

# HTML

## 1. (百度)基于搜索引擎做项目的话 需要注意哪些问题

- ① 合理的 title、description、keywords：搜索对着三项的权重逐个减小，title 值强调重点即可，重要关键词出现不要超过 2 次，而且要靠前，不同页面 title 要有所不同；description 把页面内容高度概括，长度合适，不可过分堆砌关键词，不同页面 description 有所不同；keywords 列举出重要关键词即可
- ② 语义化的 HTML 代码，符合 W3C 规范：语义化代码让搜索引擎容易理解网页
- ③ 重要内容 HTML 代码放在最前：搜索引擎抓取 HTML 顺序是从上到下，有的搜索引擎对抓取长度有限制，保证重要内容一定会被抓取
- ④ 重要内容不要用 JavaScript 输出：爬虫不会执行 JavaScript 获取内容
- ⑤ 少用 iframe：搜索引擎不会抓取 iframe 中的内容
- ⑥ 非装饰性图片必须加 alt
- ⑦ 提高网站速度：网站速度是搜索引擎排序的一个重要指标
- ⑧ a 标签使用 title 属性 在不影响页面功能的情况下，可以尽量给 a 标签加上 title 属性，可以更有利于蜘蛛抓取信息。
- ⑨ 集中网站权重 由于蜘蛛分配到每个页面的权重是一定的，这些权重也将平均分配到每个 a 链接上，那么为了集中网站权重，可以使用“rel=nofollow”属性，它告诉蜘蛛无需抓取目标页,可以将权重分给其他的链接。

## 2. (百度)html 标签的语义化理解

- ① 去掉或者丢失样式的时候能够让页面呈现出清晰的结构
- ② 有利于 SEO：和搜索引擎建立良好沟通，有助于爬虫抓取更多的有效信息：爬虫依赖于标签来确定上下文和各个关键字的权重；
- ③ 方便其他设备解析（如屏幕阅读器、盲人阅读器、移动设备）以意义的方式来渲染网页；
- ④ 便于团队开发和维护，语义化更具可读性，是下一步吧网页的重要动向，遵循 W3C 标准的团队都遵循这个标准，可以减少差异化。

### 3. 介绍一下你对浏览器内核的理解

主要分成两部分：渲染引擎(layout engineer 或 Rendering Engine)和 JS 引擎。

① 渲染引擎：负责取得网页的内容（HTML、XML、图像等等）、整理讯息（例如加入 CSS 等），以及计算网页的显示方式，然后会输出至显示器或打印机。浏览器的内核的不同对于网页的语法解释会有不同，所以渲染的效果也不相同。所有网页浏览器、电子邮件客户端以及其它需要编辑、显示网络内容的应用程序都需要内核。

② JS 引擎则：解析和执行 JavaScript 来实现网页的动态效果。最开始渲染引擎和 JS 引擎并没有区分的很明确，后来 JS 引擎越来越独立，内核就倾向于只指渲染引擎。

### 4. 请描述一下 Cookies, Session Storage 和 Local Storage 的区别

Cookie 是网站为了标示用户身份而储存在用户本地终端（Client Side）上的数据（通常经过加密）。Cookie 数据始终在同源的 http 请求中携带（即使不需要），即会在浏览器和服务端间来回传递。Session Storage 和 Local Storage 不会自动把数据发给服务器，仅在本地保存。

**存储大小：**

Cookie 数据大小不能超过 4k。

Session Storage 和 Local Storage 虽然也有存储大小的限制，但比 cookie 大得多，可以达到 5M 或更大。

**有效期时间：**

Local Storage 存储持久数据，浏览器关闭后数据不丢失除非主动删除数据；

Session Storage 数据在当前浏览器窗口关闭后自动删除。

Cookie 设置的 Cookie 过期时间之前一直有效，即使窗口或浏览器关闭

### 5. iframe 有那些缺点？

① iframe 会阻塞主页面的 onload 事件；

② 搜索引擎的检索程序无法解读这种页面，不利于 SEO；

③ iframe 和主页面共享连接池，而浏览器对相同域的连接有限制，所以会影响页面的并行加载。frame 之前需要考虑这两个缺点。如果需要使用 iframe，最好是通过 JavaScript 动态给 iframe 添加 src 属性值，这样可以绕开以上两个问题。

## 6. HTML 和 XHTML 有什么区别？

① 是功能上的差别

主要是 XHTML 可兼容各大浏览器、手机以及 PDA，并且浏览器也能快速正确地编译网页

② 是书写习惯的差别

XHTML 元素必须被正确地嵌套，闭合，区分大小写，文档必须拥有根元素

## 7. DOCTYPE 作用？严格模式与混杂模式如何区分？它们有何意义？

① 声明位于 HTML 文档中的第一行，处于 标签之前。告知浏览器的解析器用什么文档标准解析这个文档。DOCTYPE 不存在或格式不正确会导致文档以兼容模式呈现。

② 标准模式的排版 和 JS 运作模式都是以该浏览器支持的最高标准运行。在兼容模式中，页面以宽松的向后兼容的方式显示，模拟老式浏览器的行为以防止站点无法工作。

## 8. 行内元素有哪些？块级元素有哪些？空(void)元素有那些？行内元素和块级元素有什么区别？

行内元素有：a b span img input select strong

块级元素有：div ul ol li dl dt dd h1 h2 h3 h4...p

空元素：<br> <hr> <img> <input> <link> <meta>

行内元素不可以设置宽高，不独占一行

块级元素可以设置宽高，独占一行

## 9. HTML 全局属性(global attribute)有哪些

class 为元素设置类标识

data-\*: 为元素增加自定义属性

draggable: 设置元素是否可拖拽

id: 元素 id，文档内唯一

lang: 元素内容的语言

style: 行内 css 样式

title: 元素相关的建议信息

页面导入样式有哪些方式及区别？

## 10. WebSocket 如何兼容低浏览器？(阿里)

Adobe Flash Socket 、ActiveX HTMLFile (IE)、基于 multipart 编码发送 XHR、基于长轮询的 XHR

## 11. 页面可见性 (Page Visibility API) 可以有哪些用途?

通过 `visibilityState` 的值检测页面当前是否可见，以及打开网页的时间等；

在页面被切换到其他后台进程的时候，自动暂停音乐或视频的播放；

## 12. ( 百度 )HTML5 audio 标签是在移动端和 pc 端什么区别

移动端浏览器大部分是禁用 video 和 audio 的 autoplay 功能，并且，很多移动浏览器也不支持首次 JavaScript 调用 play 方法进行播放

## 13. ( 百度 )HTML 中的 meta 标签都有那些

① HTML 文档的字符编码：`<meta charset="UTF-8">`

② 网页信息描述：`<meta name="keywords" content="蝉壳学院,蝉壳,前端,前端开发">`

Keywords 为页面关键字，即该网页的的关键字有哪些。

`<meta name="description" content="蝉壳学院拥有全网最豪华的教学团队，IT 培训就要选择蝉壳">`

description 为页面描述，content 写入对该页面的信息描述。

③ 浏览器的渲染方式：`<meta name="renderer" content="webkit|ie-comp|ie-stand">`

渲染速度快；兼容模式兼容问题少，适合银行政府网站。标准模式则是 IE8

## 14. ( 百度 )基于 SEO 的话，一个页面里的 h1 标签应该控制在多少个

首页最佳的 H1 标题数是出现一次，基本上都用在文章的名称或者商品的名称上，出现 H1

标题的次数过多会被搜索引擎认为是作弊而被降权，严重了还会被删除。