

TLDでの1文字IDN 及び SLDでの日本語ドメインラベル について

2022年10月20日

株式会社日本レジストリサービス (JPRS)
堀田 博文 <hotta@jprs.co.jp>

- 1 TLDでの1文字IDN
 - 2 SLDでの日本語ドメインラベル
- おまけ RZ-LGR celebration

1 TLDでの1文字IDN

1文字だけのTLDもOK？

- これまで、ASCII、IDNいずれにも、1文字だけのTLDは許されていなかった
- 次の新gTLD募集で1文字TLDを許すべきか、許す場合の条件は、という検討がGNSOのIDNs EPDP WGで始まった
 - EPDP : Expedited PDP (迅速なポリシー開発プロセス)



- あらゆるスクリプトの中で、表意文字である漢字だけが1文字TLDとして許されるべきとの中間的判断がWGで下された
- 漢字の中でも、使用が許される漢字と許されない漢字を決めるための検討がCJKの言語コミュニティに託され、その初期打合せがICANN75内で持たれた

ICANN75での議論

- IDNs EPDP WG及びICANNの呼びかけにより、CJKの旧生成パネル^(注)代表が集まり、漢字1文字TLDのあり方に関する意見交換を行った
- この意見交換では、次の方向がラフコンセンサスとなった
 - a. 漢字1文字TLDは可能とすべきである
 - b. 1文字TLDとして許すべきでない文字をリストアップする
 - c. 記号等の表意文字でない文字(例えば「々」)は許すべきでない
 - d. 他スクリプト内の文字と酷似している文字は許すべきでない
- 今後の進め方として、以下が合意された
 - b. c.についてはCJKの各言語コミュニティが検討
 - EPDP WGより正式に検討要請が来る予定
 - d.については今後の進め方をさらに検討

(注) 生成パネル: 各言語のルートゾーンLGRを検討するチーム(言語ごとに存在)

日本語から見た考え方(案)

- b. 1文字TLDとして許すべきでない文字をリストアップする
 - c. 記号等の表意文字でない文字(例えば「々」)は許すべきでない
 - d. 他スクリプト内の文字と酷似している文字は許すべきでない
- 日本語コミュニティとしての考え方は？
 - 記号(MARK や SYMBOL) は1文字TLDとして許すべきでない
 - UNICODE で記号に分類されている文字の例
 - U+3005 々 IDEOGRAPHIC ITERATION MARK
 - U+3006 々 IDEOGRAPHIC CLOSING MARK
 - それ以外、適切な根拠で「不適切な文字」を抽出することは困難
 - UNICODE consortiumの文書では、漢数字零「〇」は「々」「々」と並んでCJK Symbols and Punctuationとして記されている
 - しかし、「〇」はMARKではないし、「一」～「九」とともに漢数字を構成するので、漢数字零「〇」はb. c.の基準では1文字TLDとして許されるべき
 - ただし、d.の基準で許されない可能性あり

2 SLDでの日本語ドメインラベル

ICANN側から提示されたルール案

- 今、ICANNでは、セカンドレベルでのIDNラベルに関する参照ルール(セカンドレベルに登録するIDNの実質的標準仕様)を作る作業が進行中
- ICANNは、原則としてRZ-LGRをベースとするルールをSLDにも適用すべきという方針で作業
 - SSACのドキュメントSAC060でも、その方針が明示されている
- この度、ICANNからセカンドレベルでの日本語ラベルの参照ルール案が非公式に旧JGPチェアに提示された
- その仕様は、詳細を除き大まかに言うと
 - レポートリー：RZ-LGRのレポートリーに英数字を加えたもの
 - 異体字：RZ-LGRと同様に、中国語RZ-LGRと韓国語RZ-LGRで定義された異体字を取入れたもの

日本語コミュニティからの反応案

- しかし、RZ-LGRで行った中国語や韓国語からの異体字取入れは、日本語の使い勝手を阻害している
 - たとえば、辺と邊と邊、面と麵、髪と発、乾と幹と干、働と動、堺と界は中国語LGRで異体字であり、片方を含むラベルが登録されていると、他方の文字に入れ替えたラベルは登録できない
 - つまり、「渡辺」と「渡邊」と「渡邊」は、異体字ラベルなので、「渡邊」という日本語ラベルが登録されていると、「渡辺」も「渡邊」も登録できない
- そこで、ICANNに対し、次のように2つのタイプのLGRがSLD用に必要という提案をすることを考えている
 - type-1: 日本語を含む複数言語が共存するSLDに適用するルール
 - ICANNが提案しているもの
 - type-2: 日本語のみが存在するSLDに適用する参照ルール
 - 中点「・」と漢数字零「〇」をレパートリーに加える
 - 日常的にも使われ、商標や商号にも使われている
 - 漢字間の異体字は存在しないものとする

おまけ RZ-LGR celebration

RZ-LGR : TLD文字列に関するルール

- TLDには様々な言語・scriptのラベルが存在
- 様々な言語・scriptのTLDラベルにおいて、以下の標準ルール (Root Zone Label Generation Rules; RZ-LGR) が必要
 - 使える文字の集合
 - 文字間の異体字関係と異体字ラベル(例:”大学”と”大學”)の使用可否
 - ラベルの適切さに関するルール (例:”っん”はダメ)



- 2014年より標準ルール策定活動を開始
- 新gTLDの次の申請ラウンドへの適用が目標

- 言語ごとに生成パネル(GP: Generation Panel)を編成し、それぞれの言語用の標準ルールを策定



- ICANNは、2022年5月26日に第5版のRZ-LGRを制定、公開し、RZ-LGRを構成すべきほぼすべての言語が出揃った

RZ-LGR 第5版 に含まれる言語・スクリプト

Arabic	Kannada
Armenian	Khmer
Bangla	Korean
Chinese	Lao
Cyrillic	Latin
Devanagari	Malayalam
Ethiopic	Myanmar
Georgian	Oriya
Greek	Sinhala
Gujarati	Tamil
Gurmukhi	Telugu
Hebrew	Thai
Japanese	

2022年5月に第5版に組み込まれて公開

Celebrating GP's Achievement on RZ-LGR



in Kuala Lumpur Convention Center

ご質問? ご意見?