

ルートゾーンにおける日本語ラベル生成ルール に関する生成パネルの提案

1. 概説

現代の日本語を書き表すために主に使われている文字は漢字、平仮名、片仮名、英数字である。漢字は5世紀ごろに中国から輸入された。平仮名や片仮名は10世紀ごろに漢字を崩したり漢字の一部を使って日本で作られた音節文字である。アルファベットは16世紀以降に西欧から入ってきたが、一般に使われるようになったのは19世紀以降である。

平仮名や片仮名はそれらが作られた当初から漢字と入り混じって使われており、日本語の書記体系として定着している。さらに、近代になってからは漢字・平仮名・片仮名に混じって英数字も使われるようになり、たとえば英数字と漢字が混ざった商号が普通に使われている。このような複数の用字系の入り混じった書記体系が日本では常態化し定着している。

1.1. 本生成パネルが対象とする文字

日本語の書記体系で使う用字系は前述の通り漢字、平仮名、片仮名、英数字である。

以下に、日本語ドメイン名として使用する文字の範囲を示す。

ISO 15924 では、日本語の用字系の範囲は主に以下の仕様で定義されている。

ISO 15924 code: Jpan

ISO 15924 no.: 413

English name: Japanese (alias for Han + Hiragana + Katakana)

Unicode で定義される漢字の範囲

CJK Symbols and Punctuation: 3005, 3006

CJK Unified Ideographs: 4E00-9FCC

CJK Unified Ideographs Extension A: 3400-4DB5

CJK Unified Ideographs Extension B: 20000-2A6D6

CJK Unified Ideographs Extension C: 2A700-2B734

CJK Unified Ideographs Extension D: 2B740-2B81D

(CJK は中国語、日本語、韓国語を表す)

Unicode で定義される平仮名・片仮名の範囲

Hiragana: 3041-3096, 309D-309E

Katakana: 30A1-30FA, 30FC-30FE

ISO 15924 で定義される英数字の範囲

ISO 15924 code: Latn

ISO 15924 no.: 215

English name: Latin

Unicode で定義される英字の範囲

Basic Latin: 0061-007A

1.2. 日本で使われる用字系を使用する主要な言語

日本語においては、漢字、平仮名・片仮名、英字の4つの用字系が使われる。

漢字は中国、台湾、香港、澳門、マレーシア、シンガポール、韓国、日本等で使われてきた。日本語で使われる漢字は当初中国から輸入されたものだが、千年以上の歴史の中で日本固有の文字(国字)が追加されたり、日本固有の簡略化を施した文字(新字)が作られたりしている。公的文書、新聞、一般書籍等で使われる漢字の集合として「常用漢字」が定義され、そこでは2,100文字あまりが規定されている。また、常用漢字に含まれない人名用漢字として860文字あまりが規定されている。日本国内で広く普及している文字の規格であるJIS X 0208は、これら常用漢字および人名用漢字、さらには地名や歴史的名称等を表す文字を含み、全体で6,300文字あまりの漢字を規定している。日本語においては簡略化前の旧字体と簡略化後の新字体が互いに異体字として対応しているが、人名や組織名等名前に対しては、旧字体と新字体は別の文字として認識されることが多い。

中国語で使われる漢字には簡体字(中国本土、シンガポール等で使用)と繁体字(台湾、香港、澳門等で使用)がある。簡体字は、繁体字の中でよく使われるものを書きやすくしたものと捉えることができ、簡体字と繁体字は可換な異体字であると認識されている。なお、簡体字を使う国・地域で繁体字を使うこと、およびその逆はほぼない。また、簡体字と繁体字を混ぜて使うこともほぼない。

Hanjaとは韓国語で漢字を指す。韓国で使われる漢字はほぼ繁体字であるが、現代韓国語の書記体系としてはほぼハングルが用いられており、最近の公的文書、新聞、一般書籍等での漢字の使用は補助的である。

平仮名と片仮名は日本で発明された文字であり、日本でのみ使用されており、他の国や地域では使用されていない(パラオ共和国アンガウル州では日本語が公用語とされているが日常での利用実態はない)。平仮名、片仮名はともに音節文字であり、同じ発音に対してそれぞれ対応する文字があるが、それらは異体字とはみなされていない。平仮名は主に漢字の送り仮名や副詞、接続詞に、また、難読漢字を平易に書き表す場合等に使われる。片仮名は主に外来語や擬音語で使われる。

英字はアラビア数字と同様に世界中の多数の国や地域で使われている。日本語における英字は主に外来語や略語を表すために使われる。さらに、会社名や商標名等で漢字、平仮名、片仮名と入り混じって使われることがある。

1.3. 日本で使われる用字系を使用する国々

漢字は、中国語を使用する中国、台湾、香港、澳門、マレーシア、シンガポール等で使われてきた。さらに、日本、韓国等では、漢字が輸入され使われてきた。平仮名、片仮名は日本で発明された文字であり、日本でのみ使用されている。

英字は、英語圏の国々で使われている。また、アクセント記号のついた(英字を含む)ラテン文字は、ラテン系言語を使う国々(フランス、ドイツ等)で使われている。

漢字を使う国々については、以下の図が参考になる。(この図は中国語生成パネル(CGP)の提案内容を参考に作成したものである。)



	漢字が平仮名・片仮名とともに同一言語内で使われている(日本)
	漢字がハングルとともに同一言語内で使われている(韓国)
	ほぼ簡体字のみが使われている(中国)
	ほぼ繁体字のみが使われている(台湾、澳門、香港)
	公的には簡体字が使われているがまだ繁体字も使われている(マレーシア)

2. パネルの初期メンバー案

2.1. パネルのチェアとメンバー（専門性の背景含む）

堀田博文（議長）

Proposal for Japanese Generation Panel

- 株式会社日本レジストリサービス (JPRS) 取締役
- ICANN ccNSO 評議委員
- APTLD 理事、財務担当
- 一般財団法人インターネット協会 (IAJapan) 理事
- 元) Asia&Pacific Internet Association (APIA) 議長
- 元) JPNIC 国際関係ワーキンググループ 議長
- 元) CJK Joint Engineering Team (JET) メンバー
- 元) JPNIC IDN Taskforce (IDN-TF) メンバー

前村昌紀 (副議長)

- 一般社団法人日本ネットワークインフォメーションセンター (JPNIC) インターネット推進部 部長
- Asia Pacific Network Information Centre (APNIC) 理事会議長
- NETmundial Initiative 調整評議会 委員
- 一般社団法人 JPCERT コーディネーションセンター (JPCERT/CC) 理事
- 元) NETmundial マルチステークホルダー実行委員会 委員
- 元) 社団法人日本ネットワークインフォメーションセンター (JPNIC) 理事
- 元) 日本ネットワーク・オペレーターズ・グループ (JANOG) 運営委員会

後藤滋樹

- 早稲田大学 理工学術院 基幹理工学部 情報理工学科 教授
- 一般社団法人 日本ネットワークインフォメーションセンター(JPNIC) 理事長
- Asia-Pacific Advanced Network (APAN) アドバイザー
- JP ドメイン名諮問委員会 委員長
- 元) 総務省情報通信審議会 委員
- 元) インターネットドメイン名評議会 副議長
- 元) 日本語ドメイン名協会(JDNA) 会長
- 元) CJK Joint Engineering Team (JET) メンバー
- 元) JPNIC IDN タスクフォース(IDN-TF) メンバー

小西和憲

- Asia-Pacific Advanced Network (APAN) NOC Director
- 元) サイバー大学 教授

Proposal for Japanese Generation Panel

- 元) 社団法人 日本ネットワークインフォメーションセンター(JPNIC) 理事
- 元) 日本語ドメイン名協会(JDNA) 理事
- 元) CJK Joint Engineering Team (JET) メンバー
- 元) JPNIC IDN タスクフォース(IDN-TF) メンバー
- JET Guidelines (RFC 3743)共著者

久保次三

- 専修大学 法科大学院 教授
- Panel of Experts appointed by WIPO in the WIPO Internet Domain Name Process 1999
- 一般財団法人インターネット協会 (IAJapan) 評議員
- 元) 一般社団法人日本知的財産協会 (JIPA) 商標委員会委員長
- 元) 日本商標協会 (JTA)常務理事
- 元) JPNIC 運営委員、評議員
- 元) JPNIC ドメイン名検討部会 (DOM-WG)主査
- 元) JPNIC ドメイン名紛争処理方針タスクフォース (DRP-TF)委員長
- 元) JPRS 都道府県型 JP ドメイン名優先登録 DRP パネリスト

村上嘉隆

- 株式会社ブライツコンサルティング コンサルタントグループ シニアコンサルタント
- 株式会社ブライツコンサルティング 新 gTLD プロジェクトチーム プロジェクトリーダー
- Accredited Paraprofessional Interpreter/Paraprofessional Translator (Australia)

田代秀一

- 独立行政法人 情報処理推進機構 技術本部 国際標準推進センター 参与 (技術本部国際標準推進センター長)
- ISO/IEC JTC1/SC2/WG2/IRG メンバー
- IPA 文字情報基盤整備事業 事務局
- 元) CJK Joint Engineering Team (JET)メンバー
- 元) JPNIC ドメイン名検討部会(DOM-WG) メンバー
- 元) JPNIC IDN タスクフォース(IDN-TF) メンバー

米谷嘉朗

- 株式会社日本レジストリサービス(JPRS) 技術研究部 主任研究員
- IETF precis WG 共同議長
- 日本 DNS オペレーターズグループ 幹事
- 元) ICANN IDN VIP 中国語スタディチーム メンバー
- 元) 日本語ドメイン名協会(JDNA) 代表幹事
- 元) CJK Joint Engineering Team (JET) メンバー
- 元) JPNIC IDN タスクフォース(IDN-TF) 主査
- downgrade mechanism for EAI (RFC 5504)共著

2.2. パネルメンバーの多様性

日本語生成パネルの初期メンバーはさまざまな背景を持つ専門家が集まっている。いずれのメンバーも IDN の規格策定や JET(Joint Engineering Team : 2.3 節を参照)における CJK での議論、日本国内における日本語ドメイン名ルールの制定、ICANN における多様な分野からの参加を通じたポリシー策定の経験を持っている。

日本語生成パネルの初期メンバーはいずれも日本人であるが、インターネット国際化(多言語化)を世界的な範囲で実施していた経験者が多く、インターネットの国際化における日本語や他の言語(特にアジア系言語)に造詣が深い。

2.3. 過去の活動やワーキンググループとの関係

JP ドメイン名においては、2001 年 2 月より、IDN の登録サービスを開始した。それは、汎用 JP ドメイン名と呼ばれる第 2 レベルへの ASCII ラベルおよび IDN ラベルの登録サービスの中で開始された。また、2012 年 11 月からは、都道府県型と呼ばれる第 2 レベルに都道府県名を持ち第 3 レベルに ASCII ラベルおよび IDN ラベルを登録させるサービスを開始した。

汎用 JP ドメイン名においては、2014 年 11 月 1 日現在、全 93.9 万ドメイン名登録のうち約 13%が IDN であり、また、都道府県型 JP ドメイン名においては、全 1.2 万ドメイン名登録のうち約 25%が IDN である。

また、それら IDN のうち、約 4%は英数字交じりの IDN である。

Proposal for Japanese Generation Panel

この JP における日本語ドメイン名のルールは IANA レポジトリに ja-JP として登録されており、それは、.ASIA の登録規則でも使われている。

この日本語ドメイン名のルールは、2000 年 9 月より JPNIC が中心となり、ドメイン名の専門家、商標専門家や文字コードの専門家等、多様な専門家を集めて議論し、作ったものである。その過程において、コミュニティに仕様を公開し、コメントも受け付けた。そのルールの主だった特徴は、次のものである。

- 漢字、ひらがな、カタカナが少なくとも 1 文字入った、漢字、ひらがな、カタカナ、英数字の列
- 漢字は、JIS X 0208 第 1 水準と第 2 水準からなる 6,355 個の範囲。平仮名は 85 字、片仮名は 89 字、漢字に準ずる文字は 5 字。
- 異体字は存在しない

その約 14 年間に及ぶサービス提供において、この日本語ドメイン名のルールに対する苦情や反対意見は受けていない。

CJK は、同じ漢字を共通に使う可能性がある。しかし、漢字の範囲や異体字の定義は、CJK それぞれで異なる。

IETF での IDNA(Internationalizing Domain Names in Applications)の検討を通して、異体字をどこで処理すべきかという議論があった。議論の中で、プロトコルのレベルで異体字を処理すべきという意見もあったが、プロトコルのレベルでは言語毎に異なる扱いを要求する異体字を処理できないという問題がある。

このため、CJK の文字を使う国や地域が JET(Joint Engineering Team)と呼ぶチームを作り、異体字の扱い方を協力して検討し、プロトコルのレベルでなくドメイン名の登録や利用のレベルで処理すべきということを合意し、RFC3743 を作り出した。

3. 活動計画

3.1. 日本語 LGR の特徴

1.1 節で定義された文字の中で、日本語 LGR によって定義される文字の範囲は MSR-1 の中で定義された中から選ばれる。文字の範囲は、文字識別語 Jpan

Proposal for Japanese Generation Panel

(日本語コードのために設定された ISO 15924)の範囲に制限され、これは ASCII セットを含まない。

日本語は、1.3 節と 2.3 節に記述されているように、中国語コミュニティおよび韓国語コミュニティと漢字を共有している。しかし、言語文字の範囲および異体字定義は、CJK 間で異なっている。このため、JGP(日本語生成パネル)は、CGP(中国語生成パネル)、KGP(韓国語生成パネル)と協力・協調して日本語 LGR を定義し、それが中国語 LGR、韓国語 LGR と整合がとれたものとする必要がある。

3.2. スケジュールとマイルストーン

2014 年

8-12 月 JGP 設立準備
日本語 LGR 検討
CJK 調整内容検討

2015 年

1 月 日本語 LGR 検討
CJK 調整内容検討
2 月 JGP 設立申請[対 ICANN]
JGP 設立申請したことを発表[対国内一般]
CJK 調整委員会
3-5 月 日本語 LGR 検討
CJK 調整
5 月 日本語 LGR の ICANN への提出

その後 IP と調整

※ 本活動予定は、アクションアイテムの増加/減少、CJK での調整状況などに応じ、時宜に即して更新する。

3.3. 会合と電話会議のスケジュール

毎月 1～2 回の会合を予定

3.4. 移動及び運営の資金

事務局である JPRS もしくは JPNIC が会議場所および必要に応じリモート参加環境を提供。また、会議場への交通費やリモート参加に要する機材等は、個々のメンバーもしくはメンバーが所属する個々の組織が負担する。

3.5. ICANN からのアドバイザーについて

パネル創設当初は、特に想定していない。ただし、検討の進展中において、必要に応じアドバイスを依頼することはある。

参考

- 日本語について
 - Japanese language
<http://en.wikipedia.org/wiki/Japanese_language>
- 日本の商標と商号に関する法律
 - About the designation of the standard characters to prescribe in Trademarks Law Article 5 Clause 3 (In Japanese)
<http://www.jpo.go.jp/shiryoku/kijun/kijun2/pdf/syouhyoubin/19_01.pdf>
 - Japan's trademark system
<http://www.jetro.go.jp/en/invest/setting_up/laws/section5/page2.html>
 - About using Roman characters for trade names (In Japanese)
<<http://www.moj.go.jp/MINJI/minji44.html>>
- 文字に関する標準
 - JIS X 0208
<http://en.wikipedia.org/wiki/JIS_X_0208>
 - Unicode
<<http://www.unicode.org/>>
- インターネットに関する標準
 - Joint Engineering Team (JET) Guidelines for Internationalized Domain Names (IDN) Registration and Administration for Chinese, Japanese, and Korean
<<http://tools.ietf.org/html/rfc3743>>
- ICANN における関連文書
 - Root Zone LGR Project
<<https://community.icann.org/display/croscomlgrprocedure/Root+Zone+LGR+Project>>
 - Report on Chinese Variants in Internationalized Top-Level Domains
<<http://archive.icann.org/en/topics/new-gtlds/chinese-vip-issues-report-03oct11-en.pdf>>

Proposal for Japanese Generation Panel

Proposal for the Generation Panel for the Chinese Script Label
Generation

Ruleset for the Root Zone

<<https://www.icann.org/en/system/files/files/chinese-script-lgr-proposal-24sep14-en.pdf>>

2015年2月27日