UTF8 in OpenCA
(version 0.9.2.4-plus as of October 14, 2005)

Julia Dubenskaya and Sergei Vyshenski,

Cryptocom Ltd, Moscow Russia
julia@cryptocom.ru, svysh@cryptocom.ru

Munich, October 17-18, 2005

- UTF8: what for?
- Curious multi-lingual Certificate
- System bonuses of adding UTF8 support
- System problems solved by adding UTF8
- How to enable/ disable UTF8 mode
- Features in UTF8 mode
- Features in LATIN1 mode
- Backward compatibility
- Model for user/admin who prefers LATIN1 or UTF8 modes
- Known problems

UTF8: what for?

- Go with RFCs
- My lang in reqs and certs
- My lang in config.xml to name my organization and location
- Data in *different* langs on one and the same web page (certs list, langs list)
- Simultaneous support of *all* langs at a database level
- *Multi-lingual* reqs and certs

Multi-lingual Certificate:

- Issuer:

  - O: ооо Криптоком

  - L: Москва Российская Федерация

  - CN: Βάση δεδομένων

- Owner:

  - O: Technische Universität München

  - L: München Bundesrepublik Deutschland

  - CN: 署名情報

System bonuses of adding UTF8 support:

- **System is predictable**: remote lang selection is absent

- **System is simple**: no decoding, no change of charset

- **System is robust**: all supported interface langs are fully functional and cross-platform compatible with data in any lang

- **System is universal**: work with (multi-lingual?) reqs and certs in ***any*** langs. Supported ***or not supported*** by OpenCA interface

System problems solved
by adding UTF8:

- Non-controllable language switching in web interfaces

- Races between (non-UTF8) field names and (UTF8) field content in web forms

- Unstable macros processing for emails

How to enable (disable) UTF8 mode:

- Enable (disable?) UTF8 support in database..
- Pre-configure OpenCA with OpenSSL-0.9.8+ (OpenSSL-0.9.7+)

- For *__each__* OpenCA server:

  - Change cert_chars → UTF8 (LATIN1) in config.xml

  - Run configure_etc.sh

  - Restart the server

Features in UTF8 mode:

- UTF8 chars are allowed in config.xml and reqs and certs

- UTF8 support for OpenSSL is switched on. Thus OpenSSL-0.9.8+ is required.

- Web interface uses UTF8 for any supported langs (not a problem with modern browsers)

- Emails are sent **obligatory** in UTF8 for any supported langs (email may be *multi-lingual*)

http://demo2.openca.info

Features in LATIN1 mode:

- Only LATIN1 chars are allowed in config.xml and reqs and certs
- UTF8 support for OpenSSL is switched off. Thus OpenSSL-0.9.7 is ok
- Web interface uses UTF8 for any supported langs (not a problem with modern browsers)
- Emails are sent in UTF8 for any supported langs (discussible here in Munich)

Backward compatibility

- With default setting for cert_chars option in config.xml, the system goes into usual LATIN1-oriented mode
- OpenSSL-0.9.7 is supported in LATIN1 mode

------------------(non compatibility:)----------------

- Web interfaces in all langs are shown in UTF8
- Emails in all langs are sent in UTF8 (?)

Model for user/admin who prefers…

LATIN1 mode:

- does not need UTF8 reqs and certs
- likes OpenSSL-0.9.7
- does not have UTF8-enabled text editor

- has UTF8-enabled browser
- has UTF8-enabled mail client

UTF8 mode:

- needs UTF8 reqs and certs
- has OpenSSL-0.9.8+
- has UTF8-enabled text editor

- has UTF8-enabled browser
- has UTF8-enabled mail client

Known problems

- OpenCA-0.9.2 still calls a number of obsolete OpenSSL functions which does not support UTF8 (help needed)
- System will become more stable if default language "C" is coupled to UTF8 charset
- MIME::Words::encode_mimewords() has buggy UTF8 support, and should be avoided everywhere