SenML Data Value Content-Format Indication
draft-ietf-core-senml-data-ct-01

Abstract

The Sensor Measurement Lists (SenML) media type supports multiple types of values, from numbers to text strings and arbitrary binary data values. In order to simplify processing of the data values this document proposes to specify a new SenML field for indicating the Content-Format of the data.

Status of This Memo

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

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at https://datatracker.ietf.org/drafts/current/.

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."

This Internet-Draft will expire on May 7, 2020.

Copyright Notice

Copyright (c) 2019 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 (https://trustee.ietf.org/license-info) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of
1. Introduction

The Sensor Measurement Lists (SenML) media type [RFC8428] can be used
to send various different kinds of data. In the example given in
Figure 1, a temperature value, an indication whether a lock is open,
and a data value (with SenML field "vd") read from an NFC reader is
sent in a single SenML pack.

```json
[
  { "bn": "urn:dev:ow:10e2073a01080063:", "n": "temp", "u": "Cel", "v": 7.1 },
  { "n": "open", "vb": false },
  { "n": "nfc-reader", "vd": "aGkgCg" }
]
```

Figure 1: SenML pack with unidentified binary data

The receiver is expected to know how to interpret the data in the
"vd" field based on the context, e.g., name of the data source and
out-of-band knowledge of the application. However, this context may
not always be easily available to entities processing the SenML pack.
To facilitate automatic interpretation it is useful to be able to
indicate an Internet media type and content-coding right in the SenML
Record. The CoAP Content-Format (Section 12.3 in [RFC7252]) provides
just this information; enclosing a Content-Format number (in this
case number 60 as defined for content-type application/cbor in
[RFC7049]) in the Record is illustrated in Figure 2. All registered
CoAP Content-Formats are listed in the Content-Formats subregistry of
the CoRE Parameters registry [IANA.core-parameters].
In this example SenML Record the data value contains a string "foo" and a number 42 encoded in a CBOR [RFC7049] array. Since the example above uses the JSON format of SenML, the data value containing the binary CBOR value is base64-encoded. The data value after base64 decoding is shown with CBOR diagnostic notation in Figure 3.

82        # array(2)
63        # text(3)
666F6F    # "foo"
18 2A     # unsigned(42)

Figure 3: Example Data Value in CBOR diagnostic notation

2. Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

Readers should also be familiar with the terms and concepts discussed in [RFC8428]. Awareness of terminology issues discussed in [I-D.bormann-core-media-content-type-format] can also be very helpful.

3. SenML Content-Format ("ct") Field

When a SenML Record contains a Data Value field ("vd"), the Record MAY also include a Content-Format indication field. The Content-Format indication uses label "ct" and a string value with either a CoAP Content-Format identifier in decimal form with no leading zeros except for the value "0" itself (representing an unsigned integer in the range of 0-65535, similar to the CoRE Link Format [RFC6690] "ct" attribute) or with a string containing a Content-Type and optionally a Content-Coding (see below).

The CoAP Content-Format identifier provides a simple and efficient way to indicate the type of the data. Since some Internet media types and their content coding and parameter alternatives do not have assigned CoAP Content-Format identifiers, using Content-Type and Content-Coding is also allowed. Both methods use a string value in the "ct" field to keep its data type consistent across uses. When
the "ct" field contains only digits, it is interpreted as a CoAP
Content-Format identifier.

To indicate that a Content-Coding is used with a Content-Type, the
Content-Coding value (e.g., "deflate" [RFC7230]) is appended to the
Content Type (media type and parameters, if any), separated by a "@"
sign. For example: "text/plain; charset=utf-8@deflate". If Content-
Coding is not specified with a Content-Type (no "@" sign is present
outside any media type parameters), the identity (i.e., no)
transformation is used.

4. SenML Base Content-Format ("bct") Field

The Base Content-Format Field, label "bct", provides a default value
for the Content-Format Field (label "ct") within its range. The
range of the base field includes the Record containing it, up to (but
not including) the next Record containing a "bct" field, if any, or
up to the end of the pack otherwise. Resolution (Section 4.6 of
[RFC8428]) of this base field is performed by adding its value with
the label "ct" to all Records in this range that carry a "vd" field
but do not already contain a Content-Format ("ct") field.

5. Mandatory to Understand Content-Format

If the Content-Format field needs to be understood by all processors
of the SenML Pack, the mandatory to understand versions of the
fields, "ct_" and "bct_", can be used. These fields have identical
semantics to the "ct" and "bct" fields respectively except that a
SenML processor that does not support this specification would reject
a SenML Pack with such fields and generate an error (see Section 4.4
of [RFC8428]).

Using the regular Content-Format indication enables to use this
extension in a backward compatible way to indicate information that
is not critical to be understood. The choice between the two methods
is application dependent.

If both a "ct_" field and a "ct" field are present in a resolved
Record (i.e., from fields in the Record or from base fields), the
"ct_" field overrides the "ct" field. Using both "ct" and "ct_" in
the same Record is NOT RECOMMENDED as it MAY be treated as an error
by the recipient.

6. Examples

The following examples are valid values for the "ct" and "bct" fields
(explanation/comments in parenthesis):
7. Security Considerations

The indication of a media type in the data does not exempt a consuming application from properly checking its inputs. Also, the ability for an attacker to supply crafted SenML data that specify media types chosen by the attacker may expose vulnerabilities of handlers for these media types to the attacker.

8. IANA Considerations

IANA is requested to assign new labels in the "SenML Labels" subregistry of the SenML registry [IANA.senml] (as defined in [RFC8428]) for the Content-Format indication as per Table 1:

<table>
<thead>
<tr>
<th>Name</th>
<th>Label</th>
<th>JSON Type</th>
<th>XML Type</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base Content-Format</td>
<td>bct</td>
<td>String</td>
<td>string</td>
<td>RFC-AAAA</td>
</tr>
<tr>
<td>Content-Format</td>
<td>ct</td>
<td>String</td>
<td>string</td>
<td>RFC-AAAA</td>
</tr>
</tbody>
</table>

Table 1: IANA Registration for new SenML Labels

Acknowledgements

The authors would like to thank Sergio Abreu for the discussions leading to the design of this extension and Isaac Rivera for reviews and feedback.
10. References

10.1. Normative References


10.2. Informative References


Authors’ Addresses

Ari Keranen
Ericsson
Jorvaa 02420
Finland

Email: ari.keranen@ericsson.com

Carsten Bormann
Universitaet Bremen TZI
Postfach 330440
Bremen D-28359
Germany

Phone: +49-421-218-63921
Email: cabo@tzi.org