Text Record Type Definition

Technical Specification

NFC Forum™

RTD-Text 1.0

NFCForum-TS-RTD_Text_1.0

2011-08-01

RESTRICTIONS ON USE

NFC Forum, Inc.
401 Edgewater Place, Suite 600
Wakefield, MA, USA 01880

Last published August 14, 2013

Contents

Tables

1 Overview

The Text Record Type Description defines an NFC Forum Well Known Type [NFC RTD] for plain text data. It may be used as free form text descriptions of other objects on an RFID tag.

1.1 Objectives

The objective of this document is to function as a normative reference to the Text RTD.

1.2 Purpose

The Text RTD was designed to be used as a general purpose text field to add metadata to things such as URLs. It needs to provide a lightweight component with clearly defined semantics.

The goal is not to replace text/plain, but to define a clear subset that can be used in cases where there is not much space to be used, and to cover the most probable use cases.

The Text RTD must work well for non-western languages also, and it needs to include the language information for localization purposes so that the language can be identified and served to the user.

1.3 Applicable Documents and References

[NDEF]	NFC Data Exchange Format, Version 1.0, NFC Forum
[RTD]	NFC Record Type Definition (RTD), Version 1.0, NFC Forum
[RFC2119]	Key words for use in RFCs to Indicate Requirement Levels, RFC 2119, S. Bradner, March 1997, Internet Engineering Task Force
[RFC3066bis]	Tags for the Identification of Languages, IETF Draft A. Phillips, M. Davis, October 2005, Internet Engineering Task Force
[RFC5646]	Tags for Identifying Languages, RFC 5646, A. Phillips and M. Davis, September 2009, Internet Engineering Task Force NOTE: Was RFC3066.
[UNICODE]	The Unicode Standard, Version 4.1 The Unicode Consortium, Addison-Wellesley

1.4 Administration

The Text Record Type Definition is an open specification supported by the Near Field Communication Forum, Inc., located at:

401 Edgewater Place, Suite 600
Wakefield, MA, 01880

Tel.: +1 781-876-8955
Fax: +1 781-224-1239

http://www.nfc-forum.org

The Reference Applications Framework technical working group maintains this specification. Comments, errors, and other feedback can be submitted at http://www.nfc-forum.org/public_group_terms?document_id=5080.

1.5 Name and Logo Usage

The Near Field Communication Forum's policy regarding the use of the trademarks *NFC Forum* and the NFC Forum logo is as follows:

- Any company MAY claim compatibility with NFC Forum specifications, whether a member of the NFC Forum or not.
- Permission to use the NFC Forum logos is automatically granted to designated members only as stipulated on the most recent Membership Privileges document, during the period of time for which their membership dues are paid.
- Member's distributors and sales representatives MAY use the NFC Forum logo in promoting member's products sold under the name of the member.
- The logo SHALL be printed in black or in color as illustrated on the Logo Page that is available from the NFC Forum at the address above. The aspect ratio of the logo SHALL be maintained, but the size MAY be varied. Nothing MAY be added to or deleted from the logos.
- Since the NFC Forum name is a trademark of the Near Field Communication Forum, the following statement SHALL be included in all published literature and advertising material in which the name or logo appears:

  NFC Forum and the NFC Forum logo are trademarks of the Near Field Communication Forum.

1.6 Intellectual Property

The Text Record Type Definition Technical Specification conforms to the Intellectual Property guidelines specified in the NFC Forum's *Intellectual Property Rights Policy*, as outlined in the NFC Forum *Rules of Procedure*. These documents are available on the NFC Forum website.

1.7 Special Word Usage

The Text Record Type Definition conforms to the Intellectual Property guidelines specified in the NFC Forum's *Intellectual Property Rights Policy*, as outlined in the NFC Forum *Rules of Procedure*. These documents are available on the NFC Forum website.

1.8 Acronyms

This table defines all relevant terms and acronyms used in this specification.

Table 1: Abbreviations

Abbreviation	Description
BOM	Unicode Byte-Order-Mark
CRLF	Carriage Return – Line Feed
IETF	Internet Engineering Task Force
lsb	least significant bit
NDEF	NFC Data Exchange Format
RFU	Reserved for Future Use
RTD	Record Type Description
URI	Uniform Resource Identifier (e.g., http://, ftp://, mailto:, news:)
URL	Uniform Resource Locator (special case of a URI)

2 Text Record

2.1 Introduction

The “Text” record contains freeform plain text. It can be used to describe a service or the contents of the tag.

The Text record MAY appear as the sole record in an NDEF message [NDEF], but in this case the behavior is undefined and left to the application to handle. Typically, the Text record should be used in conjunction with other records to provide explanatory text.

2.2 Dependencies

There are no dependencies for the Text element.

2.3 Security Considerations

It is possible to write different text on the Text record than what the tag actually does, and thus spoof the user into doing something else than what he actually wanted (i.e., *phishing*). Thus it is a good idea for the user interface to use the Text field only as an informative field.

3 NDEF Structure

3.1 Messaging Sequence

There is no particular messaging sequence available for this RTD.

3.2 Records Mapping

3.2.1 Syntax

The NFC Forum Well Known Type [NDEF], [RTD] for the Text record is “T” (in NFC binary encoding: 0x54).

See Table 2 for the data content:

Table 2: Text Record Content Syntax

Offset (bytes)	Length (bytes)	Content
0	1	Status byte. See Table 3.
1	<n>	ISO/IANA language code. Examples: “fi”, “en-US”, “fr-CA”, “ja”. The encoding is US-ASCII.
n+1	<m>	The actual text. Encoding is either UTF-8 or UTF-16, depending on the status bit.

The Status bit encodings are described in Table 3. Any value marked RFU SHALL be ignored, and any software writing these bits SHALL use zero as the value for these bits.

Table 3: Status Byte Encodings

Bit number (0 is LSB)	Content
7	0: The text is encoded in UTF-8 1: The text is encoded in UTF16
6	RFU (MUST be set to zero)
5..0	The length of the IANA language code.

The contents of the text field MAY be shown to the user. If multiple 'T' records exist, the one with the closest matching language to the user preference SHOULD be displayed. To have multiple text elements within a single application, context with the same language code SHOULD be considered an error.

Control characters (0x00-0x1F in UTF-8) should be removed prior to display, except for newline, line feed (0x0D, 0x0A), and tab (0x08) characters. Markup MUST NOT be embedded (use the "text/xhtml" or other suitable MIME types). The Text record should be considered to be equal to the MIME type "text/plain; format=fixed".

Line breaks in the text MUST be represented using the CRLF (so-called DOS convention, the sequence 0x0D,0x0A in UTF-8). The device may deal with the tab character as it wishes.

White space other than newline and tab SHOULD be collapsed (i.e., multiple space characters are to be considered a single space character).

To find the length of the actual text in bytes, you calculate the length with "m=(length of the payload – length of the IANA language code – 1)"

3.2.2 Structure

If the Text record describes an element, it SHOULD occur in the NDEF record list before the element it is describing. This makes it faster to find and display to the user if the element is very large.

3.3 Language Codes

All language codes MUST be done according to [RFC5646]. The language code MAY NOT be omitted.

The language code length is encoded in the six least significant bits of the status byte. Thus it is easy to find by masking the status byte with the value 0x3F.

The language code is typically either two characters or five characters, though in the future, it is likely that it will be possible to have longer codes. At this time, IETF is considering an extension to [RFC5646] which will cover language codes up to 33 bytes in length [RFC3066bis]. The two-character version disregards any dialects, and thus is used most often; for example, "fi" for Finnish, "ja" for Japanese, "fr" for French. However, in some cases you might want to differentiate between variants of the same language, such as providing US-English and British English versions via "en-US" and "en-UK" respectively.

3.4 UTF-16 Byte Order

The Unicode Byte-Order-Mark (BOM) in the actual string MUST be tolerated (i.e., no error condition). When generating a Text record, the BOM MAY be omitted. If the BOM is omitted, the byte order shall be Big Endian (UTF-16 BE).

A. Example UTF-8 Encoding

Here's an example of how the English phrase "Hello, world!" would be encoded in UTF-8.

Table 4: Example: "Hello, world!"

Offset	Content	Explanation	Syntactical info
0	N/A	IL flag = 0 (no ID field), SF=1 (Short format)	NDEF record header
1	0x01	Length of the record name	
2	0x10	The length of the payload data (16 bytes)	
3	"T"	The binary encoding of the name, as defined in [1]	
4	0x02	Status byte: This is UTF-8, and has a two-byte language code	Payload
5	"en"	"en" is the ISO code for "English"	
7	"Hello, world!"	UTF-8 string "Hello, world!"	The actual body text.

B. Revision History

The following table outlines the revision history of the Text RTD Technical Specification.

Table 5: Revision History

Document Name	Revision and Release Date	Status	Change Notice	Supersedes
Text Record Type Definition	Version 1.0, July 2006		None	
Text Record Type Definition	Version 1.0, August 2011		Minor edits, implement CR	Version 1.0, July 2006