"Comparable" JSON (JSONCOMP)
draft-rundgren-comparable-json-00

Abstract

This application note describes how JCS [JCS] can be utilized to support applications needing canonicalization beyond the core JSON [RFC8259] level, with comparisons as the primary target.

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 August 13, 2019.

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 the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.
1. Introduction

The purpose of JCS [JCS] is creating "Hashable" JSON [RFC8259] data intended for cryptographic solutions. JCS accomplishes this by combining normalization of the native JSON String and Number primitives with a deterministic property sorting scheme. That is, JCS provides canonicalization at the core JSON level. For interoperability reasons JCS also constrains data to the I-JSON [RFC7493] subset.

However, if you rather would like to compare JSON data from different sources or runs, JCS would in many cases be inadequate since the JSON String type is commonly used for holding subtypes like "DateTime" or "BigNumber" objects.

This application note outlines how JCS in spite of having a limited canonicalization scope still may be utilized by applications like above.

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.

3. String Subtype Normalization

Assume you want to compare productions of JSON data where the schema designer assigned the property "big" for holding a "BigInteger" subtype and "time" for holding a "DateTime" subtype, while "val" is supposed to be a JSON Number compliant with JCS. The example below shows such a string:
{"time": "2019-01-28T07:45:10Z", "big": "055", "val": 3.5}

A problem here is that "055" clearly is not a canonical form for a "BigInteger" while a "DateTime" object like "2019-01-28T07:45:10Z" might as well be expressed as "2019-01-28T08:45:10.000+01:00" making comparisons based on JCS canonicalization fail.

To resolve this issue using JCS the following measures MUST be taken:

- The community or standard utilizing a specific JSON schema defines a strict normalized form for each of the used subtypes.
- Compatible serializers are created for each subtype.

A positive side effect of this arrangement is that it enforces strict definitions of subtypes which improves interoperability in general as well.

Below is an example how such a serializer could be expressed in ECMAScript [ECMASCRIPT] for a "DateTime" subtype:

```javascript
Date.prototype.toJSON = function() {
    let date = this.toISOString();
    // In this particular case we selected a UTC notation
    // yyyy-mm-ddThh:mm:ssZ
    return date.substring(0, date.indexOf('.')) + 'Z';
};
```

Defining specific subtypes and their normalized form is out of scope for this application note.

4. IANA Considerations

This document has no IANA actions.

5. Security Considerations

TBD.

6. Acknowledgements

TBD.

7. References
7.1. Normative References


7.2. Informal References


Author’s Address

Anders Rundgren
Independent
Montpellier
France

Email: anders.rundgren.net@gmail.com
URI: https://www.linkedin.com/in/andersrundgren/