

Internet draft
Network Working Group
Request for Comments: XXXX

Gia Shervashidze
Georgian Internet Avenue
November 2001

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/lid-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 meeting 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 Committee 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.
a Coded Character Set

New Charsets:

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

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 Committee of Georgia in May 1998. Order No.70

SSP 18.1-98 -- Information Technologies -- 8-bit single-byte coded graphic character set -- Georgian alphabet -- Part II

Approved by Ministry of Justice of Georgia - No. 120.013.003.270

Code-point assignments for the proposed charset can be viewed at <http://www.gia.ge/g13n/geostd8.gif>

The reference table between GEOSTD8 and Unicode is available at <http://www.gia.ge/g13n/standard.htm#RefTable> - GeoWIN (Gov.)
<http://www.gia.ge/g13n/geostd8u.gif>

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, '#')

0x00	0x0000	#NULL
0x01	0x0001	#START OF HEADING
0x02	0x0002	#START OF TEXT
0x03	0x0003	#END OF TEXT
0x04	0x0004	#END OF TRANSMISSION
0x05	0x0005	#ENQUIRY
0x06	0x0006	#ACKNOWLEDGE
0x07	0x0007	#BELL
0x08	0x0008	#BACKSPACE
0x09	0x0009	#HORIZONTAL TABULATION
0x0A	0x000A	#LINE FEED
0x0B	0x000B	#VERTICAL TABULATION
0x0C	0x000C	#FORM FEED
0x0D	0x000D	#CARRIAGE RETURN
0x0E	0x000E	#SHIFT OUT
0x0F	0x000F	#SHIFT IN
0x10	0x0010	#DATA LINK ESCAPE
0x11	0x0011	#DEVICE CONTROL ONE
0x12	0x0012	#DEVICE CONTROL TWO
0x13	0x0013	#DEVICE CONTROL THREE
0x14	0x0014	#DEVICE CONTROL FOUR
0x15	0x0015	#NEGATIVE ACKNOWLEDGE
0x16	0x0016	#SYNCHRONOUS IDLE
0x17	0x0017	#END OF TRANSMISSION BLOCK
0x18	0x0018	#CANCEL
0x19	0x0019	#END OF MEDIUM
0x1A	0x001A	#SUBSTITUTE
0x1B	0x001B	#ESCAPE
0x1C	0x001C	#FILE SEPARATOR
0x1D	0x001D	#GROUP SEPARATOR
0x1E	0x001E	#RECORD SEPARATOR
0x1F	0x001F	#UNIT SEPARATOR
0x20	0x0020	#SPACE

0x21	0x0021	#EXCLAMATION MARK
0x22	0x0022	#QUOTATION MARK
0x23	0x0023	#NUMBER SIGN
0x24	0x0024	#DOLLAR SIGN
0x25	0x0025	#PERCENT SIGN
0x26	0x0026	#AMPERSAND
0x27	0x0027	#APOSTROPHE
0x28	0x0028	#LEFT PARENTHESIS
0x29	0x0029	#RIGHT PARENTHESIS
0x2A	0x002A	#ASTERISK
0x2B	0x002B	#PLUS SIGN
0x2C	0x002C	#COMMA
0x2D	0x002D	#HYPHEN-MINUS
0x2E	0x002E	#FULL STOP
0x2F	0x002F	#SOLIDUS
0x30	0x0030	#DIGIT ZERO
0x31	0x0031	#DIGIT ONE
0x32	0x0032	#DIGIT TWO
0x33	0x0033	#DIGIT THREE
0x34	0x0034	#DIGIT FOUR
0x35	0x0035	#DIGIT FIVE
0x36	0x0036	#DIGIT SIX
0x37	0x0037	#DIGIT SEVEN
0x38	0x0038	#DIGIT EIGHT
0x39	0x0039	#DIGIT NINE
0x3A	0x003A	#COLON
0x3B	0x003B	#SEMICOLON
0x3C	0x003C	#LESS-THAN SIGN
0x3D	0x003D	#EQUALS SIGN
0x3E	0x003E	#GREATER-THAN SIGN
0x3F	0x003F	#QUESTION MARK
0x40	0x0040	#COMMERCIAL AT
0x41	0x0041	#LATIN CAPITAL LETTER A
0x42	0x0042	#LATIN CAPITAL LETTER B
0x43	0x0043	#LATIN CAPITAL LETTER C
0x44	0x0044	#LATIN CAPITAL LETTER D
0x45	0x0045	#LATIN CAPITAL LETTER E
0x46	0x0046	#LATIN CAPITAL LETTER F
0x47	0x0047	#LATIN CAPITAL LETTER G
0x48	0x0048	#LATIN CAPITAL LETTER H
0x49	0x0049	#LATIN CAPITAL LETTER I
0x4A	0x004A	#LATIN CAPITAL LETTER J
0x4B	0x004B	#LATIN CAPITAL LETTER K
0x4C	0x004C	#LATIN CAPITAL LETTER L
0x4D	0x004D	#LATIN CAPITAL LETTER M
0x4E	0x004E	#LATIN CAPITAL LETTER N
0x4F	0x004F	#LATIN CAPITAL LETTER O
0x50	0x0050	#LATIN CAPITAL LETTER P
0x51	0x0051	#LATIN CAPITAL LETTER Q
0x52	0x0052	#LATIN CAPITAL LETTER R
0x53	0x0053	#LATIN CAPITAL LETTER S
0x54	0x0054	#LATIN CAPITAL LETTER T
0x55	0x0055	#LATIN CAPITAL LETTER U
0x56	0x0056	#LATIN CAPITAL LETTER V
0x57	0x0057	#LATIN CAPITAL LETTER W
0x58	0x0058	#LATIN CAPITAL LETTER X
0x59	0x0059	#LATIN CAPITAL LETTER Y

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

0xCC	0x10DB	#GEORGIAN LETTER MAN
0xCD	0x10DC	#GEORGIAN LETTER NAR
0xCE	0x10F2	#GEORGIAN LETTER HIE
0xCF	0x10DD	#GEORGIAN LETTER ON
0xD0	0x10DE	#GEORGIAN LETTER PAR
0xD1	0x10DF	#GEORGIAN LETTER ZHAR
0xD2	0x10E0	#GEORGIAN LETTER RAE
0xD3	0x10E1	#GEORGIAN LETTER SAN
0xD4	0x10E2	#GEORGIAN LETTER TAR
0xD5	0x10F3	#GEORGIAN LETTER WE
0xD6	0x10E3	#GEORGIAN LETTER UN
0xD7	0x10E4	#GEORGIAN LETTER PHAR
0xD8	0x10E5	#GEORGIAN LETTER KHAR
0xD9	0x10E6	#GEORGIAN LETTER GHAN
0xDA	0x10E7	#GEORGIAN LETTER QAR
0xDB	0x10E8	#GEORGIAN LETTER SHIN
0xDC	0x10E9	#GEORGIAN LETTER CHIN
0xDD	0x10EA	#GEORGIAN LETTER CAN
0xDE	0x10EB	#GEORGIAN LETTER JIL
0xDF	0x10EC	#GEORGIAN LETTER CIL
0xE0	0x10ED	#GEORGIAN LETTER CHAR
0xE1	0x10EE	#GEORGIAN LETTER XAN
0xE2	0x10F4	#GEORGIAN LETTER HAR
0xE3	0x10EF	#GEORGIAN LETTER JHAN
0xE4	0x10F0	#GEORGIAN LETTER HAE
0xE5	0x10F5	#GEORGIAN LETTER HOE
0xE6		#UNDEFINED
0xE7		#UNDEFINED
0xE8		#UNDEFINED
0xE9		#UNDEFINED
0xEA		#UNDEFINED
0xEB		#UNDEFINED
0xEC		#UNDEFINED
0xED		#UNDEFINED
0xEE		#UNDEFINED
0xEF		#UNDEFINED
0xF0		#UNDEFINED
0xF1		#UNDEFINED
0xF2		#UNDEFINED
0xF3		#UNDEFINED
0xF4		#UNDEFINED
0xF5		#UNDEFINED
0xF6		#UNDEFINED
0xF7		#UNDEFINED
0xF8		#UNDEFINED
0xF9		#UNDEFINED
0xFA		#UNDEFINED
0xFB		#UNDEFINED
0xFC		#UNDEFINED
0xFD	0x2116	#NUMERO SIGN
0xFE		#UNDEFINED
0xFF		#UNDEFINED

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>