

JSON
Internet Draft
[draft-crockford-jsonorg-json-01.txt](http://www.ietf.org/internet-drafts/draft-crockford-jsonorg-json-01.txt)
Intended status: Informational
Expires: June 10, 2006

D. Crockford
JSON.org
February, 2006

JSON

Status of this Memo

This document may not be modified, and derivative works of it may not be created, except to publish it as an RFC and to translate it into languages other than English.

By submitting this Internet-Draft, each author represents that any applicable patent or other IPR claims of which he or she is aware have been or will be disclosed, and any of which he or she becomes aware will be disclosed, in accordance with [Section 6 of 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/lid-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 June 10, 2006.

Copyright Notice

Copyright (C) The Internet Society (2006).

Abstract

JSON (JavaScript Object Notation) is a light-weight, text-based, language-independent, data interchange format. It was derived from the ECMAScript Programming Language Standard. JSON defines a small set of formatting rules for the portable representation of structured data.

Conventions used in this document

The 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](#).

The syntax rules in this document are to be interpreted as described in [RFC-2234](#).

[1](#). Introduction

JSON, or JavaScript Object Notation, is a text format for the serialization of structured data. It is derived from the object literals of JavaScript, as defined in the ECMAScript Programming Language Standard [[ECMA](#)].

JSON can represent four primitive types (strings, numbers, booleans, and null) and two structured types (objects and arrays).

A string is a sequence of zero or more Unicode characters.

An object is an unordered collection of zero or more name/value pairs, where a name is a string, and a value is a string, number, boolean, null, object, or array.

An array is an ordered sequence of zero or more values.

The terms "object" and "array" come from the conventions of JavaScript.

JSON's design goals were to be minimal, portable, textual, and a subset of JavaScript.

[2](#). JSON Grammar

A JSON text is a sequence of tokens. The set of tokens includes six structural characters, strings, numbers, and three literal names.

A JSON text is a serialized object or array.

JSON-text = object / array

These are the six structural characters:

begin-array	= %x5B	; [left square bracket
begin-object	= %x7B	; { left curly bracket
end-array	= %x5D	;] right square bracket
end-object	= %x7D	; } right curly bracket

name-separator = %x3A ; : colon

value-separator = %x2C ; , comma

2.1. Whitespace

The tokens MAY be separated by any combination of these insignificant whitespace characters:

space	U+0020	Space
HT	U+0009	Horizontal tab
LF	U+000A	Line feed or New line
CR	U+000D	Carriage return

Insignificant whitespace MUST NOT be placed within a multicharacter token (a literal name, number, or string). A space character in a string is significant.

2.2. Values

A JSON value MUST be a object, array, number, or string, or one of the three literal names:

false null true

The literal names MUST be in lower case. No other literal names are allowed.

value = false / null / true / object / array / number / string

false = %x66.61.6c.73.65 ; false

null = %x6e.75.6c.6c ; null

true = %x74.72.75.65 ; true

2.3. Objects

An object structure is represented as a pair of curly brackets surrounding zero or more name/value pairs (or members). A name is a string. A single colon comes after each name, separating the name from the value. A single comma separates a value from a following name.

object = begin-object [member *(value-separator member)]
 end-object

member = string name-separator value

2.4. Arrays

An array structure is represented as square brackets surrounding zero or more values (or elements). Elements are separated by commas.

```
array = begin-array [ value *( value-separator value ) ]
        end-array
```

2.5. Numbers

The representation of numbers is similar to that used in most programming languages. A number contains an integer component which may be prefixed with an optional minus sign, which may be followed by a fraction part and/or an exponent part.

Octal and hex forms are not allowed. Leading zeros are not allowed.

A fraction part is a decimal point followed by one or more digits.

An exponent part begins with the letter E in upper or lower case, which may be followed by a plus or minus sign. The E and optional sign are followed by one or more digits.

Numeric values that cannot be represented as sequences of digits (such as Infinity and NaN) are not permitted.

```
number = [ minus ] int [ frac ] [ exp ]

decimal-point = %x2E          ; .

digit1-9 = %x31-39           ; 1-9

e = %x65 / %x45              ; e E

exp = e [ minus / plus ] 1*DIGIT

frac = decimal-point 1*DIGIT

int = zero / ( digit1-9 *DIGIT )

minus = %x2D                  ; -

plus = %x2B                   ; +

zero = %x30                   ; 0
```

2.6. Strings

The representation of strings is similar to conventions used in the C family of programming languages. A string begins and ends with quotation marks. All Unicode characters may be placed within

the quotation marks except for the characters which must be escaped: quotation mark, reverse solidus, and the control characters (U+0000 through U+001F).

Any character may be escaped. If the character is in the Basic Multilingual Plane (U+0000 through U+FFFF) then it may be represented as a six-character sequence: a reverse solidus followed by the lower case letter u followed by four hexadecimal digits which encode the character's code point. The hexadecimal letters A through F can be in upper or lower case. So, for example, a string containing only a single reverse solidus character may be represented as `"\u005C"`.

Alternatively, there are two-character sequence escape representations of some popular characters. So, for example, a string containing only a single reverse solidus character may be represented more compactly as `"\\"`.

To escape an extended character that is not in the Basic Multilingual Plane, then the character is represented as a twelve-character sequence, encoding the UTF-16 surrogate pair. So, for example, a string containing only the G clef character (U+1D11E) may be represented as `"\uD834\uDD1E"`.

```
string = quotation-mark *char quotation-mark

char = unescaped /
      escape (
          %x22 /           ; "      quotation mark  U+0022
          %x5C /           ; \      reverse solidus U+005C
          %x2F /           ; /      solidus          U+002F
          %x62 /           ; b      backspace        U+0008
          %x66 /           ; f      form feed        U+000C
          %x6E /           ; n      line feed        U+000A
          %x72 /           ; r      carriage return  U+000D
          %x74 /           ; t      tab              U+0009
          %x75 4HEXDIG )   ; uXXXX          U+XXXX

escape = %x5C             ; \

quotation-mark = %x22     ; "

unescaped = %x20-21 / %x23-5B / %x5D-10FFFF
```

[3. Encoding](#)

JSON text SHOULD be encoded in Unicode. The default encoding is UTF-8.

Since the first two characters of a JSON text will always be ASCII characters, it is possible to determine if an octet stream is

UTF-8, UTF-16 (BE or LE), or UTF-32 (BE or LE) by looking at the pattern of nulls in the first four octets.

00	00	00	xx	UTF-32BE
00	xx	00	xx	UTF-16BE
xx	00	00	00	UTF-32LE
xx	00	xx	00	UTF-16LE
xx	xx	xx	xx	UTF-8

[4. Parsers](#)

A JSON parser transforms a JSON text into another representation. A JSON parser **MUST** accept all texts that conform to the JSON grammar. A JSON parser **MAY** accept non-JSON forms or extensions.

An implementation may set limits on the size of texts that it accepts. An implementation may set limits on the maximum depth of nesting. An implementation may set limits on the range of numbers. An implementation may set limits on the length and character contents of strings.

[5. Generators](#)

A JSON generator produces JSON text. The resulting text **MUST** strictly conform to the JSON grammar.

[6. IANA Considerations](#)

The MIME media type for JSON text is text/json.

[7. Security Considerations](#)

Generally there are security issues with scripting languages. JSON is a subset of JavaScript, but it is a safe subset that excludes assignment and invocation.

A JSON text can be safely passed into JavaScript's eval() function (which compiles and executes a string) if all of the characters not enclosed in strings are in the set of characters which form JSON tokens. This can be quickly determined in JavaScript with two regular expressions and calls to the test and replace methods.

```
var my_JSON_object = !(/[^\,:{}\[\]0-9.\-+Eaeflnr-u \n\r\t]/.test(
    text.replace(/"(\.|\\"|\\")*/g, ''))) &&
    eval('(' + text + ')');
```

[8. References](#)

[8.1 Normative References](#)

[ECMA] European Computer Manufacturers Association, "ECMAScript

Language Specification 3rd Edition", December 1999,
<<http://www.ecma-international.org/publications/files/ecma-st/ECMA-262.pdf>>.

[UNICODE] The Unicode Consortium, "The Unicode Standard
Version 4.0", 2003,
<<http://www.unicode.org/versions/Unicode4.1.0/>>.

Author's Address

Douglas Crockford
JSON.org
Contact Email: douglas@crockford.com

Intellectual Property Statement

The IETF takes no position regarding the validity or scope of any Intellectual Property Rights or other rights that might be claimed to pertain to the implementation or use of the technology described in this document or the extent to which any license under such rights might or might not be available; nor does it represent that it has made any independent effort to identify any such rights. Information on the procedures with respect to rights in RFC documents can be found in [BCP 78](#) and [BCP 79](#).

Copies of IPR disclosures made to the IETF Secretariat and any assurances of licenses to be made available, or the result of an attempt made to obtain a general license or permission for the use of such proprietary rights by implementers or users of this specification can be obtained from the IETF on-line IPR repository at <http://www.ietf.org/ipr>.

The IETF invites any interested party to bring to its attention any copyrights, patents or patent applications, or other proprietary rights that may cover technology that may be required to implement this standard. Please address the information to the IETF at ietf-ipr@ietf.org.

Disclaimer of Validity

This document and the information contained herein are provided on an "AS IS" basis and THE CONTRIBUTOR, THE ORGANIZATION HE/SHE REPRESENTS OR IS SPONSORED BY (IF ANY), THE INTERNET SOCIETY AND THE INTERNET ENGINEERING TASK FORCE DISCLAIM ALL WARRANTIES, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO ANY WARRANTY THAT THE USE OF THE INFORMATION HEREIN WILL NOT INFRINGE ANY RIGHTS OR ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Copyright Statement

Copyright (C) The Internet Society (2006). This document is subject to the rights, licenses and restrictions contained in [BCP 78](#), and except as set forth therein, the authors retain all their rights.

This Internet-Draft will expire on June 10, 2006.