

日本語ルートゾーンLGRに関する報告

2022年7月28日

日本語生成パネル チェア
堀田 博文 <hotta@jprs.co.jp>

IDN TLD文字列に関するルール

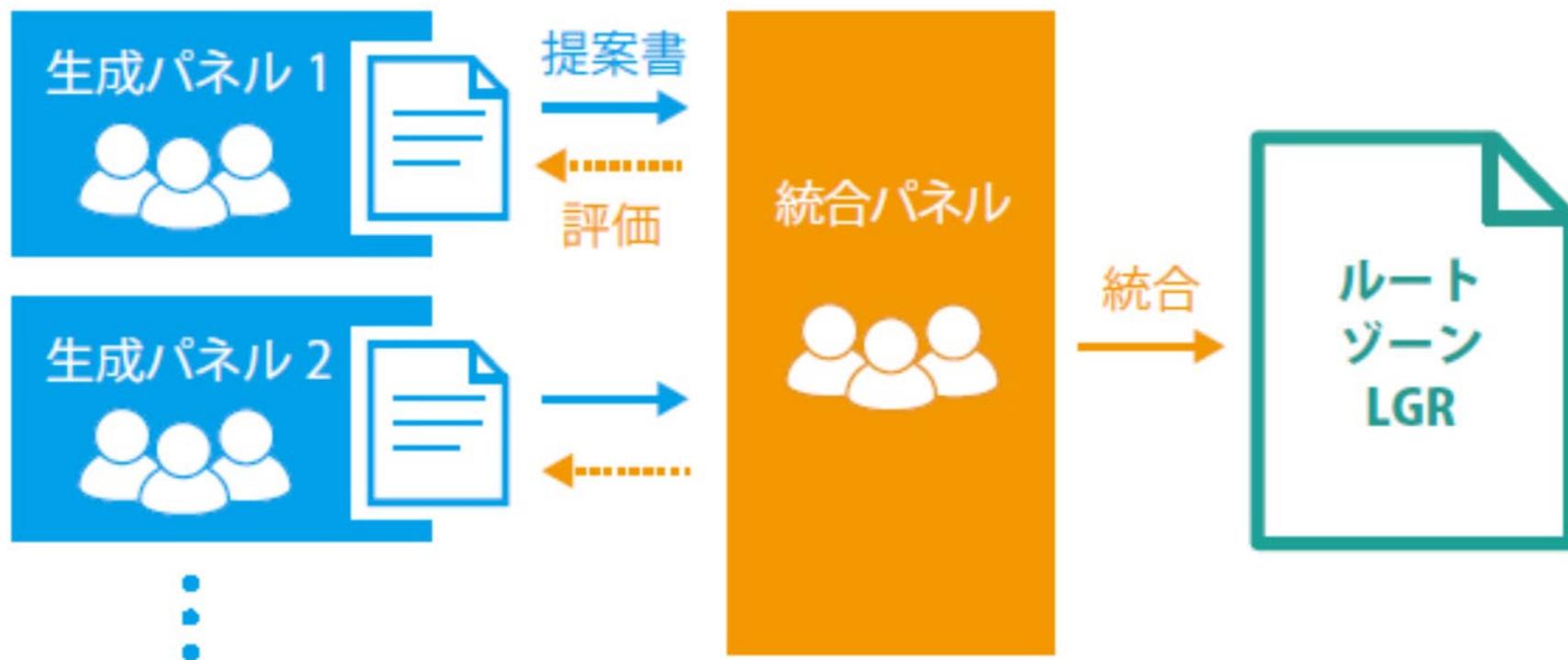
- TLDには様々な言語・scriptのラベルが存在
- 様々な言語・scriptのTLDラベルにおいて、以下の標準ルール (Root Zone Label Generation Rules; RZ-LGR) が必要
 - 使える文字の集合
 - 文字間の異体字関係と異体字ラベル(例:”大学”と”大學”)の使用可否
 - ラベルの適切さに関するルール (例:”っん”はダメ)



- 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
- scriptを共有する言語同士は標準ルール作成において調整
 - 例: CJK(中国語/日本語/韓国語)は漢字を共有
 - 中国語生成パネル(CGP)、日本語生成パネル(JGP)、韓国語生成パネル(KGP)が調整

RZ-LGR作成プロセス



GP : Generation Panel

IP : Integration Panel

日本語生成パネル(JGP)の活動経緯

- 2014年 8月29日(金) JGP第1回準備会合(準備会合 計2回)
- 2014年 9月24日(水) JGP第1回会合
- 2015年 2月 6日(金) ●ICANNがJGPを正式な生成パネルとして認定
- 2015年 5月15日(金)～ ●CJK調整会合第1回
- 2015年 9月29日(火) ●日本語RZ-LGRの基本方針を合意
 - + レパートリー(文字集合)
 - + Jは異体字なし + CKの異体字定義をインポート
 - + 独自WLE(文字順序等のルール)は定義しない

* 記号文字の文字位置に関するものは除く
- 2017年 3月 ●CK間で異体字の定義に合意
- ~2021年3月 ●文字対の視覚的同一性に関し2件のフィールド調査
- 2021年 ●IPとの間で視覚的類似文字定義を合意
- 2021年 9月 ●IPとの間で「申請文字列のみ登録可能」を合意
 - + 試行錯誤・IPとの議論・国内意見聴取ののち判断
- 2022年 5月26日(木) ●日本語RZ-LGRが完成しRZ-LGR-5に組み込まれICANNから公開
- 2022年 5月31日(火) JGP解散 (JGP会合 計14回)

日本語RZ-LGRの概要

- 使用可能な文字の範囲(レパートリー)
 - JIS X 0208:2012の第一水準・第二水準の範囲(漢字、平仮名、片仮名、漢字および仮名に準ずる一部記号文字からなる約6,300文字)
- 異体字の定義
 - 日本語独自の異体字
 - 字形は異なるが読み・意味が同じ文字(例:”学”と”學”)は異体字としない
 - 視覚的に同一(もしくは酷似)な文字同士は異体字とする
 - Unicodeコンソーシアム公開の「錯視が起こりやすい文字同士リスト」に基づく10組
 - 酷似する1ストロークの記号2組: U+ [30FC(一) 4E00(一)], [30FD(丶) 4E36(丶)]
 - 中国語/韓国語LGRで定義された異体字
 - 日本語LGRに取り入れる(日本語LGRでも異体字として扱う)
- ラベル文字列全体に関するルールの定義
 - 長音記号、繰り返し記号は、TLDの先頭文字として許されない
 - 捨て仮名(拗音・促音)もTLDの先頭文字として許されない
- 異体字ラベルの使用可否
 - 申請されたラベルのみが使用可能で、その異体字ラベルは使用不可

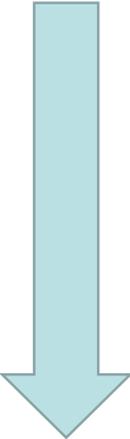
視覚的同一文字による異体字ラベルの扱い

- 視覚的同一文字

- | | | | |
|------|------|------|------|
| ・ へへ | ・ ハハ | ・ べべ | ・ トト |
| ・ ペペ | ・ ロロ | ・ ニニ | ・ タタ |
| ・ カカ | ・ エエ | | |

- 視覚的同一文字による異体字ラベルは区別がつかない/つきづらい

- 「.ヘリコプター」と「.ヘリコプター」 注) 2つ目はひらがな「へ」と漢字「タ」
- 「.ソニー」と「.ソニー」 注) 2つ目は漢字「ニ」と漢字「一」



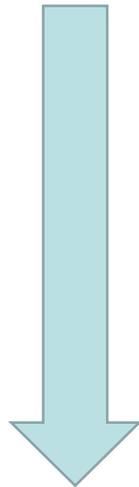
基本的考え方

- 異体字ラベルは、別登録者には与えない
- 多様な利用者が目にするTLDとして、錯視されやすいものは共存すべきでない

- 申請ラベルのみを使用可能とする

異体漢字による異体字ラベルの扱い

- 日本語LGR独自の異体字ラベルは存在しない
- 中国語/韓国語LGRで定義された異体字は日本語LGRに取り入れる(日本語LGRでも異体字として扱う)
- 結果として、異体字ラベルが定義される
 - たとえば、「.国立大学」を申請すると、付随する異体字ラベルとして、「.國立大学」、「.国立大學」、「.國立大學」が存在する



- 異体字ラベルは、別登録者には与えない
- 異体字ラベルのうちTLDとして使えるものは最小限(せいぜい1~2個)
- 複数個生成される異体字ラベルのうち、申請者が申請ラベル以外にどれを使いたいかを第三者が一般ルールとして静的に決めることは不可能

- 申請ラベルのみを使用可能とする

日本語生成パネル(JGP)メンバー

- メンバー (2021年8月時点)
 - 堀田博文 議長
 - Registry/Registrar (esp. Policy/business aspects)
 - 前村昌紀 副議長
 - Policy (esp. Internet governance and domain name in general)
 - 後藤滋樹
 - Community (esp. Internet technology in general)
 - 小西和憲
 - Community (esp. Internet technology in general)
 - 久保次三
 - Community (esp. Trademarks and DRP)
 - 高松百合
 - Registry/Registrar (esp. Policy and International relationship aspects)
 - 田代秀一
 - Language expert (esp. Character codes)
 - 村上嘉隆
 - Registry/Registrar (esp. Trademarks and gTLD market)
 - 米谷嘉朗
 - Registry/Registrar (esp. technical aspects of IDN and LGR)

ご質問？

もう少し知りたい方は、JPRS トピックス & コラム No.025 をご覧ください
<<https://jprs.jp/related-info/guide/topics-column/no25.html>>