



今号のテーマ

国際化 ドメイン名

今回の10分講座は、国際化ドメイン名 (IDN) について解説します。

[関連記事] P.27 「日本語ドメイン名協会 (JDNA) の活動の「軌跡」が書籍として刊行されました」

アジア太平洋地域のインターネット関連会合であるAPNG、APTLD、APRICOTなどの場で意見交換、検討が進められ、試験的な実装とサービスも幾つか登場してきました。

シンガポール国立大学で開発された技術は「ゼロレベルドメイン (ZLD)」という方式を採用していました。これは、IDNであることを識別するために、専用のドメイン名空間を用意するものです。しかし、このゼロレベルドメイン名をはじめ、当時用いられていた方式は、IDN登録管理サービス提供者毎に異なっていたため、同じドメイン名文字列でも利用環境によって異なる実体を指すという問題がありました。これは、どこからでも同じ名前で見えるという均質性が重要であるインターネットにとって、分断につながる危険なものだったのです。

- IDNはASCII文字列へ変換して取り扱う
迅速に安定的な導入を進めるため、ASCIIをベースとした既存のDNSシステムの変更を必要としない方式を採用し、その仕組みはアプリケーションプログラムでの変換により実装する

- ZLDは利用しない
ASCII変換された文字列でIDNであることを識別するために、IDN用の接頭語 (prefix) を定める

- UNICODE文字集合を用いる
IETFではIDN用文字集合を検討せず、UNICODEで規定される文字集合を用いて世界中の言語の文字を均質な方式でドメイン名として使用可能にする

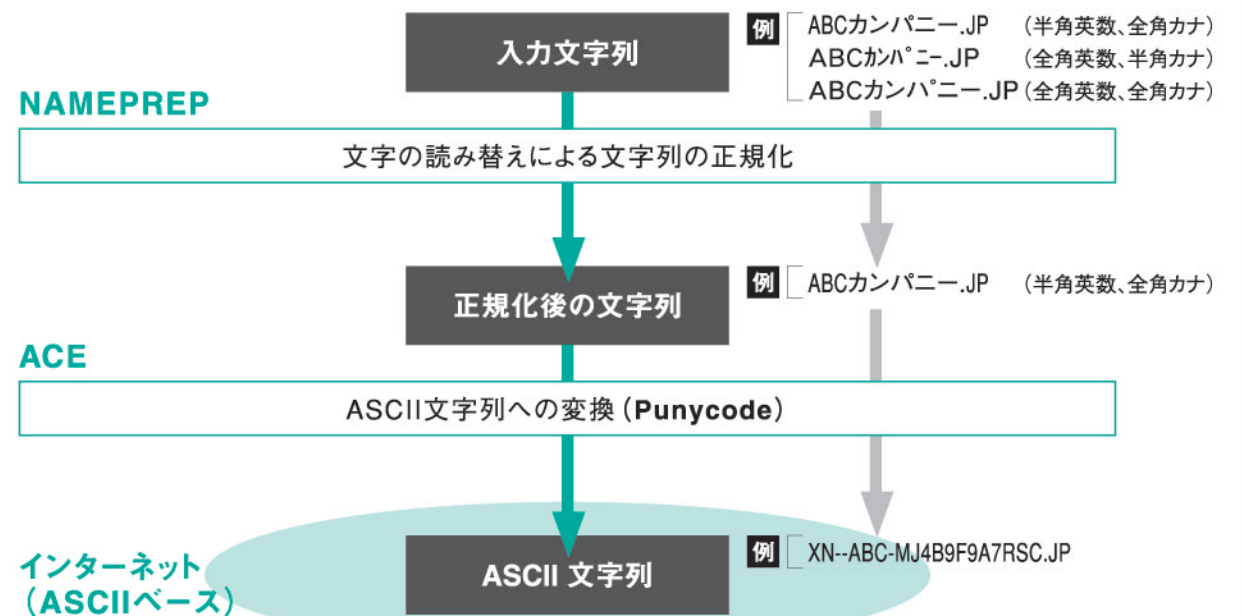
このIETFでの標準化は、2000年7月に結成された、日 (JP)、韓 (KR)、中 (CN)、台 (TW) のccTLDレジストリを中心とする技術集団JET (Joint Engineering Team) が牽引しました。そして、紆余曲折を経ながらも、次のように着実に成果を生み出しました。

- 2001年12月、IETF IDN WGが技術仕様を合意
- 2002年10月、IESGがRFCとしての発行を承認

2.3 IDNプロトコルの標準化

この問題を避けるため、IDNの検討は、IETFに舞台を移し、標準方式の作成に努力が傾けられることになりました。この転換は、日本の技術者の「インターネットは単一空間であるべき」という強い信念に基づくイニシアティブにより実現しました。そして、2000年には、次の基本的な枠組みが世界的に合意されました。

図1 基本的なIDN処理方式



1 はじめに

本稿では、国際化ドメイン名 (Internationalized Domain Names: 以下、IDNといいます) の技術およびその動向に関して解説します。技術内容のみでなく、IDNと日本の利用者との関わりや日本が行ってきた技術標準化への貢献についても、紹介しています。

2 IDNの発展経緯

2.1 インターネットの国際化

コンピュータもインターネットも、従来米国を中心に育ってきたため、基本的にはASCII文字 (英数字) で利用することが前提となっていました。しかし、インターネットの広がりとともに、当然のごとく、それぞれの文化に根ざした言語でインターネットを利用したいという要求が増加してきました。

一般には、各国言語を使えるようにするということが中心になります。これは、単に各国が独自に自国用にシステムを変更するというのではなく、実際には、次の2つが必要となります。

- 国際化 (Internationalization)
ASCII以外の種々の文字が使えるように、基本的枠組みを拡張する

- 地域化 (Localization)
国際化された基本的枠組みに、各地域で用いる文字集合を適用する

たとえば、電子メールでは、次の各ステップで国際化と地域化が行われ、日本語文字が使えるようになってきました。なお、(3-2) (3-3) はこれから実施が待たれているものです。

- (1) メール本文を日本語文字で
- (2) 題名を日本語文字で
- (3) 送信元、送信先メールアドレスを日本語文字で
 - (3-1) メールアドレスの左に付ける表示名 (display name) を日本語文字で
 - (3-2) ドメイン名を日本語文字で
 - (3-3) メールアドレス全体を日本語文字で

2.2 ドメイン名の国際化

IDNの起源は、1990年代後半に遡ります。当初は、「国際化ドメイン名」と呼ばれず「多言語ドメイン名」と呼ばれていました。つまり、英語以外の言語でドメイン名を表現しようということです。しかし、検討が進むにつれ、必要なのは、「英字以外の文字」を使用可能とすることであり、「英語以外の言語」を使用可能とすることではないことが認識されてきました。したがって、本稿では、特に区別する必要がない限り、以後は「国際化」という言葉を使います。

IDNの誕生は、シンガポール国立大学において、Tan Tin-Wee氏、James Seng氏らにより進められた研究に端を発します。そして、1998年～1999年には、基本的な技術仕様案が組み立てられ、プロトタイプ実装も登場しました。これらは、

- ・2003年3月、IDNに関連するRFCとして以下を発行
 - IDNA (RFC3490)
アプリケーションプログラム内でIDNをASCIIコードに変換して扱うという枠組み
 - NAMEPREP (RFC3491)
IDNを正規化するためのマッピングとアルゴリズム
 - Punycode (RFC3492)
IDNからASCIIコードへの変換と逆変換のアルゴリズム
- IDNの処理概要を図1に示します。

2.4 プロトコル以外の標準化

前述のようにIDNのプロトコル仕様の標準化は進められましたが、検討が必要なのは、プロトコルだけではありませんでした。つまり、「文字」と「言語」の関係が新たな課題としてクローズアップされてきたのです。

特徴的な例を一つ紹介しましょう、CJK(中国語(Chinese)、日本語(Japanese)、韓国語(Korean))は多くの漢字を共通して使っています。たとえば、文字「国」と「國」は、どちらも中国語の文字でもあり日本語の文字でもあります。この2つの文字は、中国では同じ文字、日本では違う文字として扱います。このとき、

- ・「国.cn」と「國.cn」(.cnは中国本土のccTLD)は同じドメイン名とすべき?
- ・「国.jp」と「國.jp」は違うドメイン名とすべき?
- ・「国.com」と「國.com」は?

というような課題が存在するのです。

この課題に対し、前出のJETが、CJKの文字集合を例題に、解決に向け取り組みました。その結果は、通称「JETガイドライン」と呼ばれるものにまとめられ、IETFのRFC3743としても採用されました。その主な内容は次のものです。

- ・IDNが対象とする言語に対し、使用可能な文字を規定すべきである
- ・IDNが対象とする言語に対し、使用可能な文字の中に、等価とみなすべきものがあれば、それを定義すべきである
- ・個々のIDN登録管理サービスでは、対象とする言語を定義すべきである

そして、その思想は、ICANNが2003年6月に制定したドメイン名レジストリ向けガイドラインに引き継がれました。このガイド

ラインが解きたかった課題は、そもそも2000年11月にVeriSignが.com等で始めたIDN登録管理サービスにおいて、英語圏以外でも使いやすきたいという本来の目的からかけ離れた無意味なIDNを数多く生んだことに端を発します。このとき採用した「UNICODE表のすべての文字を任意に組み合わせることができる文字列すべてをIDNとみなす」というルールが、この問題のきっかけとなったのです。また、等価な複数の文字(たとえば中国での繁体字と簡体字)を違うものとして扱うことによりサイバースクワッティング等の問題を増加させるという問題も持っていました。

ICANN-IDNガイドラインは、「健全なIDN環境のためには、IDNの利用が広がる前に、これらの問題の解決を図る仕組みを作っておく必要がある」という共通認識が生まれたというところに端を発しています。このガイドラインは、さらにセキュリティを高めるべく、2005年11月と2006年2月に改訂されています。また、このガイドラインをICANNコミュニティ外にも広く使ってもらおうべく、IETFのBCP(Best Current Practice)文書にする作業も進められています。

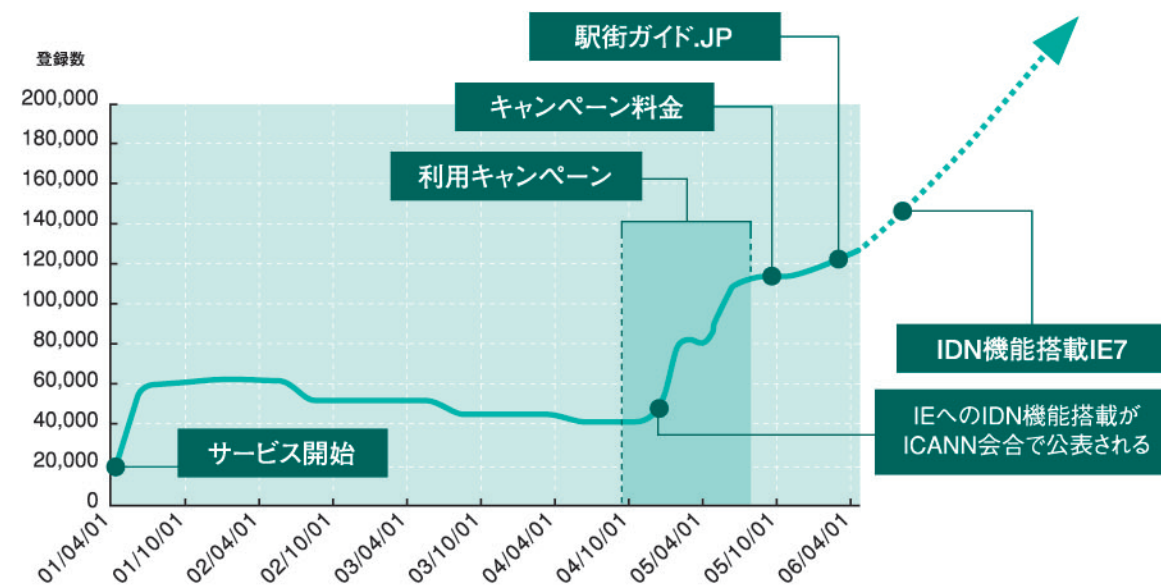
3 日本語JPDメイン名

3.1 日本語JPDメイン名の導入経緯

日本語JPDメイン名は、.JPの直下(第2レベル)にドメイン名を登録する汎用JPDメイン名の枠組みにおいて、ASCIIドメイン名と同時に、日本語の漢字、平仮名、片仮名などが利用可能なドメイン名として導入されました。その経緯は、次のようなものでした。

- ・2000年10月、導入方針公開
- ・2000年11月、技術仕様書公開
公的名称や一般的な用語を、一般には登録できない予約ドメイン名として定義
- ・2001年2月、優先登録申請期間開始
知的財産権に関する紛争を抑止するため、商標、商号、個人名、大学名を先行して登録受付
- ・2001年4月、同時登録申請期間開始
登録開始時の大量の登録要求を混乱なく処理するため、期間内の申請に対して抽選で登録者を決定

図2 日本語JPDメイン名登録数推移



- ・2001年5月、先願登録申請期間開始
早い者勝ちの登録申請受付を開始

この経緯に示すように、日本語ドメイン名が、より社会生活に直結していることも考慮し、導入期に混乱を防止する処置がとられました。これらの処置は、以後、諸外国でのドメイン名登録管理においても参考とされています。

3.2 日本語JPDメイン名の登録数

日本語JPDメイン名全体の登録累計数は、2006年6月時点では、約12万件あり、JPDメイン名全体の約15%を占めています。

また、2001年2月の登録管理サービス開始以降、日本語JPDメイン名の登録累計数の推移は、図2のようになっています。登録数は、2年半ほど継続して漸減した後、2005年初頭から急激に増加し始めました。漸減は、IDNを利用可能なアプリケーション環境が不十分であったことによると考えられます。特に、Webブラウザの大きなシェアを占めるマイクロソフトのIE(Internet Explorer)が、IDN用プラグインなしではIDNを処理できなかったことが大きな影響を与えていたと考えられています。2005年初頭からの急増は、IEがIDNをサポートするということが2004

年12月のICANN会合で表明されたことに加え、それに呼応する形でJPRSが行ったキャンペーンが功を奏したものと考えられます。

3.3 日本語JPDメイン名の推進

(1) 政府機関ドメイン名での日本語JPDメイン名利用
2006年2月に公開された、政府の「第1次情報セキュリティ基本計画」で、政府機関のドメイン名にはgo.jpもしくは、日本語JPDメイン名で予約ドメイン名となっているものの利用を推進することが記されました。これにより、日本での日本語JPDメイン名の信頼度が向上するものと考えられます。

(2) 人名辞典や駅街ガイドでの利用法例示
JPRSは、著名人の名を日本語JPDメインとして用いているWebページを紹介したり、日本に約9000ある駅名すべてを日本語JPDメイン名として用いて駅周辺情報を案内するWebページを提供するなどにより、日本語JPDメイン名の便利さの認知を推進しています。

4 IDNに関する最近の話題

4.1 TLDへのIDN導入 (IDN TLD)

これまで、IDNというと、TLDレジストリが登録管理する対象であるドメイン名について中心的に議論されてきました。たとえば、〇〇.jpや〇〇.comのような形式のものです。しかし、アラブ語圏や中国等、国や地域によっては、そのTLDラベルも含めてドメイン名の一部にでもASCII文字を使うということ自体が無理があるといわれています。つまり、日常生活で接する新聞や雑誌、看板等で使われる文字に全くASCII文字が含まれない、といった国や地域では、TLDラベルもその国や地域の文字でないと、一般の人はドメイン名を認識できないとのこと。

ICANNでは、このような需要に応え、IDN TLD等に関する検討を加速するため、昨年12月にIDNに関する事務総長諮問委員会 (IDN-PAC : President's Advisory Committee for IDNs) が設けられました。現在、検討項目の洗い出し中で、3月のICANN会合にて、その中間状況が種々のコミュニティと共有されたという状況です。以下、これまでにいわれている主な検討項目を紹介します。

(a) 技術検討項目

(a1) 次の2方式について、その実現可能性を純技術的に実験すべき

- ・DNSにNSレコードを追加する
全く新しいTLDとして、IDNラベルをルートDNSに登録する (例:「.日本」と「.jp」は別空間とする)
- ・DNSにDNAMEを設定する
ルートDNSにDNAMEレコードを登録し、IDNラベルと既存のTLDラベルを等価とみなす (例:「.日本」を導入し、それが参照されるときは「.jp」に読み替える)

(a2) 実験により検証すべき技術試験項目として他に何かあるか

(b) ポリシー検討項目

(b1) IDN TLDを既存のASCII TLDと同一空間とみなすか
みなさないか

(b2) (b1) で同一空間とみなす場合、ASCII TLDに対応するIDN TLDを幾つ作るか、誰がその文字列を決めるか、複数のASCII TLDが同一IDN TLDに対応するという主張があったときどう解決するか

(b3) (b1) で同一空間とみなさない場合、IDN TLDレジストリをどうやって決めるか、ASCIIと同じ通常の新gTLD導入プロセスに従うか

(b4) ccTLDを先に導入するか、gTLDを先に導入するか、両者同時か

(b5) SLD以降は、TLDと同一の言語/スクリプトでなければならないとするか

これら課題については、多方面からコメントを求めつつ、2006年7月に、技術実験の提案を募集するという計画になっています。

4.2 国際化電子メールアドレス

ドメイン名を使う代表的なアプリケーションの一つは、電子メールです。電子メールについては、@の左側 (ローカルパートと呼ばれます) も必ず書かねばならず、この部分にASCII以外が使えるという「国際化電子メールアドレス」への要望が大きくなっています。しかし、ローカルパートはドメイン名ではないため、ASCII以外が使えるようにするには新たな技術標準が必要になります。さらに、ローカルパートの記法とその解釈は、その名の通り (たとえばメールサーバを運用する企業単位で) ローカルに決められている場合もあり、IDNのように標準化が簡単ではありません。

本件に関しても、JETのメンバーが検討を主導し、その必要性をアピールするとともに、基本技術をIETFに提案し、技術検討のベースを作っています。IDNが本格的にインターネット全体で受け入れられ始めていること、電子メールには関連するプロトコルが数多くありIDNを扱えるようにするためには広範囲の技術者の協力が必要であること、などから、最初から世界中の多くの技術者を巻き込んで検討が行われています。

2005年8月IETFでのBoF準備会合、2005年11月IETFでのBoFを経て、2006年3月にIETFのEAI WG (Email Address Internationalization Working Group) が設立されました。その直後の2006年3月IETFで、数十名を集めて第1回WG会合が開かれました。主な論点は、国際化電子メールアドレスを扱うための、

- ・全体フレームワーク
- ・電子メールヘッダの構成
- ・SMTPの拡張
- ・国際化電子メールアドレスに対応したソフトウェアと対応しないソフトウェアの通信方法

に関する技術です。

現在の予定では、2007年初頭までに初期技術検討を終わり、2007年3月から、正式な技術標準化を目指した活動に移行することになっています。電子メールに関連するプロトコルも数多く、また、それを実装しているサーバやクライアントソフトウェアも数多くあるため、技術標準となった後も、インターネットユーザーが一般的に国際化電子メールアドレスを利用できるまでには、まだ時間がかかるものと思われます。

5 まとめ

本稿では、インターネットの国際化の一つ、IDNについて経緯と今後の展開について述べました。2006年後半に予定されているIEのIDNサポートにより、IDNは世界的に爆発的に使われる可能性を秘めており、その兆候が出始めています。

また、本稿により、IDN発展の歴史において占めてきた日本の重要な位置付けを知っていただき、これを一例として、技術が「使える」状態と「使い物になる」状態は大きく違うものであることを再認識し、IDNに限らず、今後のインターネット技術そして活用法の開拓に熱意を持つ方が一人でも増えれば幸いです。

(株式会社日本レジストリサービス 取締役 堀田博文)