

Network Working Group	E. Wilde	
Internet-Draft	UC Berkeley	
Intended status: Standards Track	A. Vaha-Sipila	
Expires: April 27, 2010	Nokia	
	October 24, 2009	

[TOC](#)

URI Scheme for GSM Short Message Service draft-wilde-sms-uri-20

Status of this Memo

This Internet-Draft is submitted to IETF in full conformance with the provisions of BCP 78 and BCP 79.

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>.

This Internet-Draft will expire on April 27, 2010.

Copyright Notice

Copyright (c) 2009 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents in effect on the date of publication of this document (<http://trustee.ietf.org/license-info>). Please review these documents carefully, as they describe your rights and restrictions with respect to this document.

Abstract

This memo specifies the Uniform Resource Identifier (URI) scheme "sms" for specifying one or more recipients for an SMS message. SMS messages are two-way paging messages that can be sent from and received by a mobile phone or a suitably equipped networked device.

Table of Contents

- [1.](#) Introduction
 - [1.1.](#) What is GSM?
 - [1.2.](#) What is SMS?
- [2.](#) The "sms" URI Scheme
 - [2.1.](#) Applicability
 - [2.2.](#) Formal Definition
 - [2.3.](#) Processing an "sms" URI
 - [2.4.](#) Comparing "sms" URIs
 - [2.5.](#) Examples of Use
 - [2.6.](#) Using "sms" URIs in HTML Forms
- [3.](#) URI Scheme Registration
 - [3.1.](#) URI Scheme Name
 - [3.2.](#) Status
 - [3.3.](#) URI Scheme Syntax
 - [3.4.](#) URI Scheme Semantics
 - [3.5.](#) Encoding Considerations
 - [3.6.](#) Applications/Protocols that use this URI Scheme Name
 - [3.7.](#) Interoperability Considerations
 - [3.8.](#) Security Considerations
 - [3.9.](#) Contact
- [4.](#) Security Considerations
- [5.](#) IANA Considerations
- [6.](#) Change Log
 - [6.1.](#) From -19 to -20
 - [6.2.](#) From -18 to -19
 - [6.3.](#) From -17 to -18
 - [6.4.](#) From -16 to -17
 - [6.5.](#) From -15 to -16
 - [6.6.](#) From -14 to -15
 - [6.7.](#) From -13 to -14
 - [6.8.](#) From -12 to -13
 - [6.9.](#) From -11 to -12
 - [6.10.](#) From -10 to -11
 - [6.11.](#) From -09 to -10
 - [6.12.](#) From -08 to -09
 - [6.13.](#) From -07 to -08
 - [6.14.](#) From -06 to -07
 - [6.15.](#) From -05 to -06
 - [6.16.](#) From -04 to -05
 - [6.17.](#) From -03 to -04
 - [6.18.](#) From -02 to -03
 - [6.19.](#) From -01 to -02
 - [6.20.](#) From -00 to -01
 - [6.21.](#) Change Log of draft-wilde-sms-service
- [7.](#) Acknowledgements
- [8.](#) References

8.1.	Normative References
8.2.	Informative References
Appendix A.	Syntax of 'telephone-subscriber'
§	Authors' Addresses

1. Introduction

[TOC](#)

The capitalized key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [[RFC2119](#)] ([Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels," March 1997.](#)).

1.1. What is GSM?

[TOC](#)

GSM (Global System for Mobile Communications) is a digital mobile phone standard which is used extensively in many parts of the world. First named after its frequency band around 900 MHz, GSM-900 has provided the basis for several other networks utilizing GSM technology, in particular GSM networks operating in the frequency bands around 1800 MHz and 1900 MHz. When referring to "GSM" in this document, we mean any of these GSM-based networks that operate a short message service.

1.2. What is SMS?

[TOC](#)

The Short Message Service (SMS) [[SMS](#)] ([European Telecommunications Standards Institute, "3GPP TS 23.040 V7.0.1 \(2007-03\): 3rd Generation Partnership Project; Technical Specification Group Core Network and Terminals; Technical realization of the Short Message Service \(SMS\) \(Release 7\)," March 2007.](#)) is an integral part of the GSM network technology. It has been very successful and currently is a major source of revenue for many GSM operators. SMS as a service is so successful that other Global Switched Telephone Network (GSTN) technologies have adopted it as well, in particular the Integrated Services Digital Network (ISDN). Because of this development, this memo uses the term "SMS client" to refer to user agents that are able to send and/or receive SMS messages.

[TOC](#)

1.2.1. SMS content

GSM SMS messages are alphanumeric paging messages that can be sent to and from SMS clients. SMS messages have a maximum length of 160 characters (7-bit characters from the GSM character set [\[SMS-CHAR\] \(European Telecommunications Standards Institute, "TS 100 900 \(GSM 03.38 version 7.2.0 Release 1998\): Digital Cellular Telecommunications System \(Phase 2+\); Alphabets and language-specific information," July 1999.\)](#)), or 140 octets. Other character sets (such as UCS-2 16-bit characters, resulting in 70 character messages) MAY also be supported [\[SMS-CHAR\] \(European Telecommunications Standards Institute, "TS 100 900 \(GSM 03.38 version 7.2.0 Release 1998\): Digital Cellular Telecommunications System \(Phase 2+\); Alphabets and language-specific information," July 1999.\)](#), but are defined as being optional by the SMS specification. Consequently, applications handling SMS messages as part of a chain of character processing applications MUST make sure that character sets are correctly mapped to and from the character set used for SMS messages.

While the 160 character variety for SMS messages is by far the most widely used one, there are numerous other content types for SMS messages, such as small bitmaps ("operator logos") and simple formats for musical notes ("ring tones"). However, these formats are proprietary and are not considered in this memo.

SMS messages are limited in length (140 octets), and the first versions of the SMS specification did not specify any standardized methods for concatenating SMS messages. As a consequence, several proprietary methods were invented, but the current SMS specification does specify message concatenation. In order to deal with this situation, SMS clients composing messages SHOULD use the standard concatenation method based on the header in the TP-User Data field as specified in the SMS specification [\[SMS\] \(European Telecommunications Standards Institute, "3GPP TS 23.040 V7.0.1 \(2007-03\): 3rd Generation Partnership Project; Technical Specification Group Core Network and Terminals; Technical realization of the Short Message Service \(SMS\) \(Release 7\)," March 2007.\)](#). When sending a message to an SMS recipient whose support for concatenated messages is unknown, the SMS client MAY opt to use the backwards-compatible (text-based) concatenation method defined in the SMS specification [\[SMS\] \(European Telecommunications Standards Institute, "3GPP TS 23.040 V7.0.1 \(2007-03\): 3rd Generation Partnership Project; Technical Specification Group Core Network and Terminals; Technical realization of the Short Message Service \(SMS\) \(Release 7\)," March 2007.\)](#). Proprietary concatenation methods SHOULD NOT be used except in closed systems, where the capabilities of the recipient(s) are always known.

1.2.2. SMS infrastructure

SMS messages can be transmitted over an SMS client's network interface using the signaling channels of the underlying GSTN infrastructure, so there is no delay for call setup. Alternatively, SMS messages may be submitted through other front-ends (for example Web-based services), which makes it possible for SMS clients to run on computers which are not directly connected to a GSTN network supporting SMS.

SMS messages sent with the GSTN SMS service MUST be sent as class 1 SMS messages, if the client is able to specify the message class.

1.2.2.1. SMS Centers

[TOC](#)

For delivery within GSTN networks, SMS messages are stored by an entity called SMS Center (SMSC), and sent to the recipient when the subscriber connects to the network. The number of a cooperative SMSC must be known to the SMS sender (i.e., the entity submitting the SMS message to a GSTN infrastructure) when sending the message (usually, the SMSC's number is configured in the SMS client and specific for the network operator to which the sender has subscribed). In most situations, the SMSC number is part of the sending SMS client's configuration. However, in some special cases (such as when the SMS recipient only accepts messages from a certain SMSC), it may be necessary to send the SMS message over a specific SMSC. The scheme specified in this memo does not support the specification of SMSC numbers, so in case of scenarios where messages have to be sent through a certain SMSC, there must be some other context establishing this requirement, or message delivery may fail.

1.2.3. Uniform Resource Identifiers

[TOC](#)

One of the core specifications for identifying resources on the Internet is RFC 3986 [\[RFC3986\]](#) (Berners-Lee, T., Fielding, R., and L. Masinter, "Uniform Resource Identifier (URI): Generic Syntax," January 2005.), specifying the syntax and semantics of a Uniform Resource Identifier (URI). The most important notion of URIs are "schemes", which define a framework within which resources can be uniquely identified and addressed. URIs enable users to access resources, and are used for very diverse schemes such as access protocols (HTTP, FTP), broadcast media (TV channels [\[RFC2838\]](#) (Zigmond, D. and M. Vickers, "Uniform Resource Identifiers for Television Broadcasts," May 2000.)), messaging (email [\[RFC2368\]](#) (Hoffmann, P., Masinter, L., and J. Zawinski, "The mailto URL scheme," June 1998.)),

and even telephone numbers (voice [\[RFC3966\]](#) ([Schulzrinne, H., "The tel URI for Telephone Numbers," December 2004.](#))).

URIs often are mentioned together with Uniform Resource Names (URNs) and/or Uniform Resource Locators (URLs), and it often is unclear how to separate these concepts. For the purpose of this memo, only the term URI will be used, referring to the most fundamental concept. The World Wide Web Consortium (W3C) has issued a note [\[uri-clarification\]](#) ([World Wide Web Consortium, "URIs, URLs, and URNs: Clarifications and Recommendations 1.0," September 2001.](#)) discussing the topic of URIs, URNs, and URLs in detail.

1.2.4. SMS Messages and the Internet

[TOC](#)

One of the important reasons for the universal access of the Web is the ability to access all information through a unique interface. This kind of integration makes it easy to provide information as well as to consume it. One aspect of this integration is the support of user agents (in the case of the Web, commonly referred to as browsers) for multiple content formats (such as HTML, GIF, JPEG) and access schemes (such as HTTP, HTTPS, FTP).

The "mailto" scheme has proven to be very useful and popular, because most user agents support it by providing an email composition facility when the user selects (e.g., clicks on) the URI. Similarly, the "sms" scheme can be supported by user agents by providing an SMS message composition facility when the user selects the URI. In cases where the user agent does not provide a built-in SMS message composition facility, the scheme could still be supported by opening a Web page which provides such a service. The specific Web page to be used could be configured by the user, so that each user could use the SMS message composition service of his choice.

The goal of this memo is to specify the "sms" URI scheme, so that user agents (such as Web browsers and email clients) can start to support it. The "sms" URI scheme identifies SMS message endpoints as resources. When "sms" URIs are dereferenced, implementations MAY create a message and present it to be edited before being sent, or they MAY invoke additional services to provide the functionality necessary for composing a message and sending it to the SMS message endpoint. In either case, simply activating a link with an "sms" URI SHOULD NOT cause a message to be sent without prior user confirmation.

1.2.4.1. SMS Messages and the Web

[TOC](#)

SMS messages can provide an alternative to "mailto" URIs [\[RFC2368\]](#) ([Hoffmann, P., Masinter, L., and J. Zawinski, "The mailto URL scheme,"](#)

[June 1998.](#)), or "tel" or "fax" URIs [\[RFC3966\] \(Schulzrinne, H., "The tel URI for Telephone Numbers," December 2004.\)](#). When a "sms" URI is activated, the user agent MAY start a program for sending an SMS message, just as "mailto" may open a mail client. Unfortunately, most browsers do not support the external handling of internally unsupported URI schemes in the same generalized way as most of them support external handling of content for media types which they do not support internally. Ideally, user agents should implement generic URI parsers and provide a way to associate unsupported schemes with external applications (or Web-based services).

The recipient of an SMS message need not be a mobile phone. It can be a server that can process SMS messages, either by gatewaying them to another messaging system (such as regular electronic mail), or by parsing them for supplementary services.

SMS messages can be used to transport almost any kind of data (even though there is a very tight size limit), but the only standardized data formats are character-based messages in different character encodings. SMS messages have a maximum length of 160 characters (when using 7-bit characters from the SMS character set), or 140 octets. However, SMS messages can be concatenated to form longer messages. It is up to the user agent to decide whether to limit the length of the message, and how to indicate this limit in its user interface, if necessary. There is one exception to this, see [Section 2.6 \(Using "sms" URIs in HTML Forms\)](#).

1.2.4.2. SMS Messages and Forms

[TOC](#)

The Hypertext Markup Language (HTML) [\[HTML401\] \(Raggett, D., Le Hors, A., and I. Jacobs, "HTML 4.01 Specification," December 1999.\)](#) provides a way to collect information from a user and pass it to a server for processing. This functionality is known as "HTML forms". A filled-in form is usually sent to the destination using the Hypertext Transfer Protocol (HTTP) or email. However, SMS messages can also be used as the transport mechanism for these forms. Depending on the network configuration, the sender's telephone number may be included in the SMS message, thus providing a weak form of authentication.

2. The "sms" URI Scheme

[TOC](#)

Syntax definitions are given using the Augmented BNF (ABNF) for syntax specifications [\[RFC5234\] \(Crocker, D., Ed. and P. Overell, "Augmented BNF for Syntax Specifications: ABNF," January 2008.\)](#).

2.1. Applicability

[TOC](#)

This URI scheme provides information that can be used for sending SMS message(s) to specified recipient(s). The functionality is comparable to that of the "mailto" URI, which (as per RFC 2368 [\[RFC2368\]](#) (Hoffmann, P., Masinter, L., and J. Zawinski, "The mailto URL scheme," June 1998.)) can also be used with a comma-separated list of email addresses.

The notation for phone numbers is taken from [\[RFC3966\]](#) (Schulzrinne, H., "The tel URI for Telephone Numbers," December 2004.) and its Erratum 203. [Appendix A \(Syntax of 'telephone-subscriber'\)](#) provides a corrected syntax of the telephone number. Refer to this document for information on why this particular format was chosen.

How SMS messages are sent to the SMSC or other intermediaries is outside the scope of this specification. SMS messages can be sent over the GSM air interface, by using a modem and a suitable protocol, or by accessing services over other protocols, such as a Web-based service for sending SMS messages. Also, SMS message service options like deferred delivery and delivery notification requests are not within the scope of this document. Such services MAY be requested from the network by the user agent if necessary.

SMS messages sent as a result of this URI MUST be sent as class 1 SMS messages, if the user agent is able to specify the message class.

2.2. Formal Definition

[TOC](#)

The URI scheme's keywords specified in the following syntax description are case-insensitive. The syntax of an "sms" URI is formally described as follows, where the URI base syntax is taken from RFC 3986 [\[RFC3986\]](#) (Berners-Lee, T., Fielding, R., and L. Masinter, "Uniform Resource Identifier (URI): Generic Syntax," January 2005.):

```
sms-uri      = scheme ":" sms-hier-part [ "?" sms-fields ]
scheme       = "sms"
sms-hier-part = sms-recipient *( "," sms-recipient )
sms-recipient = telephone-subscriber ; defined in RFC 3966
sms-fields    = sms-field *( "&" sms-field )
sms-field     = sms-field-name "=" escaped-value
sms-field-name = "body" / sms-field-ext ; "body" MUST only appear once
sms-field-ext  = 1*( unreserved )
escaped-value  = *( unreserved / pct-encoded ) ; defined in RFC 3986
```

Some illustrative examples using this syntax are given in [Section 2.5 \(Examples of Use\)](#).

The syntax definition for <telephone-subscriber> is taken from RFC 3966 [\[RFC3966\]](#) (Schulzrinne, H., "The tel URI for Telephone Numbers,"

[December 2004.](#)) (Section 5.1). Please consider Erratum 203 in that specification. For the reader's convenience, [Appendix A \(Syntax of 'telephone-subscriber'\)](#) contains a fixed syntax of the telephone number URI scheme including Erratum 203, but RFC 3966 (plus all applicable errata) is the normative reference. The description of phone numbers in RFC 3966 states (quoted from RFC 3966, Section 5.1): "The 'telephone-subscriber' part of the URI indicates the number. The phone number can be represented in either global (E.164) or local notation. All phone numbers MUST use the global form unless they cannot be represented as such. Numbers from private numbering plans, emergency ('911', '112'), and some directory-assistance numbers (e.g., '411') and other 'service codes' (numbers of the form N11 in the United States) cannot be represented in global (E.164) form and need to be represented as a local number with a context. Local numbers MUST be tagged with a 'phone-context'."

This specification defines a single <sms-field>: "body". Extensions to this specification MAY define additional fields. Extensions MUST NOT change the semantics of the specifications they are extending. Unknown fields encountered in "sms" URIs MUST be ignored by implementations. The "body" <sms-field> is used to define the body of the SMS message to be composed. It MUST not appear more than once in an "sms" URI. It consists of percent-encoded UTF-8 characters. Implementations MUST make sure that the "body" <sms-field> characters are converted to a suitable character encoding before sending, the most popular being the 7-bit SMS character encoding, another variant (though not as universally supported as 7-bit SMS) is the UCS-2 character encoding (both specified in [\[SMS-CHAR\] \(European Telecommunications Standards Institute, "TS 100 900 \(GSM 03.38 version 7.2.0 Release 1998\): Digital Cellular Telecommunications System \(Phase 2+\); Alphabets and language-specific information," July 1999.\)](#)). Implementations MAY choose to discard (or convert) characters in the <sms-body> that are not supported by the SMS character set they are using to send the SMS message. If they do discard or convert characters, they MUST notify the user.

The syntax definition for <escaped-value> refers to the text of an SMS where all <reserved> (as per RFC 3986 [\[RFC3986\] \(Berners-Lee, T., Fielding, R., and L. Masinter, "Uniform Resource Identifier \(URI\): Generic Syntax," January 2005.\)](#)) characters in the SMS text are percent-encoded, please refer to RFC 3986 [\[RFC3986\] \(Berners-Lee, T., Fielding, R., and L. Masinter, "Uniform Resource Identifier \(URI\): Generic Syntax," January 2005.\)](#) for the definition of <unreserved> and <pct-encoded>, and the details about percent-encoding.

User agents SHOULD support multiple recipients, and SHOULD make it clear to users what the entire list of recipients is, before committing the user to sending all the messages.

2.3. Processing an "sms" URI

The following list describes the steps for processing an "sms" URI:

1. The phone number of the first <sms-recipient> is extracted. It is the phone number of the final recipient and it MUST be written in international form with country code, unless the number only works from inside a certain geographical area or a network. Note that some numbers may work from several networks but not from the whole world - these SHOULD be written in international form. According to RFC 3966 [\[RFC3966\]](#) ([Schulzrinne, H., "The tel URI for Telephone Numbers," December 2004.](#)), all international numbers MUST begin with a "+" character. Hyphens, dots, and parentheses (referred to as "visual separators" in RFC 3966) are used only to improve readability and MUST NOT convey any other meaning.
2. The "body" <sms-field> is extracted, if present.
3. The user agent SHOULD provide some means for message composition, either by implementing this itself, or by accessing a service providing it. Message composition SHOULD start with the body extracted from the "body" <sms-field>, if present.
4. After message composition, a user agent SHOULD try to send the message first using the default delivery method employed by that user agent. If that fails, the user agent MAY try another delivery method.
5. If the URI contains a comma-separated list of recipients (i.e., it contains multiple <sms-recipient> parts), all of them are processed in this manner. Exactly the same message SHOULD be sent to all of the listed recipients, which means that the message resulting from the message composition step for the first recipient is used unaltered for all other recipients as well.

2.4. Comparing "sms" URIs

[TOC](#)

Two "sms" URIs are equivalent according to the following rules. Since the definition of the <telephone-subscriber> is taken from RFC 3966 [\[RFC3966\]](#) ([Schulzrinne, H., "The tel URI for Telephone Numbers," December 2004.](#)), equivalence of individual values of <telephone-subscriber> is based on the rules defined in Section 4 of RFC 3966

[\[RFC3966\] \(Schulzrinne, H., "The tel URI for Telephone Numbers," December 2004.\)](#), repeated here for convenience:

- *Both must be either a 'local-number' or a 'global-number', i.e., start with a '+'.
- *The 'global-number-digits' and the 'local-number-digits' must be equal, after removing all visual separators.
- *For mandatory additional parameters and the 'phone-context' and 'extension' parameters defined in RFC 3966 [\[RFC3966\] \(Schulzrinne, H., "The tel URI for Telephone Numbers," December 2004.\)](#), the 'phone-context' parameter value is compared as a host name if it is a 'domainname' or digit by digit if it is 'global-number-digits'. The latter is compared after removing all 'visual-separator' characters.
- *Parameters are compared according to 'pname', regardless of the order they appeared in the URI. If one URI has a parameter name not found in the other, the two URIs are not equal.
- *URI comparisons are case-insensitive.

Since "sms" URIs can contain multiple <telephone-subscriber>s as well as <sms-fields>, in addition to adopting the rules defined for comparing <telephone-subscriber> as defined by RFC 3966 [\[RFC3966\] \(Schulzrinne, H., "The tel URI for Telephone Numbers," December 2004.\)](#), two "sms" URIs are only equivalent if their <sms-fields> are identical, and if all <telephone-subscriber>s, compared pairwise as a set (i.e., without taking sequence into consideration), are equivalent.

2.5. Examples of Use

[TOC](#)

sms:+15105550101

This indicates an SMS message capable recipient at the given telephone number. The message is sent using the user agent's default SMS delivery method.

sms:+15105550101,+15105550102

This indicates SMS message capable recipients at the given telephone numbers. The identical message should be sent to both recipients using the user agent's default SMS delivery method.

sms:+15105550101?body=hello%20there

In this case, a message (initially being set to "hello there", which may be modified by the user before sending) will be sent via SMS using the user agent's default SMS delivery method.

2.6. Using "sms" URIs in HTML Forms

[TOC](#)

When using a "sms" type URI as an action URI for HTML form submission [[HTML401](#)] ([Raggett, D., Le Hors, A., and I. Jacobs, "HTML 4.01 Specification," December 1999.](#)), the form contents MUST be packaged in the SMS message just as they are packaged when using a "mailto" URI [[RFC2368](#)] ([Hoffmann, P., Masinter, L., and J. Zawinski, "The mailto URL scheme," June 1998.](#)), using the "application/x-www-form-urlencoded" media type (as defined by HTML [[HTML401](#)] ([Raggett, D., Le Hors, A., and I. Jacobs, "HTML 4.01 Specification," December 1999.](#))), effectively packaging all form data into URI compliant syntax [[RFC3986](#)] ([Berners-Lee, T., Fielding, R., and L. Masinter, "Uniform Resource Identifier \(URI\): Generic Syntax," January 2005.](#)). The SMS message MUST NOT contain any HTTP header fields, only the form data. The media type is implicit. It MUST NOT be transferred in the SMS message. Since the SMS message contains the form field values, the body <sms-field> of an "sms" type URI used for an HTML form will be ignored.

The character encoding used for form submissions MUST be UTF-8 [[RFC3629](#)] ([Yergeau, F., "UTF-8, a transformation format of ISO 10646," November 2003.](#)). It should be noted, however, that user agents MUST percent-encode form submissions before sending them (this encoding is specified by the URI syntax [[RFC3986](#)] ([Berners-Lee, T., Fielding, R., and L. Masinter, "Uniform Resource Identifier \(URI\): Generic Syntax," January 2005.](#))).

The user agent SHOULD inform the user about the possible security hazards involved when submitting the form (it is probably being sent as plain text over an air interface).

If the form submission is longer than the maximum SMS message size, the user agent MAY either concatenate SMS messages, if it is able to do so, or it MAY refuse to send the message. The user agent MUST NOT send out partial form submissions.

3. URI Scheme Registration

[TOC](#)

This memo requests the registration of the Uniform Resource Identifier (URI) scheme "sms" for specifying one or more recipients for an SMS message. The registration request complies with RFC 4395 [[RFC4395](#)]

(Hansen, T., Hardie, T., and L. Masinter, "Guidelines and Registration Procedures for New URI Schemes," February 2006.).

3.1. URI Scheme Name

[TOC](#)

sms

3.2. Status

[TOC](#)

Permanent

3.3. URI Scheme Syntax

[TOC](#)

See [Section 2 \(The "sms" URI Scheme\)](#).

3.4. URI Scheme Semantics

[TOC](#)

The "sms" URI scheme defines a way how a message may be composed which is then transmitted using the SMS message transmission method. This scheme can thus be compared to be "mailto" URI scheme [\[RFC2368\]](#) (Hoffmann, P., Masinter, L., and J. Zawinski, "The mailto URL scheme," June 1998.). See [Section 2.3 \(Processing an "sms" URI\)](#) for the details of operation.

3.5. Encoding Considerations

[TOC](#)

The optional body field of "sms" URIs may contain a message text, but this text uses percent-encoded UTF-8 characters and thus can always be represented using URI characters. See [Section 2 \(The "sms" URI Scheme\)](#) for the details of encoding.

[TOC](#)

3.6. Applications/Protocols that use this URI Scheme Name

The "sms" URI scheme is intended to be used in a similar way as the "mailto" URI scheme [\[RFC2368\] \(Hoffmann, P., Masinter, L., and J. Zawinski, "The mailto URL scheme," June 1998.\)](#). By using "sms" URIs, authors can embed information into documents which can be used as a starting point for initiating message composition. Whether the client is sending the message itself (for example over a GSM air interface) or redirecting the user to a third party for message composition (such as a Web service for sending SMS messages) is outside of the scope of the URI scheme definition.

3.7. Interoperability Considerations

[TOC](#)

No interoperability issues have been identified.

3.8. Security Considerations

[TOC](#)

See [Section 4 \(Security Considerations\)](#).

3.9. Contact

[TOC](#)

Erik Wilde
School of Information
UC Berkeley
Berkeley, CA 94720-4600
U.S.A.
tel:+1-510-6432252
mailto:dret@berkeley.edu

4. Security Considerations

[TOC](#)

SMS messages are transported without any provisions for privacy or integrity, so SMS users should be aware of these inherent security problems of SMS messages. Unlike electronic mail, where additional mechanisms exist to layer security features on top of the basic infrastructure, there currently is no such framework for SMS messages.

SMS messages very often are delivered almost instantaneously (if the receiving SMS client is online), but there is no guarantee for when SMS messages will be delivered. In particular, SMS messages between different network operators sometimes take a long time to be delivered (hours or even days) or are not delivered at all, so applications SHOULD NOT make any assumptions about the reliability and performance of SMS message transmission.

In most networks, sending SMS messages is not a free service.

Therefore, SMS clients MUST make sure that any action that incurs costs is acknowledged by the end user, unless explicitly instructed otherwise by the end user. If an SMS client has different ways of submitting an SMS message (such as a Web service and a phone line), then the end user MUST have a way to control which way is chosen.

SMS clients often are limited devices (typically mobile phones), and the sending SMS client SHOULD NOT make any assumptions about the receiving SMS client supporting any non-standard services, such as proprietary message concatenation or proprietary content types. However, if the sending SMS client has prior knowledge about the receiving SMS client, then he MAY use this knowledge to compose non-standard SMS messages.

There are certain special SMS messages defined in the SMS specification [\[SMS\] \(European Telecommunications Standards Institute, "3GPP TS 23.040 V7.0.1 \(2007-03\): 3rd Generation Partnership Project; Technical Specification Group Core Network and Terminals; Technical realization of the Short Message Service \(SMS\) \(Release 7\)," March 2007.\)](#) that can be used, for example, to turn on indicators on the phone display, or to send data to certain communication ports (comparable to UDP ports) on the device. Certain proprietary systems (for example, the Wireless Application Protocol [\[WAP\] \(WAP Forum, "Wireless Application Protocol - Architecture Specification \(WAP-210-WAPArch-20010712\)," July 2001.\)](#)) define configuration messages that may be used to reconfigure the devices remotely. Any SMS client SHOULD make sure that malicious use of such messages is not possible, for example by filtering out certain SMS User Data header fields. Gateways that accept SMS messages (e.g., in e-mail messages or Web forms) and pass them on to an SMSC SHOULD implement this kind of "firewalling" approach as well.

Because the narrow bandwidth of the SMS communications channel, there should also be checks in place for excessively long concatenated messages. As an example, it may take two minutes to transfer thirty concatenated text messages.

Unchecked input from a user MUST NOT be used to populate any other fields in an SMS message other than the User Data field (not including the User Data Header field). All other parts, including the User Data Header, of the short message should only be generated by trusted means. By including "sms" URIs in unsolicited messages (a.k.a. "spam") or other types of advertising, the originator of the "sms" URIs may attempt to reveal an individual's phone number and/or to link the identity (i.e., e-mail address) used for messaging with the identity (i.e., phone number) used for the mobile phone. This attempt to collect

information may be a privacy issue, and user agents may make users aware of that risk before composing or sending SMS messages. Users agents which do not provide any feedback about this privacy issue make users more vulnerable to this kind of attack.

A user agent SHOULD NOT send out SMS messages without the knowledge of the user, because of associated risks, which include sending masses of SMS messages to a subscriber without his consent, and the costs involved in sending an SMS message.

As suggested functionality, the user agent MAY offer a possibility for the user to filter out those phone numbers that are expressed in local format, as most premium-rate numbers are expressed in local format, and because determining the correct local context (and hence the validity of the number to this specific user) may be very difficult.

When using "sms" URIs as targets of forms (as described in [Section 2.6 \(Using "sms" URIs in HTML Forms\)](#)), the user agent SHOULD inform the user about the possible security hazards involved when submitting the form (it is probably being sent as plain text over an air interface).

5. IANA Considerations

[TOC](#)

Upon publication of this memo as an RFC, IANA has registered the "sms" URI scheme, using the template in [Section 3 \(URI Scheme Registration\)](#), in accordance with RFC 4395 [[RFC4395](#)] ([Hansen, T., Hardie, T., and L. Masinter, "Guidelines and Registration Procedures for New URI Schemes," February 2006.](#)).

6. Change Log

[TOC](#)

This section will not be part of the final RFC text, it serves as a container to collect the history of the individual draft versions. To the editor: Please remove this section before publication as RFC.

6.1. From -19 to -20

[TOC](#)

*Added [Section 2.4 \(Comparing "sms" URIs\)](#) defining how to compare SMS URIs.

*Minor editorial changes.

6.2. From -18 to -19

[TOC](#)

- *Changed phone numbers in [Section 2.5 \(Examples of Use\)](#) to match the rules for reserved phone numbers in the +1 country code (as given by <http://www.nanpa.com/nas/public/form555MasterReport.do?method=display555MasterReport>).
 - *Changed "sms-field-ext = *(unreserved)" to "sms-field-ext = 1*(unreserved)" (i.e., field name cannot be empty).
 - *Changed the registration request from "Provisional" to "Permanent".
 - *Renamed [\[8.2\]](#) from "Non-Normative References" to "Informative References".
 - *Minor text changes.
-

6.3. From -17 to -18

[TOC](#)

- *Added normative reference to [\[HTML401\] \(Raggett, D., Le Hors, A., and I. Jacobs, "HTML 4.01 Specification," December 1999.\)](#) for the "application/x-www-form-urlencoded" media type.
 - *Changed <sms-body> to become <sms-field> (with "body" as the only predefined field name) and <escaped-sms> to become <escaped-value>, to allow extensions and defined extension model to be mustIgnore.
-

6.4. From -16 to -17

[TOC](#)

- *Changed phone numbers in [Section 2.5 \(Examples of Use\)](#) to be clearly identifiable example phone numbers.
- *Changed syntax of <sms-body> to use <escaped-sms> (defined internally) instead of <query> (which had been reused from RFC 3986).
- *Removed RFC 2822 from the references (unused).
- *Moved [Section 2 \(The "sms" URI Scheme\)](#) outside of [Section 1 \(Introduction\)](#), becoming a top-level section.

*Renamed <hier-part> to <sms-hier-part> to avoid confusion with RFC 3986 tokens.

*Removed unused tokens from [Appendix A \(Syntax of 'telephone-subscriber'\)](#).

6.5. From -15 to -16

[TOC](#)

*Changed phone numbers to be defined by the <telephone-subscriber> part of RFC 3966 [\[RFC3966\] \(Schulzrinne, H., "The tel URI for Telephone Numbers," December 2004.\)](#) rather than being defined here.

*Removed the intermediate <sms-number> part and defined <sms-recipient> to directly use <telephone-subscriber>.

*Replaced reference to [\[RFC5234\] \(Crocker, D., Ed. and P. Overell, "Augmented BNF for Syntax Specifications: ABNF," January 2008.\)](#) in [Section 5 \(IANA Considerations\)](#) with correct reference to [\[RFC4395\] \(Hansen, T., Hardie, T., and L. Masinter, "Guidelines and Registration Procedures for New URI Schemes," February 2006.\)](#).

6.6. From -14 to -15

[TOC](#)

*RFC 4234 (ABNF) has been obsoleted by [\[RFC5234\] \(Crocker, D., Ed. and P. Overell, "Augmented BNF for Syntax Specifications: ABNF," January 2008.\)](#).

6.7. From -13 to -14

[TOC](#)

*Updated abstract of [Section 3 \(URI Scheme Registration\)](#) to the new abstract of the complete document (mentioning of gateway functionality removed).

*Removed the "smsc" qualifier which allowed the specification of an SMSC to be used for SMS message delivery.

- *Removed the section about interfacing between a user agent processing an "sms" URI, and a Web-based service for SMS composition and sending.
 - *Replaced "gstn-phone" with "sms-number" in [Section 2.3 \(Processing an "sms" URI\)](#) (inconsistency from earlier syntax revision).
 - *Replaced all quotes around ABNF symbols within text with <...> to clearly identify them as ABNF.
 - *Updated all references to their complete XML version.
 - *Added [Section 5 \(IANA Considerations\)](#).
-

6.8. From -12 to -13

[TOC](#)

- *Updated [\[SMS\] \(European Telecommunications Standards Institute, "3GPP TS 23.040 V7.0.1 \(2007-03\): 3rd Generation Partnership Project; Technical Specification Group Core Network and Terminals; Technical realization of the Short Message Service \(SMS\) \(Release 7\)," March 2007.\)](#) to point to the latest version of the SMS spec (3GPP TS 23.040 V7.0.1).
 - *Removed support for telematic interworking from the draft. This feature of SMS messaging is hard to understand, poorly supported by clients as well as network operators, and for these reasons would have been likely to see poor adoption in implementations anyway.
 - *Added some text making it more clear that all recipients SHOULD be sent the exact same message.
 - *Several small editorial changes.
-

6.9. From -11 to -12

[TOC](#)

- *Integrated draft-wilde-sms-service-11 and draft-wilde-sms-uri-registration-00 into this draft. This draft is now self-contained.
- *Changed the phone number notation to use the definition from RFC 3966 [\[RFC3966\] \(Schulzrinne, H., "The tel URI for Telephone](#)

[Numbers," December 2004.\)](#) rather than RFC 3601 (RFC 3966 is a simpler notation disallowing some of the special characters allowed by RFC 3601).

*Rephrased various parts of [Section 4 \(Security Considerations\)](#).

*Removed Section "Author/Change Controller".

*RFC 2806 (tel URI scheme) has been obsoleted by [\[RFC3966\] \(Schulzrinne, H., "The tel URI for Telephone Numbers," December 2004.\)](#).

6.10. From -10 to -11

[TOC](#)

*RFC 2234 (ABNF) has been obsoleted by RFC4234.

*Updated reference information for [\[SMS\] \(European Telecommunications Standards Institute, "3GPP TS 23.040 V7.0.1 \(2007-03\): 3rd Generation Partnership Project; Technical Specification Group Core Network and Terminals; Technical realization of the Short Message Service \(SMS\) \(Release 7\)," March 2007.\)](#) and [\[SMS-CHAR\] \(European Telecommunications Standards Institute, "TS 100 900 \(GSM 03.38 version 7.2.0 Release 1998\): Digital Cellular Telecommunications System \(Phase 2+\); Alphabets and language-specific information," July 1999.\)](#).

*Minor textual changes.

6.11. From -09 to -10

[TOC](#)

*Added security consideration about filtering local format phone numbers.

*Changed IPR clause from RFC 3667 to RFC 3978 (updated version of RFC 3667).

[TOC](#)

6.12. From -08 to -09

*Fixed syntax error in hier-part and sms-recipient non-terminals, which allowed sms-recipients to be concatenated without comma separation.

6.13. From -07 to -08

[TOC](#)

*URIs are now defined by RFC 3986 [\[RFC3986\] \(Berners-Lee, T., Fielding, R., and L. Masinter, "Uniform Resource Identifier \(URI\): Generic Syntax," January 2005.\)](#), so the text (including the syntax definitions) and the references have been updated.

6.14. From -06 to -07

[TOC](#)

*Changed IPR clause from RFC 2026 to RFC 3667 (updated version of RFC 2026).

6.15. From -05 to -06

[TOC](#)

*Updated reference from draft-allocchio-gstn to RFC 3601.

6.16. From -04 to -05

[TOC](#)

*Updated reference to SMS spec to the version referenced in the SMS service draft.

6.17. From -03 to -04

[TOC](#)

*Updated reference to draft-allocchio-gstn (to revision -05).

6.18. From -02 to -03

[TOC](#)

- *Changed ordering of "change Log" section (descending to ascending).
- *Clarified the wording at the beginning of [Section 2.2 \(Formal Definition\)](#) about only the keywords of the scheme being case-insensitive.
- *Changed "sms-body" to be a URI query string.
- *Added some text describing "sms" URIs as addressing resources.

6.19. From -01 to -02

[TOC](#)

- *Changed the sms-body field to percent-encoded UTF-8 characters.

6.20. From -00 to -01

[TOC](#)

- *Added the "sms-body" field and its processing rules.
- *Added section about using "sms" URIs as query strings for SMS Web services.
- *Fixed typo in ABNF (said "global-phone" instead of "gstn-phone").
- *Added some explanatory text about form submissions using email telematic interworking.
- *Added some text about character encoding in form submissions.

6.21. Change Log of draft-wilde-sms-service

[TOC](#)

This section contains the change log of draft-wilde-sms-service-11 before it was incorporated into this document at version draft-wilde-sms-uri-12.

6.21.1. From -10 to -11

[TOC](#)

*RFC 2234 (ABNF) has been obsoleted by RFC4234.

*Added new security issue in [Section 4 \(Security Considerations\)](#).

*Updated reference information for [\[SMS\] \(European Telecommunications Standards Institute, "3GPP TS 23.040 V7.0.1 \(2007-03\): 3rd Generation Partnership Project; Technical Specification Group Core Network and Terminals; Technical realization of the Short Message Service \(SMS\) \(Release 7\)," March 2007.\)](#) and [\[SMS-CHAR\] \(European Telecommunications Standards Institute, "TS 100 900 \(GSM 03.38 version 7.2.0 Release 1998\): Digital Cellular Telecommunications System \(Phase 2+\); Alphabets and language-specific information," July 1999.\)](#).

*Minor textual changes.

6.21.2. From -09 to -10

[TOC](#)

*Changed IPR clause from RFC 3667 to RFC 3978 (updated version of RFC 3667).

6.21.3. From -08 to -09

[TOC](#)

*No changes, re-release for alignment with draft-wilde-sms-uri.

6.21.4. From -07 to -08

[TOC](#)

*No changes, re-release for alignment with draft-wilde-sms-uri.

[TOC](#)

6.21.5. From -06 to -07

*Changed IPR clause from RFC 2026 to RFC 3667 (updated version of RFC 2026).

6.21.6. From -05 to -06

[TOC](#)

*Updated reference from draft-allocchio-gstn-05 to RFC 3601.

6.21.7. From -04 to -05

[TOC](#)

*No changes, re-release for alignment with draft-wilde-sms-uri.

6.21.8. From -03 to -04

[TOC](#)

*Updated reference to draft-allocchio-gstn (to revision -05).

6.21.9. From -02 to -03

[TOC](#)

*Changed ordering of "change Log" section (descending to ascending).

*Fixed some spelling errors.

6.21.10. From -01 to -02

[TOC](#)

*Removed address specification for X.400 SMS from ABNF (surprisingly not part of the SMS spec).

*Added some explanatory text about character set mapping for SMS messages.

*Added text requiring the use of message class 1 for sending SMS messages.

6.21.11. From -00 to -01

[TOC](#)

*Added a number of new security considerations.

*Added the "PID" qualif-type1 keyword and the section about "SMS Telematic Interworking" (removed in -13, available as Section 1.2.3 in -12).

7. Acknowledgements

[TOC](#)

This document has been prepared using the IETF document DTD described in RFC 2629 [\[RFC2629\]](#) (Rose, M., "Writing I-Ds and RFCs using XML," June 1999.).

Thanks to (listed alphabetically) Claudio Allocchio, Derek Atkins, Nevil Brownlee, John Cowan, Leslie Daigle, Lisa Dusseault, Miguel Garcia, Vijay Gurbani, Alfred Hoenes, Cullen Jennings, Graham Klyne, Larry Masinter, Alexey Melnikov, Michael Patton, and Magnus Westerlund for their comments.

8. References

[TOC](#)

8.1. Normative References

[TOC](#)

[HTML401]	Raggett, D., Le Hors, A., and I. Jacobs, " HTML 4.01 Specification ," W3C REC-html401, December 1999.
[RFC2119]	Bradner, S. , " Key words for use in RFCs to Indicate Requirement Levels ," BCP 14, RFC 2119, March 1997 (TXT , HTML , XML).
[RFC3629]	Yergeau, F., " UTF-8, a transformation format of ISO 10646 ," STD 63, RFC 3629, November 2003 (TXT).
[RFC3966]	Schulzrinne, H., " The tel URI for Telephone Numbers ," RFC 3966, December 2004 (TXT).
[RFC3986]	

	Berners-Lee, T., Fielding, R., and L. Masinter, "Uniform Resource Identifier (URI): Generic Syntax," STD 66, RFC 3986, January 2005 (TXT , HTML , XML).
[RFC4395]	Hansen, T., Hardie, T., and L. Masinter, " Guidelines and Registration Procedures for New URI Schemes ," BCP 115, RFC 4395, February 2006 (TXT).
[RFC5234]	Crocker, D., Ed. and P. Overell, "Augmented BNF for Syntax Specifications: ABNF," RFC 5234, January 2008.
[SMS]	European Telecommunications Standards Institute, " 3GPP TS 23.040 V7.0.1 (2007-03): 3rd Generation Partnership Project; Technical Specification Group Core Network and Terminals; Technical realization of the Short Message Service (SMS) (Release 7) ," March 2007.
[SMS-CHAR]	European Telecommunications Standards Institute, " TS 100 900 (GSM 03.38 version 7.2.0 Release 1998): Digital Cellular Telecommunications System (Phase 2+); Alphabets and language-specific information ," July 1999.

8.2. Informative References

[TOC](#)

[RFC2368]	Hoffmann, P., Masinter, L., and J. Zawinski, " The mailto URL scheme ," RFC 2368, June 1998.
[RFC2629]	Rose, M., "Writing I-Ds and RFCs using XML," RFC 2629, June 1999 (TXT , HTML , XML).
[RFC2838]	Zigmond, D. and M. Vickers, " Uniform Resource Identifiers for Television Broadcasts ," RFC 2838, May 2000 (TXT).
[WAP]	WAP Forum, "Wireless Application Protocol - Architecture Specification (WAP-210-WAPArch-20010712)," July 2001.
[uri-clarification]	World Wide Web Consortium, " URIs, URLs, and URNs: Clarifications and Recommendations 1.0 ," W3C uri-clarification, September 2001.

Appendix A. Syntax of 'telephone-subscriber'

[TOC](#)

The following syntax is reproduced from Section 3 of RFC 3966 [\[RFC3966\]](#) ([Schulzrinne, H., "The tel URI for Telephone Numbers," December 2004.](#)). It defines the <telephone-subscriber> part used in the "sms" URI scheme syntax. Please note that it includes Erratum 203 for RFC 3966, which changes the definition of <isdn-subaddress>.

```

telephone-subscriber = global-number / local-number
global-number       = global-number-digits *par
local-number        = local-number-digits *par context *par
par                 = parameter / extension / isdn-subaddress
isdn-subaddress     = ";isub=" 1*paramchar
extension           = ";ext=" 1*phonedigit
context             = ";phone-context=" descriptor
descriptor          = domainname / global-number-digits
global-number-digits = "+" *phonedigit DIGIT *phonedigit
local-number-digits =
    *phonedigit-hex (HEXDIG / "*" / "#")*phonedigit-hex
domainname          = *( domainlabel "." ) toplabel [ "." ]
domainlabel         = alphanum
                    / alphanum *( alphanum / "-" ) alphanum
toplabel            = ALPHA / ALPHA *( alphanum / "-" ) alphanum
parameter           = ";" pname ["=" pvalue ]
pname               = 1*( alphanum / "-" )
pvalue              = 1*paramchar
paramchar           = param-unreserved / unreserved / pct-encoded
unreserved          = alphanum / mark
mark                = "-" / "_" / "." / "!" / "~" / "*" /
                    "'" / "(" / ")"
pct-encoded         = "%" HEXDIG HEXDIG
param-unreserved    = "[" / "]" / "/" / ":" / "&" / "+" / "$"
phonedigit          = DIGIT / [ visual-separator ]
phonedigit-hex      = HEXDIG / "*" / "#" / [ visual-separator ]
visual-separator    = "-" / "." / "(" / ")"
alphanum            = ALPHA / DIGIT

```

Authors' Addresses

[TOC](#)

	Erik Wilde
	UC Berkeley
	Berkeley, CA 94720-4600
	U.S.A.
Phone:	+1-510-6432253
Email:	dret@berkeley.edu
URI:	http://dret.net/netdret/
	Antti Vaha-Sipila
	Nokia
Email:	antti.vaha-sipila@nokia.com
URI:	http://www.iki.fi/avs/