Ukrainian Code Page KOI8-U
<draft-rfced-info-tkachuk-00.txt>

Status of This Memo

This document is an Internet-Draft. Internet-Drafts are working
documents of the Internet Engineering Task Force (IETF), its
areas, and its working groups. Note that other groups may also
distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six
months and may be updated, replaced, or obsoleted by other
documents at any time. It is inappropriate to use Internet-
Drafts as reference material or to cite them other than as
"work in progress."

To learn the current status of any Internet-Draft, please check
the "1id-abstracts.txt" listing contained in the Internet-
Drafts Shadow Directories on ftp.is.co.za (Africa),
nic.nordu.net (Europe), munnari.oz.au (Pacific Rim),
ds.internic.net (US East Coast), or ftp.isi.edu (US West Coast).

Distribution of this document is unlimited.

2. Abstract

   This memo is the specification of the code page, which was originally
   developed for the transmission of texts in the Ukrainian language
   through networks, and in particular the networks, employing Unix
   operating systems. It is, however, operating-system-neutral, and is
   used in various situations, just as its Russian language counterpart
   KOI-8R.

   With some minor exceptions, this specification can be considered
   to be an extension of KOI8-R (RFC-1489). The present edition of the
   specification is the result of the wide concensus reached after many
   prolonged exchanges and discussions, which took place in Ukraine and
   on some Usenet conferences covering Ukraine and read by the public
   who employs the Ukrainian language.

3. Relation to other RFCs

   This memo should be used in conjunction with RFC-1489, which is an
   official description of a Cyrillic character set, adapted and
   adopted for the Russian language, as well as with RFC-1345,

ISO10646 (Unicode) and ISO 8859-5.

[4](). Specification

The lower (7-bit) part of the Code Page KOI8-U is a perfect copy
of ASCII, just as is the case with KOI-8R and the vast majority of
other codepages.

   The upper (8-bit) part of the KOI-8U Code Page is compatible with
   all the symbols of Cyrillic alphabet used in KOI-8R, and adds
   four (4) pairs of symbols, specific to the Ukrainian alphabet.

   <hex code> <Unicode> <KOI8-U>
        AD   0491   CYRILLIC SMALL LETTER GHE WITH UPTURN
        A4   0454   CYRILLIC SMALL LETTER UKRAINIAN IE
        A6   0456   CYRILLIC SMALL LETTER BELORUSSIAN-UKRAINIAN I
        A7   0457   CYRILLIC SMALL LETTER UKRAINIAN YI
        B4   0404   CYRILLIC CAPITAL LETTER UKRAINIAN IE
        B6   0406   CYRILLIC CAPITAL LETTER BELORUSSIAN-UKRAINIAN I
        B7   0407   CYRILLIC CAPITAL LETTER UKRAINIAN YI
        BD   0490   CYRILLIC CAPITAL LETTER GHE WITH UPTURN

   Official two-letter abbreviations for three pairs of these symbols
   are defined in RFC 1345, "Character Mnemonics & Character Sets",
   June 1992

   IE   0404   CYRILLIC CAPITAL LETTER UKRAINIAN IE
   II   0406   CYRILLIC CAPITAL LETTER BYELORUSSIAN-UKRAINIAN I
   YI   0407   CYRILLIC CAPITAL LETTER YI (Ukrainian)
   ie   0454   CYRILLIC SMALL LETTER UKRAINIAN IE
   ii   0456   CYRILLIC SMALL LETTER BYELORUSSIAN-UKRAINIAN I
   yi   0457   CYRILLIC SMALL LETTER YI (Ukrainian)

   These Ukrainian symbols replace the following rarely used symbols
   present in the KOI8-R table:

   <hex code>   <KOI8-R>
        A4   FORMS DOWN DOUBLE AND RIGHT SINGLE
        A6   FORMS DOWN SINGLE AND LEFT DOUBLE
        A7   FORMS DOWN DOUBLE AND LEFT SINGLE
        AD   FORMS UP DOUBLE AND LEFT SINGLE
        B4   FORMS VERTICAL DOUBLE AND LEFT SINGLE
        B6   FORMS DOWN SINGLE AND HORIZONTAL DOUBLE
        B7   FORMS DOWN DOUBLE AND HORIZONTAL SINGLE
        BD   FORMS VERTICAL DOUBLE AND HORIZONTAL SINGLE

5. Background

   The specification given in this standard was adopted officially
   by a conference of postmasters of Ukrainian Internet Service
   Providers in the town of Slavsk (a resort in the Carpathian
   Mountains) during the autumn of 1992, from submissions by Igor
   Sviridov from Kiev and Stas Vorony from Xarkiv. The original
   specification lacked the symbol pair for the letter GHE WITH
   UPTURN; this was added by Igor Sviridov in June 1995.
   The generic KOI-8 Code Page, or, more precisely, KOI-8/KOI-7
   character sets, trace their origin from the mainframe and Unix

worlds of the 1970's. The three letter mnemonic stands for "Code
of Information Exchange" (Kod Obminu Informacieu). One of the
rationales for its creation was that, as the KOI-8 character set
becomes KOI-7 if the 8th bit is lost, the transmitted text remains
somewhat readable in KOI-7, with an arguable degree of convenience.

6. References

[1] A. Chernov, "Registration a Cyrillic Character Set", RFC 1489,
    July 1993
[2] K. Simonsen, "Character Mnemonics & Character Sets" RFC 1345,
    June 1992

7. Author's Address

Roman A. Tkacuk
BitterNet
Ternopil, Ukraine

E-mail: roman@bit.ternopil.ua

Mykola Sereda
Gerelo Intercontinental
Monreal, Canada

E-mail: gerelo@cam.org