

Baltic Region and Eastern Europe International Seminar, 4 Oct 2006, Riga, Latvia
*“The Internet and the post-WSIS Environment:
enhancing dialog among the stakeholders”*

Multilingualism, Multiculturalism and Internationalized Domain Names

Cary Karp

Director of Internet Strategy and Technology
Swedish Museum of Natural History

President and CEO

Museum Domain Management Association

There are approximately
6,500 languages
in the world today.

Each language reflects a unique world-view and culture complex, mirroring the manner in which a speech community has resolved its problems in dealing with the world, and has formulated its thinking, its system of philosophy and understanding of the world around it. In this, each language is the means of expression of the intangible cultural heritage of people ...

The Atlas of the World's Languages in Danger of Disappearing

Stephen Wurm, UNESCO, 2001

The threat to language diversity goes beyond its obvious cultural implications.

Many of the most clearly endangered languages enshrine unique knowledge that is essential for the survival of the natural environment in which the speech community lives.

The number of languages in which content is provided is large and growing, but significant obstacles must still be overcome before the corresponding language communities can derive full benefit from it.

One such issue currently being addressed is devising ways to identify a resource with the same languages that is used for its content.

Each such identifier includes
a domain name and
Internationalized Domain Names
have been available since 2003.

The deployment of IDN and the problems that it entails can be addressed on national, regional, and global levels.

This can be illustrated with the situation in Sweden, which extends directly into the broader contexts.

But first —

A Tale of Two Cities

with apologies to Charles Dickens

Riga

riga

rīga

xn--rga-uta

rīga

xn--rga-uta

rīga

xn--rga-jua94u

ru

ru

ru

ru

py

xn--p1ag

py

xn--p1ag

py

py

россия

The Swedish language is written using the Latin script, but requires several letters that are not available in the restricted segment of that script that is supported by the Domain Name System

å ä ö
é ü

This situation is shared by many other European languages that use extended Latin alphabets.

This document gives guidelines for a character set – with certain accompanying obligations – containing the letters and other characters that governmental agencies must be able to use for writing personal names, corporate names and geographic names for computerized processing.

*Character set guidelines for writing names
in Swedish electronic administration*

Date 2005-10-12 Dnr 2004/15-3

á LATIN SMALL LETTER A WITH ACUTE (U+00E1)
à LATIN SMALL LETTER A WITH GRAVE (U+00E0)
ã LATIN SMALL LETTER A WITH BREVE (U+0103)
â LATIN SMALL LETTER A WITH CIRCUMFLEX (U+00E2)
å LATIN SMALL LETTER A WITH RING ABOVE (U+00E5)
ä LATIN SMALL LETTER A WITH DIAERESIS (U+00E4)
ã LATIN SMALL LETTER A WITH TILDE (U+00E3)
ą LATIN SMALL LETTER A WITH OGONEK (U+0105)
ā LATIN SMALL LETTER A WITH MACRON (U+0101)
ă LATIN SMALL LETTER A WITH CARON (U+01CE)
æ LATIN SMALL LETTER AE (U+00E6)
ạ LATIN SMALL LETTER A WITH DOT BELOW (U+1EA1)
ǎ LATIN SMALL LETTER A WITH HOOK ABOVE (U+1EA3)
Ǻ LATIN SMALL LETTER A WITH BREVE AND ACUTE (U+1EAF)
ą̣ LATIN SMALL LETTER A WITH BREVE AND DOT BELOW (U+1EB7)
ǻ LATIN SMALL LETTER A WITH BREVE AND GRAVE (U+1EB1)
Ǽ LATIN SMALL LETTER A WITH BREVE AND HOOK ABOVE (U+1EB3)
ǽ LATIN SMALL LETTER A WITH BREVE AND TILDE (U+1EB5)
ǿ LATIN SMALL LETTER A WITH CIRCUMFLEX AND ACUTE (U+1EA5)
ậ LATIN SMALL LETTER A WITH CIRCUMFLEX AND DOT BELOW (U+1EAD)
ǻ̂ LATIN SMALL LETTER A WITH CIRCUMFLEX AND GRAVE (U+1EA7)
Ǽ̂ LATIN SMALL LETTER A WITH CIRCUMFLEX AND HOOK ABOVE (U+1EA9)
ǽ̂ LATIN SMALL LETTER A WITH CIRCUMFLEX AND TILDE (U+1EAB)
ć LATIN SMALL LETTER C WITH ACUTE (U+0107)
č LATIN SMALL LETTER C WITH CARON (U+010D)
ċ LATIN SMALL LETTER C WITH DOT ABOVE (U+010B)
ç LATIN SMALL LETTER C WITH CEDILLA (U+00E7)
ď LATIN SMALL LETTER D WITH CARON (U+010F)
đ LATIN SMALL LETTER D WITH STROKE (U+0111)

é LATIN SMALL LETTER E WITH ACUTE (U+00E9)
è LATIN SMALL LETTER E WITH GRAVE (U+00E8)
ê LATIN SMALL LETTER E WITH CIRCUMFLEX (U+00EA)
ě LATIN SMALL LETTER E WITH CARON (U+011B)
ë LATIN SMALL LETTER E WITH DIAERESIS (U+00EB)
ę LATIN SMALL LETTER E WITH OGONEK (U+0119)
ē LATIN SMALL LETTER E WITH MACRON (U+0113)
è LATIN SMALL LETTER E WITH DOT ABOVE (U+0117)
ẹ LATIN SMALL LETTER E WITH DOT BELOW (U+1EB9)
è LATIN SMALL LETTER E WITH HOOK ABOVE (U+1EBB)
ě LATIN SMALL LETTER E WITH TILDE (U+1EBD)
é̇ LATIN SMALL LETTER E WITH CIRCUMFLEX AND ACUTE (U+1EBF)
ê̇ LATIN SMALL LETTER E WITH CIRCUMFLEX AND DOT BELOW (U+1EC7)
è̇ LATIN SMALL LETTER E WITH CIRCUMFLEX AND GRAVE (U+1EC1)
ě̇ LATIN SMALL LETTER E WITH CIRCUMFLEX AND HOOK ABOVE (U+1EC3)
ẽ̇ LATIN SMALL LETTER E WITH CIRCUMFLEX AND TILDE (U+1EC5)
ə LATIN SMALL LETTER SCHWA (U+0259)
ğ LATIN SMALL LETTER G WITH BREVE (U+011F)
ġ LATIN SMALL LETTER G WITH CARON (U+01E7)
ĝ LATIN SMALL LETTER G WITH DOT ABOVE (U+0121)
ġ LATIN SMALL LETTER G WITH CEDILLA (U+0123)
g̃ LATIN SMALL LETTER G WITH STROKE (U+01E5)
ĥ LATIN SMALL LETTER H WITH STROKE (U+0127)
i LATIN SMALL LETTER I (U+0069)
ı LATIN SMALL LETTER DOTLESS I (U+0131) 2)
í LATIN SMALL LETTER I WITH ACUTE (U+00ED)
ì LATIN SMALL LETTER I WITH GRAVE (U+00EC)
î LATIN SMALL LETTER I WITH CIRCUMFLEX (U+00EE)

ı LATIN SMALL LETTER I WITH OGONEK (U+012F)
ī LATIN SMALL LETTER I WITH MACRON (U+012B)
ÿ LATIN SMALL LETTER I WITH CARON (U+01D0)
İ LATIN SMALL LETTER I WITH DOT BELOW (U+1ECB)
ï LATIN SMALL LETTER I WITH HOOK ABOVE (U+1EC9)
ĩ LATIN SMALL LETTER I WITH TILDE (U+0129)
j LATIN SMALL LETTER J (U+006A)
k LATIN SMALL LETTER K (U+006B)
ķ LATIN SMALL LETTER K WITH CARON (U+01E9)
ķ LATIN SMALL LETTER K WITH CEDILLA (U+0137)
l LATIN SMALL LETTER L (U+006C)
ĺ LATIN SMALL LETTER L WITH ACUTE (U+013A)
ļ LATIN SMALL LETTER L WITH CARON (U+013E)
ļ LATIN SMALL LETTER L WITH CEDILLA (U+013C)
ł LATIN SMALL LETTER L WITH STROKE (U+0142)
m LATIN SMALL LETTER M (U+006D)
n LATIN SMALL LETTER N (U+006E)
ń LATIN SMALL LETTER N WITH ACUTE (U+0144)
ň LATIN SMALL LETTER N WITH CARON (U+0148)
ñ LATIN SMALL LETTER N WITH TILDE (U+00F1)
ņ LATIN SMALL LETTER N WITH CEDILLA (U+0146)
ŋ LATIN SMALL LETTER ENG (U+014B)
ó LATIN SMALL LETTER O WITH ACUTE (U+00F3)
ò LATIN SMALL LETTER O WITH GRAVE (U+00F2)
ô LATIN SMALL LETTER O WITH CIRCUMFLEX (U+00F4)
ö LATIN SMALL LETTER O WITH DIAERESIS (U+00F6)
õ LATIN SMALL LETTER O WITH DOUBLE ACUTE (U+0151)
õ LATIN SMALL LETTER O WITH TILDE (U+00F5)
ō LATIN SMALL LETTER O WITH MACRON (U+014D)

œ LATIN SMALL LIGATURE OE (U+0153)
ø LATIN SMALL LETTER O WITH DOT BELOW (U+1ECD)
ø LATIN SMALL LETTER O WITH HOOK ABOVE (U+1ECF)
ó LATIN SMALL LETTER O WITH CIRCUMFLEX AND ACUTE (U+1ED1)
ô LATIN SMALL LETTER O WITH CIRCUMFLEX AND DOT BELOW (U+1ED9)
ò LATIN SMALL LETTER O WITH CIRCUMFLEX AND GRAVE (U+1ED3)
õ LATIN SMALL LETTER O WITH CIRCUMFLEX AND HOOK ABOVE (U+1ED5)
õ LATIN SMALL LETTER O WITH CIRCUMFLEX AND TILDE (U+1ED7)
ơ LATIN SMALL LETTER O WITH HORN (U+01A1)
ớ LATIN SMALL LETTER O WITH HORN AND ACUTE (U+1EDB)
ợ LATIN SMALL LETTER O WITH HORN AND DOT BELOW (U+1EE3)
ờ LATIN SMALL LETTER O WITH HORN AND GRAVE (U+1EDD)
ở LATIN SMALL LETTER O WITH HORN AND HOOK ABOVE (U+1EDF)
ỡ LATIN SMALL LETTER O WITH HORN AND TILDE (U+1EE1)
í LATIN SMALL LETTER R WITH ACUTE (U+0155)
ř LATIN SMALL LETTER R WITH CARON (U+0159)
ŗ LATIN SMALL LETTER R WITH CEDILLA (U+0157)
ś LATIN SMALL LETTER S WITH ACUTE (U+015B)
š LATIN SMALL LETTER S WITH CARON (U+0161)
ş LATIN SMALL LETTER S WITH CEDILLA (U+015F)
ß LATIN SMALL LETTER SHARP S (U+00DF)
ť LATIN SMALL LETTER T WITH CARON (U+0165)
‡ LATIN SMALL LETTER T WITH CEDILLA (U+0163)
‡ LATIN SMALL LETTER T WITH STROKE (U+0167)
ú LATIN SMALL LETTER U WITH ACUTE (U+00FA)
ù LATIN SMALL LETTER U WITH GRAVE (U+00F9)
û LATIN SMALL LETTER U WITH CIRCUMFLEX (U+00FB)
û LATIN SMALL LETTER U WITH RING ABOVE (U+016F)

ů LATIN SMALL LETTER U WITH DOUBLE ACUTE (U+0171)
ų LATIN SMALL LETTER U WITH OGONEK (U+0173)
ū LATIN SMALL LETTER U WITH MACRON (U+016B)
ů LATIN SMALL LETTER U WITH CARON (U+01D4)
ȳ LATIN SMALL LETTER U WITH DOT BELOW (U+1EE5)
ǔ LATIN SMALL LETTER U WITH HOOK ABOVE (U+1EE7)
ũ LATIN SMALL LETTER U WITH TILDE (U+0169)
ǖ LATIN SMALL LETTER U WITH HORN (U+01B0)
ǘ LATIN SMALL LETTER U WITH HORN AND ACUTE (U+1EE9)
Ǚ LATIN SMALL LETTER U WITH HORN AND DOT BELOW (U+1EF1)
ǚ LATIN SMALL LETTER U WITH HORN AND GRAVE (U+1EEB)
ǜ LATIN SMALL LETTER U WITH HORN AND HOOK ABOVE (U+1EED)
ǝ LATIN SMALL LETTER U WITH HORN AND TILDE (U+1EEF)
ŵ LATIN SMALL LETTER W WITH ACUTE (U+1E83)
Ẁ LATIN SMALL LETTER W WITH GRAVE (U+1E81)
Ẃ LATIN SMALL LETTER W WITH CIRCUMFLEX (U+0175)
Ẅ LATIN SMALL LETTER W WITH DIAERESIS (U+1E85)
ÿ LATIN SMALL LETTER Y WITH ACUTE (U+00FD)
ÿ LATIN SMALL LETTER Y WITH GRAVE (U+1EF3)
ÿ LATIN SMALL LETTER Y WITH CIRCUMFLEX (U+0177)
ÿ LATIN SMALL LETTER Y WITH DIAERESIS (U+00FF)
ÿ LATIN SMALL LETTER Y WITH DOT BELOW (U+1EF5)
ÿ LATIN SMALL LETTER Y WITH HOOK ABOVE (U+1EF7)
ÿ LATIN SMALL LETTER Y WITH TILDE (U+1EF9)
ż LATIN SMALL LETTER Z WITH ACUTE (U+017A)
ż LATIN SMALL LETTER Z WITH CARON (U+017E)
ż LATIN SMALL LETTER Z WITH DOT ABOVE (U+017C)
ż LATIN SMALL LETTER EZH (U+0292)
ż LATIN SMALL LETTER EZH WITH CARON (U+01EE)

The legally recognized minority languages in Sweden are Sami, Finnish, Meänkieli (Tornedal Finnish) Romany Chib and Yiddish.

Special laws have been adopted which entitle individuals to use Sami, Finnish and Meänkieli in dealings with administrative authorities and courts of law in the geographical areas (administrative areas) in which these languages have traditionally been used and are still widely used today.

Samiland extends across northern Norway, Sweden, Finland, and into the Kola Peninsula.

Sami is not a single language, but is a group of languages written with both Latin and Cyrillic script.

The operators of the Nordic ccTLDs are collaborating to develop IDN support for these languages, as well as the other languages that are shared across the national borders.

The .museum gTLD has its center of operation in Stockholm and is contributing its experience in dealing with even broader international multicultural requirements to this effort.

افغانستانicom.museum.	الجزائرicom.museum.	österreich.icom.museum
বাংলাদেশ.icom.museum	беларусь.icom.museum	belgië.icom.museum
българия.icom.museum	تشادرicom.museum.	中国 .icom.museum
القمرicom.museum.	кутпроç.icom.museum	مصرicom.museum.
ελλάδα.icom.museum	ísland.icom.museum	भारत.icom.museum
ایرانicom.museum.	éire.icom.museum	מ"ם.ישראלmuseum.
日本 .icom.museum	الأردنicom.museum.	қазақстан.icom.museum
한국 .icom.museum	кыргызстан.icom.museum	ᄒᆞᆫ.icom.museum
البنانicom.museum.	македонија.icom.museum	اموربتانياicom.museum.
méxico.icom.museum	монголулс.icom.museum	المغربicom.museum.
नेपाल.icom.museum	عمانicom.museum.	قطرicom.museum.
românia.icom.museum	россия.иком.museum	இலங்கை.icom.museum
españa.icom.museum	ไทย.icom.museum	تونسicom.museum.
türkiye.icom.museum	украина.icom.museum	việtnam.icom.museum
česká republika .icom.museum	македонија .icom.museum	енбүлүмүстөрү .иком.museum

Further info

- ck@nrm.museum
- <http://about.museum>
- <http://icom.museum>