Registration of a Georgian Character Set
draft-giasher-geostd8-00.txt

Status of this Memo

This document is an Internet-Draft and is in full conformance with all provisions of Section 10 of RFC2026.

This memo provides information for the Internet community. This memo does not specify an Internet standard of any kind. Distribution of this memo is unlimited.

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."

The list of current Internet-Drafts can be accessed at http://www.ietf.org/ietf/1id-abstracts.txt

The list of Internet-Draft Shadow Directories can be accessed at http://www.ietf.org/shadow.html

Abstract

This document provides information about character encoding GEOSTD8 (Georgian standard single byte encoding) which is a de-facto standard in Georgia for information interchange and presentation of WWW information resources in Georgian language.

This charset is suitable for use in MIME text body parts.

Introduction

Though the proposed character set "geostd8" is not currently an international standard, there is very large user community (including most of Georgian ISP and governmental organizations) supporting it. Factually, "geostd8" widely used in Georgian Internet community for mail and news exchange as well as for presentation of WWW information resources in Georgian language.

Originally, specification of proposed standard "geostd8" was officially adopted by the meeteng of representatives of different Georgian organizations at Autumn of 1999 presented by workgroup which was
created earlier by the meeting of the same representatives. Later, this standard was adopted by Standard's Commitee of Georgia.

GEOSTD8 is developed as a single byte encoding for all contemporary and old Georgian characters used in the Georgian language. It is developed for use in Microsoft Windows environments, but is also being incorporated in GNU libc.

Requirements

Characteristics:

- The proposed charset conforms to the definition of a "charset" as defined in 2.4 of RFC 2278.
- The proposed charset is suitable for use in MIME.
- The proposed charset is specified in a stable, openly available specification.

New Charsets:

- The proposed charset is NOT new charset. It is a Georgian national standard published on May 15, 1998 by Standardization Commitee of Georgia.
- The charset is suitable for use in MIME text body parts.
- The proposed charset is specified in a stable, openly available specification.

Name:

- proposed name for the charset is "geostd8"

Functionality:

- The proposed charset functions as an actual charset.
- The name conforms to the ABNF definition in sec 3.3 of RFC 2978.

Usage and Implementation:

GEOSTD8 (Georgian standard) is a de-facto standard in Georgia and used in different operation systems and Internet user applications including encoding tables, fonts, locales for many operating systems and environments.

Examples of implementation of the GEOSTD8 charset can be viewed at http://www.gia.ge/windows_9x/windows_9x.htm and at http://www.gia.ge/scrshots/ns_6/ns6.htm

Publication:

The proposed charset was published by Standardization Commitee of Georgia in May 1998. Order No.70
MIBenum:

A MIBenum value for the proposed charset will be assigned by IANA at the time of registration

Unicode mapping:

General notes: #  
Column #1 is the Geostd8 code (in hex)  
Column #2 is the Unicode (in hex as 0xXXXX)  
Column #3 is the Unicode name (follows a comment sign, '#')

<table>
<thead>
<tr>
<th>Code-point</th>
<th>Unicode</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>0x00</td>
<td>0x0000</td>
<td>NULL</td>
</tr>
<tr>
<td>0x01</td>
<td>0x0001</td>
<td>START OF HEADING</td>
</tr>
<tr>
<td>0x02</td>
<td>0x0002</td>
<td>START OF TEXT</td>
</tr>
<tr>
<td>0x03</td>
<td>0x0003</td>
<td>END OF TEXT</td>
</tr>
<tr>
<td>0x04</td>
<td>0x0004</td>
<td>END OF TRANSMISSION</td>
</tr>
<tr>
<td>0x05</td>
<td>0x0005</td>
<td>ENQUIRY</td>
</tr>
<tr>
<td>0x06</td>
<td>0x0006</td>
<td>ACKNOWLEDGE</td>
</tr>
<tr>
<td>0x07</td>
<td>0x0007</td>
<td>BELL</td>
</tr>
<tr>
<td>0x08</td>
<td>0x0008</td>
<td>BACKSPACE</td>
</tr>
<tr>
<td>0x09</td>
<td>0x0009</td>
<td>HORIZONTAL TABULATION</td>
</tr>
<tr>
<td>0x0A</td>
<td>0x000A</td>
<td>LINE FEED</td>
</tr>
<tr>
<td>0x0B</td>
<td>0x000B</td>
<td>VERTICAL TABULATION</td>
</tr>
<tr>
<td>0x0C</td>
<td>0x000C</td>
<td>FORM FEED</td>
</tr>
<tr>
<td>0x0D</td>
<td>0x000D</td>
<td>CARRIAGE RETURN</td>
</tr>
<tr>
<td>0x0E</td>
<td>0x000E</td>
<td>SHIFT OUT</td>
</tr>
<tr>
<td>0x0F</td>
<td>0x000F</td>
<td>SHIFT IN</td>
</tr>
<tr>
<td>0x10</td>
<td>0x0010</td>
<td>DATA LINK ESCAPE</td>
</tr>
<tr>
<td>0x11</td>
<td>0x0011</td>
<td>DEVICE CONTROL ONE</td>
</tr>
<tr>
<td>0x12</td>
<td>0x0012</td>
<td>DEVICE CONTROL TWO</td>
</tr>
<tr>
<td>0x13</td>
<td>0x0013</td>
<td>DEVICE CONTROL THREE</td>
</tr>
<tr>
<td>0x14</td>
<td>0x0014</td>
<td>DEVICE CONTROL FOUR</td>
</tr>
<tr>
<td>0x15</td>
<td>0x0015</td>
<td>NEGATIVE ACKNOWLEDGE</td>
</tr>
<tr>
<td>0x16</td>
<td>0x0016</td>
<td>SYNCHRONOUS IDLE</td>
</tr>
<tr>
<td>0x17</td>
<td>0x0017</td>
<td>END OF TRANSMISSION BLOCK</td>
</tr>
<tr>
<td>0x18</td>
<td>0x0018</td>
<td>CANCEL</td>
</tr>
<tr>
<td>0x19</td>
<td>0x0019</td>
<td>END OF MEDIUM</td>
</tr>
<tr>
<td>0x1A</td>
<td>0x001A</td>
<td>SUBSTITUTE</td>
</tr>
<tr>
<td>0x1B</td>
<td>0x001B</td>
<td>ESCAPE</td>
</tr>
<tr>
<td>0x1C</td>
<td>0x001C</td>
<td>FILE SEPARATOR</td>
</tr>
<tr>
<td>0x1D</td>
<td>0x001D</td>
<td>GROUP SEPARATOR</td>
</tr>
<tr>
<td>0x1E</td>
<td>0x001E</td>
<td>RECORD SEPARATOR</td>
</tr>
<tr>
<td>0x1F</td>
<td>0x001F</td>
<td>UNIT SEPARATOR</td>
</tr>
<tr>
<td>0x20</td>
<td>0x0020</td>
<td>SPACE</td>
</tr>
</tbody>
</table>
0x5A 0x005A  #LATIN CAPITAL LETTER Z
0x5B 0x005B  #LEFT SQUARE BRACKET
0x5C 0x005C  #REVERSE SOLIDUS
0x5D 0x005D  #RIGHT SQUARE BRACKET
0x5E 0x005E  #CIRCUMFLEX ACCENT
0x5F 0x005F  #LOW LINE
0x60 0x0060  #GRAVE ACCENT
0x61 0x0061  #LATIN SMALL LETTER A
0x62 0x0062  #LATIN SMALL LETTER B
0x63 0x0063  #LATIN SMALL LETTER C
0x64 0x0064  #LATIN SMALL LETTER D
0x65 0x0065  #LATIN SMALL LETTER E
0x66 0x0066  #LATIN SMALL LETTER F
0x67 0x0067  #LATIN SMALL LETTER G
0x68 0x0068  #LATIN SMALL LETTER H
0x69 0x0069  #LATIN SMALL LETTER I
0x6A 0x006A  #LATIN SMALL LETTER J
0x6B 0x006B  #LATIN SMALL LETTER K
0x6C 0x006C  #LATIN SMALL LETTER L
0x6D 0x006D  #LATIN SMALL LETTER M
0x6E 0x006E  #LATIN SMALL LETTER N
0x6F 0x006F  #LATIN SMALL LETTER O
0x70 0x0070  #LATIN SMALL LETTER P
0x71 0x0071  #LATIN SMALL LETTER Q
0x72 0x0072  #LATIN SMALL LETTER R
0x73 0x0073  #LATIN SMALL LETTER S
0x74 0x0074  #LATIN SMALL LETTER T
0x75 0x0075  #LATIN SMALL LETTER U
0x76 0x0076  #LATIN SMALL LETTER V
0x77 0x0077  #LATIN SMALL LETTER W
0x78 0x0078  #LATIN SMALL LETTER X
0x79 0x0079  #LATIN SMALL LETTER Y
0x7A 0x007A  #LATIN SMALL LETTER Z
0x7B 0x007B  #LEFT CURLY BRACKET
0x7C 0x007C  #VERTICAL LINE
0x7D 0x007D  #RIGHT CURLY BRACKET
0x7E 0x007E  #TILDE
0x7F 0x007F  #DELETE
0x80 0x20AC  #EURO SIGN
0x81  #UNDEFINED
0x82 0x201A  #SINGLE LOW-9 QUOTATION MARK
0x83  #UNDEFINED
0x84 0x201E  #DOUBLE LOW-9 QUOTATION MARK
0x85 0x2026  #HORIZONTAL ELLIPSIS
0x86 0x2020  #DAGGER
0x87 0x2021  #DOUBLE DAGGER
0x88  #UNDEFINED
0x89 0x2030  #PER MILLE SIGN
0x8A  #UNDEFINED
0x8B 0x2039  #SINGLE LEFT-POINTING ANGLE QUOTATION MARK
0x8C  #UNDEFINED
0x8D  #UNDEFINED
0x8E  #UNDEFINED
0x8F  #UNDEFINED
0x90  #UNDEFINED
0x91 0x2018  #LEFT SINGLE QUOTATION MARK
0x92 0x2019  #RIGHT SINGLE QUOTATION MARK
0x93 0x201C #LEFT DOUBLE QUOTATION MARK
0x94 0x201D #RIGHT DOUBLE QUOTATION MARK
0x95 0x2022 #BULLET
0x96 0x2013 #EN DASH
0x97 0x2014 #EM DASH
0x98 #UNDEFINED
0x99 #UNDEFINED
0x9A #UNDEFINED
0x9B 0x203A #SINGLE RIGHT-POINTING ANGLE QUOTATION MARK
0x8C #UNDEFINED
0x9D #UNDEFINED
0x9E #UNDEFINED
0x9F #UNDEFINED
0xA0 0x00A0 #NO-BREAK SPACE
0xA1 0x00A1 #INVERTED EXCLAMATION MARK
0xA2 0x00A2 #CENT SIGN
0xA3 0x00A3 #POUND SIGN
0xA4 0x00A4 #CURRENCY SIGN
0xA5 0x00A5 #YEN SIGN
0xA6 0x00A6 #BROKEN BAR
0xA7 0x00A7 #SECTION SIGN
0xA8 0x00A8 #DIAERESIS
0xA9 0x00A9 #COPYRIGHT SIGN
0xAA 0x00AA #FEMININE ORDINAL INDICATOR
0xAB 0x00AB #LEFT-POINTING DOUBLE ANGLE QUOTATION MARK
0xAC 0x00AC #NOT SIGN
0xAD 0x00AD #SOFT HYPHEN
0xAE 0x00AE #REGISTERED SIGN
0xAF 0x00AF #MACRON
0xB0 0x00B0 #DEGREE SIGN
0xB1 0x00B1 #PLUS-MINUS SIGN
0xB2 0x00B2 #SUPERSCRIPT TWO
0xB3 0x00B3 #SUPERSCRIPT THREE
0xB4 0x00B4 #ACUTE ACCENT
0xB5 0x00B5 #MICRO SIGN
0xB6 0x00B6 #PILCROW SIGN
0xB7 0x00B7 #MIDDLE DOT
0xB8 0x00B8 #CEDILLA
0xB9 0x00B9 #SUPERSCRIPT ONE
0xBA 0x00BA #MASCULINE ORDINAL INDICATOR
0xBB 0x00BB #RIGHT-POINTING DOUBLE ANGLE QUOTATION MARK
0xBC 0x00BC #VULGAR FRACTION ONE QUARTER
0xBD 0x00BD #VULGAR FRACTION ONE HALF
0xBE 0x00BE #VULGAR FRACTION THREE QUARTERS
0xBF 0x00BF #INVERTED QUESTION MARK
0xC0 0x10D0 #GEORGIAN LETTER AN
0xC1 0x10D1 #GEORGIAN LETTER BAN
0xC2 0x10D2 #GEORGIAN LETTER GAN
0xC3 0x10D3 #GEORGIAN LETTER DON
0xC4 0x10D4 #GEORGIAN LETTER EN
0xC5 0x10D5 #GEORGIAN LETTER VIN
0xC6 0x10D6 #GEORGIAN LETTER ZEN
0xC7 0x10F1 #GEORGIAN LETTER HE
0xC8 0x10D7 #GEORGIAN LETTER TAN
0xC9 0x10D8 #GEORGIAN LETTER IN
0xCA 0x10D9 #GEORGIAN LETTER KAN
0xCB 0x10DA #GEORGIAN LETTER LAS
Security Considerations

Security issues are not discussed in this memo.
References

[1]

Authors' Address

Gia Shervashidze
Georgian Internet Avenue
Tbilisi
Georgia
Fax: +995 32 958313
Phone: +995 32 959825
EMail: giasher@wanex.net
WWW: http://www.gia.ge