

# 国際化関連/ APPエリアレビューチームの紹介

2013年4月18日

IETF86 Update Meeting

米谷嘉朗 <yoshiro.yoneya@jprs.co.jp>

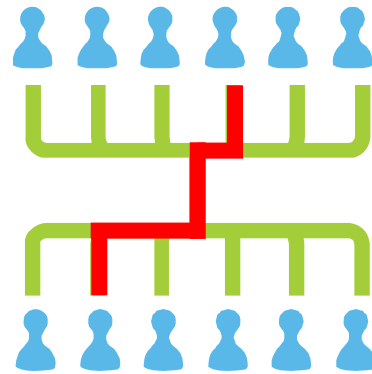
# 目次

- 国際化関連
  - プロトコルとは
  - 標準化とは
  - プロトコルの国際化とは
  - プロトコルの国際化における問題とその解決
  - プロトコルでは解決できない問題
  - IETF precis WG
  - 日本からの貢献
- APPエリアレビューチーム
  - IETF Applications Area Directorate
  - 実際の仕事
  - レビューチームメンバー
  - 日本からの貢献

# 国際化関連

# プロトコルとは

- 複数の人がコミュニケーションをするときの共通の取り決め

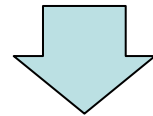


- 取り決めを守らないと情報が正しく伝わらずコミュニケーションが成立しない

「こんにちは」 $\leftrightarrow$ 「縛薙s縛オ縛。縛ッ」

# 標準化とは

- インターネットでは世界中の不特定多数の人がコミュニケーションする



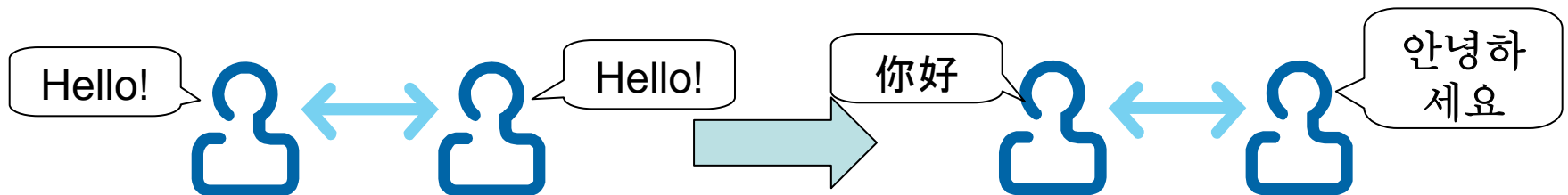
誰もが共通に守るプロトコルが必要

## 標準化

- インターネットのプロトコル標準化を行う  
IETF(Internet Engineering Task Force)
  - IP(v4、v6)、TCP、SMTP、HTTP、DNS他
  - RFC(Request For Comments)

# プロトコルの国際化とは(1/2)

- プロトコルで使える文字を各種言語の文字に拡張すること
  - 初期のインターネットではASCII(英数字)のみ



- プロトコルの中でどのような文字をどのような形式で使えるようにするのかを決めること
  - ドメイン名、メールアドレス、URIなど

# プロトコルの国際化とは(2/2)

- 国際化ドメイン名
  - Internationalized Domain Name; IDN
  - 2003年にRFC 3454,3490,3491,3492で規定され、2010年にRFC 5890,5891,5892,5893,5894,5895で更新された
- 国際化メールアドレス
  - Email Address Internationalization; EAI
  - 2007年～2008年にRFC 4952,5335,5336,5337,5504,5721,5738,5825,5983で実験的に規定され、2012年～2013年にRFC 6530,6531,6532,6533,6855,6856,6857,6858で標準化された
- 国際化URI
  - Internationalized Resource Identifier; IRI
  - 2005年にRFC 3987で規定され、現在改定作業が進められている

# プロトコルの国際化における問題と その解決 (1/3)

- プロトコルが伝達するコンテンツは文字セットとその表現形式が決まっていればよい



– Ex: UnicodeでUTF-8を使う



# プロトコルの国際化における問題と その解決 (2/3)

- 相手を特定するなど、通信を制御する部分を国際化する場合はプロトコル要素(文字列)の比較一致を正確に行う必要がある
  - 大文字小文字 (A  $\leftrightarrow$  a)
  - 合成文字 (が  $\leftrightarrow$  か)
  - 全角半角 (ア  $\leftrightarrow$  ア)
- 利用者からは同一に見えるプロトコル要素の比較一致が否となる
  - コミュニケーションの不成立
  - 不正アクセスの誘導

# プロトコルの国際化における問題と その解決 (3/3)

- 比較一致を正確に行うために
  - 文字の正規化や使える文字の定義を行う
    - 大文字を小文字に変換したり全角を半角に変換したり
    - 記号文字は使えなくしたり
  - プロトコルごとに正規化の定義や使える文字の定義は異なる
    - ただし基本的な考え方は共通のはず
    - IETF precis WGでそれを行っている

# プロトコルでは解決できない問題

- 似た文字問題
  - ロとロなど
- 異体字問題
  - 国と國など
- 運用で解決しようとしている
  - IDNにおけるJETガイドラインやICANNガイドラインなど

# IETF precis WG(1/3)

- precisとは
  - Preparation and Comparison of Internationalized Strings
  - 国際化された文字列の前処理と比較
- precisの背景
  - IETFにおける国際化された文字列の前処理と比較の枠組みはstringprep(RFC 3454)が起源
    - IDN標準化の成果の一つ
  - IDNの更新の際にstringprepは対象外とされた
    - Unicodeバージョン依存性の排除
    - 国際化表示とASCII互換表示の一意性の保証
  - SASLやLDAPなどstringprepを使うIDN以外のプロトコルも更新が必要という認識

# IETF precis WG(2/3)

- precis WGの歴史
  - newprep BoF (IETF77@Anaheim, 2010/03)
    - IDNの更新を受けて、stringprepを使うIDN以外のプロトコルも更新が必要かが議論された
    - Unicodeバージョン依存性の排除、アラビア文字など右から左に書く文字へのよりよい対応などが必要だということが共通認識となった
  - precis WG設立(2010/6)
  - 最初のRFC発行(2013/3)
    - RFC 6885 (Problem Statement)
    - Milestoneから2年ほど遅れたが、この後は順調に行きそう

# IETF precis WG(3/3)

- precisでの作業
  - 共通の仕組みの規定(framework)
    - RFC 5892をベースに一般化
  - frameworkを補助する処理の推奨(mappings)
    - RFC 5895をベースに補助的な前処理を推奨
  - stringprepを使用するプロトコルのprecisへの移行(profiles)
    - precis frameworkの適用
  - stringprepからprecisへの移行ガイドライン
    - framework完成後に作業開始予定

# 日本からの貢献

- もっと増やしたい (@\_@)
- 人材募集中 (^\_^)/

# APPエリアレビューチーム



# IETF Applications Area Directorate

- AppエリアADをサポートし、以下を提供するチーム
  - ドキュメントのレビュー
  - 他エリアでAppエリアと関連する作業の特定
  - Appエリアの成果が他エリアで適切に使われているかの確認
  - 他エリアの成果がAppエリアにもたらす影響の特定
  - アプリケーションプロトコル間の層分けと責任分解の維持
  - アプリケーション層の高いセキュリティの維持、国際化や地域化への密接な配慮

[出典] <http://trac.tools.ietf.org/area/app/trac/wiki/ApplicationsAreaDirectorate>

# 実際の仕事

- 主に、IETF Last CallのかかったAppエリア以外のI-Dのレビュー
  - Appエリア的視点から
  - 1～2ヶ月に1本程度の頻度
  - レビュー結果はappsawg ML、I-D著者、IESGにフィードバック
- 2009年から開始、2011年から本格化、2011年は42本、2012年は116本のレビューを実施

# レビューチームメンバー

- Apps AreaのAD、WGチェア、優れた貢献のある人から選ばれる
  - 現在35名
  - 特別なメンバー選出プロセスはなく、レビューチームのリーダーが主にIETF期間中にリクルート活動を行っている
  - 全員が常時スタンバイというわけではないので、チームメンバーはいつでも募集中

# 日本からの貢献

- もっと増やしたい (@\_@)
- 人材募集中 (^\_^)/