



JavaScript Object Notation (JSON) and Extensible Markup Language (XML)

by Sophia



WHAT'S COVERED

In this lesson, you will be introduced to two of the common forms of text-based data storage that can be used to store complex data objects as plain text. You will learn about JSON and XML standards and how they are used to store and communicate complex data objects.

Specifically, this lesson will cover the following:

1. [Introduction to Web Data Formats](#)
2. [XML](#)
3. [JSON](#)

1. Introduction to Web Data Formats

The transmission of information on the web primarily uses text-based data, which is why the protocol is named HTTP (Hyper *Text* Transfer Protocol). As such, developers needed a way to transmit simple and complex data structures using some type of text-based format. **XML (Extensible Markup Language)** and **JSON (JavaScript Object Notation)** address this need, providing frameworks for representing and exchanging data in a structured manner that goes beyond the capabilities of HTML, which is primarily designed for document markup.

XML was originally designed and specified by the W3C (World Wide Web Consortium) back in 1998 as a free and open-source framework for specifying complex object structures and easily serializing, transmitting, and reconstructing data structures.

JSON was designed a little later in the early 2000s as an open standard for a human-readable data interchange format, similar to XML but using a completely different syntax. JSON was derived from the syntax format of a JavaScript object but has been adopted by a wide range of other programming languages.

Both of these languages were designed to provide a format for data exchange on the web. However, JSON provides less overhead and can be used for real-time communications between the client and the server. While

JSON is smaller, the drawback is that JSON data is less expressive and more restrictive, providing no method for including metadata or comments.



TERMS TO KNOW

XML (Extensible Markup Language)

A markup language that provides rules to define simple and complex data structures.

JSON (JavaScript Object Notation)

An open standard file format and data interchange format that uses human-readable text to store and transmit data.

2. XML

XML is an excellent option for transmitting and storing simple and complex data structures to and from clients and servers. XML's syntax resembles an HTML document in that it uses sets of tags to surround and mark up data values and executable code. The XML tags themselves are organized to indicate the structure and relationship of the content as well as describe the semantic meaning of each piece.

The “extensibility” of XML means that the user has complete control of tag names and organization. Unlike an HTML document, there are no predefined structures or patterns that developers have to follow other than their own. This provides a lot of freedom to the developer as to how XML can be structured, organized, and used. XML is also a common storage data type for database management systems (DBMS).

The benefits of XML include the following:

- Human-readable
- Highly extensible
- Supported by common DBMS
- Able to describe complex data structures

The drawbacks of XML include the following:

- Larger file sizes due to syntax
- Slower to process
- More complex structures

As mentioned above, XML does not use any predefined tags. Instead, the developer chooses the most appropriate names for the different tags and structures them as needed. For example, if we wanted to create a directory of movies in a collection, we might structure it using the outermost container with the tag name “moviecollection,” where each movie could be contained within a “movie” tag, and then there would be sets of tags for the movie’s title, director, year, star, genre, IMDb rating, run-time, description, and so on.

🔗 **EXAMPLE** Movie collection sample of XML data

```
<?xml version="1.0" encoding="UTF-8"?>
<moviecollection>

  <movie>
    <title>Willy Wonka & the Chocolate Factory</title>
    <director>Mel Stuart</director>
    <year>1971</year>
    <star>Gene Wilder</star>
    <genre>Mixed</genre>
    <IMDb>7.8</imdb>
  </movie>

  <movie>
    <title>Predator</title>
    <director>John McTiernan</director>
    <year>1987</year>
    <star>Arnold Schwarzenegger</star>
    <star>Carl Weathers</star>
    <runtime>107</runtime>
    <IMDb>7.8</imdb>
  </movie>

</moviecollection>
```

While the above example shows a relatively shallow data structure of movies, XML supports complex structures that will be able to meet the needs of different types of documents and storage.

3. JSON

JSON was derived from JavaScript's object syntax and has become a common communication format for client-server communications and interactions. Compared to XML, JSON's syntax is minimal, and the files tend to be much smaller as well. The drawback to JSON is that it does not support as many data types as XML does and is not as extensible as XML, resulting in less flexibility and variety.

The benefits of JSON include the following:

- Human-readable
- Easier to write
- Lightweight (fast and low bandwidth overhead)
- Highly compatible with JavaScript

The drawbacks of JSON include the following:

- Is less flexible than XML
- Supports fewer data types than XML
- Does not allow commenting, attributes, namespaces, or metadata

JSON syntax consists of arrays and objects combined in any manner that meets the needs of the developer. A JSON object consists of key–value pairs; the key must be a string surrounded by quotation marks. The key and value are separated by a colon, and each key–value pair is separated by a comma. The last key–value pair in an object should not have a trailing comma. Commas are used to separate key–value pairs within an object, but a comma should not follow the last key–value pair in the object. The value of any key can be a simple data value or another object or array.

⇒ EXAMPLE Sample JSON object syntax

```
{
  "first_name": "John",
  "last_name": "Doe",
  "is_alive": true,
  "age": 27,
  "address": {
    "street_address": "21 2nd Street",
    "city": "New York",
    "state": "NY",
    "postal_code": "10021-3100"
  }
}
```

If you examine the JSON object in the image above, you will see that there are key (brown text) and value (blue text) pairs for each piece of information related to a person. Furthermore, notice how the value for the address key is also an object with its own set of key–value pairs describing the parts of the address.

JSON arrays are comma-separated sequences of individual values, the value of which can be a simple data value, an array, or an object.

⇒ EXAMPLE Sample JSON array syntax

```
{
  "first_name": "John",
  "last_name": "Doe",
  "is_alive": true,
  "age": 27,
  "address": {
    "street_address": "21 2nd Street",
    "city": "New York",
```

```
"state": "NY",
"postal_code": "10021-3100"
},
"phone_numbers": [
{
"type": "home",
"number": "212 555-1234"
},
{
"type": "office",
"number": "646 555-4567"
}
],
"children": [
"Catherine",
"Thomas",
"Trevor"
]]
```

We expanded the “person” example to include an array of phone numbers and children. Each element within the phone numbers array is an object itself. This is so we can store the type of phone number (i.e., home, work, mobile, etc.) and the number itself. If we look at the “children” example, the value of the “children” key is a simple array of names.



SUMMARY

In this lesson, you were introduced to two popular text-based **data formats** for long-term storage of complex data structures and objects. You learned about the **XML** standard and how it can be used to store complex data structures such as documents. Next, you learned about the **JSON** standard, which was based on the JavaScript object syntax. You also learned about the differences and ideal uses for both standards.

Source: This Tutorial has been adapted from "The Missing Link: An Introduction to Web Development and Programming " by Michael Mendez. Access for free at <https://open.umn.edu/opentextbooks/textbooks/the-missing-link-an-introduction-to-web-development-and-programming>. License: **Creative Commons attribution: CC BY-NC-SA**.



TERMS TO KNOW

JSON (JavaScript Object Notation)

An open standard file format and data interchange format that uses human-readable text to store and transmit data.

XML (Extensible Markup Language)

A markup language that provides rules to define simple and complex data structures.