



Display and Usage of Internationalized Registration Data

Dave Piscitello

ICANN Sr. Security Technologist

The Problem Space

- Internet applications must accommodate a global audience using characters in local languages and scripts
- Whois is an Internet application, therefore
- Whois must accommodate a global audience
- Issue is distinct from IDN TLDs
 - Registration data are different from internationalized domain names

More than IDN

- IDN standards and guidelines define how domain names will be composed and displayed
 - Whois application and web developers can apply these for submission and display of domain names

WHOIS Treatment of IDN Labels

- User experience in the field varies
- Internet users might attempt to submit A-labels or U-labels to query WHOIS services
 - over Port 43 using OS command line programs and WHOIS client applications
 - as input to web-based WHOIS query and domain name registration submission forms
 - WHOIS operators might display domain labels in A-label encoding or in characters from the local script used by the registrant

sac037.pdf (17 pages)

Drawer Previous Next Page Back/Forward Zoom In Zoom Out Tool Mode

Sample WHOIS Port 43 queries and responses (courtesy of JPRS)

Example 1: IDN input, Japanese output

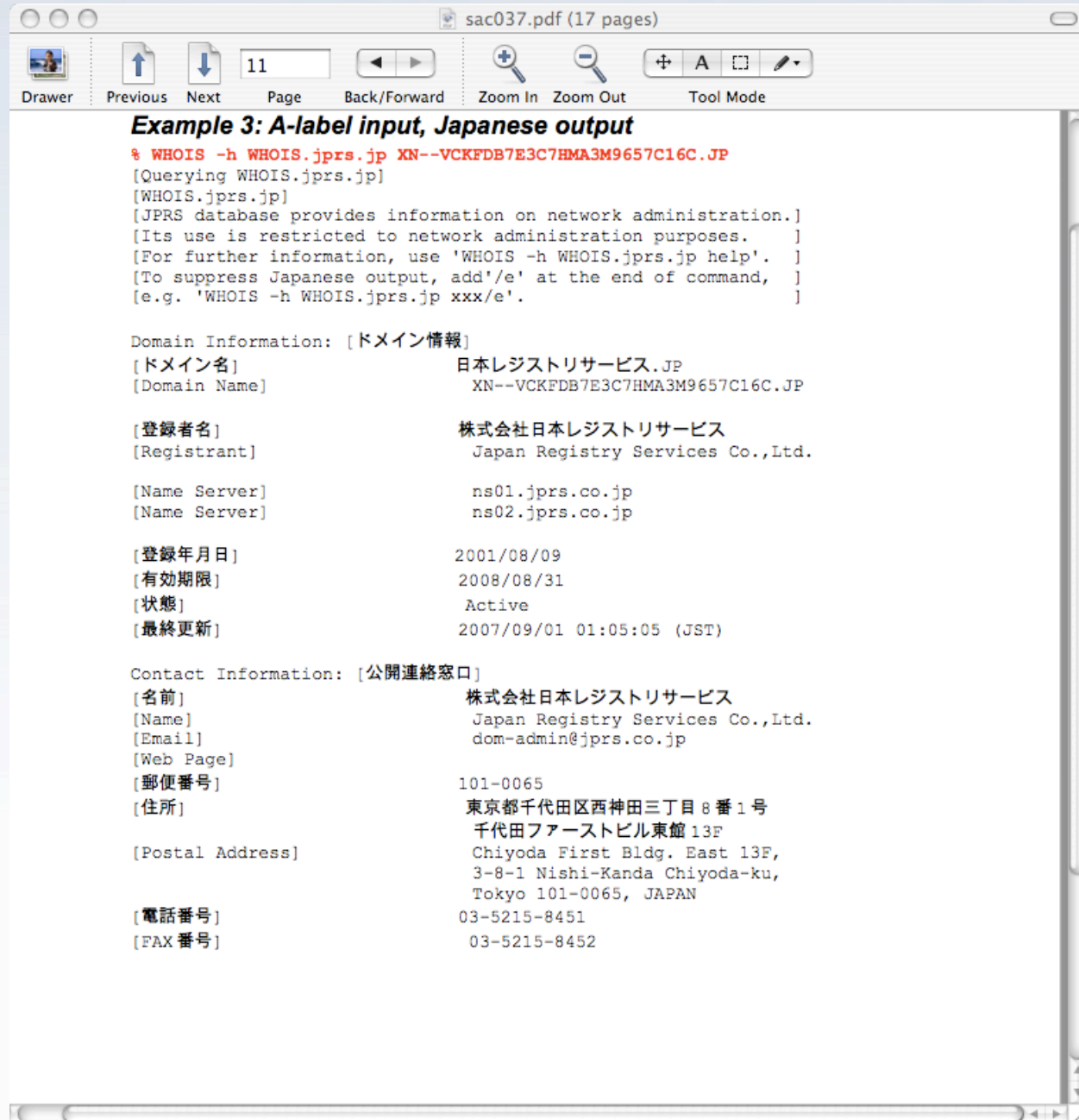
```
% WHOIS -h WHOIS.jp rs.jp 日本レジストリサービス.jp
[Querying WHOIS.jp rs.jp]
[WHOIS.jp rs.jp]
[JPRS database provides information on network administration.]
[Its use is restricted to network administration purposes. ]
[For further information, use 'WHOIS -h WHOIS.jp rs.jp help'. ]
[To suppress Japanese output, add '/e' at the end of command, ]
[e.g. 'WHOIS -h WHOIS.jp rs.jp xxx/e'. ]
```

Domain Information: [ドメイン情報]

[ドメイン名]	日本レジストリサービス.JP
[Domain Name]	XN--VCKFDB7E3C7HMA3M9657C16C.JP
[登録者名]	株式会社日本レジストリサービス
[Registrant]	Japan Registry Services Co.,Ltd.
[Name Server]	ns01.jp rs.co.jp
[Name Server]	ns02.jp rs.co.jp
[登録年月日]	2001/08/09
[有効期限]	2008/08/31
[状態]	Active
[最終更新]	2007/09/01 01:05:05 (JST)

Contact Information: [公開連絡窓口]

[名前]	株式会社日本レジストリサービス
[Name]	Japan Registry Services Co.,Ltd.
[Email]	dom-admin@jp rs.co.jp
[Web Page]	
[郵便番号]	101-0065
[住所]	東京都千代田区西神田三丁目8番1号 千代田ファーストビル東館13F
[Postal Address]	Chiyoda First Bldg. East 13F, 3-8-1 Nishi-Kanda Chiyoda-ku, Tokyo 101-0065, JAPAN
[電話番号]	03-5215-8451
[FAX 番号]	03-5215-8452



The screenshot shows a PDF viewer window titled "sac037.pdf (17 pages)". The content is a terminal output of a WHOIS query. The command used is "% WHOIS -h WHOIS.jp XN--VCKFDB7E3C7HMA3M9657C16C.JP". The output is in Japanese and provides detailed information about the domain, including its name, registrant, name servers, registration dates, and contact information.

```
Example 3: A-label input, Japanese output
% WHOIS -h WHOIS.jp XN--VCKFDB7E3C7HMA3M9657C16C.JP
[Querying WHOIS.jp]
[WHOIS.jp]
[JPRS database provides information on network administration.]
[Its use is restricted to network administration purposes. ]
[For further information, use 'WHOIS -h WHOIS.jp help'. ]
[To suppress Japanese output, add '/e' at the end of command, ]
[e.g. 'WHOIS -h WHOIS.jp xxx/e'. ]

Domain Information: [ドメイン情報]
[ドメイン名]                日本レジストリサービス.JP
[Domain Name]                XN--VCKFDB7E3C7HMA3M9657C16C.JP

[登録者名]                  株式会社日本レジストリサービス
[Registrant]                 Japan Registry Services Co.,Ltd.

[Name Server]               ns01.jpns.co.jp
[Name Server]               ns02.jpns.co.jp

[登録年月日]                2001/08/09
[有効期限]                  2008/08/31
[状態]                       Active
[最終更新]                  2007/09/01 01:05:05 (JST)

Contact Information: [公開連絡窓口]
[名前]                      株式会社日本レジストリサービス
[Name]                       Japan Registry Services Co.,Ltd.
[Email]                       dom-admin@jprs.co.jp
[Web Page]
[郵便番号]                  101-0065
[住所]                       東京都千代田区西神田三丁目 8 番 1 号
                               千代田ファーストビル東館 13F
[Postal Address]             Chiyoda First Bldg. East 13F,
                               3-8-1 Nishi-Kanda Chiyoda-ku,
                               Tokyo 101-0065, JAPAN

[電話番号]                  03-5215-8451
[FAX 番号]                   03-5215-8452
```

No Registration Data Standards or Guidelines

- What do we mean by “registration data”
 - Sponsoring registrar, contact information, DNS configuration data, ...
- Recognizable display of Registration Data is application dependent
 - Display may be affected by character sets available to Whois clients and browsers
- Registrars and 3rd party Whois providers may choose encodings that best represent registration data to local users

Domainabfrage-Ergebnis

Domaindaten

Domain häufig.de
Letzte Aktualisierung 12.06.2009

Domaininhaber

Der Domaininhaber ist der Vertragspartner der DENIC und damit der an der Domain materiell Berechtigte.

Domaininhaber: Marcus Janke
Organisation: Kröativität - Synapsen im Dialog
Adresse: Hans-Sachs-Str. 12
PLZ: 22765
Ort: Hamburg
Land: DE

Administrativer Ansprechpartner

Der administrative Ansprechpartner (admin-c) ist die vom Domaininhaber benannte natürliche Person, die als sein Bevollmächtigter berechtigt und gegenüber DENIC auch verpflichtet ist, sämtliche die Domain häufig.de betreffenden Angelegenheiten verbindlich zu entscheiden.

Name: Marcus Janke
Organisation: Kröativität - Synapsen im Dialog
Adresse: Hans-Sachs-Str. 12
PLZ: 22765
Ort: Hamburg
Land: DE

Questions to Consider:

- User experience issues
 - What features will benefit Internet users?
 - What can we do to minimize the "Babel effect"?
- Data reliability, accuracy, operational issues
 - Impact on applications that automate DNS and Whois queries?
 - Impact on law enforcement, brand protection, analysis of malicious conduct
- Security & standardization issues

Status

- ICANN's Board of Directors tasked GNSO, ALAC, CCNSO, GAC and SSAC to form a WG to study internationalized registration data submission and display specifications
- ICANN staff prepared WG charter and implementation plan
- ICANN staff and SSAC jointly studying
 - Current practices among CCTLDs
 - Requirements for standard functionality
 - Conventions and standards from other communications media (e.g., postal service)

Staff & SSAC Member Study/Exercise

Should the maintenance and display of certain registration data be required in US- ASCII to ensure a common denominator for core information display?

What Do Current WHOIS Practices Tell Us?

- Solicited current practices from 16 CCTLDs:
 - Does your registry allow users to register domain names using characters from local scripts?
 - Does your registry collect and store registration data in US-ASCII7 in addition to characters from local scripts?
 - Can users of the web interface choose the display language? What languages does your web interface support?
 - Does your registry provide access to registration information via WHOIS/port 43? Can users choose the display language?

Current Whois Practices (Responses)

- 10 of 16 CCTLDs registry allow users to register domain names using characters from local scripts
- 10 of 16 CCTLDs support “English” (US-ASCII7) and a local language/script, including
 - ✧ Arabic, Chinese, German, Japanese, Lithuanian, Portuguese, Spanish, Swedish
- All CCTLDs support WHOIS/Port 43
- Character set dependencies affect WHOIS client dependencies submission and display
 - ✧ Must support UTF-8, UTF-16, or ISO-8859

Staff and SSAC Member Study/Exercise

Are there any general principles that registry operators and registrars could adopt to minimize the "Babel effect" on Registration Data query services and to ensure some uniformity of information display?

Other Conventions and Standards

- Universal Postal Union, International Address Formatting

"The addressee's address shall be written **legibly in roman letters and Arabic numerals**.

"If other letters and figures are used in the country of destination, it shall be recommended that the **address be given also in these letters and numerals**."

- Is "must support US-ASCII7, may also <local script>" appropriate for registration data?

Other Conventions and Standards

Example 1: the address of an item posted in Great Britain to the United Arab Emirates will therefore, in principle, be written once in roman characters and once in Arabic characters:

MR. OMAR HUSSAIN
P.O. BOX 111
DUBAI
UNITED ARAB EMIRATES

السيد عمر حسين
ص.ب. 111
دبي
الإمارات العربية المتحدة

In conclusion, we would therefore recommend that:

- you write addresses in roman letters and Arabic numerals; if other letters and numerals are used in the destination country, the address should also be written using these letters and numerals;
- you comply with the recommendations of the postal authority of the country in which you post your items as regards indication of the destination country (element ❶);
- you follow the recommendations of the destination country's postal authority for the remainder of the address (elements ❷);
- you always write the name of the destination country in one of the languages used in the country in which you post your items; if this language is not an internationally known language, you should add the name of the destination country in an internationally known language.

http://www.upu.int/post_code/en/formatting_an_international_address_en.pdf

Staff and SSAC Member Study/Exercise

What information, and in what languages and scripts, should be permitted when collecting and displaying registration data for a (set of) domain name(s)?

Data Schema for Successor to WHOIS Protocol

- SAC027, SAC033, and SAC037 encourage ICANN community to consider a successor to the WHOIS protocol that can improve accuracy and availability of domain registration data
- Staff and SSAC studying IRIS/CRISP/DREG standards and prior WHOIS studies to inventory service requirements
- Before enumerating requirements, taking stock of the data we collect to understand what data need to be added to support additional/future services

SAC 037: Display and Usage of Internationalized Registration Data

Reference document can be found at
[http://www.icann.org/committees/
security/sac037.pdf](http://www.icann.org/committees/security/sac037.pdf)

Questions?