

いまさら聞けない日本語ドメイン名

民田雅人

日本レジストリサービス株式会社

2008年1月24日 JANOG 21@熊本

本プレゼンテーションの目的

- JANOGの皆さんに、是非とも日本語ドメイン名に関する正しい知識を覚えていただきたい
 - どうも、変な誤解が流通しているらしい☹
 - 技術者だったら、このぐらいは知っておいてもらいたいという極めて個人的希望☺

日本語ドメイン名クイズ(1/2)

- 日本語ドメイン名は、
 - 漢字コードにEUC、シフトJIS、Unicode(UTF-8、UTF-16)などがあって運用が大変だ
 - ブラウザへのアドレス入力で、全角半角の変換を操作しないとイケないので面倒だ
(「**日本語。JP**」と**全部全角**ではアクセスできない)
 - 日本語以外のOSを使っている人はアクセスできない
 - M Root Serverを利用して動いている
(Mが日本にあるのでなんとかなっている)

日本語ドメイン名クイズ(2/2)

- 日本語ドメイン名は、
 - JPのDNSに特別な細工を施している
 - DNSサーバに対応プログラムが必要だ
 - WEBサーバに対応プログラムが必要だ
 - WEBブラウザに対応プログラムが必要だ

国際化ドメイン名(IDN)と 日本語ドメイン名

- IDN(Internationalized Domain Name)はUnicode文字集合を用いて、世界中の言語の文字をドメイン名として使用可能にする
 - IDNではUnicodeをPunycode(ASCII文字列)に変換して取り扱うため、DNS等の変更は不要
- 日本語ドメイン名は、IDNの技術を使用して、日本語で使われる文字で表現したドメイン名
 - IDNに準拠したアプリケーションで利用可能(アプリケーションが変換を実装)

Webアクセスのイメージ

② ブラウザが日本語.jpをPunycodeに変換し
xn--wgv71a119e.jpでDNSへリクエスト



DNSサーバ

"xn--" はPunycodeのプリフィックス
(通常のASCIIドメイン名との違い)

③ IPアドレスを返す

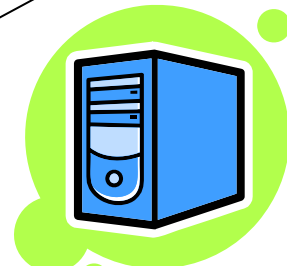


User

④ Webサーバへリクエスト
GET / HTTP/1.1
Host: xn--wgv71a119e.jp

① URLを入力
http://日本語.jp/

⑤ コンテンツを返す

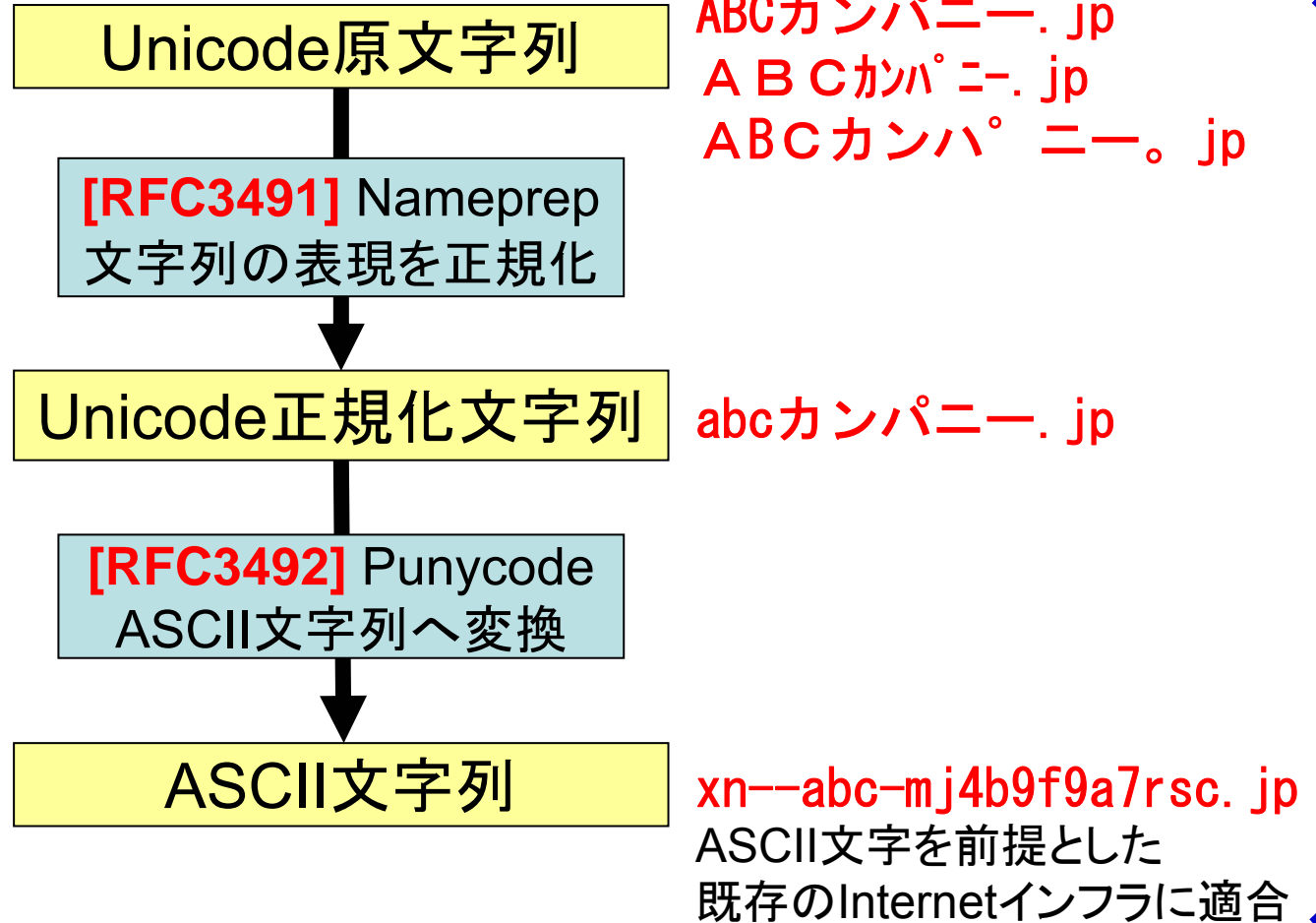


Webサーバ

国際化ドメイン名の技術構成図

[RFC3490] IDNA

ネットワーク上で国際化ドメイン名を使用する場合の正規化方式と符号化方式を指定



日本語JPドメイン名の主な仕様

登録可能文字	従来のASCII文字 + 平仮名・片仮名・漢字・一部記号
文字数	全半角問わず15文字まで
表現の統一 (正規化)	以下の表現は区別せず取り扱う 英数字の全角/半角、大文字/小文字 片仮名の全角/半角
「。」句点「.」(全角)	ピリオド(半角)「.」に変換

注) 関連するRFCの仕様によるものを含む

汎用JPドメイン名登録等に関する技術細則

<http://jprs.jp/doc/rule/saisoku-1-wideusejp.html>

日本語ドメイン名が登録できるTLD (2008年1月現在)

- jpの他には、com / net / bizが登録可
 - com / net / bizは、中国語、韓国語等も登録可
- bizの日本語ドメイン名
 - 登録可能文字はjpと同じ(jpの仕様を流用)
- com / netの日本語ドメイン名
 - 登録可能文字の明確な規定がない(みたい)
 - 漢字、ひらがな、カタカナは登録可
 - 記号類はレジストラによって変わる(みたい)

ホモグラフ(homograph)の問題

- 一見同じドメイン名

<http://日本レジストリサービス.jp/>

⇒ <http://xn--vckfdb7e3c7hma3m9657c16c.jp/>

<http://日本レジストリサービス.jp/>

⇒ <http://xn--vckfdb7e3c7hma5118ds68b16c.jp/>

実は「一」(長音)と「一」(漢数字)の違い

- 英数字でも"l"と"1"、"O"と"0"等で問題となる
- レジストリは淡々と登録を受け付ける

日本語ドメイン名を Punycodeに変換する

- Web CGIを利用して変換 (少数向け)
 - <http://punycode.jp/> (Punycode⇒日本語も可)
- ツールを利用して変換 (大量・バッチ向け)
 - idnkitを使う
<http://www.nic.ad.jp/ja/idn/idnkit/download/>
コマンドラインツール idnconv
変換ライブラリ libidnkit等
 - GNU libidnを使う
変換ライブラリとEmacs Lispのツール類
(Perlモジュール等も別パッケージで存在)

よくある質問とその回答(1/2)

- SSLはどうするの？
 - CSR生成時、Common NameにPunycodeで設定する
 - Common NameにはUTF8で設定し、subjectAltNameのDNSにPunycodeで設定する方法もある
- <http://日本語.日本語.jp/> は使えるの？
 - 問題ありません
- 検索で十分なんじゃないの？
 - ドメイン名はSEO対策の必要が無いなど、一長一短
- 未対応のブラウザからのアクセス手段は？
 - <http://ja.jp/> (JANOG14で紹介)を使ってね

よくある質問とその回答(2/2)

- メールアドレスはどのような?
 - ドメイン名部分だけなら対応メーラですでに利用可能
 - ローカル部分(@の左側)も国際化したい!
ということで、IETFのEAI WGでプロトコルレベルから策定中 (EAI: Email Address Internationalization)
- TLDの国際化は?
 - ICANNで現在11の言語を使って試験運用中
<http://idn.icann.org/> (以下は抜粋)

http://例え.テスト/	http://실례.테스트/
http://مثال.إختبار/	http://ביישפיל.טעסט/
http://例子.测试/	

参考

- 「いまずぐ使える国際化ドメイン名の理論と実践」
<http://www.atmarkit.co.jp/fnetwork/tanpatsu/11idn/idn1.html>
 - Punycode変換アルゴリズムの簡単な説明等
- <http://日本語.jp/> 日本語ドメイン名ポータルサイト
 - 日本語ドメイン名対応ブラウザ等
- 国際化ドメイン名に関するRFC
 - RFC 3490 Internationalizing Domain Names in Applications (IDNA)
 - RFC 3491 Nameprep: A Stringprep Profile for Internationalized Domain Names (IDN)
 - RFC 3492 Punycode: A Bootstring encoding of Unicode for Internationalized Domain Names in Applications (IDNA)

御清聴ありがとうございました!

