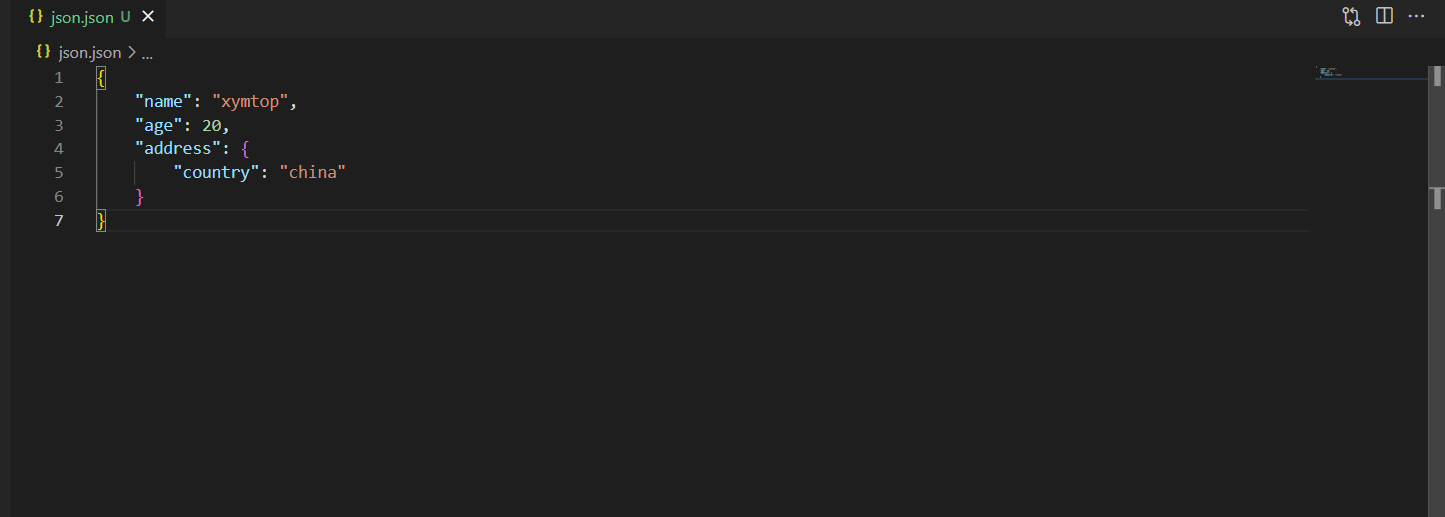
# JSON与XML

JSON（[JavaScript](https://baike.baidu.com/item/JavaScript" \t "https://baike.baidu.com/item/JSON/_blank) Object Notation, JS对象简谱）是一种轻量级的数据交换格式。它基于 [ECMAScript](https://baike.baidu.com/item/ECMAScript" \t "https://baike.baidu.com/item/JSON/_blank)（European Computer Manufacturers Association, 欧洲计算机协会制定的js规范）的一个子集，采用完全独立于编程语言的文本格式来存储和表示数据。简洁和清晰的层次结构使得 JSON 成为理想的数据交换语言。 易于人阅读和编写，同时也易于机器解析和生成，并有效地提升网络传输效率。

JSON（JavaScript Object Notation）是一种轻量级的数据交换格式。易于人阅读和编写，可以在多种语言之间进行数据交换。同时也易于机器解析和生成。它基于JavaScript Programming Language, Standard ECMA-262 3rd Edition - December 1999的一个子集。



JSON是Douglas Crockford在2001年开始推广使用的数据格式，在2005年-2006年正式成为主流的数据格式，雅虎和谷歌就在那时候开始广泛地使用JSON格式。

任何支持的类型都可以通过 JSON 来表示，例如字符串、数字、对象、数组等。但是对象和数组是比较特殊且常用的两种类型。

对象在 JS 中是使用花括号包裹 {} 起来的内容，数据结构为 {key1：value1, key2：value2, ...} 的键值对结构。在面向对象的语言中，key 为对象的属性，value 为对应的值。键名可以使用整数和字符串来表示。值的类型可以是任意类型。

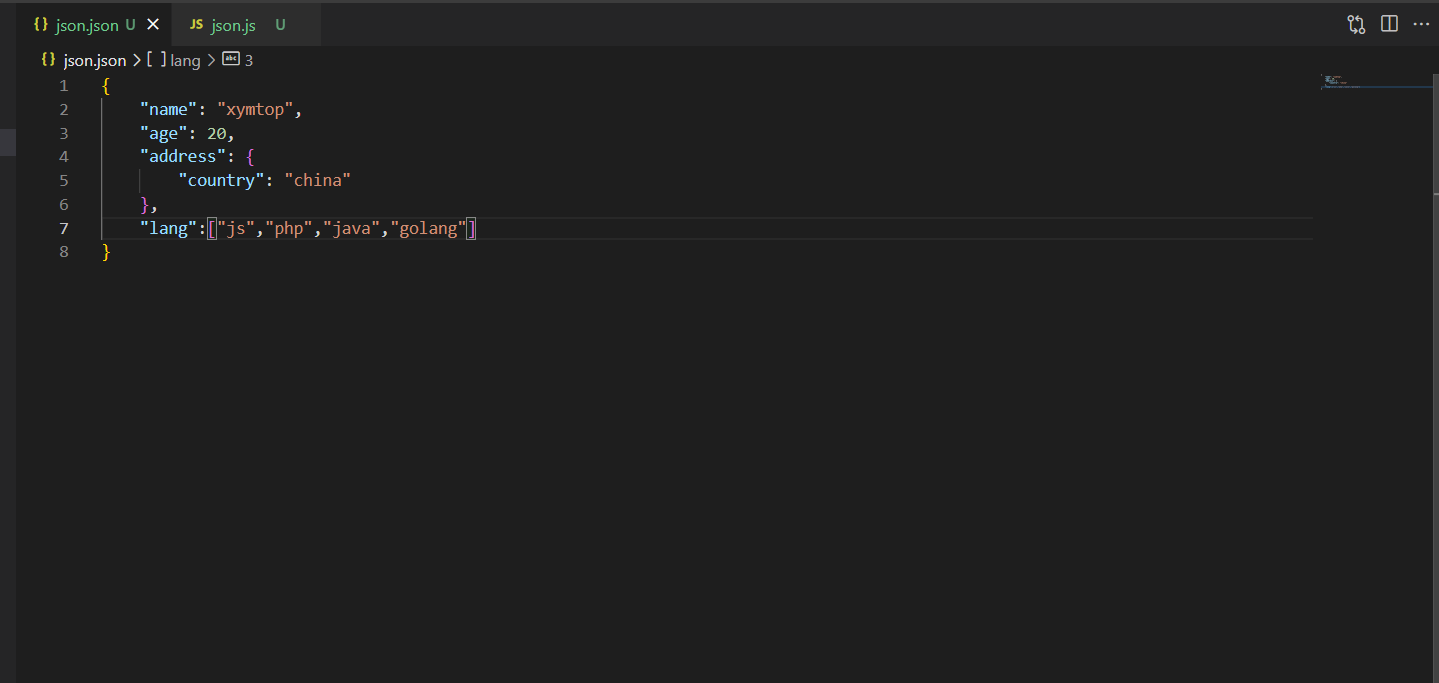


数组：数组在 JS 中是方括号 [] 包裹起来的内容，数据结构为 ["java", "javascript", "vb", ...] 的索引结构。在 JS 中，数组是一种比较特殊的数据类型，它也可以像对象那样使用键值对，但还是索引使用得多。同样，值的类型可以是任意类型。

简单地说，JSON 可以将 JavaScript 对象中表示的一组数据转换为字符串，然后就可以在网络或者程序之间轻松地传递这个字符串，并在需要的时候将它还原为各编程语言所支持的数据格式，例如在 PHP 中，可以将 JSON 还原为数组或者一个基本对象。在用到AJAX时，如果需要用到数组传值，这时就需要用JSON将数组转化为字符串。

对象是一个无序的“‘名称/值’对”集合。一个对象以{左括号开始，}右括号结束。每个“名称”后跟一个:冒号；“‘名称/值’ 对”之间使用,逗号分隔。

和普通的 JS 数组一样，JSON 表示数组的方式也是使用方括号 []。



在处理 JSON 格式的数据时，没有需要遵守的预定义的约束。所以，在同样的数据结构中，可以改变表示数据的方式，也可以使用不同方式表示同一事物。

如前面所说，除了对象和数组，你也可以简单地使用字符串或者数字等来存储简单的数据，但这样并没有多大意义。

JSON和[XML](https://baike.baidu.com/item/XML" \t "https://baike.baidu.com/item/JSON/_blank)的可读性可谓不相上下，一边是简易的语法，一边是规范的标签形式，很难分出胜负。

XML天生有很好的扩展性，JSON当然也有，没有什么是XML可以扩展而JSON却不能扩展的。不过JSON在Javascript主场作战，可以存储Javascript复合对象，有着xml不可比拟的优势。



XML有丰富的[编码工具](https://baike.baidu.com/item/%E7%BC%96%E7%A0%81%E5%B7%A5%E5%85%B7" \t "https://baike.baidu.com/item/JSON/_blank)，比如Dom4j、Dom、SAX等，JSON也有提供的工具。无工具的情况下，相信熟练的开发人员一样能很快的写出想要的xml文档和JSON[字符](https://baike.baidu.com/item/%E5%AD%97%E7%AC%A6" \t "https://baike.baidu.com/item/JSON/_blank)串，不过，xml文档要多很多结构上的字符。



XML的解析方式有两种：

一是通过文档模型解析，也就是通过父标签索引出一组标记。例如：xmlData.getElementsByTagName("tagName")，但是这样是要在预先知道文档结构的情况下使用，无法进行通用的封装。

另外一种方法是遍历节点（document 以及 childNodes）。这个可以通过[递归](https://baike.baidu.com/item/%E9%80%92%E5%BD%92" \t "https://baike.baidu.com/item/JSON/_blank)来实现，不过解析出来的数据仍旧是形式各异，往往也不能满足预先的要求。

凡是这样可扩展的结构数据解析起来一定都很困难。

JSON也同样如此。如果预先知道JSON结构的情况下，使用JSON进行数据传递简直是太美妙了，可以写出很实用美观可读性强的代码。如果你是纯粹的前台开发人员，一定会非常喜欢JSON。但是如果你是一个应用开发人员，就不是那么喜欢了，毕竟xml才是真正的结构化[标记语言](https://baike.baidu.com/item/%E6%A0%87%E8%AE%B0%E8%AF%AD%E8%A8%80" \t "https://baike.baidu.com/item/JSON/_blank)，用于进行数据传递。

而如果不知道JSON的结构而去解析JSON的话，那简直是噩梦。费时费力不说，代码也会变得冗余拖沓，得到的结果也不尽人意。但是这样也不影响众多前台开发人员选择JSON。因为json.js中的toJSONString()就可以看到JSON的字符串结构。当然对于不是经常使用这个字符串的人，这样做仍旧是噩梦。常用JSON的人看到这个字符串之后，就对JSON的结构很明了了，就更容易的操作JSON。

以上是在Javascript中仅对于数据传递的xml与JSON的解析。在Javascript地盘内，JSON毕竟是主场作战，其优势当然要远远优越于xml。如果JSON中存储Javascript复合对象，而且不知道其结构的话，我相信很多程序员也一样是哭着解析JSON的。

除了上述之外，JSON和[XML](https://baike.baidu.com/item/XML" \t "https://baike.baidu.com/item/JSON/_blank)还有另外一个很大的区别在于有效数据率。JSON作为数据包格式传输的时候具有更高的效率，这是因为JSON不像XML那样需要有严格的闭合标签，这就让有效数据量与总数据包比大大提升，从而减少同等数据流量的情况下，网络的传输压力。

可扩展标记语言 (Extensible Markup Language, XML) ，[标准通用标记语言](https://baike.baidu.com/item/%E6%A0%87%E5%87%86%E9%80%9A%E7%94%A8%E6%A0%87%E8%AE%B0%E8%AF%AD%E8%A8%80/6805073" \t "https://baike.baidu.com/item/_blank)的子集，可以用来标记数据、定义数据类型，是一种允许用户对自己的标记语言进行定义的源语言。 XML是标准通用标记语言 可扩展性良好,内容与形式分离,遵循严格的语法要求,保值性良好等优点 。

在电子计算机中，标记指[计算机](https://baike.baidu.com/item/%E8%AE%A1%E7%AE%97%E6%9C%BA/140338" \t "https://baike.baidu.com/item/_blank)所能理解的信息符号，通过此种标记，计算机之间可以处理包含各种的信息比如文章等。它可以用来标记数据、定义数据类型，是一种允许用户对自己的标记语言进行定义的源语言。 它非常适合[万维网](https://baike.baidu.com/item/%E4%B8%87%E7%BB%B4%E7%BD%91/215515" \t "https://baike.baidu.com/item/_blank)传输，提供统一的方法来[描述](https://baike.baidu.com/item/%E6%8F%8F%E8%BF%B0/8928757" \t "https://baike.baidu.com/item/_blank)和交换独立于[应用程序](https://baike.baidu.com/item/%E5%BA%94%E7%94%A8%E7%A8%8B%E5%BA%8F/5985445" \t "https://baike.baidu.com/item/_blank)或供应商的[结构化数据](https://baike.baidu.com/item/%E7%BB%93%E6%9E%84%E5%8C%96%E6%95%B0%E6%8D%AE/5910594" \t "https://baike.baidu.com/item/_blank)。是[Internet](https://baike.baidu.com/item/Internet/272794" \t "https://baike.baidu.com/item/_blank)环境中跨平台的、依赖于内容的技术，也是当今处理分布式结构信息的有效工具。早在1998年，W3C就发布了XML1.0规范，使用它来简化Internet的文档信息传输。

可扩展[标记语言](https://baike.baidu.com/item/%E6%A0%87%E8%AE%B0%E8%AF%AD%E8%A8%80/5964436" \t "https://baike.baidu.com/item/_blank)与[Access](https://baike.baidu.com/item/Access/10275" \t "https://baike.baidu.com/item/_blank)，[Oracle](https://baike.baidu.com/item/Oracle/301207" \t "https://baike.baidu.com/item/_blank)和[SQL Server](https://baike.baidu.com/item/SQL Server/245994" \t "https://baike.baidu.com/item/_blank)等数据库不同，数据库提供了更强有力的数据存储和分析能力，例如：数据索引、排序、查找、相关一致性等，可扩展[标记语言](https://baike.baidu.com/item/%E6%A0%87%E8%AE%B0%E8%AF%AD%E8%A8%80/5964436" \t "https://baike.baidu.com/item/_blank)仅仅是存储数据。事实上它与其他数据表现形式最大的不同是：可扩展标记语言极其简单，这是一个看上去有点琐细的优点，但正是这点使它与众不同。

XML的简单易于在任何应用程序中读/写数据，这使XML很快成为数据交换语言（此类语言主要包括XML，JSON等，常用于接口调用，配置文件，数据存储等场景），虽然不同的应用软件也支持其他的数据交换格式，但不久之后它们都将支持XML，那就意味着程序可以更容易的与[Windows](https://baike.baidu.com/item/Windows/165458" \t "https://baike.baidu.com/item/_blank)、[Mac OS](https://baike.baidu.com/item/Mac OS/2840867" \t "https://baike.baidu.com/item/_blank)、[Linux](https://baike.baidu.com/item/Linux/27050" \t "https://baike.baidu.com/item/_blank)以及其他平台下产生的信息结合，然后可以很容易加载XML数据到程序中并分析它，并以XML格式输出结果。

XML有两个先驱：[SGML](https://baike.baidu.com/item/SGML/2901416" \t "https://baike.baidu.com/item/_blank)和[HTML](https://baike.baidu.com/item/HTML/97049" \t "https://baike.baidu.com/item/_blank)，这两个语言都是非常成功的标记语言，但是都有一些与生俱来的缺陷。XML正是为了解决它们的不足而诞生的。

早在Web未发明之前，SGML(Standard Generalized Markup Language，标准通用标记语言)就已存在，正如它的名称所言，SGML是国际上定义电子文件结构和内容描述的标准。SGML具有非常复杂的文档结构，主要用于大量高度结构化数据的访问和其他各种工业领域，在分类和索引数据中非常有用。

虽然SGML的功能很强大，但是它不适用于Web数据描述，而且SGML软件的价格非常昂贵；另外，SGML十分庞大，既不容易学，又不容易使用，在计算机上实现也十分困难：不仅如此，几个主要的浏览器厂商都明确拒绝支持SGML，这无疑是SGML在网上传播遇到的最大障碍。鉴于这些因素，Web的发明者——欧洲核子物理研究中心的研究人员，根据当时（1989年）的[计算机技术](https://baike.baidu.com/item/%E8%AE%A1%E7%AE%97%E6%9C%BA%E6%8A%80%E6%9C%AF/1127562" \t "https://baike.baidu.com/item/_blank)，发明并推出了HTML。

1989年，HTML诞生，它抛弃了SGML复杂庞大的缺点，继承了[SGML](https://baike.baidu.com/item/SGML/2901416" \t "https://baike.baidu.com/item/_blank)的很多优点。HTML最大的特点是简单性和跨平台性。

HTML是一种界面技术，它只使用了SGML中很少的一部分标记，例如HTML 4.0中只定义了70余种标记。为了便于在计算机上实现，HTML规定的标记是固定的，即HTML语法是不可扩展的。HTML这种固定的语法使它易学易用，在计算机上开发HTML的浏览器也十分容易。正是由于HTML的简单性，使得基于HTML的Web应用得到了极大的发展。

随着Web应用的不断发展，HTML的局限性也越来越明显地显现了出来，如HTML无法描述数据、可读性差、搜索时间长等。人们又把目光转向SGML，再次改造SGML使之适应现在的网络需求。随着先辈的努力，1998年2月10日，[W3C](https://baike.baidu.com/item/W3C/216888" \t "https://baike.baidu.com/item/_blank)(World Wide Web Consortium，万维网联盟)公布XML 1.0标准，XML诞生了。

XML最初的设计目的是为了[EDI](https://baike.baidu.com/item/EDI/1049267" \t "https://baike.baidu.com/item/_blank)(Electronic Data Interchange，电子数据交换)，确切地说是为EDI提供一个标准数据格式。

当前的一些网站内容建设者们已经开始开发各种各样的XML扩展，比如数学标记语言MathML、化学标记语言CML等。此外，一些著名的IT公司，如[Oracle](https://baike.baidu.com/item/Oracle/301207" \t "https://baike.baidu.com/item/_blank)、[IBM](https://baike.baidu.com/item/IBM/9190" \t "https://baike.baidu.com/item/_blank)以及微软等都积极地投入人力与财力研发XML相关软件与服务支持，这无疑确定了XML在IT产业的重要地位。

XML具有以下特点：

(1) XML可以从HTML中分离数据。即能够在[HTML文件](https://baike.baidu.com/item/HTML%E6%96%87%E4%BB%B6/7176861" \t "https://baike.baidu.com/item/_blank)之外将数据存储在XML文档中，这样可以使开发者集中精力使用HTML做好数据的显示和布局，并确保数据改动时不会导致HTML文件也需要改动，从而方便维护页面。XML也能够将数据以“[数据岛](https://baike.baidu.com/item/%E6%95%B0%E6%8D%AE%E5%B2%9B/2897494" \t "https://baike.baidu.com/item/_blank)”的形式存储在HTML页面中，开发者依然可以把精力集中到使用HTML格式化和显示数据上。

(2) XML可用于交换数据。基于XML可以在不兼容的系统之间交换数据，计算机系统和数据库系统所存储的数据有多种形式，对于开发者来说，最耗时间的工作就是在遍布网络的系统之间交换数据。把数据转换为XML格式存储将大大减少交换数据时的复杂性，还可以使这些数据能被不同的程序读取。

(3) XML可应用于B2B中。例如在网络中交换金融信息， 目前XML正成为遍布网络的商业系统之间交换信息所使用的主要语言，许多与B2B有关的完全基于XML的应用程序正在开发中。

(4)利用XML可以共享数据。XML数据以纯文本格式存储，这使得XML更易读、更便于记录、更便于调试，使不同系统、不同程序之间的数据共享变得更加简单。

(5) XML可以充分利用数据。XML是与软件、硬件和应用程序无关的，数据可以被更多的用户、设备所利用，而不仅仅限于基于HTML标准的浏览器。其他客户端和应用程序可以把XML文档作为数据源来处理，就像操作数据库一样，XML的数据可以被各种各样的“阅读器”处理。

(6) XML可以用于创建新的语言。比如，WAP和WML语言都是由XML发展来的。WML(Wireless Markup Language，无线标记语言)是用于标识运行于手持设备上（比如手机）的Internet程序的工具，它就采用了XML的标准。

总之，XML使用一个简单而又灵活的标准格式，为基于Web的应用提供了一个描述数据和交换数据的有效手段。但是，XML并非是用来取代HTML的。HTML着重如何描述将文件显示在浏览器中，而XML与SGML相近，它着重描述如何将数据以结构化方式表示。

XML文件格式是纯文本格式，在许多方面类似于HTML，XML由XML元素组成，每个XML元素包括一个开始标记（），一个结束标记（）以及两个标记之间的内容，例如，可以将XML元素标记为价格、订单编号或名称。标记是对文档存储格式和逻辑结构的描述。在形式上，标记中可能包括注释、引用、字符数据段、起始标记、结束标记、空元素、文档类型声明( DTD)和序言。

具体规则如下：

1、必须有声明语句

XML声明是XML文档的第一句，其格式如下：



2、注意大小写

在XML文档中，大小写是有区别的。“A”和“a”是不同的标记。注意在写元素时，前后标记的大小写要保持一致。最好养成一种习惯，或者全部大写，或者全部小写，或者大写第一个字母，这样可以减少因为大小写不匹配而产生的文档错误。

3、XML文档有且只有一个根元素

良好格式的XML文档必须有一个根元素，就是紧接着声明后面建立的第一个元素，其他元素都是这个根元素的子元素，根元素完全包括文档中其他所有的元素。根元素的起始标记要放在所有其他元素的起始标记之前；根元素的结束标记要放在所有其他元素的结束标记之后。

4、属性值使用引号

在HTML代码里面，属性值可以加引号，也可以不加。但是XML规定，所有属性值必须加引号（可以是单引号，也可以是双引号，建议使用双引号），否则将被视为错误。

5、所有的标记必须有相应的结束标记

在HTML中，标记可以不成对出现，而在XML中，所有标记必须成对出现，有一个开始标记，就必须有一个结束标记，否则将被视为错误。

6、所有的空标记也必须被关闭

空标记是指标记对之间没有内容的标记，比如“”等标记。在XML中，规定所有的标记必须有结束标记。

(1)可扩展性方面：HTML不允许用户自行定义他们自己的标识或属性，而在XML中，用户能够根据需要自行定义新的标识及属性名，以便更好地从语义上修饰数据。

(2)结构性方面：HTML不支持深层的结构描述，XML的文件结构嵌套可以复杂到任意程度，能表示面向对象的等级层次。

(3)可校验性方面：HTML没有提供规范文件以支持应用软件对[HTML文件](https://baike.baidu.com/item/HTML%E6%96%87%E4%BB%B6/7176861" \t "https://baike.baidu.com/item/_blank)进行结构校验，而XML文件可以包括一个语法描述，使应用程序可以对此文件进行结构校验。

虽然XML标准本身简单，但与XML相关的标准却种类繁多，W3C制定的相关标准就有二十多个，采用XML制定的重要的[电子商务标准](https://baike.baidu.com/item/%E7%94%B5%E5%AD%90%E5%95%86%E5%8A%A1%E6%A0%87%E5%87%86/1803796" \t "https://baike.baidu.com/item/_blank)就有十多个。这一方面说明XML确实是一种非常实用的结构化通用标记语言，并且已经得到广泛应用；另一方面，这又为了解这些标准带来一定的困难，除了标准种类繁多外，标准之间通常还互相引用，特别是应用标准，它们的制定不仅仅使用的是XML标准本身，还常常用到了其他很多标准。XML标准的体系与SGML标准的体系非常相似，XML相关标准也可分为元语言标准、基础标准、应用标准三个层次。

描述的是用来描述标准的元语言。在XML标准体系中就是XML标准，是整个体系的核心，其他XML相关标准都是用它制定的或为其服务的。

这一层次的标准是为XML的进一步实用化制定的标准，规定了采用XML制定标准时的一些公用特征、方法或规则。如 XML Schema描述了更加严格地定义XML文档的方法，以便可以更自动化处理XML文档；XMLNamespace用于保证 XML DTD中名字的一致性，以便不同的DTD中的名字在需要时可以合并到一个文档中；[XSL](https://baike.baidu.com/item/XSL/838292" \t "https://baike.baidu.com/item/_blank)是描述XML文档样式与转换的一种语言；[XLink](https://baike.baidu.com/item/XLink/10832769" \t "https://baike.baidu.com/item/_blank)用来描述XML文档中的超链接；[XPointer](https://baike.baidu.com/item/XPointer/10832704" \t "https://baike.baidu.com/item/_blank)描述了定位到XML文档结构内部的方法；[DOM](https://baike.baidu.com/item/DOM/50288" \t "https://baike.baidu.com/item/_blank)定义了与平台和语言无关的接口，以便程序和脚本动态访问和修改文档内容、结构及样式等。

XML已开始被广泛接受，大量的应用标准，特别是针对因特网的应用标准，纷纷采用XML进行制定。有人甚至认为，XML标准是因特网时代的ASCII标准。在这因特网时代，几乎所有的行业领域都与因特网有关。而它们旦与因特网发生关系，都必然要有其行业标准，而这些标准往往采用XML来制定。