

2015年2月20日  
JGP会合(7)  
資料5

# 日本語LGRの検討課題と 解決方針

# 日本語LGRで決めるべきこと

- a. 文字範囲(Repertoire)
- b. 異体字(Variants)
- c. 処置(Disposition) (\*1)
- d. ラベル評価ルール(Whole Label Evaluation (WLE) Rules) (\*2)

(\*1) Rootゾーンにおける異体字の取り扱いを決めるもので、個々の文字に対して定義される。処置には「登録可能」、「ブロック」などがある。

(\*2) 個々の文字が組み合わされた文字列(ラベル)の有効性を評価するルールで、例えば、中国語では簡体字と繁体字の混在は登録不可能とする、などが定義される。

# a. 文字範囲

- 考えられる漢字の選択肢(\*1)

- ① 常用漢字・人名用漢字の範囲(2998文字)

日本国内の地名や歴史的名称を表現するには不十分である。例えば奈良県橿原市の「橿」はこの範囲に含まれない

- ② JIS X 0208:2012の第一水準・第二水準の範囲(6358文字)

常用漢字(JIS X 0213:2012の第三水準に含まれる4文字(叱剝墳頰)を除く)・人名用漢字を包含し、かつ日本国内の地名や歴史的名称を表記でき、JPドメイン名でも14年にわたる実績があり有用性が確認されている(文字の多寡について苦情や要望がない)

- ③ JIS X 0213:2012の第三水準・第四水準までの範囲(10053文字)

平均的な日本人が通常使用しておらず、また理解していない(読み書きできない)第三水準・第四水準の文字を入れるのはリスクが大きい

- ④ IICORE(\*2)の範囲(9810文字)

JIS X 0203:2012と同様にリスクが大きい

仮  
合  
意

(\*1) 平仮名(83文字)、片仮名(86文字)、平仮名・片仮名に準ずる文字(5文字)はいずれの選択肢でも入るものとする

(\*2) ISO/IEC 10646:2003/ Amendment 1:2005の一部として発行された国際標準で、漢字圏で共通に使える日常生活の用を満たす漢字集合

## b. 異体字

- 考えられる選択肢(\*1)

- ① 異体字を定義しない

ドメイン名のコンテキストは自由なため、文字列としては固有名詞も一般名詞も使われる。日本語において、特に固有名詞では異体字を区別して(個々の文字は独立した文字として)使う事例が散見される。異体字なしはJPドメイン名で14年にわたる実績があり有用性が確認されている(異体字の有無について苦情や要望がない)

- ② 異体字を定義する

ドメイン名というコンテキストに適切な異体字を、適切な根拠に基づいて定義しオーソリティを得ることは困難である

(\*1) ここではstep1での選択について考察する

## c. 処置

- 考えられる選択肢(\*1)

① 日本語として申請可能なすべての文字の異体字は「登録可能」とする

ドメイン名では固有名詞と一般名詞が組み合わされて同時に使われることがあり、異体字同士が1つのドメイン名の中に同時に存在することがある

② 個々の異体字について処置を決める

ドメイン名というコンテキストに適切な異体字の処置を、適切な根拠に基づいて定義しオーソリティを得ることは困難である

(\*1) ここではstep1での選択について考察する

## d. ラベル評価ルール

- 考えられる選択肢

① 日本語として不自然な文字の並びを排除する  
(例えば長音や繰り返し文字が文字列の先頭に  
来ないようにする)ルールを積極的に定義する

② 日本語用のラベル評価ルールは定義しない

仮  
合  
意

Rootゾーンに登録されるTLD文字列は審査が行われるので積極的に定義しなくとも不自然な文字列は排除され则认为られる。子細なルールを適切な根拠に基づいて定義しオーソリティを得ることは困難である