように統合するか

漢字での異体字

- 中国語の異体字
 - 中国語は複数の国や地域で使われており、中国本土では簡体字(例えば国)、それ以外の多くの地域では繁体字(例えば國)が使われている
 - 繁体字と簡体字が入り混じって使われることはほぼない
 - グローバルに使われる中国語ドメイン名という観点では、異体字を定義して等価な文字という扱いをしないと混乱が生じる
 - CNNIC, TWNIC等により作られた2LD, 3LD, ...用の言語ルールが公開されている
- 日本語の異体字
 - 日本語でも異体字の概念はあるが明確な使い分けの基準はJISでも定義されていない
 - 文脈によって異体字は入り混じって使われる(例えば「四国在住の國中さん」)
 - そのため、日本語JPドメイン名では異体字を定義していない
 - JPRS(元はJPNIC)により作られた2LD, 3LD, ...用の言語ルールが公開されている

漢字に関する検討状況

- 生成パネル
 - 中国・台湾等は中国語生成パネルを設立
 - 韓国は韓国語生成パネルをほぼ設立
 - 日本は日本語生成パネルの設立準備中(参加候補者に相談中)
- CJK間の協力
 - TLDでは文字列申請時以外に言語の区別ができないため、漢字に関しては中国語も日本語も韓国語も統ルールで扱う必要がある
 - 例えば、
 - 「.国中」と「.國中」のレジストリが別にならないようにする
 - 「.国中」と「.國中」のどちらが(もしくは両方が)TLDとして使えるかを定める
 - そのための協力体制を中国・日本・韓国・台湾のccTLDレジストリが中心となって開始した
 - 各言語パネルが協力して統ルールの基礎作りを実施する
 - これらはすべて、ICANN IDN TLDに関連する活動の一環として行われている